QUANTUM

MAY/JUNE 1998

US $7.50/CAN $10.50

Springer

GALLERY

Purchased by subscription. Courtesy, Museum of Fine Arts, Boston

The Whirlwind: Ezekiel's Vision of the Cherubim and the Eyed Wheels (Ezekiel 1, 4–28), England (c. 1803–1805) by William Blake

THIS INTERPRETATION OF EZEKIEL'S VISION takes the prophet's description literally. However, knowing the vagaries of perception, we might wonder if Ezekiel's description is more a manifestation of what he thought he saw than an accurate depiction of the actual images that focused on his retinas.

What we think we see and what we are looking at are often very different things. For example, a fishing boat captain is sure he sees a school of fish but disappointedly realizes it is the shadow of a small cloud.

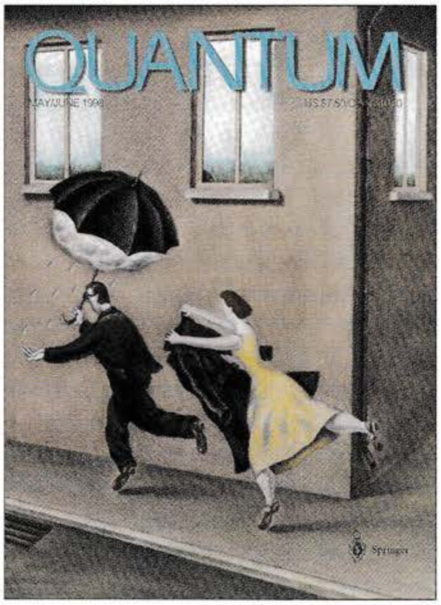
The problems of perception multiply when we encounter something unexpected or new. Those seeking a physical rather than a metaphysical explanation of Ezekiel's vision might agree with D. H. Menzel, who thinks Ezekiel saw an unfamiliar optical phenomenon and attempted to classify what he saw in terms of familiar images.

For more on Ezekiel's vision and a look at the atmospheric phenomena known as halos, turn to page 21.

QUANTUM

MAY/JUNE 1998

VOLUME 8, NUMBER 5



Cover art by Marguerite Renais

To run or not to run. That is the question we are faced with when caught in an unexpected shower.

Are you a runner, fleeing the raindrops in a panicky sprint, aghast at the prospect of getting wet? Or are you a walker, content to saunter smugly toward shelter, pleased at your ability to outwit Mother Nature?

To find out who's right, turn to page 38 for an in-depth analysis and a definitive answer.

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

Enough nerdiness

Thoughts on why the geek stereotype is so uncool

ON THE SCREEN, THE VISAGE of a lovely, purple-skinned woman, Gal 2000, the *Cassini* spacecraft computer, stares down at Dave. Her soft, sexy voice states matter-of-factly, "Analysis shows a 99.987 percent probability that the antenna will be struck and destroyed." With supreme focus and intensity, the main character, Dave, pounds his head, commanding himself to "Think! Think! Think!—One last chance!" Our hero springs into action ...

I tossed down the script and shook my head. How could this be a science education show? It didn't fit the usual mold. The focus of the show was clearly science entertainment, so shouldn't the hero be a lab-coated nerd with thick glasses? Yet some poor, misguided writer had chosen a cool character that couldn't possibly be mistaken for a geek! I racked my brain. Historically, the nerd would be the default choice, and for a very simple reason. Uh, what was that reason? I knew that the nerd image had been recognized as a barrier to achieving national science literacy, so why do we use it to promote science? Perplexed, I asked a producer of a national science television show why the geek angle is so strongly pursued. He answered simply, "It works." Hmmm. Works? Works for whom?

Unsatisfied, I continued to dig. Perhaps the nerd image benefits our tech industry—you know, the 400-billion-dollar technology industry that had 200,000 tech jobs unfilled due to a shortage of qualified college graduates. This shortage has defied the laws of supply and demand by persisting for 17 years. This shortage, according to employment experts, is due in great part to the nerd image promoted in every sector of our society (*USA Today*, 2/16/98). Well, I guess the nerd image doesn't work for industry.

How about in colleges? Maybe the nerd image serves a purpose here. According to surveys made throughout colleges, students believe that "techies are nerds" and, of course, that "no one wants to be a nerd." Contrary to the new rash of Sprite commercials, image does seem to be everything, and the image of techies—whether they are physicists, engineers, or computer scientists—is not flattering.

Well, if at first you don't succeed, look elsewhere. Perhaps the nerd image is good for young children. In a recent *CIO* magazine survey, 10-year-olds in Sherborn, Massachusetts, were asked to draw pictures of tech workers. They turned in pictures like "Dr. Vaun Squawshnut," "a bald guy with glasses, and a tie that ends halfway down his chest," and "a buck-toothed, big-nosed, high-water trouser guy with black-rimmed glasses." For a moment I thought I had succeeded: They all drew nerds. But then I read further. None of them wanted to be the nerd he or she had drawn.

Could it be that promoting the nerd image is something we should avoid like a cup of *Ebola* tea? Our nation's experts think so. They believe the nerd image will have dire consequences for the U.S. economy. Deans and CEOs throughout the country agree (*USA Today*, 2/17/98), and yet within the scientific community, we continue to promote the geek message. Have we become our own worst enemy?

Having attended a recent national meeting for our best and brightest upcoming scientists and engineers, I would have to say yes. The keynote speaker who addressed these young, vibrant kids was none other than a professional, real-life geek—or at least he plays one on TV. As I observed the crowd, I noticed that the audience clearly did not relate to this personification of their worst nightmare. Despite the speaker's antics and nerdy charm, the audience's response was polite but cold. Apparently the geek image doesn't even work for our best and brightest—they are willing to join the guild of scientists and engineers despite the negative image.

Perplexed, I investigated the National Science Education Standards (NSES), in which I found a clear statement that stereotypical images of science and scientists should be avoided. Evidently, the geek angle doesn't work for the policy makers of the science education elite either. So, it might be concluded that if "it works," the nerd image most likely works for the science illiterate and those who make a living by promoting the stereotype.

Fortunately, in my travels, I

have discovered a growing number of individuals throughout the fields of science who have reached the same conclusion. These folks have taken the directives of the NSES to heart and seek to reshape our image. Even now coalitions of Hollywood producers, TV networks, and serious science educators and scientists are attempting to improve the image of science and scientists. This group has proclaimed that geeks are out and science-literate heroes are in. In the vein of *Star Trek*, these groups are considering action/adventure productions with positive role models and science shows with brilliant, sexy hosts who can solve an integral while doing 160 in their open-topped Viper. We need these groups and they need our support.

I also became aware that Uncle Sam is providing some incentive for networks to air shows like these. The Federal Communications Commission enacted a ruling last September that mandates all networks to broadcast three hours of educational programming each week during family viewing hours. Within a historical perspective this could hardly be greeted as good news by the networks. Educational programming just hasn't supported itself with a significant market share. However, Hollywood has begun to utter the word "edutainment" with increasing regularity and seems to be willing to create a new genre of television. We should make it very clear that we encourage such TV shows.

A handsome, classic hero, huh? I picked up the script again. This Dave character is pretty cool. I wonder who they will get to play him. Hey, how about John Travolta? Yeah, right. *That* will be the day.

—Dennis R. Harp and Harry Kloor

Dennis R. Harp *(Ph.D. physics) and* **Harry Kloor** *(Ph.D.'s in physics and chemistry) are producers of the animated, science fiction education special* 2004, *starring John Travolta and James Earl Jones. Dennis is also the outreach coordinator for the Physics Department at Purdue, and Harry is a Hollywood writer for such shows as* Star Trek: Voyager, *and* Earth: Final Conflict.

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.

Quantum (ISSN 1048-8820) is published bimonthly by the National Science Teachers Association in cooperation with Springer-Verlag New York, Inc. Volume 8 (6 issues) will be published in 1997–1998. *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; electronic mail: quantum@nsta.org. **Production offices:** Springer-Verlag New York, Inc., 175 Fifth Avenue, New York NY 10010-7858.

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

Subscription Information:

North America: Student rate: $17.70; 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: $45. 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:

Advertising Representatives: (Washington) Paul Kuntzler (703) 243-7100; (New York) Brian Skepton (212) 460-1575; 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.

Is Bingo fair?

It's all in the cards

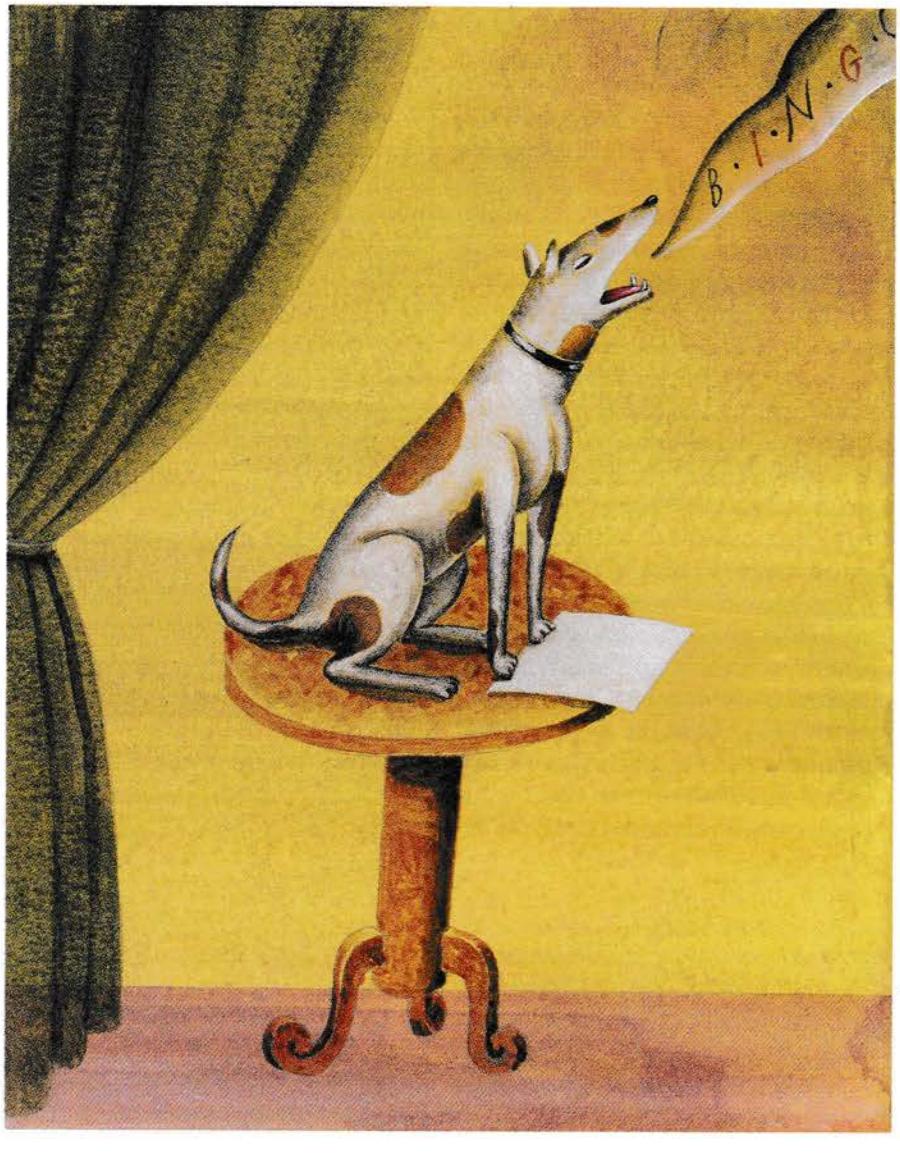
by Mark Krosky

In the game of Bingo, each player has one or more 5×5 cards. Balls numbered 1 to 75 are drawn randomly from a bin one at a time without replacement, until all the numbers in a row, column, or main diagonal on some player's card are drawn. Then, the player who holds that card calls "Bingo!" and wins.

In the event of a tie, the player who calls "Bingo!" first wins. Only numbers 1–15 appear in the first column, only 16–30 appear in the second column, 31–45 in the third, 46–60 in the fourth, and 61–75 in the fifth. When a ball is drawn, both the number of the ball and the letter of the column are read aloud, for easy reference. The center square on each card is given as a "free space." Figure 1 shows a sample Bingo card.

B	*I*	*N*	*G*	*O*
1	17	31	48	74
4	23	33	57	68
11	26	*free*	46	62
8	19	44	50	71
14	24	41	60	65

Figure 1



Art by Jose Garcia

One might not think a mathematical treatment of Bingo is possible since it appears as if a player doesn't get to make any strategic decisions. However, a player does get to make one decision: choosing a card. Of course, without any information about the other players' cards, one choice of a card is as good as another. In a Bingo parlor, players usually place their cards face up on the table without attempting to hide them, so it is reasonable to assume that players can see each other's cards. These two rules make Bingo interesting: Players can choose their cards, and they can see the other players' cards.

Rather than analyzing 5×5 Bingo, we will study a Bingo game where each card has only one winning combination, consisting of two numbers. This simplified Bingo is much easier to analyze and has many of the key properties of normal Bingo.

2×1 Bingo

In 2×1 Bingo each card has two distinct numbers on it, and a card wins when both numbers listed on that card are drawn. To simplify things further, we drop the restriction that certain numbers only appear in certain columns. A sample card is shown in figure 2.

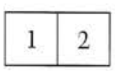


Figure 2

Problem 1. If there are 10 balls, how many possible 2×1 Bingo cards are there whose entries are in the set $\{1, ..., 10\}$?

Problem 2. If there are n balls, how many possible 2×1 Bingo cards are there whose entries are in the set $\{1, ..., n\}$?

Since there is only one winning combination per card, it is much easier to analyze 2×1 Bingo games than to analyze normal 5×5 Bingo games. As another simplification, we assume that all players are equally likely to call "Bingo!" first in the event of a tie. Thus, the tied cards will share the prize equally. For example, if two cards get a Bingo at the same time, each card gets credit for half a win.

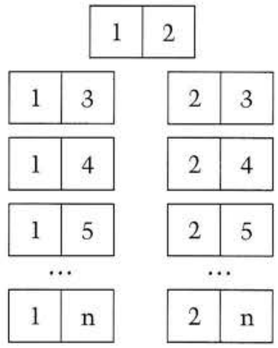


Figure 3

Consider the collection of cards in figure 3. This collection consists of every card with either a 1 or a 2 on it. There are n balls in total, and $n \geq 4$. In the following discussion, the symbol $\{a, b\}$ denotes the card with the numbers a and b on it. In our analysis, the cards $\{1, 2\}$ and $\{2, 1\}$ are considered to be the same card.

Problem 3. How many cards are there in this collection that contain the number 3?

Problem 4. How many cards are there in this collection that contain the number 1? How many cards contain 2? Any other particular number?

Problem 5. Show that there are $2n - 3$ cards in total.

Now we give each card to a different person and play. Is this a fair game? Since there are $2n - 3$ cards, if the game were fair, each player would have a winning probability of $1/(2n - 3)$. Since the numbers 1 and 2 appear on a card more frequently than any other number, the card $\{1, 2\}$ seems to be a good card to study. Let's compute the exact winning chance of this card.

We look at the first two balls drawn. There are three cases to consider: the first two balls drawn are both 1 and 2, just one of the first two balls drawn is 1 or 2, and neither of the first two balls drawn is 1 or 2.

- If the first two balls drawn are the numbers 1 and 2 in some order, then the card $\{1, 2\}$ wins.
- If exactly one of the first two balls drawn are 1 or 2, then the card $\{1, 2\}$ loses. For example, if the first two balls drawn are 1 and 3, then the card $\{1, 3\}$ wins and $\{1, 2\}$ loses.
- If neither a 1 nor a 2 are drawn among the first two balls, then as soon as either a 1 or a 2 is drawn, several cards will win simultaneously, none of which will be $\{1, 2\}$. For example, if the balls are drawn in the order 3, 5, 8, 6, 2, then the four cards $\{2, 3\}$, $\{2, 5\}$, $\{2, 6\}$, and $\{2, 8\}$ win and each card gets credit for 1/4 of a win.

So, the card $\{1, 2\}$ wins if and only if the first two balls drawn are 1 and 2.

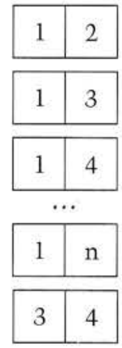


Figure 4

Problem 6. Show that the winning probability of the card $\{1, 2\}$ is $2/[n(n - 1)]$. Compare this with a fair winning probability of $1/(2n - 3)$.

Now it is natural to ask what the winning probabilities of the other cards are.

Problem 7. Which card has a greater winning probability: $\{1, 3\}$ or $\{1, 4\}$? What about any pair of cards, neither of which is $\{1, 2\}$?

Problem 8. Compute the winning probability of every card other than $\{1, 2\}$. Compare this with a fair winning probability of $1/(2n - 3)$.

Since the winning probability of the card $\{1, 2\}$ is much less than it would be in a fair game, we call this card the "big loser." Note that as the number of balls n increases, the game becomes more and more unfair for the "big loser" card.

In the above example, the "big loser" card was at an extreme disadvantage. Now we give an example (fig. 4) where one card is at an extreme advantage. This diagram shows every card with a 1 on it, along with the card $\{3, 4\}$.

Problem 9. Show that there are

$n-1$ cards containing the number 1.

Problem 10. Show that there are two cards containing the number 3. Show that there are two cards containing the number 4. Show that every number other than 1, 3, or 4 appears on only one card.

Problem 11. Show that there are n cards in total.

Ball number 1 is very important because it appears on every card but one. The two next most important balls are 3 and 4.

Problem 12. Show that no card other than {3, 4} can win until 1 is drawn.

Problem 13. Show that no card can win until 1 or both 3 and 4 are drawn. Show that if at least two balls have been drawn and either 1 or both 3 and 4 have been drawn, some card wins.

Since 1, 3, and 4 are the only balls that determine how long the game lasts, we restrict our attention to these three balls. This time, the order counts, unlike the situation in the first example.

Problem 14. Note that there are six possible permutations in which three balls can be drawn, and that the probability that 3 and 4 are drawn before 1 is 1/3.

So, if 3 and 4 are drawn before 1, then the card {3, 4} wins. Once the ball numbered 1 is drawn, then the card {3, 4} cannot win, since the cards {1, 3} and {1, 4} are held by somebody else. The card {3, 4} wins exactly when 3 and 4 are drawn before 1, which occurs with probability 1/3. If the game were fair, this card's chance of winning would be $1/n$.

Problem 15. Which card has a better probability of winning, {1, 2} or {1, 3}? Which card has a better probability of winning, {1, 2} or {1, 5}? What about other pairs of cards?

Problem 16. Compute the exact winning probability of every card in the example given in figure 4. Compare this with a fair winning probability of $1/n$. Remember that tied cards share the prize equally. The solution is difficult, because there are many cases to check.

Since the card {3, 4} has a winning probability of 1/3, we call the card {3, 4} the "big winner." Note that the value of 1/3 is independent of n. We could construct 2×1 Bingo games with arbitrarily many cards, where one card has a winning probability of 1/3.

Problem 17. Add one card to the example given in figure 3 so that there is both a "big winner" and a "big loser."

Graphing 2×1 Bingo

It is often useful to restate a problem in a different form. We can transform a collection of 2×1 Bingo cards into a graph (a collection of vertices and line segments connecting the vertices). The balls correspond to vertices, and the cards correspond to line segments. For example, the collection of cards {1, 2}, {1, 3}, {1, 4}, and {4, 5} has the graph shown in figure 5.

The graph for the example given in figure 3 is shown in figure 6, when $n = 11$.

The graph for the example given in figure 4 is shown in figure 7, when $n = 11$.

An informal analysis of these graphs helps us understand the corresponding 2×1 Bingo games. Note that in figure 6, the graph is very crowded near the "big loser" card. In figure 7, there are relatively few edges near the "big winner" card. This suggests that the "big loser" card has many potential wins stolen by other cards and that the "big winner" card has very few potential

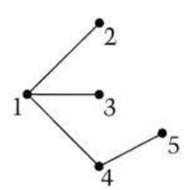


Figure 5

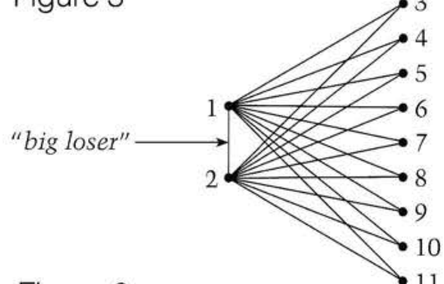


Figure 6

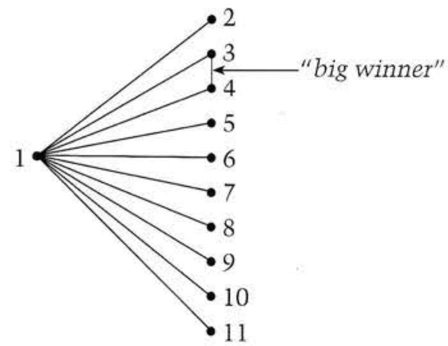


Figure 7

wins stolen by other cards. By looking at the graph, we can see which cards are likely to be at an advantage or a disadvantage.

Unfortunately, it is not always true that the best card is the card where the graph is the least crowded (that is, the card whose numbers are duplicated the least often). Consider the 2×1 Bingo game represented by the graph in figure 8.

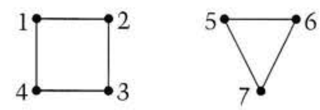


Figure 8

Problem 18. Which card has its numbers duplicated the least often?

Problem 19. What is the winning probability of each card? Hint: Three balls are usually enough to determine a winner.

We conclude that picking your card optimally is harder than just finding the card whose numbers are duplicated the least.

Analyzing the picture rather than a list of numbers helps us understand what is happening. This idea lets us relate Bingo to graph theory, an area that has already been extensively studied.

Using symmetry

Yet another benefit of viewing 2×1 Bingo as a graph is that it helps us recognize symmetry. We can use symmetry to reduce the amount of work when computing the winning chance of a card. Consider the two graphs in figure 9. Note that these two graphs are congruent. A congruence between two graphs is a way to match the vertices so that the edges also match. So, these two graphs show that cards {1, 2} and {1, 3} must have the same winning probability in the Bingo game consisting of cards {1, 2}, {1, 3}, {1, 4}, and {4, 5}.

Problem 20. Show that matching the vertices {1, 2, 3, 4, 5} to {1, 3, 2, 4, 5} gives a graph congruence between the two graphs in figure 9.

We can use graph congruences to show that two cards have the exact same winning probability. Since a graph congruence shows how to match vertices, it also shows how to match orders in which the balls can be drawn. For example, the graph congruence given above shows that cards {1, 2} and {1, 3} have the same winning probability. If the balls are drawn in the order 152, then {1, 2} wins, and in the matching order 153, {1, 3} wins. Similarly if the balls are drawn in the order 421, then card {1, 2} wins in a two-way tie, and in the matching order 431, card {1, 3} wins in a two-way tie.

Every permutation of the balls that leads to a win for card {1, 2} is matched with a permutation that leads to a win for card {1, 3}. Every permutation of the balls that leads to a two-way tie for card {1, 2} is matched with a permutation that leads to a two-way tie for card {1, 3}. So, we conclude that these two cards have the same winning chance.

Problem 21. Consider the graph in figure 10. Show that the 2×1 Bingo game represented by this graph is fair.

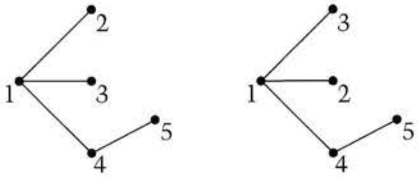


Figure 9

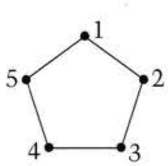


Figure 10

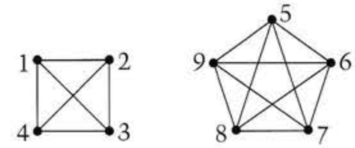


Figure 11

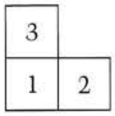


Figure 12

Problem 22. Consider the graph in figure 11. Compute the exact winning chance of each card in the corresponding 2×1 Bingo game. Hint: use symmetry and the fact that at most three balls will be drawn.

L-shaped Bingo

A 2×1 Bingo card has only one winning combination. In normal 5×5 Bingo, a card has more than one winning combination. The simplification to 2×1 Bingo ignores this property of Bingo. We can increase the complexity of our model by considering an L-shaped Bingo card, which has two winning combinations.

Figure 12 shows a sample card. The two winning combinations are the horizontal pair {1, 2} and the vertical pair {1, 3}. The diagonal consisting of 2 and 3 is not a winning combination. This is important. Without this combination, the corner number 1 plays a different role than the other two numbers. If we included the diagonal winning combination, we would add symmetry to L-shaped Bingo.

Problem 23. On the above L-shaped Bingo card, how many winning combinations contain the number 1? the number 2? the number 3?

Problem 24. On a normal 5×5 Bingo card, how many winning combinations contain the center square? the other squares on a main diagonal? the other numbers?

Problem 25. How many L-shaped Bingo cards are possible if there are five balls? If there are n balls?

Problem 26. Consider the pair of cards in figure 13. Which card has an advantage?

In normal 5×5 Bingo, people sometimes play more than one card. L-shaped Bingo can also be viewed as a 2×1 Bingo game where each player has chosen two cards that share a number.

Problem 27. Show that the L-shaped Bingo game given in figure 13 is equivalent to two players playing 2×1 Bingo where player A takes the cards {1, 2} and {1, 3} and player B takes the cards {2, 3} and {2, 4}.

Problem 28. (Proposed by Leon Harkleroad.) Consider the collection of L-shaped Bingo cards in figure 14. Show that if card A plays card B head-to-head, with card C not used, then card A has an advantage. Show that if card B plays card C head-to-head, with card A not used, then card B has an advantage. Show that if card C plays card A head-to-head, with card B not used, then card C has an advantage. Show that if all three cards are played, then the game is fair.

Problem 28 shows that the relation "has the advantage" or "beats" for L-shaped Bingo is not transitive: A beats B and B beats C, but C beats A.

The important difference between L-shaped Bingo and 2×1 Bingo is the presence of more than one winning combination per card. So, L-shaped Bingo has a greater complexity than 2×1 Bingo. L-shaped Bingo models more features of normal 5×5 Bingo. It models the presence of more than one winning combination per card and that players may take more than one card.

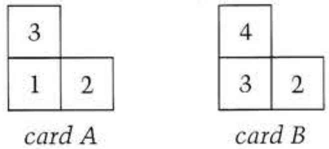


Figure 13

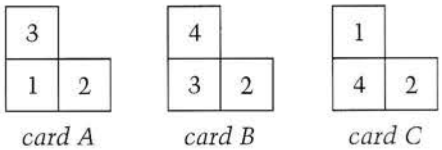


Figure 14

Normal 5×5 Bingo

Unfortunately, we cannot use the ideas contained in this paper to go to a Bingo parlor and win. Analyzing 5×5 Bingo is very hard. We can look for patterns similar to those we found in simplified Bingo, but we can't find nice results, as we could for 2×1 Bingo. Some of the similarities and differences are hinted at in the next section.

Problem 29. Try to find as many differences as possible between normal 5×5 Bingo as it is actually played and the two simplifications of Bingo discussed in this paper, 2×1 Bingo and L-shaped Bingo.

Exercises and open questions

Problem 30. Suppose seven people are playing 2×1 Bingo (the game is represented by the graph in figure 15). Another player decides to join this game, and there are two cards available from which she can

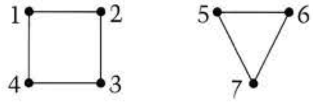


Figure 15

choose, $\{1, 3\}$ and $\{1, 5\}$. Which card should she choose? A computer would be useful to compute the exact answer, although it is possible to figure out which card is better without a computer.

Problem 31. Consider the collection of cards $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, and $\{3, 4\}$—that is, all possible cards when there are four balls. Three people play Bingo with this collection of cards, and each player takes one card. Is it possible that the resulting game is fair? What if there are four players? Five players? Six players?

Problem 32. What happens if two people try to play L-shaped Bingo with the three cards given in figure 14? Each player will change his card if it will improve his chances of winning.

Problem 33. Recall the example of an unfair 2×1 Bingo game given in figure 3. The "big loser" card had a winning probability of $2/[n(n-1)]$, and the other cards had a winning probability of

$$\frac{n+1}{2n(n-1)}.$$

Since there are $2n-3$ players, if the game were fair, each card's winning probability would be $1/(2n-3)$. Is this the most unfair 2×1 Bingo game possible involving $2n-3$ players? To answer this question, we need a formal definition of fairness. We can define fairness to be how much money a player expects to earn if he bets 1 dollar per game, as is done in *Scarne's New Complete Guide to Gambling*, by John Scarne (1986, Simon and Schuster). Do a similar analysis for the 2×1 Bingo game given in figure 4. Is this another extremal example?

Problem 34. Find as many collections of cards as possible that lead to a fair 2×1 Bingo game.

Problem 35. Find as many 2×1 Bingo games as possible where one can easily do an exact computation of the winning chance of each card. For a definition of *easy*, see *Introduction to Algorithms*, by Cormen, Leiserson, and Rivest (1986, MIT Press).

Problem 36. This is a good exercise for any variation of Bingo. Choose a collection of Bingo cards at random. What is the probability that the game will be fair? By how much do the cards' winning probabilities differ?

Problem 37. How unfair can 2×1 Bingo be if we print duplicate cards? How unfair can L-shaped Bingo be if we print cards that aren't duplicates but share a winning combination? How does this affect the analysis of normal 5×5 Bingo?

Problem 38. Viewing 2×1 Bingo as a graph was very successful. Does this work for L-shaped Bingo? Does this work for normal 5×5 Bingo?

Problem 39. In 4×1 Bingo, what would be an example that has a "big loser" card? How about a "big winner" card? How unfair can a 4×1 Bingo game be?

Problem 40. Find a "big loser" example for normal 5×5 Bingo. Find a "big winner" example. Using figure 13 as a model, find an unfair 5×5 Bingo game consisting of only two cards. Are these the most unfair examples possible?

Problem 41. Imagine that you own a Bingo parlor. How would you print the cards so that everyone had an approximately equal chance of winning? How would you print the cards if you wanted to cheat? Consider 2×1 Bingo, L-shaped Bingo, and normal 5×5 Bingo.

Problem 42. In a Bingo parlor, sometimes the goal is not covering five squares in a row. Sometimes people play to cover all four corners, which can be viewed as just 4×1 Bingo. Sometimes people play to cover all 24 numbers on a card, which is 24×1 Bingo. Forming an X or a T is another possible goal. All of these variations have something in common. Find a general model of Bingo that can handle all possible variations.

Gaming insights

Insights on many games can probably be attained using techniques similar to these. For example, Bridge with 2-card suits rather than 13-card suits may have a manageable complexity. There would be only 2,520 possible deals, and issues that are normally considered only at the championship level are easier to see. Blackjack with only Ace, 2, 3, and 4 may have a manageable enough complexity that an exact solution might be found. It may be possible to find an exact optimal strategy for Monopoly played on a board with only six squares. These ideas illustrate the value of studying a simple version of a complicated problem.

Mark Krosky *is pursuing a Ph.D. in mathematics at the University of Michigan.*

Acknowledgments

I would like to thank Robert Connelly for encouraging me to write down my thoughts on Bingo. I would also like to thank Persi Diaconis for providing some references on Bingo. I am grateful to David de la Nuez, Dan Klein, and Jeremy Bem for reading drafts and suggesting improvements.

ANSWERS, HINTS & SOLUTIONS ON PAGE 57

Just for the fun of it!

B231

Degrees of separation. What is the angle between the hands of a clock at 7:38?

B232

Victorious defeat? In a round robin soccer tournament, teams that won earned three points, teams that tied earned one point, and teams that lost earned zero points. Is it possible that the winner of this tournament could have taken last place if the points were given according to the old system (two points for a victory, one point for a draw, and zero points for a loss)?

B233

Grid kit. Consider an 8×8 grid from which one arbitrary square is missing. Show that you can dissect this grid into exactly three pieces that could be assembled to form another 8×8 grid that is missing one arbitrary square.

B234

Patch work. Mr. Wit's suit has a hole in the shape of a triangle whose sides were all different. Mr. Wit found some material and made a patch of the necessary shape. Unfortunately, he made a mistake, and the patch is turned inside out. Can he cut the patch into three parts from which to make the correct patch?

Art by Pavel Chernusky

B235

River riddle. The river near my house is shallow. I can wade across it sinking only up to my waist. The riverbed is covered with silt, and usually my feet sink in up to my ankles. Once after a storm the water rose markedly, and when I tried to wade across, the water came up to my neck. Surprisingly, my feet did not sink in the silt this time. Why?

ANSWERS, HINTS & SOLUTIONS ON PAGE 54

Turning the tides

Understanding the attraction of the Moon

by V. E. Belonuchkin

EVERYBODY KNOWS THE Moon causes the rise and fall of the tides. But how does water "feel" the Moon in the sky?

Usually we think of weight as Earth's force of gravity on a body. However, it is important to remember that any body, including the ocean, is attracted by gravity not only to Earth but also to the Moon.

But the Moon's mass is 81 times smaller than Earth's, and the ratio of the distance to the Moon to the distance to Earth's center (the Earth's radius) is about 60 to 1. Therefore, using Newton's law of universal gravitation, $F = Gm_1m_2/r^2$, the Moon's gravitational force on us (and the ocean) is almost 300,000 times less than Earth's.

Perhaps this number gives us a clue to estimating the height of the tides. But what value should this height be compared with? Two natural scales come to mind. The first is the depth of the ocean, which averages about 4 km. Dividing this depth by 300,000, we get a height of slightly more than 1 cm, which is too small for a tidal height. Another natural scale is the Earth's radius (about 6,000 km). Using this value we obtain a height of 20 m, which is too large for a tidal height.

Perhaps dimensional analysis, which has often helped us before, doesn't work in this case? Or is there more at work besides gravitation? Let's take a look at the tide-generating forces to find out.

By the way, our estimated value of 20 m is far from meaningless—it is what the tidal level would be if Earth and the Moon stopped moving. However, in this case the shape of the tide would look like figure 1, which isn't correct: There's a tidal "hump" on the side of Earth that faces the Moon and an ebb tide (a depression) on the opposite side. In reality, two high tides and two low tides take place simultaneously on Earth, and they look somewhat like figure 2 (with high tides at points A and C and low tides at points B and D).

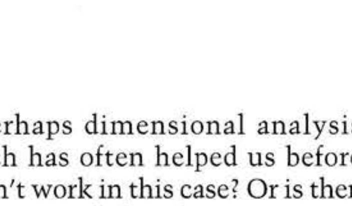


Figure 1

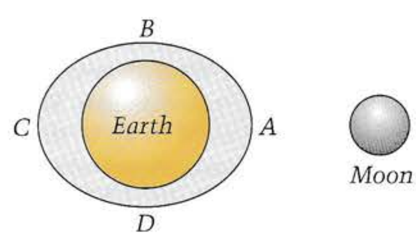


Figure 2

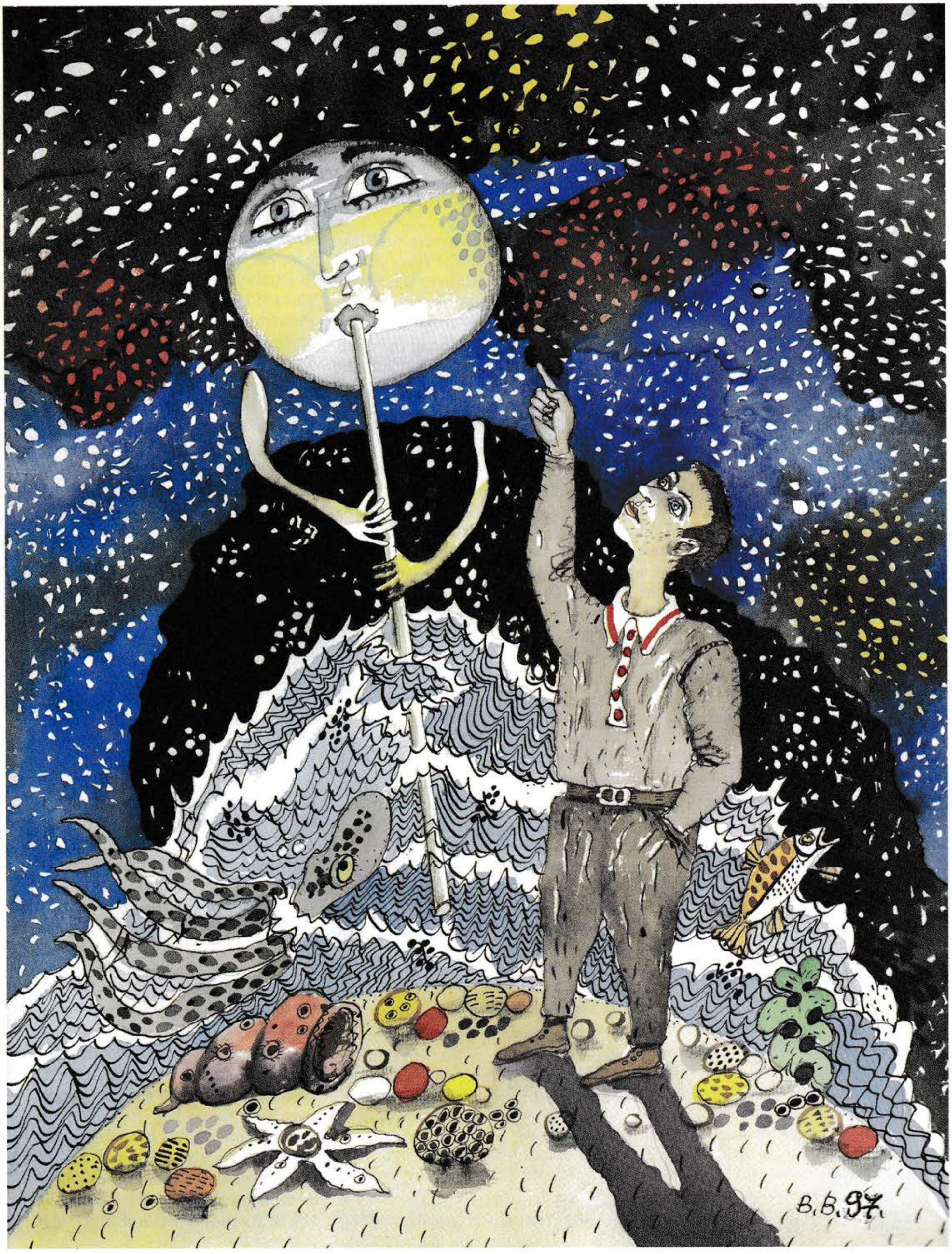
This two-tide phenomenon is explained by the fact that the areas on Earth nearer the Moon have a somewhat larger "acceleration due to the Moon's gravity." Similarly, the farthest areas of the ocean from the Moon are accelerated by the Moon to a lesser degree. The *difference* of these accelerations from the average value, the value at the Earth's center, results in the two tidal humps. Let's find this difference.

Assume the mass of the Moon to be m, and the distance from its center to Earth's center to be R. Then Earth has an acceleration toward the Moon of $a_0 = Gm/R^2$. Part of the ocean is nearer the Moon by (approximately) r (Earth's radius), and the opposite part is farther away by the same distance. The acceleration of the nearest part is $a_1 = Gm/(R-r)^2$. Therefore, the tide-generating (differential) acceleration is

$$a_t = a_1 - a_0 = \frac{Gmr(2R-r)}{R^2(R-r)^2} \cong \frac{2Gmr}{R^3}.$$

Assuming Earth's mass to be M and the acceleration due to gravity

Art by Vasily Vlasov


B.B.97

at Earth's surface to be g, we rearrange the formula to read

$$a_t = 2g\left(\frac{m}{M}\right)\left(\frac{r}{R}\right)^3.$$

In our model this is the maximum value of the tide-generating acceleration. This acceleration relates to point A in figure 2. The same value is true for point C, but in this case the tide-generating acceleration is directed away from the Moon, so it also produces a tidal hump. The middle points B and D have an acceleration equal to the average value a_0.

Substituting the ratios of m to M and r to R into the formula for a_t, we see that the maximum tidal acceleration (a_t) is less than g not by 300,000 times but by 9 million times! If we divide Earth's radius by a_t, we get what we've been looking for—the value of the tidal height, which is a little less than 3/4 m. The tidal height in the open ocean is just about this value.

But we can't give ourselves a pat on the back yet—we still haven't explained why point C in figure 2 would experience a high tide. Accurate calculations show that the absolute value of tidal acceleration is actually the same all over Earth's surface and equals

$$\frac{3}{2}g\left(\frac{m}{M}\right)\left(\frac{r}{R}\right)^3.$$

However, its direction varies as shown in figure 3. In this way, tidal acceleration produces both high tides and low tides. This result explains the shape of the ocean shown in figure 2.

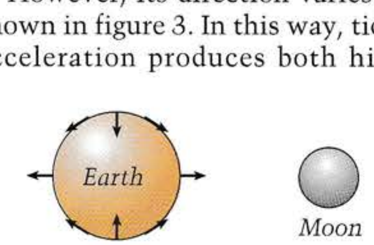


Figure 3

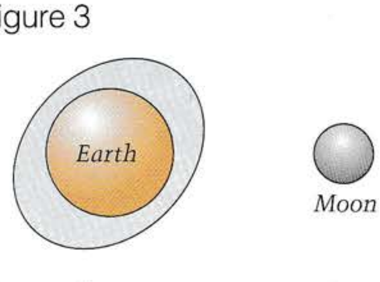


Figure 4

Alas, figure 2 is almost as wrong as figure 1! It still doesn't take into account the motions of Earth and the Moon. The true static representation of the tides is given in figure 4, which is rotated with respect to figure 2 by almost 90°!

The ocean has inertia just like any other body, so the tides should lag behind the motion of the Moon. Indeed, can the maximum height of the tide take place at point A? No, because Earth's rotation in the direction of point B, on one hand, and the run of the tides, on the other, also contribute to the elevated water surface in this region. The relation of one periodic process (Earth's rotation) to another (the tides) is described by an important characteristic of harmonic oscillations: phase.[1] So, what is the phase of the Earth's tides?

The most accurate picture of the tides is described by the dynamic model of tides. In this model the ocean's waters undergo forced oscillation. The phase of such an oscillation is determined by a relationship between the natural frequency of vibration of oceanic waters and the frequency of an external disturbance force.

An impression of the effect of the external force frequency on phase shift can be obtained by observing a swing. Let's swing it and watch how it sways back and forth when we don't interfere (when it performs its natural oscillation). As a rule, the natural period of the swing is 1 to 2 seconds. Now let's shift the swing very slowly from the equilibrium (vertical) position. For example, we'll move it to one side in 5 seconds, return it to the center in the next 5 seconds, repeat this on the other side, and so on, creating a forced period of 20 seconds. Clearly, the farther from the vertical the swing moves, the stronger the force we must apply, and this force is directed toward the side the swing is on. If we move the swing very slowly at each point along its path, we almost hold the swing against the force of gravity. This is precisely what takes place in the static model of the tides: Any deviation from the natural oscillation goes on in phase with (in the same direction as) the external force.

Quite another motion will take place if the swing is jolted so that the external force alternates its direction, say, five times per second. The swing begins to move to the left but is pushed to the right. Its direction cannot be changed immediately; it needs to first stop moving to the left, and only then will it shift to the right. However, again there is no time to increase its movement to the right because in the next instant the swing is pushed back to the left. The relationship between the phase of the swing's oscillation and the external harmonic force is not clear, but it seems that these two periodic processes are opposite in phase. (Recall that phase is the angular position of a vector that represents a harmonic oscillation, so "opposite in phase" means two vectors are opposite, which is to say their angles (phases) differ by 180°.)

Message in a jar

Now let's consider another model. Its relation to tidal forces isn't clear at first sight, but it resembles the behavior of the ocean. This model clearly shows the abrupt change in phase when the frequency of the external force gradually increases and surpasses the resonant (natural) frequency.

Pour some water into a clear jar and mark its level on the side. Suspend the jar (fig. 5) and slowly shift it from the equilibrium position in both directions, so that it swings with a frequency less than the natural frequency of such a pendulum. The water's surface will remain horizontal at all times. This causes a "high tide" on the side nearer the equilibrium position and a "low

[1] Read about waves and phases in "Wave watching," by L. Aslamazov and I. Kikoyin, in the January/February 1991 *Quantum*.

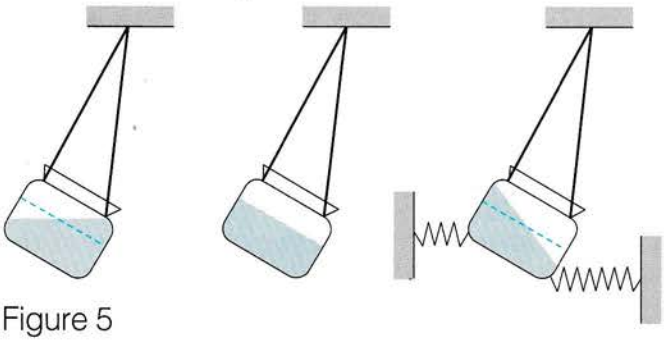


Figure 5

tide" on the opposite side.

Free the jar and let it swing by itself. After a few swings, the water's surface will rest parallel to the bottom. There will be no tides and no tide-generating force because there is no motion.

Now let's force the jar to oscillate more frequently than it "wants to." We can do this with the help of two weak springs that connect the opposing sides of the jar to anchored supports. The springs make the jar oscillate with a higher frequency than the natural frequency of the pendulum. In this periodic process water still "tries" to oscillate at the natural frequency. After a while, the natural component of the water's oscillation will be dampened, and the forced component, with the period of the oscillating jar, can be observed. This time the water's surface oscillates in a different manner: There is a high tide on the far side of the jar (with respect to the vertical position) while a low tide takes place on the nearer side. The forced increase in the water's frequency caused a 180° change in the water's phase.

Where to next? First, we need to know the natural frequency of the ocean's oscillation. Second, we should learn how a tide's phase and the tide-generating force are related. Our experiment with the swinging jar shows that the harmonic oscillations will be either in phase or opposite in phase with the external force. Which of these happens for the tides on Earth?

A natural phenomenon can help us find the answer: the tsunami. A quake occurs somewhere under the ocean. As a result, a huge mass of water swells, heaves, and generates a tsunami. It is not important that the disturbance in this example isn't generated by the Moon. What happens next? The tsunami travels across the ocean.

What is the speed of this wave? Our tried and tested tool is dimensional analysis, and it will help us again here. It looks like the wave's speed should be related to the value of the acceleration due to gravity g. Gravity does return the raised water to its resting position, after all. If we only had some natural length to multiply by g to get the square of the speed! Our only choice is the depth of the ocean. If we assume $H \cong 4$ km for the oceanic depth, the formula $v = \sqrt{gH}$ yields the correct value of 200 m/s. Therefore, if there were no continents on Earth, a tsunami would circle the planet in 56 hours. If two tsunamis on opposite sides of Earth were moving in the same direction, any place on Earth would be visited by a tsunami every 28 hours.

Why are we considering two waves? Because the Moon generates two high tides on opposite sides of Earth and drives them around the planet in a similar way. Therefore, the natural period of the tides is also 28 hours. We see the Moon reappear in the sky every 24 hours 50 minutes. You might guess that this value is determined by the daily rotation of Earth and the revolution of the Moon around it. The Moon revolves around us in the same direction as Earth's rotation. While Earth makes a full rotation in 24 hours, the Moon "runs ahead" of its previous longitudinal location, so that Earth needs an extra 50 minutes to "overtake" it each day. Therefore we know that the period of the tide-generating force is 12 hours 25 minutes because the Moon appears over each half of Earth for this duration. This period is shorter than the 28 hours needed for a tsunami to circle halfway around the Earth.

According to our jar model, the tides and the tide-generating forces should be opposite in phase because the period of the force is less than the natural period of the tides. Because 360° of harmonic tidal oscillation (a complete tidal cycle) corresponds to 180° on Earth's circumference, a phase of 180° results in a shift of 90° on Earth's surface. At the location almost under the Moon (and on the opposite side of Earth) there will be low tides, and the high tides will be at locations 90° relative to the positions of the low tides. Therefore, high tides occur at the locations where the tide-generating force drives water not from but toward Earth's center! Surprising, isn't it?

You might be wondering why we said "almost under the Moon." Forced oscillation follows the rule "in phase or opposite phase" only in idealized systems, which have no friction, and thus, no energy dissipation. If friction is almost absent, the forced oscillation will be almost in phase or almost opposite in phase.

There is another important characteristic of harmonic oscillation: amplitude. If the frequency of the external force is much less than the natural frequency of oscillation, the amplitude is simply equal to the maximum static deviation caused by this force. At the other extreme, if the frequency of the external force is much larger than the natural frequency, we can estimate the amplitude in the following way.

The acceleration of a vibrating body is proportional to the amplitude and square of the frequency (this follows from dimensional analysis). If the amplitude of the external force is constant, the amplitude of the harmonic oscillation is inversely proportional to the square of the frequency. Consequently, at high frequencies the amplitude quickly decreases.

The most interesting phenomena take place not with extreme frequencies but with ones that often take place in real life. When the frequency of the external disturbing force approaches the natural frequency of an

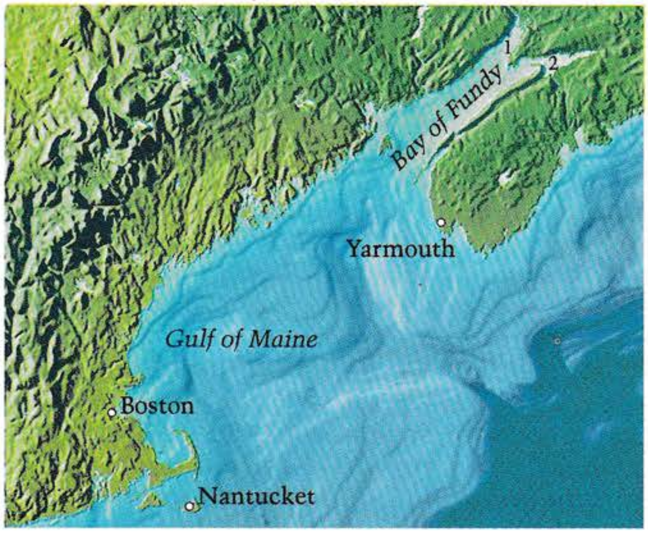


Figure 6

oscillating system, resonance occurs, which means an increase in the amplitude to infinity in systems without friction (energy dissipation). Therefore, we should think carefully about resonance phenomena when designing oscillating systems.[2]

The importance of resonance in extremely high tides can be illustrated by tidal power stations. An oceanic tide of the standard 1 m height flows through the 400 km gap between Nantucket, Massachusetts, and Yarmouth, Nova Scotia (fig. 6). After gaining speed and a 4 m height in the narrow Bay of Fundy, the tide bursts into tiny Chignecto Bay and Minas Basin (locations 1 and 2 in figure 6). In the latter the tidal amplitude is the largest on Earth—as high as 6 m. The difference in high and low tide water levels is 12 m. Wouldn't this be an ideal place to build a tidal power plant?

One project proposed to separate Shepody Bay (0.2% of the Gulf of Maine's area) from Chignecto Bay by a dam. Another project planned to partition off the Minas Basin, which would mean cutting off more than 0.5% of the Gulf of Maine.

A power station takes energy from the tides, so the tides would decrease in the first case by 20–25 cm and in the second case by 30–35 cm.[3] How would this affect the waters in the Bay of Fundy or in the entire Gulf of Maine? Distributing the water "deficiency" along these spacious bays, we get the following estimates: The Bay of Fundy would lose 0.2 mm in the first case and about 1 cm in the second project. In the Gulf of Maine we would need a microscope to detect the change in water level. However, these estimates don't take into account all the changes that would occur.

Most importantly, the resonant frequency of the water oscillations in the Gulf of Maine would change. Instead of 13 hours 20 minutes, the period would be only 13 hours 10 minutes. This value is dangerously close to the period of the tide-generating force (12 hours 25 minutes), so the entire system becomes "more resonant." In addition, in the second project a big change in the oscillation would be made by a change in the character of water flow at the neck of the Minas Basin, which results from the change in friction.

Precise calculations show that the dam in Shepody Bay would not change the tidal height in the Bay of Fundy, but would augment the tide in the Gulf of Maine by 3–4 cm, or more than 2% of the present amplitude. In the second project the tide all along the water surface from Boston to the Minas Basin would increase by 15–20 cm, which is more than 10%! This would necessitate new moorings and disturb the migratory movements of fishes, among other important consequences. This shows that even projects that are ecologically safer than thermal or nuclear power stations should be thoroughly planned to diminish their harmful effects on nature.

[2]See "The Horrors of Resonance," by A. Stasenko in the March/April 1998 *Quantum.*

[3]See the paper of D. A. Greenberg in the November 1987 *Scientific American*, vol. 257, no. 5.

HOW DO YOU FIGURE?

Challenges in physics and math

Math

M231

Sum qualification. A set of 1998 different natural numbers is given. None of these numbers can be represented as a sum of two other numbers from the set. What is the least possible value of the greatest number from such a set? (V. Y. Protasov)

M232

Angling for answers. Point M is taken inside a parallelogram $ABCD$. We know $\angle MBC = 20°$, $\angle MCB = 50°$, $\angle MDA = 70°$, and $\angle MAD = 40°$. Find the angles of the parallelogram. (M. Volchkevich)

M233

Fractional neighbors. Let all the regular irreducible fractions whose denominators don't exceed 99 be written in order of ascending value. Between which fractions does 5/8 stand? (D. I. Averianov)

M234

Sine language. Draw the set of points $M(x, y)$ on the plane whose coordinates satisfy the equation

$$\sin x \cos^2 y + \sin y \cos^2 x = 0.$$

(S. B. Gashkov)

M235

Projecting area. Find the greatest possible value of the area of an orthogonal projection of a cylinder with radius r and altitude h on a plane. (M. Volchkevich)

Physics

P231

Around the corner. A circle of maximum diameter is cut from a uniform square with side d. Where is the center of mass of an individual corner piece located? The center of mass of a semicircle of radius R is located $a = 4R/(3\pi)$ from its diameter. (A. Zilberman)

P232

Gunfire racing. A projectile was fired horizontally from a mountain at an altitude h = 1 km with a velocity v = 500 m/s. After the time t_0 = 1 s, another shell was fired in pursuit of the first. What must the minimal initial velocity of the second shell be and at what angle should it be fired in order to hit the first shell? (V. Nikiforov)

P233

Cryogenic wind tunnel. To cool the air stream in a cylindrical tube under normal conditions, identical drops of liquid nitrogen are injected in some sections of the tube, and these drops evaporate downstream. The speeds of the gas and the drops are equal everywhere (their initial values are v_0 = 10 m/s), and there is no heat transfer through the tunnel's wall. Find the speed, density, and temperature of the stream after all the drops evaporate if their initial mass flow equals that of the air. The boiling point of nitrogen at atmospheric pressure is 77 K, and its latent heat of vaporization is $L = 2 \cdot 10^5$ J/kg. Assume the properties of gaseous air and nitrogen to be identical. (A. Stasenko)

P234

Capacitor in an electric field. An amount of work W was performed in placing a charged parallel-plate capacitor into a uniform electric field oriented as shown in figure 1. By performing work W_2, the capacitor was then turned through an angle α. Using the known value of angle α, find the ratio W_2/W_1. Consider the capacitor's own field to be uniform and entirely enclosed inside the capacitor. (V. Mozhayev)

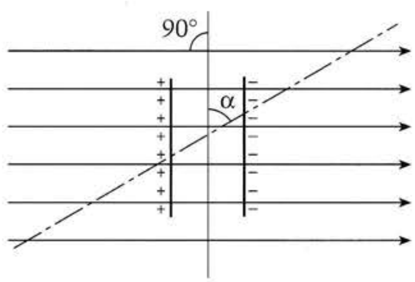


Figure 1

P235

See spot focus. A plano-convex lens is made of glass with refractive index n = 1.6. The radius of the spherical surface is R = 10 cm, and the thickness of the lens is d = 0.2 cm. A parallel beam of light is directed at the planar surface of the lens parallel to its principal axis. All but a small central part of the lens is covered by a diaphragm, and the light is focused on a screen. Then the diaphragm is taken away. Find the diameter of the spot on the screen. (A. Zilberman)

ANSWERS, HINTS & SOLUTIONS ON PAGE 50

Tied into knot theory

Unraveling the basics of mathematical knots

by O. Viro

WHEN MATHEMATICIANS try to study an everyday phenomenon, they usually replace it with a convenient mathematical object. Ordinary knots, made of rope, have shared this fate. They've turned into *mathematical* knots.

One question that prompted mathematicians to study knots was which knots could be undone without cutting the rope and which could not.

The prototype of a mathematical knot is a twisted piece of rope with fixed ends. The ends of the rope must remain stationary because if we could manipulate them, it would be possible to undo each knot by moving one of the ends of the rope through the loops.

Another, more convenient way to spare ourselves trouble with the ends of the rope is to connect the ends. This operation is shown in figure 1. As a result, the rope turns into a twisted ring.

The next step toward the mathematical knot is replacing the rope with a line (the axis of the rope).

While every rope can be modeled by a line along its axis, not every line can be regarded as the axis of a rope of constant thickness. The knots get to small for the thickness of the rope. For example, a curve with an infinite sequence of infinitely diminishing knots (fig. 2) cannot be "inflated" to make a rope with a finite thickness. Curves of this sort (with infinitely many knots on them) are also studied in mathematics. They are called *wild knots*.

We'll not discuss wild knots. To exclude them from our considerations, let's adopt the following definition: A *tame knot* is any closed, connected polygonal path in space that has no self-intersections and is comprised of a finite number of line segments.

This definition corresponds well to our intuitive concept of knots. Every tame knot can be inflated to make a rope ring, and every rope ring has an axis made of a finite number of segments.

The word *connected* in the definition of a tame knot means that the polygonal path cannot be represented as a union of several closed

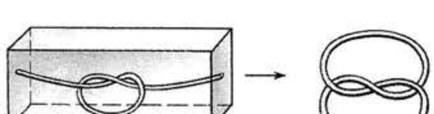


Figure 1

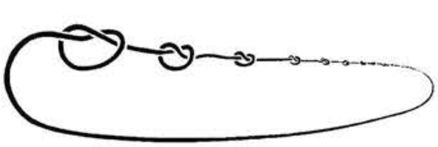


Figure 2

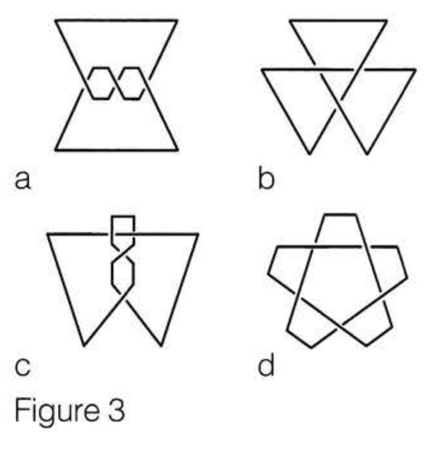


Figure 3

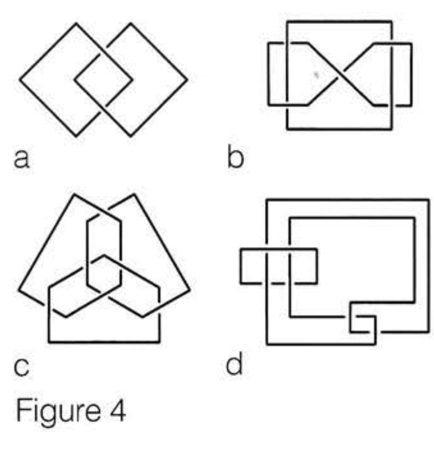
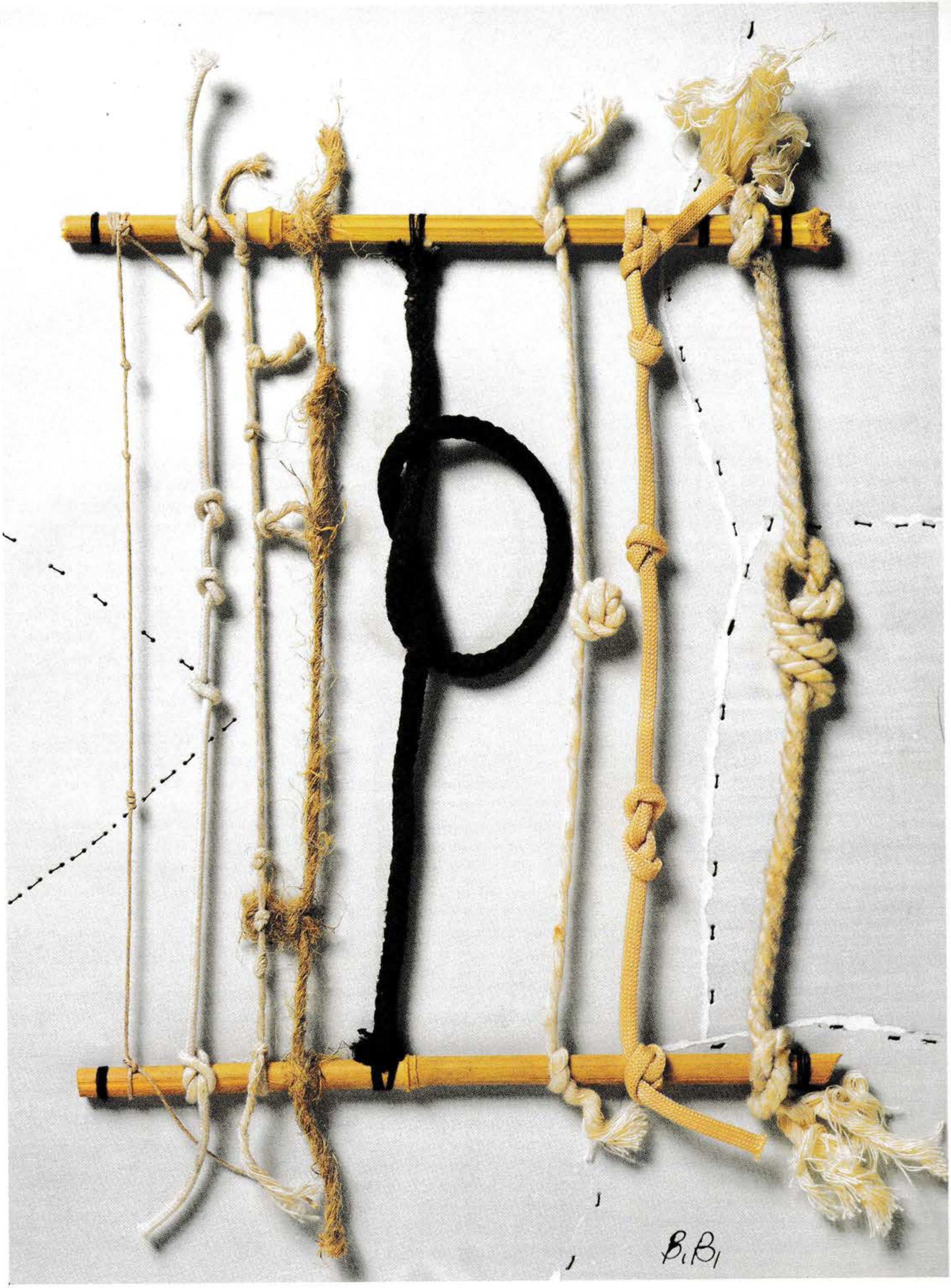


Figure 4

Art by Vasily Vlasov


B.B.

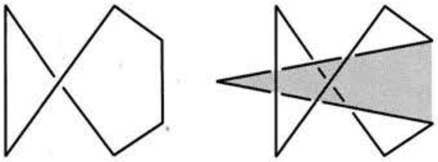


Figure 5

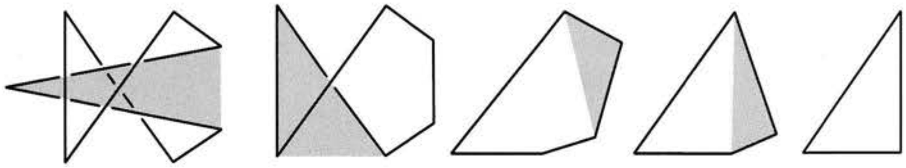
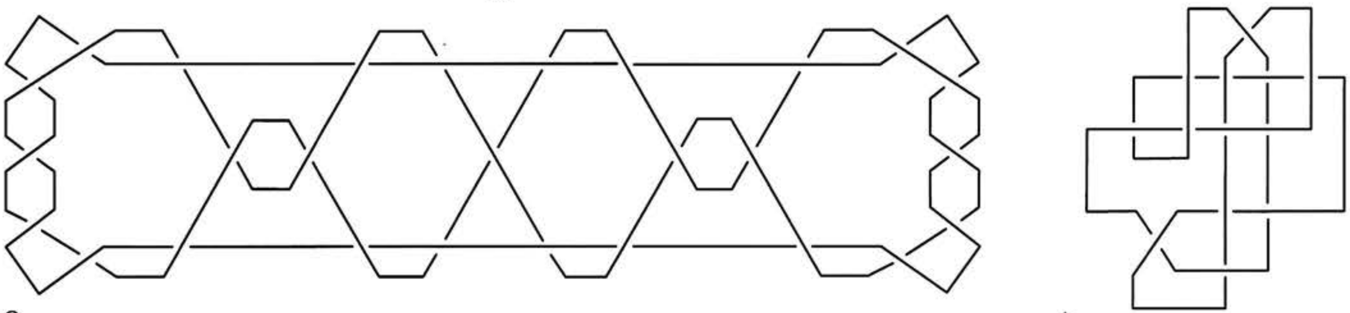


Figure 6



a

b

Figure 7

polygonal paths. If we do not insist on the connectedness of a knot, we obtain the definition of a *tame link*. Figure 3 shows some tame knots, and figure 4 shows some tame links, which are not knots. (Since all the knots and links considered here will be tame, we'll omit this word from now on.)

Drawing knots

To draw a knot, we start by choosing a viewing point. This point must satisfy the following two conditions: First, the knot should lie on one side of a plane passing through this point. Second, from this point, no three segments of the knot should look as if they are passing through one point.

We can meet the first condition if we choose a point far enough from the knot. Then we can always choose a plane such that the knot lies on one side of this plane. Then we can project the knot onto this plane. To meet the second condition, we must find such a point that each line through it meets no more than two segments of the knot. We can satisfy this condition by an arbitrarily small shift in the viewing point.

At any place where the images of two nonintersecting segments meet, we must show which of them passes nearer to the viewer. For this purpose, the image of the farther segment is interrupted, as shown in figures 3 and 4. A picture made in accordance with all these rules is called a *diagram* of the knot.

Equivalent knots

It is natural to consider as *equivalent* those knots that can be transformed into one another through twisting, stretching, shrinking, or any other continuous deformation. This intuitive concept of equivalence becomes, when we consider tame knots, the notion of *isotopes*: Two knots (or, more generally, two links) are called *isotopic* if we can pass from one of them to the other by a sequence of transformations called elementary isotopes.

An *elementary isotope* of a knot is either

(1) a substitution of one of its sides for two new segments such that these three segments form the border of a triangle that intersects the original knot only in the substituted side, or

(2) the inverse operation: substituting two adjacent sides of the knot for a new segment such that the three segments form the outline of a triangle that intersects the knot only in the substituted sides.

Figure 5 shows a pair of knots, each transformed into the other by an elementary isotope. The triangle that appears in the definition of an elementary isotope is called the *trajectory* of the deformation.

A knot that is isotopic to the outline of a triangle is a *trivial knot*. The knots in figure 5 are trivial knots. A sequence of elementary isotopes that transforms them into the outline of a triangle is presented in figure 6.

It's clear that the outlines of any two triangles are isotopic, and thus, all trivial knots are isotopic. However, some trivial knots look very complicated.

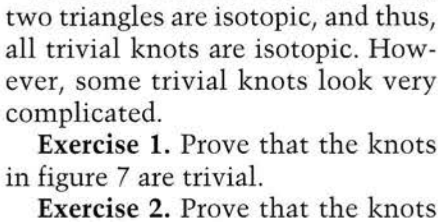
Exercise 1. Prove that the knots in figure 7 are trivial.

Exercise 2. Prove that the knots in figures 3a and 3b are isotopic.

Proving that knots are not isotopic

The definition of *isotope* immediately suggests the following problem: Are there knots that are not isotopic? Until we have answered this question, we may have the idea that all knots are essentially the same.

We know that very intricate knots can prove to be trivial. On the other hand, if we try to undo the knot in figure 8a, called the *trefoil knot*, we



a

b

Figure 8

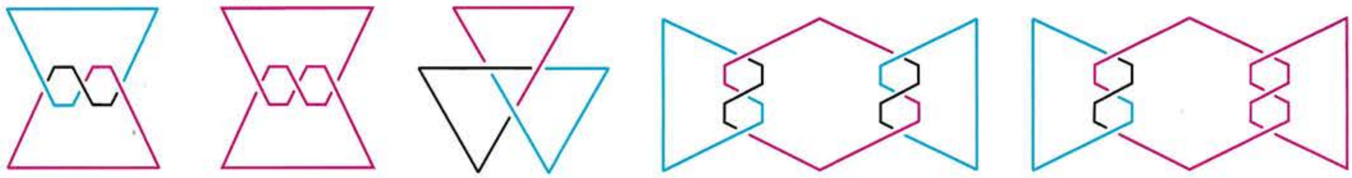


Figure 9

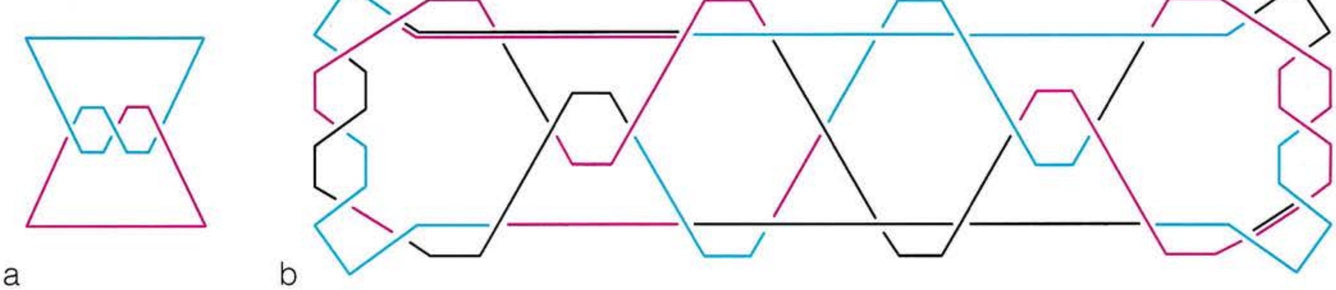


Figure 10

soon realize that the task is impossible. But how can we prove that the knot in picture 8b is nontrivial?

Remember that a knot is considered nontrivial if it is impossible to turn it into the outline of a triangle using elementary isotopes. The intricacy of the trivial knots in figure 7 shows how difficult it may be to recognize such knots. The only way to prove that a knot is nontrivial is to find some property of knots that is not changed by the elementary isotopes, and that the considered knot possesses, but that the outline of a triangle does not.

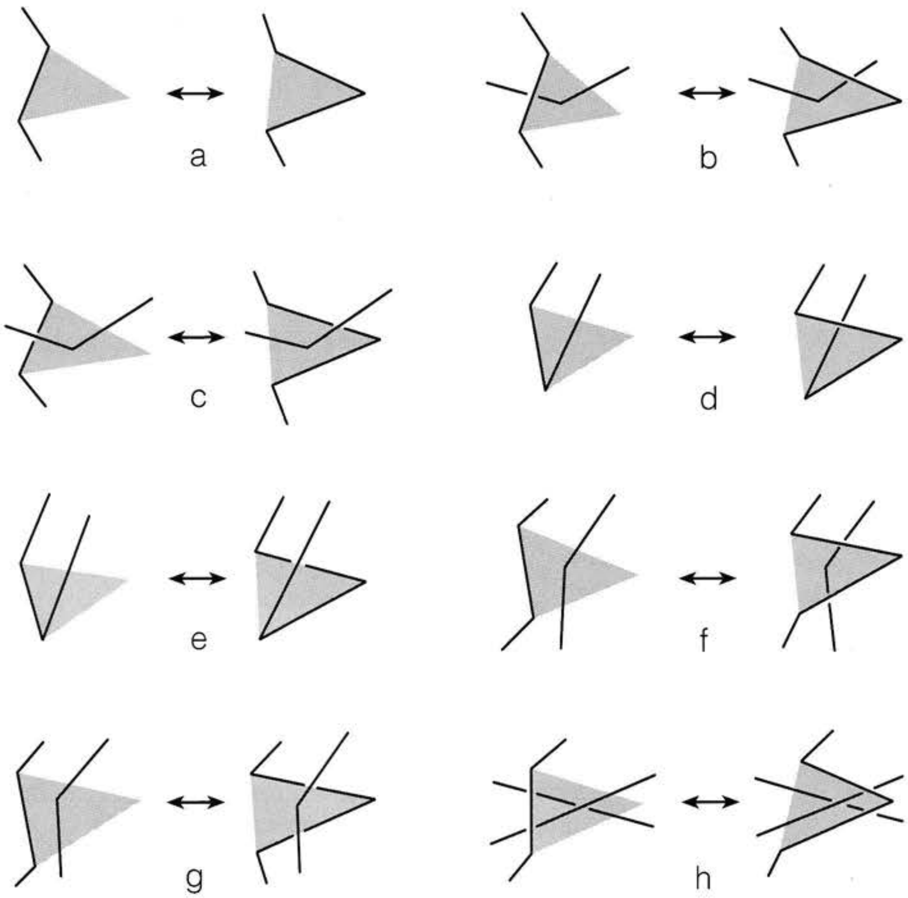


Figure 11

A property of a knot that is unchanged by elementary isotopes is called an *invariant*. Invariants of knots are the main focus of knot theory. As an example of an invariant, we can consider the triviality of a knot. It is easy to define many other invariants. For example, the fewest possible sides and fewest possible crossings on the diagrams of knots isotopic to a given knot are invariants of the knot. However, introducing these and many other analogous invariants not only fails to help answer our old questions but also poses several new questions. As a matter of fact, when we try to calculate these invariants, we run across the same difficulties we meet when we try to prove the nontriviality of a knot.

What may help prove nontriviality is an invariant that is easy to calculate and that is defined by a diagram of the specific knot and not by a whole class of isotopic knots. We will now introduce such an invariant.

Coloring knot diagrams

Remember that the diagram of a knot is simply the knot's image on a plane such that the images of any three of its sides do not pass through one point, and such that when the images of nonintersecting sides meet, the image of the farther one is interrupted. Because of these inter-

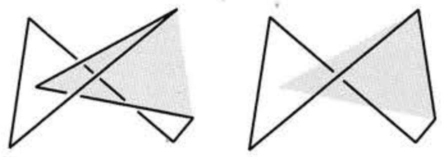


Figure 12

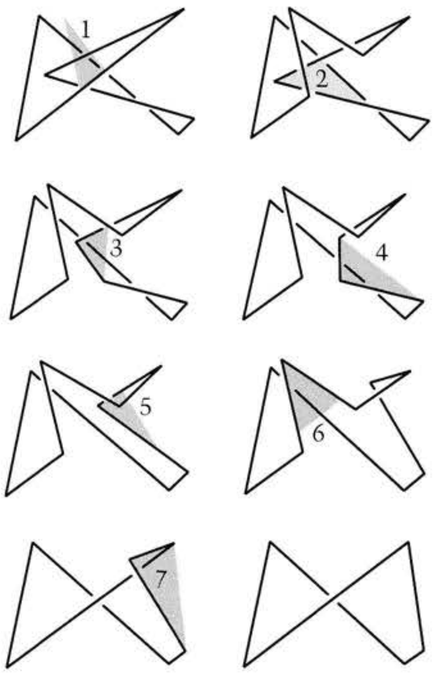


Figure 13

ruptions, the diagram of a knot consists of several nonintersecting polygonal paths, which we'll call *sections* of the diagram. The number of sections is equal to the number of interruptions (crossings).

To prove that two knots are not isotopic, it is convenient to color their diagrams in accordance with certain rules.[1] We say that the coloring of a diagram in three colors is *correct* if each section of the knot is one color, and for every crossing, either all the sections that meet are the same color or all three colors appear near it. Several examples of correct coloring are shown in figure 9, and figure 10 contains several examples of incorrect coloring.

Exercise 3. Prove that it is impossible to color the diagrams of the knots in figures 6 and 7 if all three colors must be used.

Exercise 4. Show that there are nine different correct colorings of the trefoil knot (fig. 8b).

Theorem 1. (This is the main theorem.) The number of correct three-color colorings of a diagram is an invariant of a knot.

Before we try to prove theorem 1, let's discuss some of its applications. Here is its simplest corollary:

Corollary 1. If the diagram of a knot can be colored correctly using all three colors, then the knot is not trivial.

This corollary immediately allows us to obtain the result formulated in exercise 3 (since the knots that appear there are trivial). But much more important is the fact that the nontriviality of the trefoil and many other knots follows directly from it.

Exercise 5. Prove (with the help of theorem 1) that the trefoil is not isotopic to the knots in figures 3c and 3d, and that the knot in figure 9d is not isotopic to either the trefoil or these two knots.

The knots in figures 3c and 3d are nontrivial. However, it impossible to prove this by correctly coloring them in three colors. In fact, theorem 1 gives no universal way to prove that knots are nonisotopic, although it allows us to prove it in many particular cases.

Exercise 6. Think of an infinite sequence of knots, each with a different number of correct three-color colorings. (According to theorem 1, any two knots from this sequence will be nonisotopic.)

Theorem 1 remains correct if we replace the word *knot* with the word *link*. Note: If the diagram of a link allows no correct coloring with more than one color in it, then the link can't be unlinked (that is, it is not isotopic to a link made of knots lying on different sides of a plane).

Exercise 7. Prove that the links in figure 4 cannot be unlinked.

Proving theorem 1

Theorem 1 is implied by the following two theorems, whose proofs are cumbersome but not too difficult.

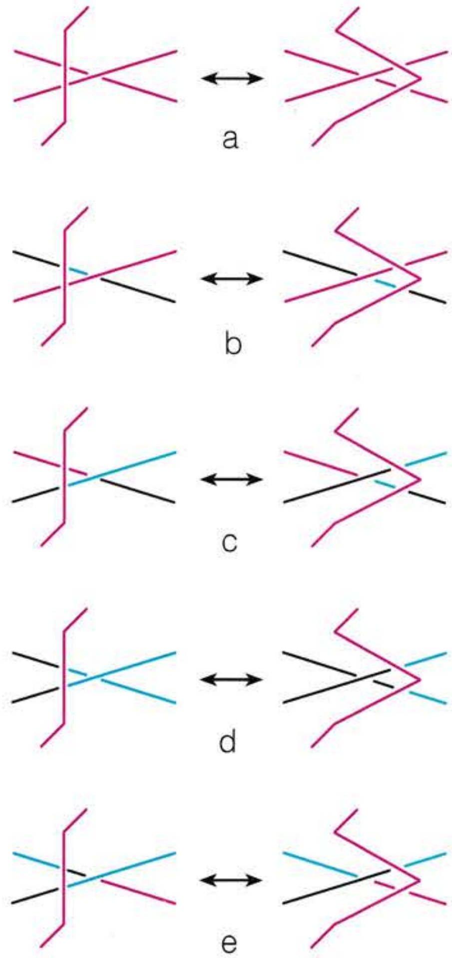


Figure 14

Theorem 2. Any elementary isotope of a link can be replaced with a finite sequence of elementary isotopes so that each of them changes the diagram of the link in one of the ways shown in figure 11.

For example, the elementary isotope in figure 12 can be replaced by the sequence of elementary isotopes in figure 13 (the numbers in fig. 13 explain the order of operations).

Theorem 3. There is a one-to-one correspondence between the correct three-color colorings of any two diagrams, obtained from one another by the operations shown in figure 11, such that the unchanged parts in the corresponding colorings are colored the same.

This statement is evident for all the transformations shown in figure 11 except for 11h. In the latter case, we derive the proof from the drawings in figure 14.

[1]The idea of coloring the diagrams of knots is due to the American mathematician Ralph Fox. See his article "Metacyclic invariance of knots and links" in the *Canadian Journal of Mathematics*, 1980, 2(22), 193–201.

Visionary science

Halos, sun dogs, and other optical phenomena

by V. Novoseltzev

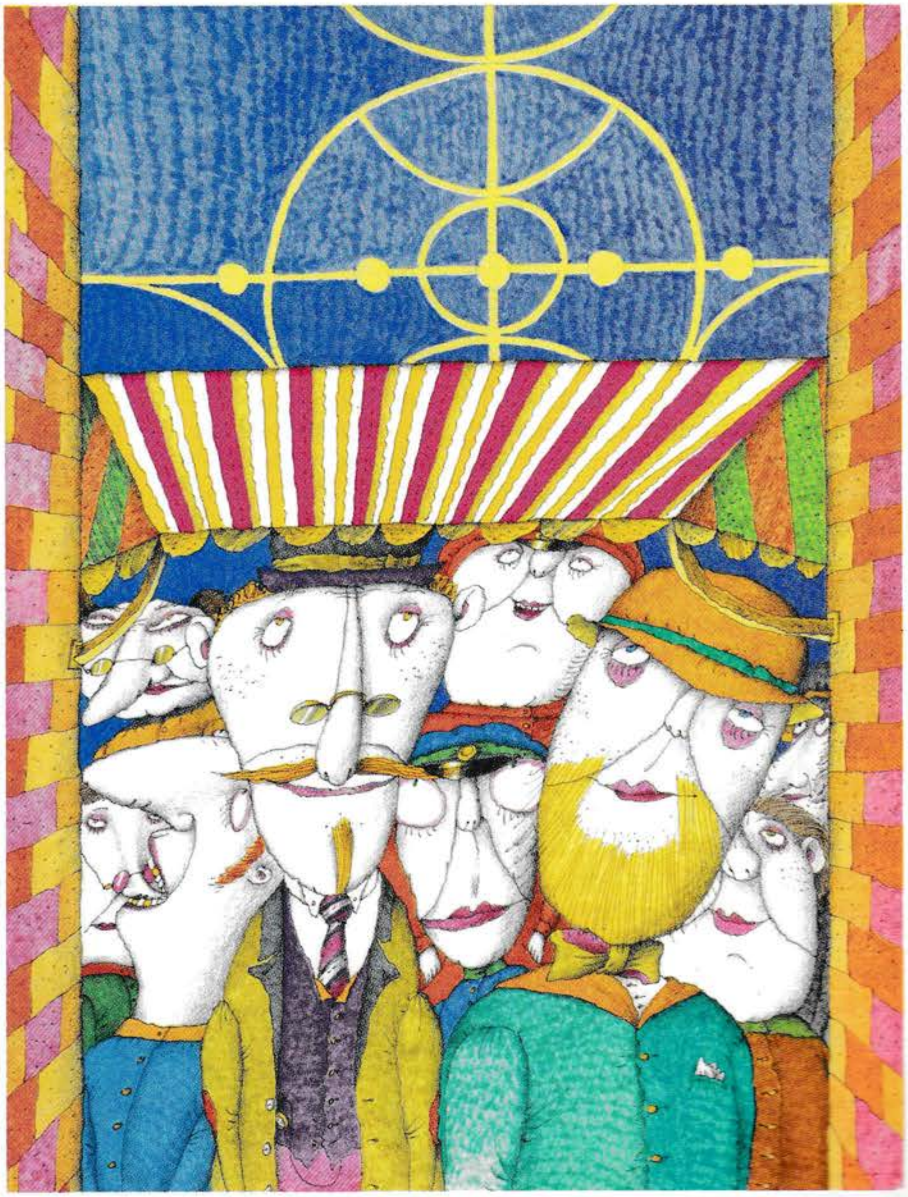
In his famous book *Atmosphere,* the great French naturalist and astronomer Camille Flammarion (1842–1925) described for the first time all the unusual natural phenomena in Earth's atmosphere and explained them to the general public. Since then descriptions of mirages, solar halos, and other optical phenomena have become familiar.

Since Flammarion's time, industry and technology have changed the face of our planet, and new technologies have caused variations of natural atmospheric phenomena. It is not a rare occasion when a phenomenon becomes unrecognizable and turns into an "unidentified object."

The honor of the first observation of a technological variant of an atmospheric phenomenon was also Flammarion's. Flying in a hot-air balloon, he was the first to observe the colorful halo around the shadow cast on the clouds by his aircraft.

Halos

We will describe in detail only one of the atmospheric optical phenomena—the halo. Halos (from the Greek *halos*—a circle, disk) are symmetrical figures—the circles, arcs, and spots located near a bright source of light (usually near the Sun or Moon)—and are observable in cold, clear weather.



Art by V. Ivanyuk

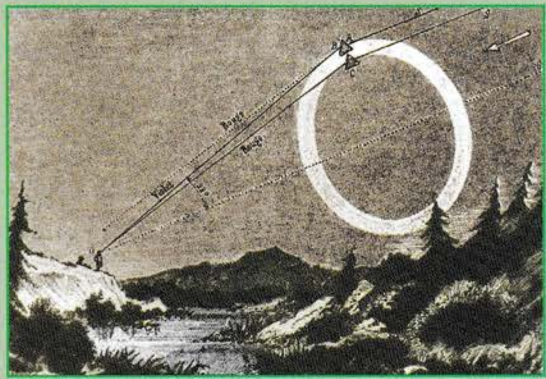


Figure 1. *The appearance of a halo. An illustration from C. Flammarion's* Atmosphere.

Halos are caused by light refracting through or reflecting off the small ice crystals suspended in the air. If the orientations of the crystals are randomized by air currents and evenly distributed, circles appear near a source of light, as illustrated by Flammarion in his book (fig. 1). If the air is still, then the flat, hexagonal type of ice crystals fall with a horizontal orientation, tilting from side to side. In this case light reflects off the flat surfaces, which results in a light pillar near the light source (fig. 2). In addition, the horizontal faces of the crystals sometimes cast horizontal bars of light, though they are rather weak. Therefore, cruciform figures can appear in the sky.

If there are few ice crystals in the air, "ears" and sun dogs appear at the sides of the Sun. If there are many crystals in the sky, the intensity of the refracted and reflected light will be large enough to form crosses, circles, and even spaced repetitions of these figures. A rather complicated pattern may be produced in this case, with a number of sun dogs in the horizontal plane (fig. 3).

As for the rainbow-colored circles observed by Flammarion from his hot-air balloon, they are formed by the refraction and diffraction of light in the suspended water droplets. These kinds of halos are always round because all the droplets have a simple spherical form.

So, there are a number of different atmospheric phenomena, and a careful observer can sometimes enjoy splendid views—especially in very cold weather. However, what the observer sees when observing an atmospheric phenomenon depends not only on the weather.

A matter of perception

How do we see? More specifically, how do we perceive and recognize surrounding objects? When an object appears in our field of vision, its image is projected on the retina, where nerve endings are excited. A complicated pattern of electrical impulses is then fed to the brain and analyzed.

The human brain is constructed such that it tries to associate a new image with one of the classes of objects it already knows. It seeks and finds a logical, noncontradictory explanation of every detail. Moreover, the interpretation of the image must be consistent with the observer's world view.

The same phenomenon observed by someone in medieval times and someone today will be interpreted on the basis of quite different notions about the world. On this problem, A. S. Gurvich, a specialist on optical phenomena, writes: "Mirages in the sky were often identified as air ships. But what kind of ships? In the Middle Ages they were sailing ships with anchors. After dirigibles were invented, new details were noted by eyewitnesses, who included portholes, searchlights, and landing gear in their descriptions."

There are various names for the underlying fabric upon which the brain strings together details to produce a final picture. Sometimes it's simply an "idea" of the perceived world—its model. Most students have come across the planetary model of the atom proposed by Rutherford: electrons spinning in varying orbits around a nucleus. And the students probably say to themselves, "What's so new about this? It's obvious." But when the data about the structure and properties of atoms were unclear and incomplete, it was indeed the insight of a genius who was able to "see" such a model.

There is a very apt term in the theory of vision: *object hypothesis*. It refers to the basic (a priori) notions of human beings about what we observe, which in every

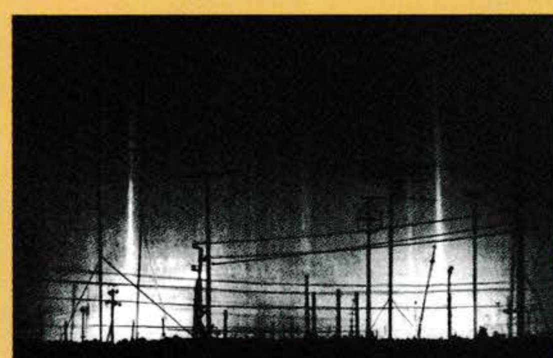


Figure 2. *Light pillars resulting from the refraction of light from street lamps on the tiny ice crystals suspended in air.*

Figure 3. *Halo and sun dogs.*

situation provides us an opportunity to quickly and surely determine what we see. When a completed object hypothesis for an observable phenomenon exists in the brain, all the details of the perceived image easily take their places and form an integral picture. On the contrary, when the brain has no suitable template, it has difficulty composing the picture.

Here is an instructive example. Nowadays people are familiar with the physical features of Saturn. Looking at it through a telescope, everyone sees a globe surrounded by a ring. However, depending upon the planet's orbital location, this ring will be positioned at different angles to the observer, producing different images.

In the seventeenth century astronomers did not know that Saturn has rings. Many times when they scrutinized it through telescopes, they didn't observe the ring. Galilei himself, who in 1610 discovered what we now call the ring, described the planet as a "triple star": "The middle star seems to be rather large, and the two others, located one in the East and another in the West, probably touch it. It looks like two servants helping old Saturn to tread its path." Thus, even with a scientific mind, the absence of a suitable template in the brain prevented the synthesis of a true picture of the planet.

Unknown or unrecognized?

The same problem arises during the observation of many anomalous atmospheric phenomena. Having no suitable model for the observed phenomenon, the brain subconsciously searches for the most closely related object hypotheses and then provides the consciousness with an interpretation based on the most appropriate template.

Records of anomalous atmospheric phenomena ("visions") can be found in the most ancient historical sources—on the pages of the Bible, Egyptian papyri, Chinese chronicles, and Russian chronicles.

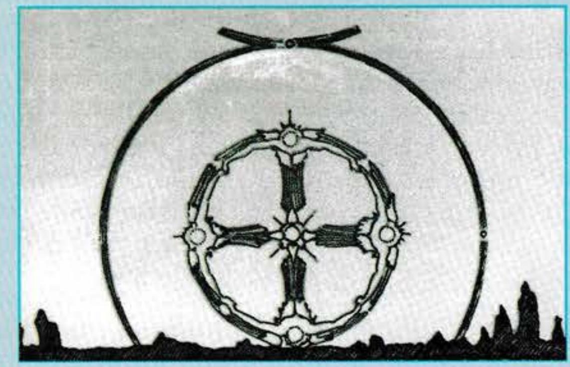


Figure 4. *A reconstruction by D. H. Menzel of the halo observed and described by Ezekiel.*

Even in modern times some stories of this kind appear in newspapers. Analyzing these wonderful and colorful descriptions is an intriguing process.

Unfortunately, it is often impossible to examine or reconstruct the details of historical or even recent events that are needed to identify the phenomenon. It's not surprising that many "wonders" are simple in nature, but occur in the "wrong place" or at the "wrong time."

A vision of the prophet Ezekiel

One of the most ancient descriptions of an unidentified atmospheric phenomenon can be found in the Bible. It is the famous wheels of Ezekiel (Ezekiel's first vision of God, Ezekiel 1:4–1:27):

"As I looked, behold, a stormy wind came out of the north, and a great cloud, with brightness round about it, and fire flashing forth continually, and in the midst of the fire, as it were gleaming bronze. And from the midst of it came the likeness of four living creatures. And this was their appearance: they had the form of men, but each had four faces, and each of them had four wings. Their legs were straight, and the soles of their feet were like the sole of a calf's foot; and they sparkled like burnished bronze. Under their wings on their four sides they had human hands. And the four had their faces and their wings thus: their wings touched one another; they went every one straight forward, without turning as they went....

"Now as I looked at the living creatures, I saw a wheel upon the earth beside the living creatures, one for each of the four of them. As for the appearance of the wheels and their construction: their appearance was like the gleaming of a chrysolite; and the four had the same likeness, their construction being as it were a wheel within a wheel. When they went, they went in any of their four directions without turning as they went. The four wheels had rims and they had spokes; and their rims were full of eyes round about....

"And above the firmament over their heads there was the likeness of a throne, in appearance like sapphire; and seated above the likeness of a throne was a likeness as it were of a human form. And upward from what had the appearance of his loins I saw as it were gleaming bronze, like the appearance of fire enclosed round about; and downward from what had the appearance of his loins I saw as it were the appearance of fire, and there was brightness round about him. Like the appearance of the bow that is in the cloud on the day of rain, so was the appearance of the brightness round about."

What was this miracle? The first scientific interpretation of Ezekiel's wheels was given by the American astrophysicist Donald Howard Menzel, known not only for his scientific research but mainly for his book about "flying saucers." He considered the phenomenon described by Ezekiel to be a solar halo (a very rare phenomenon at southern latitudes).

Ezekiel's description inspired William Blake, a famous English poet and artist and the contemporary of the French Encyclopedists, to picture it in one of his woodcuts (see inside front cover). In the central human figure with four faces, one can easily perceive the halo's cross, and the contours of the wings mimic the structure of rings. Figure 4 shows the reconstruction of Ezekiel's wheels made by Menzel.

Halos in Russian chronicles

Halos were also known to Russian bookworms. In the beginning of the twentieth century, the famous Russian astronomer D. O. Svjatsky published a picture borrowed from an ancient chronicle that showed what kind of object hypotheses existed in an ancient Russian mind (figure 5). The world of crosses and crowns was much more familiar than the modern world of physical laws.

In the Vologodskaya chronicle of the year 1171 (this year lasted from September 1, 1662 to August 31, 1663) one can read: "On November 29 after sunset there was a sign, awful and terrible, occurring at Belozersky Uezd in Cyril's district on Erge: A starlike apparition, bright and long, came and ran away with the swiftness of lightning. And the heavens gaped and the vision shone for half an hour, an indescribable bright light like fire. There stood a man in that fire, his hands and feet spread, fire around him; the air was clean and very cold." Then "a great stonefall occurred from heaven to the ground with fury and noise, and the stones were enveloped with heavenly fire and deeply pierced the frozen ground."

It turned out that this case was very well known to astronomers—D. O. Svjatsky classified it as fragments of a large bolide. The cited description had come from a detailed account given by a priest from near the town of Ustjug. This and many other corroborative sources show that the bolide flew over Vologodskaya territory for more than 200 kilometers from west to east. Those who observed the fiery trace of the approaching bolide head-on saw a stable fiery point or disk. In the clear, frosty evening this light source produced a halo. Curiously, the description of this halo as a huge human apparition with spread hands has certain features in common with the watercolor of William Blake (see front inside cover) and the drawing of a Russian chronicler.

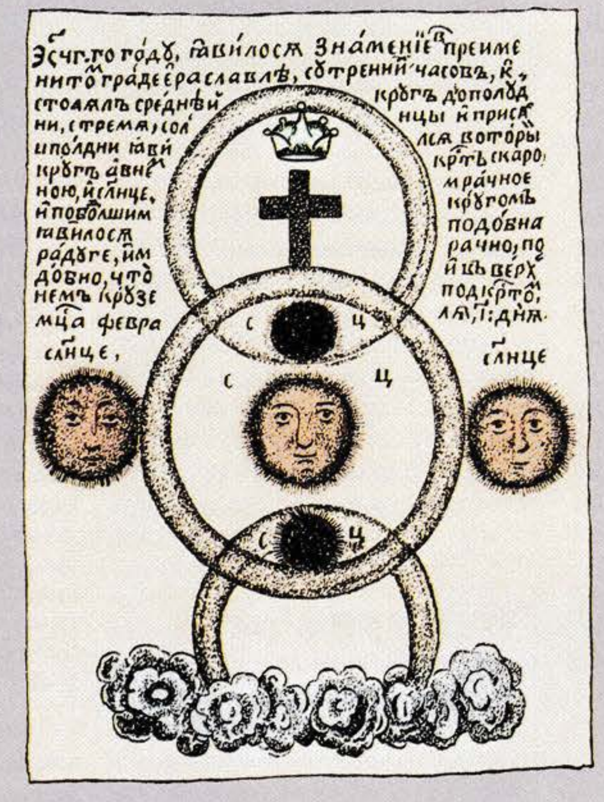


Figure 5. *Picture of a halo from one of the last Russian chronicles.*

Union Jack over Moscow

It seems that we have finished the story of transforming halos into ghosts and supreme beings. However, there is another source of halos—human technology. Sometimes such phenomena are listed among the unidentified. For example, in the book of S. Shulman, the appearance of "extraterrestrials over Russia" is recounted.

On May 15, 1981, at 1:15 A.M. a bright, spherical object was over Tula (a Russian town), moving in the direction of Moscow. Having approached Vnukovo airport, it hovered over it in the presence of hundreds of witnesses. The object was motionless for about one minute, then "a dazzling white lightning" broke away from its center and surrounded it with something like a solar crown, which quickly disintegrated into a cascade of sparks that "looked like fireworks."

When the sparks died out, "a black square appeared at that place, which was then crossed with some luminescent bars. Inside this square the bars formed something like a huge cross. Taken together, all this looked like a British flag.... Then the initial object withdrew, but the black square with the 'flag' kept its place for some time. Several minutes passed, and it became dim and disintegrated. Apparently it wasn't material—that is, it must have been an optical phenomenon because people saw stars through it."

Figure 6 shows sketches of the fiery globe and "British flag" over Vnukovo made by one of the witnesses. What was it? It seems that the first part of the observations is not disputable: People saw a multistage rocket launched hundreds of kilometers away from Vnukovo. The bright, motionless object is the description of the rocket's exhaust illuminated by the Sun, which was below the horizon. The rocket was moving directly away from the observers. When the engines of the next stage ignited, "a dazzling white lightning broke away" from the circle's center.

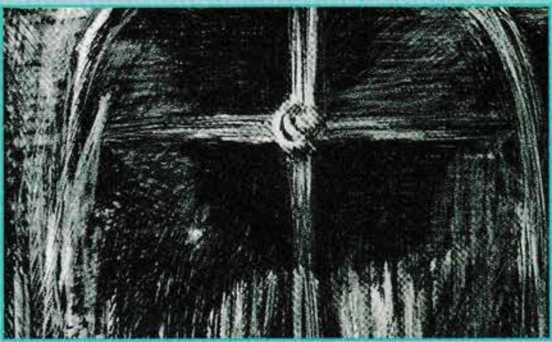
The separation of a rocket's stages is a complicated technical operation accompanied by bursts of charges that blow off the fixtures connecting the stages. So, the "solar crown disintegrating into a cascade of sparks like fireworks" would be better substituted for a more strict comparison with the bursts and sparks of melted metal produced during electric welding.

However, the "black square"

crossed by luminescent bars is indeed a unique phenomenon. It seems to be the first description of a halo caused by a humanmade light source. On the lower sketch in figure 6 we can see both the characteristic cross and the circle. The halo was observed during the several minutes of the firing of the rocket's second-stage engines. Instead of solar light, the bright point of light from the rocket's engine was refracted in the suspended crystals.

The only question to be explained is how the cloud of ice crystals could be formed on a warm May night. The combustion products of rocket fuel contain aqueous vapors. In addition, crystals can be formed due to the supercooling and crystallization of metal oxides (say, Al_2O_3), which are also present in the combustion products.

In the absence of an object hypothesis for the natural origin of the observed phenomenon, it was not possible for witnesses to recognize it, though there were conjectures about its artificial character.



a

b

Figure 6. *Halos observed over Vnukovo Airport, Moscow.*

Something to see besides halos

We considered only one example of the transformation of a classical atmospheric phenomenon into an unidentified miracle (the "artificial" halo over Vnukovo). However, there are many other examples of unrecognized and unidentified phenomena.

Some of these miracles are quite evident—for example, unusual clouds. We know that a comet's tail extends to infinity, but only its head is usually seen, where the gas density is higher and where more solar light is scattered. In the same way the rocket engine's exhaust can be seen only when it is "rather dense" near the light. If such an exhaust occurs between the Sun and an observer, it may look like a dark cloud. However, rockets have engines with round nozzles, so the cloud it produces will always have a regular form.

Moreover, an observer (located, say, aboard a plane) moving along or across the "invisible" part of a rocket's exhaust will see wonderful images depending upon the position of the Sun: a sudden appearance of a dark or fiery sphere or the transformation of this sphere into an elliptical saucer followed by its "traceless disappearance." All of these phenomena are illusions caused by human technologies that by and by become recognizable.

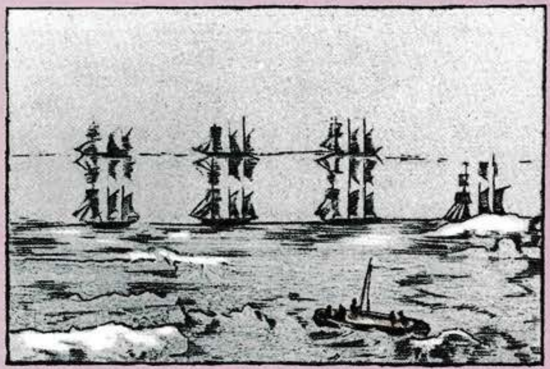


Figure 7. *The historically famous lateral mirage observed in 1869 by captain Coldway who visited the shores of Greenland aboard the ship* Germany.

Still there are some absolutely enigmatic (unidentified) phenomena that have no variations caused by human technologies. One of them is a mirage. Natural mirages are usually vertical—that is, they are located above and below the actual object. Such mirages are very well known and explained: They appear due to temperature differences between adjacent air layers that are parallel to the Earth's surface.

It is possible, however, that temperature differences are not the only causes of natural mirages. Indeed, many careful observers have seen complicated mirages in the Arctic consisting not only of vertical images but also of lateral ones.

Figure 7 shows a single ship near the coast of Greenland that produced two series of reflections—the vertical and the lateral. It is very difficult to imagine the simultaneous existence of two natural air layers having different temperatures and crossing at 90°. Therefore, this lateral mirage should have some other explanation, some other nonthermal source of the atmospheric optical heterogeneity. Perhaps it was caused by electromagnetic anomalies. Is it a coincidence that the lateral mirages were reported to occur in polar regions where the auroras (which are electromagnetic phenomena) are seen? The author couldn't find a discussion of the possible mechanisms of this phenomenon or even mention of this problem in the literature.

AT THE BLACKBOARD I

Math relay races

On your mark, get set, solve

by Don Barry

THE AMERICAN REGIONS Mathematics League (ARML) runs an annual contest among teams of 15 students representing regions that range from an entire state to a county and from a city to a single school. Approximately 1400 students from 35 states (and often guests from foreign countries) take part in the daylong competition.

There are four parts to the contest: Team, Power, Individual, and Relay. In the Team round the 15 students on a team work together to solve 10 problems in 20 minutes. Only the answer is scored. In the Power round the team has 60 minutes to write solutions or proofs to questions that explore an interesting topic. In the Individual round each student receives four pairs of questions and is given 10 minutes per pair.

The final round, the Relay Race, being rather novel, requires a longer explanation. In the Relay Race, each team is divided up into five groups of three. Each first person receives the same problem, each second the same, and each third the same. The second person's problem requires the first person's answer, and the third person's problem requires the second person's answer. Students cannot talk with one another and can pass back only unadorned answers. Each group of three receives four points for a correct answer in three minutes and two points for a correct answer in six minutes. When a team gets all three problem correct, there are high five's all around, but a wrong answer from either the first or second person will certainly provoke all sorts of frenetic activity and strange mathematics.

Here is a relay race from the 1996 ARML. Note that TNYWR stands for "The Number You Will Receive" and that solutions to all the relays can be found on page 55.

1. In trapezoid $ABCD$, $\overline{AB} \parallel \overline{CD}$ and angles B and D are acute. If $\sin\angle B = \cos\angle D$, compute $m\angle A - m\angle B$ in degrees.

2. Let T = TNYWR and set $K = T/15$. Square $LMNP$ is inscribed in right triangle ABC as shown in figure 1. If $PN = K$, compute the product $(AP)(NC)$.

3. Let T = TNYWR and set $K = T/6$. There are K positive integers in an arithmetic progression with common difference 1. If the sum of the terms is 75, compute the value of the first term.

The art of writing relays

A good relay should involve a variety of types of problems as well as various forms of mathematical thinking. The best ones provide the second and third students the opportunity to do quite a bit of analysis before they receive an answer. Writing a good relay is challenging, and I've found that it makes a great activity for a math club. Here is a favorite written by one of my students, Creence Lin.

1. A girl divided her long hair into three parts, wrapped a red ribbon around the left part, a white ribbon around the center part, and a blue ribbon around the right part. She formed a braid by bringing the left part over the center part, thereby forming braid #1 with ribbons in the following order from left to right: white, red, blue. Then she brought the right part over the center part, forming braid #2 with ribbons in the following order: white, blue, red. If the girl continued in this fashion, what would be the number of the first braid that returns to the original order of red, white, blue?

2. Let n = TNYWR. Figure 2 shows a strip of paper of width n that is formed into a circle. The strip is cut along dotted lines, some of which are shown, and the resulting six pieces are pasted onto a square as shown in figure 3. Determine the

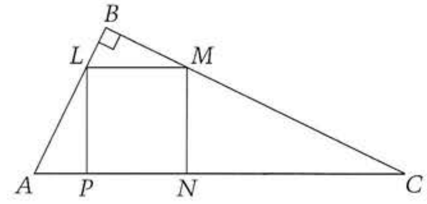


Figure 1

Figure 2

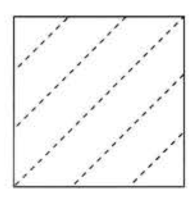


Figure 3

original circumference of the strip.

3. Let k = TNYWR and let $m = k - 100$. An aged piece of tofu is five times as old as a regular piece of tofu was when the aged tofu was as old as the regular tofu is now. The sum of their ages is m. How old is the regular tofu?

Super relay

For the 1996 ARML we thought it would be fun to challenge the students with a 15-question, full-team relay at the end of the competition. We didn't want to count the ARML Super Relay as part of the contest, but we wanted to provide an interesting activity while students were waiting for the award ceremony to start. The problems had to be easier than a normal relay, but because no scores were being kept, we felt free to put in a few tricks. By all accounts the students loved the challenge. Here are the Super Relay problems.

1. If the area of a triangle with base $2N$ and height $N - 2$ is equal to N, compute N.

2. Let T = TNYWR. Compute the slope of the line passing through $(T, 1)$ and $(1, T^2)$.

3. Let T = TNYWR. If

$$y = Tx^2 + T^2x + C$$

and the x-coordinate of the vertex equals C, compute the value of C.

4. Let T = TNYWR. If

$$(2 + Ti)^2 = a + bi,$$

compute the value of $a + b$.

5. Let T = TNYWR. Pass back the digit in the unit's place of the product $199^T \cdot 279^{T-1}$.

6. Let T = TNYWR. If

$$T^3 - 3T^2 + 3T - 1 = K^3,$$

compute the value of K.

7. Let T = TNYWR. If T is the number of sides of a regular polygon, compute in degrees the positive difference between the sum of the interior angles and the sum of the exterior angles.

8. Let T = TNYWR. Let $K = T/60$. If K is the height of an equilateral triangle, let the area of the triangle be A. Compute $A\sqrt{3}/9$.

9. Let T = TNYWR. If

$$\sin T^\circ \cos 286^\circ - \cos T^\circ \sin(-106^\circ) = \sin\theta$$

for $0^\circ \le \theta \le 180^\circ$, then if $i = \sqrt{-1}$, compute $\cos\theta + i\sin\theta$.

10. Let T = TNYWR. Compute

$$\left(\frac{|T|}{T}\right)^2.$$

11. Let T = TNYWR. Compute the value of x satisfying the following equation:

$$2x - \frac{3x - T}{2} = 7 - \frac{T + x}{5}.$$

12. Let T = TNYWR. If $\tan 9T = \cot\theta$ for $-90^\circ < \theta^\circ < 90^\circ$, compute θ.

13. Let T = TNYWR. A line passes through the point $(-2, 9)$ with slope T. Compute the x-intercept of the line.

14. Let T = TNYWR. If the ordered pair (x, y) is the solution to the system below, compute $x - y$:

$$\begin{cases} x + y = T + 7 \\ Tx - 4y = T - 8. \end{cases}$$

15. Let T = TNYWR. A square floor is covered with square tile. If the number of tiles in the two diagonals is $2T^2 + 5$, compute the number of tiles on the floor. ◘

Don Barry *is the chair of the ARML problem-writing committee and teaches mathematics at Phillips Academy in Andover, Massachusetts.*

ANSWERS, HINTS & SOLUTIONS ON PAGE 55

Want to know more about the ARML competition?
Write or e-mail

Barbara Rockow, ARML Secretary
Bronx High School of Science
75 West 205th Street
Bronx, NY 10468
73047.3156@compuserve.com.

PHYSICS CONTEST

Depth of knowledge

"Adaptability is not imitation. It means power of resistance and assimilation."
—Mahatma Gandhi

by Arthur Eisenkraft and Larry D. Kirkpatrick

IT IS A COMMON SIGHT DURing the summer months to see a car crowned with camping gear, canoes, or beach chairs as the family heads off on vacation. Such a sight causes automobile engineers responsible for aerodynamics to convulse.

Millions of dollars of research and experimentation have been invested in sleek car designs to minimize the adverse affects of air resistance. And here is the American household negating all efforts to maximize gas mileage with a bramble bush of outdoor gear strapped to the roof.

A home study of air resistance requires only a stopwatch and some coffee filters. A single coffee filter is dropped from a selected height, and its descent time is recorded. The experiment is repeated with two nested filters, three nested filters, and so on. In this way, we can find the relationship between mass and descent time. Since the filters fall with constant velocity, having reached a terminal velocity quite quickly, this is a measure of the effect of mass on terminal velocity as well. (Proving that terminal velocity is reached quickly is an important digression.)

A second experiment can be conducted to determine the effect of surface area on terminal velocity. Two coffee filters can be taped side by side, and their descent time can be compared with the time for two nested filters. This is followed by three taped filters and three nested filters. Continuing in this manner, the desired relationship is derived.

Theoretically, we can look at the effect of air resistance on an air cart moving horizontally with a piece of cardboard providing a resistive force. The cart would continue to move forward at constant velocity without this cardboard obstruction. For slow speeds the force of air resistance on the cardboard is proportional to the velocity of the cart:

$$\sum F = ma$$

$$-kv = ma$$

$$-kv = m\frac{dv}{dt}$$

$$\frac{dv}{v} = -\frac{k}{m}dt$$

$$\ln\frac{v}{v_0} = -\frac{k}{m}t$$

$$v = v_0 e^{-\frac{k}{m}t}.$$

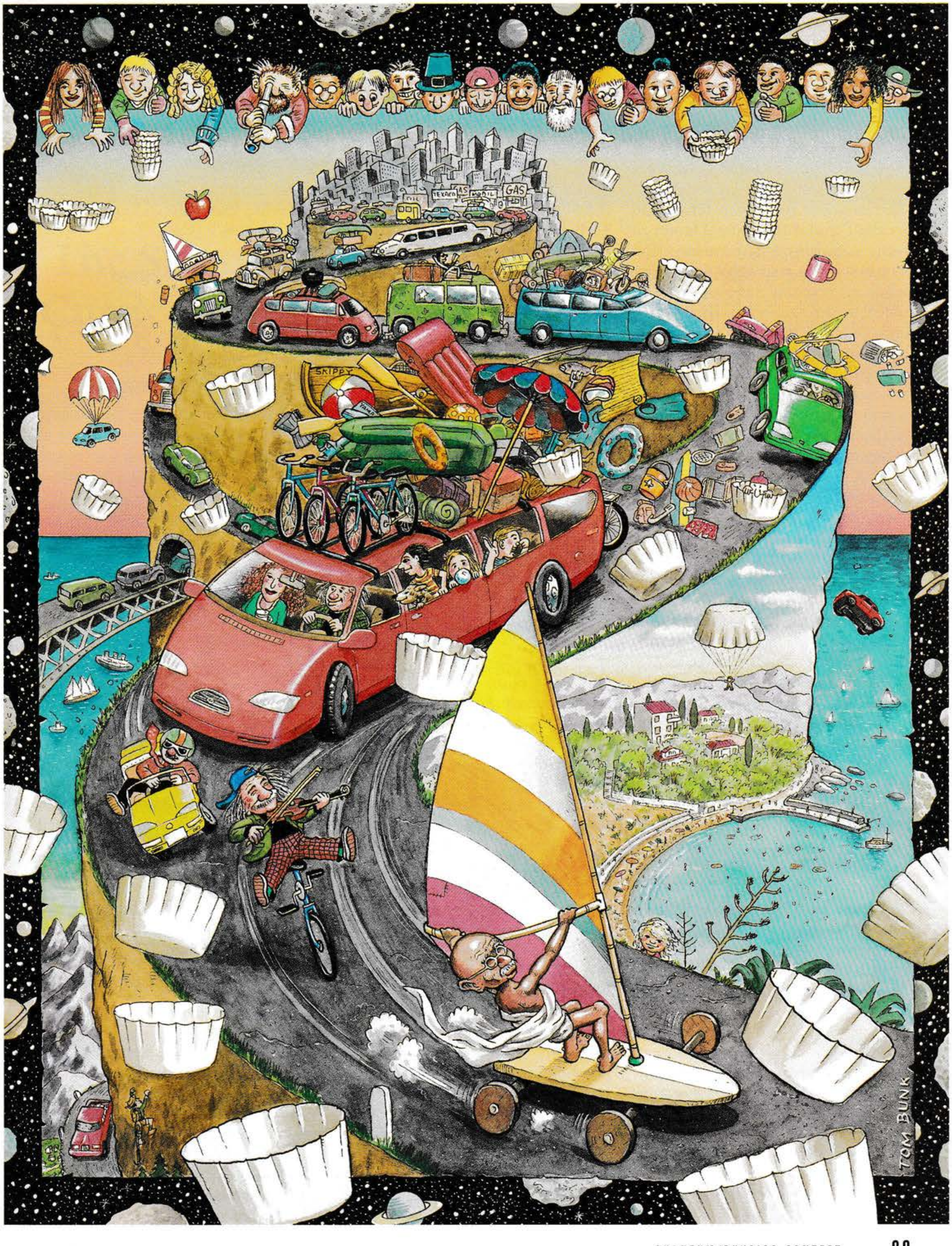
Therefore, the velocity decreases exponentially from its initial value. The proportionality constant k determines the rate of velocity decrease. It seems from the equation that the cart will never reach zero velocity in a finite time. Of course, when t is large enough, we can have a velocity that is effectively zero.

This air resistance costs us money. If there were no air resistance, the only force slowing our car on the highway would be the friction between the tires and the road. This small force would hardly retard our motion. We would glide along the ribbon of road at 60 miles per hour with little need for additional fuel.

Most of the fuel our cars consume is to counteract the air resistance. Open the car window as the car cruises at a fast clip and feel the wind on your hand. Your car is heading into quite a storm! Recognizing that the air resistance is proportional to the velocity of the car (or the square of the velocity of the car), it is easy to see that a small decrease in the speed of the car reduces the required fuel. Lower speed limits not only save lives but also save fuel.

An automobile design that lowers the air resistance so that cars could get an additional 1 mile per gallon saves an extraordinary sum of money. Let's try a quick "Fermi" calculation. There are approximately 60 million passenger cars in the United States (one for every four people). If each car travels approximately 20,000 miles per year and gets 20 miles per gallon of gas, then the fuel consumption is 1000 gallons per car or 60 billion gallons of gas. At \$1 per gallon, this is 60 billion dollars. If these cars could get 21 miles per gallon, we could save 5% of the total, or 3 billion dollars per year. How valuable is an automobile

Art by Tomas Bunk


GAS
SKIPPY
TOM BUNK

design that can save this much money every year?

Air resistance also affects sports. A baseball hit at 50 m/s (110 mph) at an angle of 45° would be able to travel 255 meters without air resistance. This is equivalent to 840 feet. Baseball would be quite a different sport if it were not for air resistance. Air resistance also affects tennis and football, and it is a crucial factor in table tennis and badminton. The symmetry of a trajectory disappears when air resistance is present. If an object is thrown vertically up with air resistance, will it take more time going up or coming down?

For this month's contest problem, we ascend from a simple problem of a stone falling with no air resistance to the more realistic situation where air resistance retards its motion.

1. A stone hits the bottom of a deep well and you hear the sound 3 seconds after the release. How deep is the well? Please assume that there is no air resistance and the speed of sound is a constant value of 340 m/s.

2. In this part, assume that the stone is affected by air resistance and that this resistive force is proportional to the velocity of the stone.

(a) Derive an expression for the velocity of the stone as it falls.

(b) Using 0.01 kg/s for the proportionality constant and 0.05 kg for the mass of the stone, determine the height of the well if the sound of the stone hitting the water arrives 3 seconds after the stone is released.

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.

Elephant ears

In the November/December 1997 issue we asked readers to solve three scaling problems from the International Physics Olympiad that was held in Canada last summer. A correct solution to the first question was submitted by Tal Carmon from Technion City, Israel.

A. The first question asked what would happen to the mean temperature T of Earth if the mean distance R between Earth and the Sun decreased by 1%. To do this we match the input radiation to the output radiation because Earth is in thermal equilibrium. If the power output of the Sun is P, the radiation reaching Earth per unit area is $P/4\pi R^2$. If we denote Earth's radius by R_E and its reflectance by r, the input power P_{in} to Earth is

$$P_{in} = (1 - r)\frac{P}{4\pi R^2}\pi R_E^2.$$

Stefan's Law gives the output power

$$P_{out} = 4\pi R_E^2 \varepsilon\sigma T^4,$$

where ε is the Earth's emissivity and σ is Stefan's constant. Although the emissivity is a function of temperature, the change in temperature is expected to be small, and we can neglect this dependence. Therefore,

$$T \propto \sqrt{\frac{1}{R}},$$

and a reduction of 1% in R gives a 0.5% rise in T. For a mean temperature of 287 K, we get a rise of 1.4 K.

B. The second question asked about the change in the density of dry air with an increase in humidity when the temperature and pressure remain the same. Let's use the subscripts d and m for dry and moist air respectively. Then the number of molecules N_d in the dry air is

$$N_d \propto \frac{M_d}{28.8},$$

where M_d is the mass of dry air in a unit volume and the mean molecular mass of dry air is 28.8 g/mol. For moist air, we must account for the proportions of dry air and water vapor. For 2% humidity, we have

$$N_m \propto 0.02\frac{M_m}{18} + 0.98\frac{M_m}{28.8},$$

where the mean molecular mass of water is 18 g/mol.

We know that identical volumes of ideal gases with the same temperature and pressure have the same number of molecules. Therefore,

$$M_d = 1.012 M_m.$$

Because the densities of equal volumes are proportional to the respective masses,

$$\frac{\rho_m}{\rho_d} = \frac{M_m}{M_d} = 0.988,$$

and using $\rho_d = 1.25$ kg/m^3, we get our answer:

$$\rho_m = 1.235 \text{ kg/m}^3.$$

C. The last question asked how the power required for a helicopter to hover depends on the size of the helicopter. The mechanical power P of the helicopter is equal to the thrust T times the downward velocity component v of the air below the blades. The thrust is given by the change in momentum of the air per unit time

$$T = v\frac{dm}{dt},$$

with

$$\frac{dm}{dt} = \rho A v,$$

where ρ is the density of the air and A is the cross-sectional area covered by the blades. Thus,

$$T = \rho A v^2.$$

When the helicopter is hovering, the thrust must be equal to the helicopter's weight. Therefore,

$$v^2 = \frac{T}{\rho A} = \frac{W}{\rho A}.$$

If the size of the helicopter is characterized by a linear dimension L, then $W \propto L^3$, $A \propto L^2$, and $v \propto L^{0.5}$. Thus,

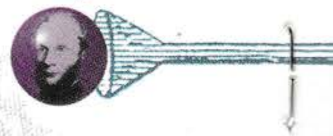
$$P = Tv = Wv \propto L^{3.5}.$$

For a half-scale helicopter, the required power is $0.5^{3.5}P = 0.0884P$. ◘

The nature o

C

I called this vapor "gas" because it hardly differs from the "chaos" of the ancients.
—Jan Baptista van Helmont



So I proposed a law which says that the molecules of various gases have an identical living force [that is, the kinetic energy] of translational motion.
—Rudolf Clausius

ALL OF THE QUOTES AT RIGHT in one way or another approach the notion of an ideal gas. Between the first statement, made by a Dutch naturalist, and the last statement, made by an Austrian physicist, lies an arduous pilgrimage of two and a half centuries.

The pathway beaten by the old atomists is far from being finished today. This road is marked by a wonderfully large number of names encompassing many countries and professions. This way was paved by, in addition to the quoted scientists, such brilliant minds as Newton, Hooke, Huygens, Laplace, Lavoisier, Boyle, Bernoulli, Joule, Maxwell, Perrin, Einstein...

Why is the ideal gas model so attractive? First of all, it provides the possibility of constructing a theory that has a broad range of consequences despite being based on the simplest of concepts. This model demonstrated the immense power of abstract thinking.

In physics the ideal gas approach was successfully applied to describe the electron "gas" in metals, radiation of electromagnetic waves, and even sound oscillation in crystals. This variety of applications testifies to the unusual universal character of the ideal gas model, which is a rare example of a fundamental theory underlying our physical world view.

Well, suppose that we are surrounded by ideal gases...

Questions and Problems

1. The force of gravity on the Moon is weaker than on Earth. However, dust hovers over Earth's surface much longer than over the Moon's. Why?

2. Do the partial pressures of nitrogen differ between places with dry soil and places with humid soil in warm, calm weather?

3. How many thermodynamic parameters determine the state of a specified ideal gas of a certain mass?

4. An ideal gas changes from state 1 to state 2 (fig. 1). How does its density vary in this process?

5. A vessel is divided into two compartments by a flexible, porous partition. One compartment is filled with hydrogen; the other with air at the same pressure. At first the partition bends toward the hydrogen compartment, then it assumes its initial position. Explain this phenomenon.

6. Two closed vessels of identical volume are filled with carbon dioxide. The height of the first vessel is half that of the second vessel. Two manometers placed at the tops of these vessels read the same value. What will the manometers show when the vessels are turned upside-down?

7. Atmospheric pressure is caused by the weight of the air. How is the pressure maintained in a spacecraft, where the air has no weight?

8. Does air exert the same pressure on the floor and the ceiling of a room?

9. Does the pressure a gas exerts on a vessel's wall depend on how the wall was machined?

10. A mixture of nitrogen and neon is placed in a vessel. Are the mean kinetic molecular energies of these gases equal or not?

11. The walls of a vessel are kept at different temperatures. Does the pressure exerted by a gas on a wall depend on the wall's temperature?

12. An ideal gas occupies half of a thermally isolated vessel, the other half of which is empty. How will the temperature of this gas change if the partition is momentarily removed?

13. A thermally isolated, cylindrical vessel filled with an ideal gas is suspended on a thread. The thread is torn and the vessel drops. Will the temperature of the gas change during the fall?

14. A moving vessel containing an ideal gas is stopped abruptly. How will the pressure of the gas change?

15. Does a strong wind increase the temperature of the air it carries?

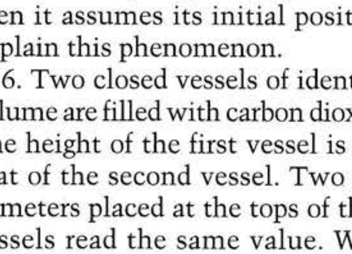


Figure 1

Microexperiment

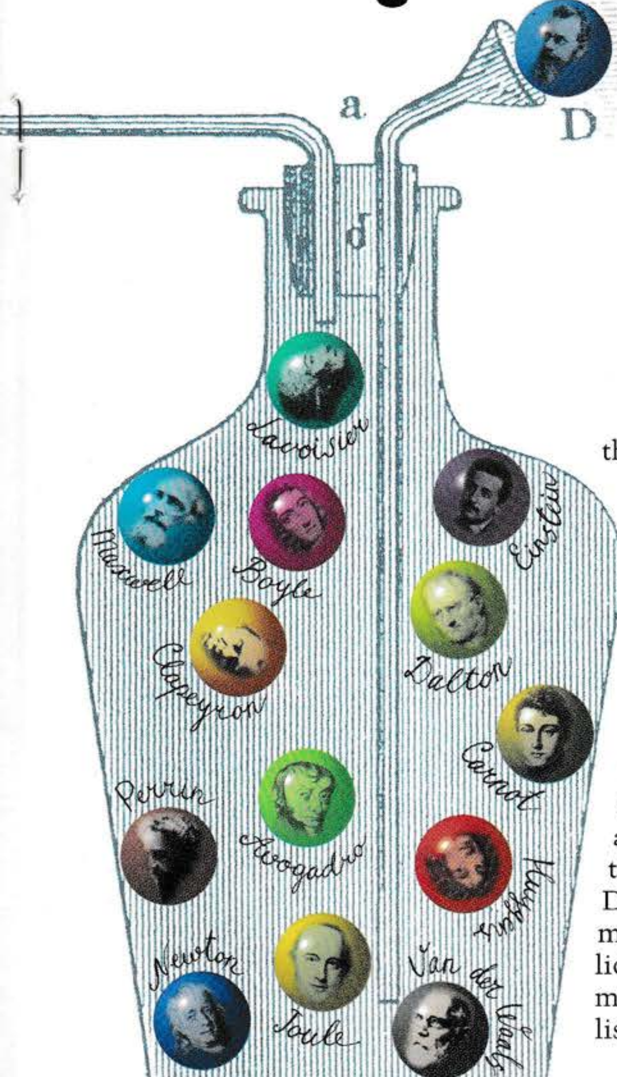
Turn a heater on in a room. After you get warm, think about what you feel—an increase in the internal energy of the air or an increase in the energy of individual molecules? Are these concepts identical?

It is interesting that...

... the word *gas*, coined at the beginning of the seventeenth century by Helmont, was unused for a rather long time, and it was reintroduced

# of an ideal gas

One takes the simplest body to investigate—that is, a gas confined within solid, absolutely elastic walls, whose molecules are hard, absolutely elastic spheres...
—Ludwig Boltzmann



by Lavoisier at the end of eighteenth century. The word became widespread at the time of the brothers Joseph and Etienne Montgolfier, who made the first hot-air balloon flight in 1783.

... at first look, Dalton's and Avogadro's laws are independent. But the former results from the latter, and both laws are the direct consequences of the kinetic theory of an ideal gas.

... the estimates of Avogadro's number made on the basis of the ideal gas approach were not as good as those given by the real gas theories, such as van der Waal's model.

... the kinetic theory of an ideal gas explains the experimentally established fact of the equality of the molar heat capacities of gases of the same type—say, monatomic or diatomic ones.

... many consequences of the kinetic theory awaited corroborative experimental evidence for a long time. It was not until 1911 that the French physicist Louis Dunoyer demonstrated that the molecules of a gas persistently collide with each other and that they move rectilinearly between the collisions.

... the widely known ideal gas law was formulated by Emile Clapeyron in a desire to "reanimate" the works of Sadi Carnot, who was undeservedly forgotten for several decades.

... the ideal gas theory makes it possible to estimate pressures and temperatures even within stars. Although they are just approximations, such estimates are quite close to those obtained from detailed calculations. Thus, according to these estimates, gas pressures inside stars are billions of times greater than normal atmospheric pressure, and the temperatures of these gases are millions of degrees.

... the ideal gas model begins to "limp" even at room temperature if the gas density increases by only 100 times its density under normal conditions.

... in his work on the kinetic theory, Maxwell was the first to use mathematical statistics to describe a physical phenomenon. Without statistics it would have been impossible to obtain a general description of gas behavior as an immense ensemble of particles.

... gas without colliding molecules is not a purely theoretical trick—there is such a gas in reality! Knudsen gas, which is so diffuse that its molecules collide only with the vessel's walls. The particular features of flow of such a gas through a small orifice are employed in the technique of gas separation.

... in recent years an ideal gas has become a virtual reality in computer simulations. Now we can see the transition from the artificially regular motion of the gas composed of identical little "balls" to stochastic motion. It is also possible to visualize the causes of "molecular chaos" and finally to describe the stochastic phenomena that had previously eluded detailed calculation.

—A. Leonovich

***Quantum* articles about ideal gas:**

A. Borovoy, "Learning about (not by) Osmosis," November/December 1991, 48–51.

A. Buzdin and V. Sorokin, "Double, Double, Toil and Trouble," May/June 1992, 52–53.

A. Stasenko, "An Ideal Gas Gets Real," September/October 1993, 42–43.

I. Vorobyov, "Cooled by the Light," September/October 1993, 20–25.

A. Eisenkraft, L. D. Kirkpatrick, "Cloud Formation," January/February 1995, 36–38.

V. Lange, "Shall We Light a Fire in the Fireplace?" January/February 1996, 40–41.

ANSWERS, HINTS & SOLUTIONS ON PAGE 54

Symmetry, part II

Polynomial equations and their roots

by Mark Saul and Titu Andreescu

IN OUR LAST COLUMN, WE explored the notion of algebraic symmetry, and used it to solve a variety of problems. In this column, we continue our exploration. Do you already know how the roots of a polynomial equation are related to its coefficients? Then skip ahead to problem 8. Otherwise, start reading below.

We start with simple quadratic equations. Every baby knows how to solve quadratic equations by factoring.

$$x^2 - 5x + 6 = 0$$
$$(x - 3)(x - 2) = 0$$
$$x = 3 \text{ or } x = 2.$$

So if $(x - 2)$ is a factor of the original polynomial, then 2 is a root of the polynomial equation. In fact, we can see that this will always work. There is nothing special about the number 2 or the factor $x - 2$:

(1) If $(x - a)$ is a factor of the quadratic polynomial $P(x)$, then $P(a) = 0$.

Is the converse of statement (1) true? Let us look at another baby example:

$$6x^2 - x - 1 = 0$$
$$(2x - 1)(3x + 1) = 0$$
$$(2x - 1) = 0 \text{ or } (3x + 1) = 0$$
$$x = 1/2 \text{ or } x = -1/3.$$

Now we see that $1/2$ is a root of the equation we started with, but $(x - 1/2)$ is not a factor. Or is it? If we had a bit more fondness for fractions, we could have solved the original equation as:

$$6x^2 - x - 1 = 0$$
$$x^2 - x/6 - 1/6 = 0$$
$$(x - 1/2)(x + 1/3) = 0$$
$$x = 1/2 \text{ or } x = -1/3.$$

Why didn't we do this? Because factoring takes some guesswork, and it's easier for us to guess about integers than about rational numbers. But this is our failing, and not that of the equation we are solving. Considering the second solution of our equation, we can see that it is true that if $x = 1/2$ is a root, then $x - 1/2$ is a factor.

Indeed, the whole story is made simpler if we consider only quadratic equations whose lead coefficient (the coefficient of x^2) is 1. We will do so for the remainder of this column, and now the converse of statement (1) is in fact true. We can state the very interesting factor theorem for quadratic polynomials:

If the lead coefficient of the quadratic polynomial $P(x)$ is 1, then $(x - a)$ is a factor of $P(x)$ if and only if $P(a) = 0$.

It is important to note that this statement is true for clumsy irrational roots as well as for neat integers or rational numbers.

Problem 1. Check that the roots of $x^2 - 3x - 5 = 0$ are

$$\frac{3 + \sqrt{29}}{2}$$

and

$$\frac{3 - \sqrt{29}}{2}.$$

Then show that the polynomial $x^2 - 3x - 5$ does indeed factor into

$$\left(x - \frac{3 + \sqrt{29}}{2}\right)\left(x - \frac{3 - \sqrt{29}}{2}\right).$$

(We promise that this is the last time you will see such complicated expressions in this column!)

Let us state problem 1 another way. Let us set

$$\alpha = \frac{3 + \sqrt{29}}{2}$$

and

$$\beta = \frac{3 - \sqrt{29}}{2}.$$

Then, noting that a quadratic polynomial can have only two linear factors, the problem actually states that

$$(x - \alpha)(x - \beta) = x^2 - 3x - 5.$$

If we multiply out the left-hand side, we get

$$x^2 - (\alpha + \beta)x + \alpha\beta.$$

Since this must equal $x^2 - 3x - 5$, it is not difficult to see (if you don't know already) that $\alpha + \beta = 3$, and $\alpha\beta = -5$.

Again, nothing is special about the silly numbers α and β. The factor theorem for quadratic polynomials has the following consequence:

The numbers α and β are roots of the polynomial equation $x^2 - px + q = 0$ if and only if $\alpha + \beta = p$ and $\alpha\beta = q$.

Problem 2. If α and β are roots of the equation $x^2 - 3x - 5 = 0$, find the numerical value of $\alpha^2 + 2\alpha\beta + \beta^2$.

Problem 3. If α and β are roots of the equation $x^2 - 3x - 5 = 0$, find the numerical value of $\alpha^2 + \beta^2$.

Problem 4. If α and β are roots of the equation $x^2 - px + q = 0$, find the numerical value of $\alpha^2 + \beta^2$.

Problem 5. If α and β are roots of the equation $x^2 - px + q = 0$, find the numerical value of $1/\alpha + 1/\beta$.

Problem 6. If α and β are roots of the equation $x^2 - px + q = 0$, find the numerical value of $\alpha^3 + \beta^3$.

Problem 7. If α and β are roots of the equation $x^2 - px + q = 0$, find the numerical value of

$$\frac{\alpha + 2\beta}{\alpha - \beta} + \frac{2\alpha + \beta}{\beta - \alpha}.$$

What has all this to do with symmetry? You can check that each of the expressions we must compute in the problems above are *symmetric* in α and β. It turns out that each of the problems above are examples of a very general statement:

Any symmetric rational function of α and β can be expressed in terms of the functions $\alpha + \beta$ and $\alpha\beta$.

Equivalently: **Any symmetric rational function of α and β can be expressed in terms of the coefficients of a quadratic equation with lead coefficient 1 and with roots α and β.**

(A rational function is just a quotient of polynomials.) This statement contains the seeds of some profound mathematical results. Let us look at a cubic equation, say

$$x^3 + 3x^2 - x - 3 = 0.$$

If we knew its factors, we could solve the equation by setting each factor equal to zero. But the converse is also true: if we knew the roots, then we could factor the equation. In fact, it is not difficult to check that 1, –1, and –3 are roots of this equation, so we know that the polynomial $x^3 + 3x^2 - x - 3 = 0$ factors as $(x + 1)(x - 1)(x - 3)$.

Problem 8. Check that the assertions made above are correct.

In short, the Factor Theorem for Quadratic Equations can be extended to polynomial equations of higher degree:

Factor theorem: The polynomial $P(x)$, with lead coefficient 1, has a factor $(x - a)$ if and only if $P(a) = 0$.

And we can once more translate this statement to a statement about the roots and the coefficients of a cubic equation. Suppose we start with a cubic polynomial $P(x)$ whose lead coefficient is 1. Then $P(x)$ factors as $(x - \alpha)(x - \beta)(x - \gamma)$ if and only if α, β, and γ are roots of the equation $P(x) = 0$. If $P(x) = x^3 - px^2 + qx - r$, then we can write

$$\begin{aligned} x^3 - px^2 + qx - r &= (x - \alpha)(x - \beta)(x - \gamma) \\ &= x^3 - (\alpha + \beta + \gamma)x^2 \\ &\quad + (\alpha\beta + \beta\gamma + \alpha\gamma)x - \alpha\beta\gamma. \end{aligned}$$

It follows that

$$\begin{aligned} \alpha + \beta + \gamma &= p, \\ \alpha\beta + \beta\gamma + \alpha\gamma &= q, \\ \alpha\beta\gamma &= r. \end{aligned}$$

Problems 9–13. If α, β, and γ are the roots of the equation

$$x^3 - px^2 + qx - r = 0,$$

express in terms of p, q, and r the values of:

9. $\alpha^2 + \beta^2 + \gamma^2$

10. $\alpha^3 + \beta^3 + \gamma^3$

11. $\alpha^4 + \beta^4 + \gamma^4$

12. $\dfrac{1}{\alpha^2} + \dfrac{1}{\beta^2} + \dfrac{1}{\gamma^2}.$

13. $\dfrac{\alpha\beta}{\gamma} + \dfrac{\alpha\gamma}{\beta} + \dfrac{\beta\gamma}{\alpha}.$

Problems 14–17. Solve the following systems.

14. $\begin{cases} \alpha + \beta = 5 \\ \alpha\beta = 6. \end{cases}$

15. $\begin{cases} \alpha + \beta = 8 \\ \alpha\beta = -7. \end{cases}$

16. $\begin{cases} \alpha + \beta + \gamma = -9 \\ \alpha\beta + \alpha\lambda + \beta\lambda = 19 \\ \alpha\beta\gamma = -11. \end{cases}$

17. $\begin{cases} \alpha + \beta + \gamma = 5 \\ \alpha^2 + \beta^2 + \gamma^2 = 29 \\ \alpha\beta\gamma = -24. \end{cases}$

Problem 18. Let a, b, c, d, and e be integers such that both $a + b + c + d + e$ and $a^2 + b^2 + c^2 + d^2 + e^2$ are divisible by an odd integer n. Prove that $a^5 + b^5 + c^5 + d^5 + e^5 - 5abcde$ is also divisible by n.

Problem 19. Let r_1 and r_2 be roots of the equation $x^2 + 2x + 3 = 0$. Compute

$$\frac{r_1^2 + 4r_1 + 5}{r_1^2 + 5r_1 + 4} + \frac{r_2^2 + 4r_2 + 5}{r_2^2 + 5r_2 + 4}.$$

ANSWERS, HINTS & SOLUTIONS ON PAGE 61

Light pressure

Are sunny days more burdensome?

by S. V. Gryslov

WHEN OBSERVING comets, medieval scientists tried to explain why the shape of the comets' tails varied with the comet's position relative to the Sun. It was Johannes Kepler who in 1604 posited that the shape of a comet's tail is determined by light pressure.

Many scientists tried to measure this pressure thereafter, including Augustin-Jean Fresnel, one of the creators of the wave theory of light. However, none of the experiments produced any results. For about three centuries the pressure of light remained just a brilliant hypothesis.

In 1865 James Clerk Maxwell formulated the electromagnetic theory of light. According to this theory, light waves are electromagnetic vibrations, and thus light can be considered an electromagnetic phenomenon. Maxwell's equations predicted the existence of light pressure. Indeed, Maxwell calculated this pressure himself. On a sunny midday the solar rays hit an absolutely reflective surface with the very small pressure of $4.7 \cdot 10^{-6}$ N/m^2.

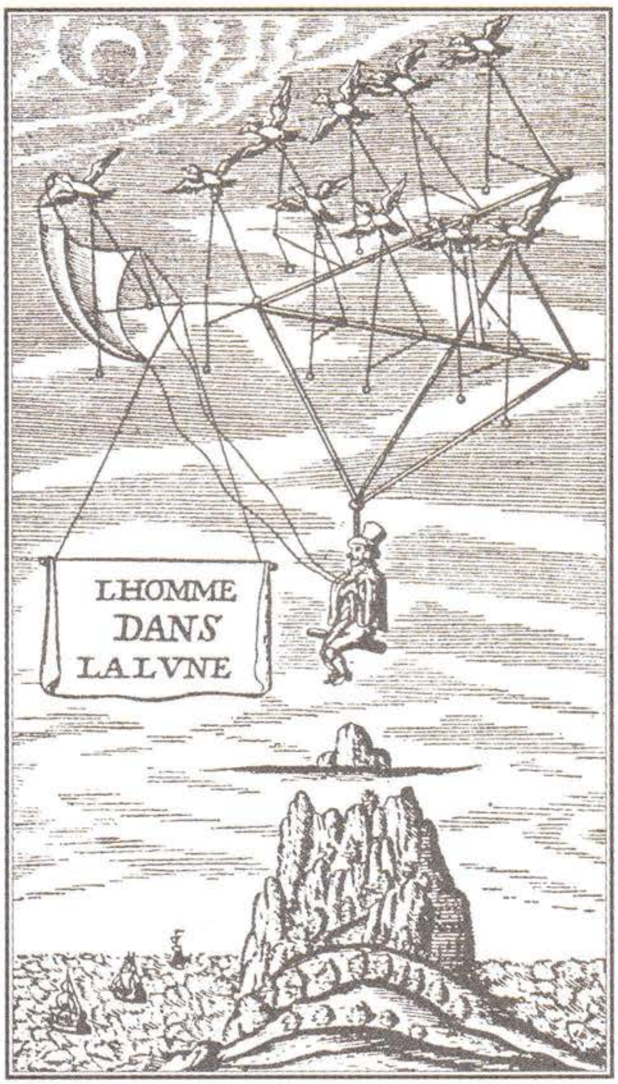


Figure 1. *This woodcut is a frontispiece from the French issue (1648) of the book* A Man on the Moon *by Godwin Francis. Perhaps the artist assumed that in the lunar heavens such flying vehicles would be propelled by the pressure of solar light!*

For a long time the validity of Maxwell's equations were a matter of vigorous scientific debate. Again there were attempts to detect the pressure of light, and again the scientists could not do it. The main problem was not the very small value of the light pressure; it was that the light itself was interfering with the experiments. The interfering effects caused by light were many times larger than the effects of light pressure.

In 1899 the famous Russian physicist Pyotr Nikolayevich Lebedev performed the first measurement of light pressure. He skillfully overcame the problems that were nightmares for experimentalists. The main part of Lebedev's setup was an arrangement of flat leaves of very small mass made of different materials (mostly of metals), which were fixed to a beam. This system of leaves was suspended by a cord in a vacuum chamber (fig. 1).

The surfaces of some leaves were blackened, and those of the others were polished. Virtually all the incident light was reflected from the polished surfaces and absorbed by the dark surfaces. As a result, the pressure of the light on the polished surfaces was almost two times that acting on the dark surfaces. This difference of forces created a torque, causing the beam to turn, twisting the cord. The pressure of the light was determined by the torsion angle.

What factors disturbed the measurements? The first problem was that the incident light heated the leaves and the surrounding air. The resulting convection flows acted like a wind on the leaves. Thus it was not clear why the beam moved—because of light pressure or air convection.

Another problem prevented the detection of the light pressure: the radiometric effect caused by the incident light. Because the light hit only one side of the leaves, the opposite sides of the leaves were not heated identically. Therefore, the opposite sides of a leaf transferred different amounts of energy (on average) to the surrounding gas molecules. After colliding with the warmer side, molecules acquired larger speeds (and momenta) than molecules that collided with the cooler side. By conservation of momentum, the molecules that collided with the warmer side imparted more momentum to the leaf upon recoiling than the molecules that collided with the cooler side.

It turned out that the radiometric forces acted in the same direction as the light pressure, but their values were a number of orders of magnitude larger than this pressure. The first step in minimizing the effect of these "destructive" forces was to perform the experiment in a vacuum. The fewer molecules in the chamber, the smaller the disturbance.

Lebedev used thin metal plates for the leaves by design. First, these thin films were good thermal conductors. Therefore, the temperature difference between the opposite sides of the leaves was smaller, and the radiometric interference decreased as well. In addition, the tiny masses of the leaves led to a small value for the moment of inertia of the entire system, which increased the accuracy of the experiment in a vacuum.

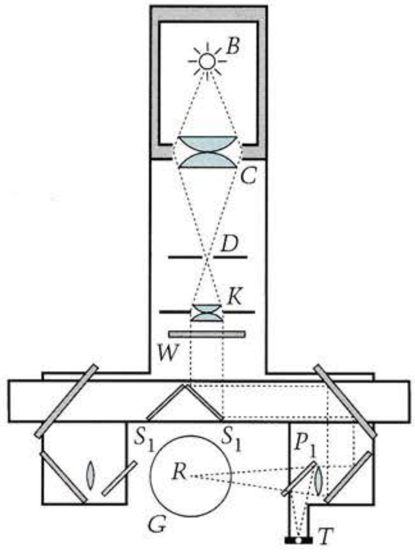


Figure 2. *Lebedev's experimental setup to measure the pressure of light. The beam* R, *with leaves of small mass suspended by a thin cord in a vacuum chamber* G, *is an extremely sensitive torsion balance. The light of an arc lamp is focused using a system of lenses and mirrors onto one of the leaves. The resulting torque on the beam is observed using a telescope and a mirror attached to the cord (not shown). By shifting the double mirror* S_1S_1, *it was possible to direct the light either onto the front or back surface of the leaf and thus change the direction of the torque. Plate* P_1 *made it possible to channel some part of the light to the thermoelement* T, *which measured the amount of incident light energy. In the experiments, different systems for fixing the leaves were used.*

Lebedev invented many clever devices and contrivances and tried many variations of experimental procedures and conditions until his efforts were rewarded. He eventually obtained a value within 20 percent of Maxwell's prediction. Lebedev thus demonstrated that light carries not only energy but also momentum.

For a long time the pressure of light was only of theoretical importance. Practical application of this phenomenon was not possible due to the extremely small power of light beams. However, the invention of lasers made it possible to apply Lebedev's discovery.

For example, it is possible to use lasers to suspend and shift small particles in the air, counterbalancing the gravitational attraction by the supporting pressure of light. In addition, particles of different masses are differently accelerated by the light pressure. Therefore, it is possible to construct special traps to sort particles of different masses.

The pressure of light can also be used to separate two gases from their mixture. Such a separation can be performed in the case when the frequency of the laser radiating the mixture coincides with the frequency corresponding to the transition of one of the gas's atoms from the ground state to an excited state. When an atom of this gas absorbs a photon, it acquires a momentum in the direction of the laser beam. After a transition from the excited state back to the ground state, the momentum vector of the emitted photon has an arbitrary direction.

In the subsequent cycles of photon absorption and radiation, the momenta caused by radiation are mutually annihilated, but the momenta resulting from photon absorption are summed. Thus, on average the resonant atoms are imparted with a nonzero momentum in the direction of the laser beam. Therefore, in passing through two successive chambers, the first of which contains a mixture of gases, the laser beam carries the atoms of one gas into the second chamber and thus separates the gases.

Light pressure can also be used to accelerate small particles in a vacuum (air resistance is too large for such a task under normal pressures). This can be used, say, to simulate the harmful effects of micrometeorites on the surfaces of a spacecraft. We may also be able to use the pressure of solar light for the orientation and acceleration of spacecrafts.

Quantum articles about light pressure:

A. Eisenkraft, L. D. Kirkpatrick, "Laser Levitation," November/December 1994, 38–39.

A. Eisenkraft, L. D. Kirkpatrick, "The Nature of Light," May/June 1997, 39.

How to escape the rain

Raindrops keep falling on my parallelepiped

by I. F. Akulich

Art by Marguerite Renais

IMAGINE THAT YOU ARE walking down a street on a cloudy day and have neither an umbrella nor a raincoat nor anything else to protect yourself from the rain. Then, all of a sudden, you are pelted by heavy rain accompanied by a strong wind. What should you do?

An overwhelming majority of people (including the author) will answer as follows: You should rush to the nearest shelter, and the faster you run there, the drier you will be. This seems like an indisputable truth.

Even so, some people will say something like, "Of course I would go to the nearest shelter. But there is no need to run, because if I run fast, I will spend less time in the rain, and thus fewer raindrops will hit me from above. But my body will run into more drops falling in front of me. My top will stay drier, but my front will get much wetter. So, what's the use of running? None, really." This type of person walks calmly in the rain, in spite of the wondering glances of passersby.

What should we think about this reasoning? That it is a mistake? Probably. But what if it is not a mistake but a brilliant intuitive revelation and they are right?

Here's another consideration supporting their idea. Suppose the wind blows toward the shelter and is so strong that the rain falls almost horizontally. Then the best solution is to run from the rain at the same speed as the wind: All the drops will fly on courses parallel to yours, and none of them will touch you. Running with less speed is ineffective, but running faster isn't good, either—it will make you wetter! Of course, it is not easy to move with the speed of a strong wind. Nevertheless, this reasoning suggests that we should further investigate the problem.

Setting the problem

Let's set the conditions of our problem:

An individual is standing in the street. Suddenly the rain starts. The individual rushes to the nearest shelter, which is situated l yards from him. How fast should he move to stay as dry as possible?

The problem is stated, but it doesn't look as if we could solve it in this form. In fact, the human body has a very complicated shape, and when we run, we move our legs and wave our hands, so the shape of our body continuously changes. Thus it is too difficult (in fact, impossible) to carry out the precise calculations in this case. We have nothing else to do but solve a simpler approximation of our problem:

A right parallelepiped whose faces have areas S_S, S_T, and S_F (the subscripts refer to "side," "top," and "front") moves with the velocity $\mathbf{u}$ perpendicular to S_F. It's raining. Every drop of rain falls with the velocity $\mathbf{v}$ (fig. 1) (the vector $\mathbf{v}$ doesn't have to be pointed vertically downward since we consider slanted rain

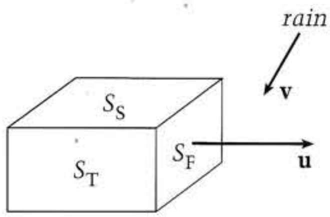


Figure 1

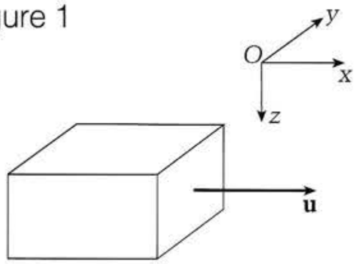


Figure 2

as well). The number of drops in a unit volume is k. How many (we denote this number by N) drops will fall on the parallelepiped while it crosses the distance l, and for what velocity $\mathbf{u}$ will the number N be as small as possible?

The work begins

Let's introduce the coordinate system $Oxyz$ in the following way: the Oz-axis is pointed vertically downward, the Ox-axis is pointed in the direction of the vector $\mathbf{u}$, and the Oy-axis is pointed perpendicular to the plane Oxz so that the projection of the drops' velocity $\mathbf{v}$ on it is positive (fig. 2).

Since the vector $\mathbf{v}$ is given, we can suppose that its projections on the coordinate axes are given, too. Let's denote them by v_x, v_y, and v_z. What can we say for sure about these projections? Clearly, $v_y \geq 0$ (that's how we chose the Oy-axis). Besides this, $v_z > 0$ (the rain must fall to the ground). And as far as the value of v_x is concerned, it can be both positive (when the rain follows you) and negative (when the rain flies in your face), or even vanish.

Let's consider the situation in the frame of reference connected to the parallelepiped—that is, the parallelepiped is stationary. Then the velocity of the raindrops becomes equal to $\mathbf{w} = \mathbf{v} - \mathbf{u}$. Projections of the velocity $\mathbf{w}$ on the coordinate axes are $w_x = v_x - u$, $w_y = v_y$, and $w_z = v_z$ (here u is the length of the vector $\mathbf{u}$). We have to determine how many drops N will fall on the parallelepiped during the time $\tau = l/u$ and for what value of u the number N is as small as possible.

Clearly a drop will fall on the parallelepiped in the time τ if and only if it lies at a distance less than or equal to $\tau\,|\mathbf{w}|$ from its faces—that is, if it lies inside the body drawn in figure 3 with red lines. What is the volume of this body? It is easy to see that the body consists of three prisms whose bases' areas are S_F, S_S, and S_T, and whose heights are the absolute values of the vector's $\tau\mathbf{w}$ projections on the axes Ox, Oy, and Oz respectively. Therefore the volume of the body is

$$\tau\left(|w_x|S_F + |w_y|S_S + |w_z|S_T\right)$$
$$= \tau\left(|v_x - u|S_F + v_yS_S + v_zS_T\right),$$

and the number of drops N is equal to

$$\tau \cdot k\left(|v_x - u|S_F + v_yS_S + v_zS_T\right).$$

Taking into consideration that $\tau = l/u$, we find how N depends on u:

$$N = kl\frac{|v_x - u|S_F + v_yS_S + v_zS_T}{u}.$$

Now we'll find $u > 0$ that corresponds to the least possible N.

The solution continues

Since the quantities k and l are constant, we will, for convenience, consider the variable

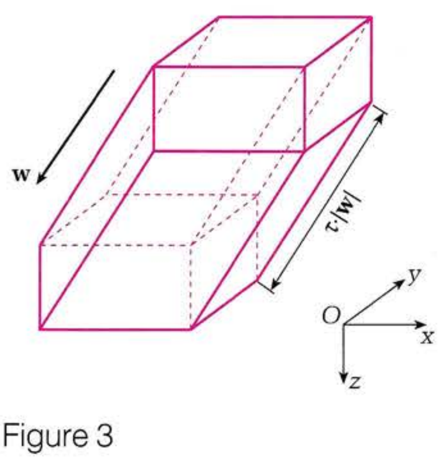


Figure 3

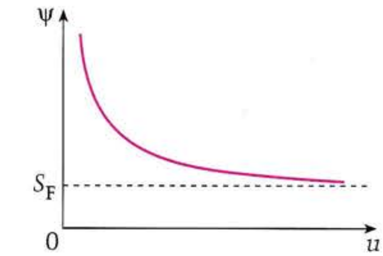


Figure 4

$$\psi = \frac{N}{kl} = \frac{|v_x - u|S_F + v_yS_S + v_zS_T}{u}.$$

Let's consider two cases:

1. $v_x \leq 0$ (the rain flies in your face).

Here $v_x - u < 0$, and we can rewrite the formula for ψ as follows:

$$\psi = \frac{(u - v_x)S_F + v_yS_S + v_zS_T}{u}$$
$$= S_F + \frac{-v_xS_F + v_yS_S + v_zS_T}{u}.$$

Since $v_x \leq 0$, the numerator of the fraction on the right is positive, and thus $\psi(u)$ is decreasing in the interval $(0, +\infty)$. The graph of the function $\psi = \psi(u)$ is shown in figure 4. We see that although ψ decreases when u increases, the inequality $\psi > S_F$ always holds, and that $\psi \to S_F$ when $\psi \to \infty$. So it seems that the supporters of the "take your time" theory are making a mistake: The faster you run, the less wet you get.

Nonetheless, here we meet with another fact, rather unexpected at first sight: Since $\psi > S_F$ for all u, $N = kl\psi$ always exceeds klS_F. This means that however fast you run (even if you fly like a bullet), you will still get the minimum amount of rain: klS_F. Thus, there is some logic to the reasoning of people who don't like to hurry.

2. $v_x > 0$ (the rain follows you).

Here we should consider two intervals:

(a) $0 < u \leq v_x$. Then $|v_x - u| = v_x - u$, and

$$\psi = \frac{v_xS_F + v_yS_S + v_zS_T}{u} - S_F.$$

This function decreases in the semi-interval $(0, v_x]$ and attains a minimum at $u = v_x$:

$$\psi\big|_{u=v_x} = \frac{v_y S_S + v_z S_T}{v_x}.$$

(b) $u > v_x$. Then $|v_x - u| = u - v_x$, and

$$\psi = \frac{-v_x S_F + v_y S_S + v_z S_T}{u} + S_F.$$

In this case we can't immediately say how ψ changes when u increases. It depends on the numerator of the fraction on the right:

$$A = -v_x S_F + v_y S_S + v_z S_T.$$

If $A > 0$, then $\psi(u)$ is a decreasing function in the interval $(v_x, +\infty)$; if $A < 0$, then it increases in the same interval; and if $A = 0$, then $\psi = S_F =$ const.

Figure 5 shows three possible graphs of the function $\psi = \psi(u)$ for all $u \in (0, +\infty)$. We see that in the interval from 0 to v_x, the curves look similar in all cases ($A > 0$, $A = 0$, and

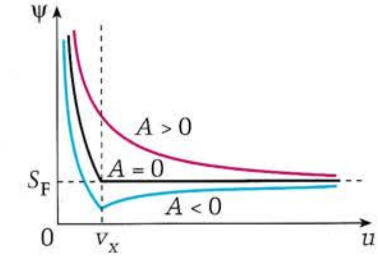


Figure 5

$A < 0$), and that they differ after the break $u = v_x$.

We conclude that, sometimes, when the wind is favorable, the arguments drawn by the supporters of the counterintuitive "take your time" theory suddenly prove to be correct and rational.

Check the wind

Taking into consideration the results of our calculations and looking at the graphs in fig. 4 and 5, we can give a well-grounded answer for the question of whether to walk or run in the rain.

If the rain flies in your face, you should run to the shelter as quickly as you can. If the rain follows you, then you should first evaluate in your mind the value

$$A = -v_x S_F + v_y S_S + v_z S_T.$$

If it turns out that $A > 0$, then you'd better hurry up again. If you see that $A = 0$, then it doesn't matter how quickly you walk: You will get as wet as ever, independent of your speed if it is greater than v_x. But when $A < 0$, you should run with the speed v_x to stay as dry as possible.

For instance, if $v_y = 0$ (the case of a favorable wind), then the inequality $A < 0$ is equivalent to $v_x S_F > v_z S_T$. For a tall and thin man, S_F is much greater than S_T, and thus this condition might be satisfied for a small "walker's" v_x. If such a person goes to the shelter so that it seems to him or her that the rain falls vertically, he or she can remain almost completely dry. That's what is called "passing between drops." ◘

Amusing electrolysis

Current thinking in chemistry

by N. Paravyan

MANY FEATURES OF ELECTROLYSIS—a process in which a substance is subjected to electric current to cause a chemical change—are presently known. Therefore we will explore not its common aspects but rather some distant corners of this electrochemical phenomenon.

For our experiments we need sources of direct and alternating currents with electromotive forces of up to 6 V. For a dc source we can use a flashlight battery (it is better to use two batteries connected in series), and for an ac source we can use a step-down transformer (110 V to 5–6 V). In addition, we need insulated connecting wires, four 6×3 cm iron plates (which can be cut from a clean tinplated can), an aluminum plate of the same size, and a small glass. We also need a small amount of table salt (sodium chloride) and baking soda (sodium bicarbonate).

Experiment 1. Fill the glass half-full with 3% sodium bicarbonate solution and insert two iron plates to serve as electrodes. The electrodes should be vertical, parallel to each other, and far apart. Next, connect the electrodes to the dc source. Electrolysis should start immediately:

$$2H_2O \rightarrow 2H_2 + O_2.$$

This equation shows that hydrogen is produced in larger quantities than oxygen, so we can easily determine the polarity of the source, which allows us to tell the anode from the cathode. We can test our determination with a burning splinter: At the cathode the splinter ignites hydrogen so that it "pops," and at the anode the flame burns more brightly because of the presence of extra oxygen.

Experiment 2. Without turning the current off, submerge a third iron electrode into the electrolyte between the cathode and the anode (it should not touch either of them). Paradoxically, gases will be produced at the middle electrode as well, and more curiously, on both sides of it. Take the burning splinter again and test to find which side oxygen or hydrogen is liberated on.

It turns out that oxygen is produced on the side of the middle electrode that faces the cathode, and hydrogen is produced on the side that faces the anode. How can we explain this? Recall that, similar to the metal electrodes used in our homemade electrolytic cell, the electrolyte is also an electric conductor, albeit with a large resistance. Therefore, let's examine the electric current in the glass (fig. 1).

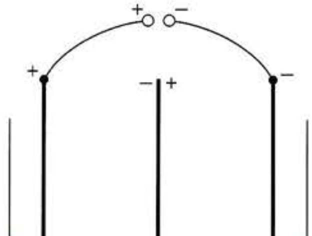


Figure 1

Current "enters" the middle electrode from the electrolyte, so the side facing the anode acquires a negative charge and becomes a cathode. As the current "leaves" the middle electrode, the side facing the cathode acquires a positive charge and becomes an anode. Therefore, hydrogen is liberated on the negatively charged side of the middle electrode, and oxygen is given off on the opposite (positively charged) side.

Is more gas (by volume) produced by three electrodes than by two? Let's do another experiment to find out.

Experiment 3. Submerge a fourth iron electrode in the electrolyte. Again, it should not touch any of the other electrodes. What's this? There is no gas at all! If we connect an ammeter in series with the battery, we see that only a negligibly small current flows in the circuit. Why is this? An iron plate with a resistance hundreds of times smaller than that of the electrolyte cannot increase the resistance of the circuit enough to stop the current almost entirely!

By sinking the third plate into the electrolyte, we made two electrolytic cells from one and connected them in series (review fig. 1). As a result the decomposition voltage—that is, the minimum voltage required for electrolysis to occur—increased by approximately twofold. Denoting the electromotive force of the source by $\mathscr{E}$ and the decomposition voltage by V, we have $(\mathscr{E} - V)$ for the voltage drop across the two electrodes in the first experiment and $(\mathscr{E} - 2V)$ for the similar value in the experiment with three electrodes. Thus, the total volume of oxygen and hydrogen produced in the two-electrode experiment equals the total volume of these gases produced in the three-electrode experiment.

When we submerged a fourth electrode, we made three electrolytic cells connected in series. In this case the decomposition voltage increased by about three times, so $\mathscr{E} - 3V < 0$. This means that the current virtually stopped flowing, and no gas was produced in the third experiment.

Experiment 4. Repeat experiment 2, but instead of the iron plate, use an aluminum plate for the middle electrode. In addition, connect an electric lamp in series with the battery. Turn the current on and notice that the lamp shines more and more dimly and dies out in 10 to 15 minutes. Take the aluminum plate out of the electrolyte—the lamp will shine again. Should we think that the aluminum plate, an excellent conductor, turned into an insulator? Let's continue the experiment.

Now replace the battery with the step-down transformer. Again submerge the aluminum plate in the electrolyte. In this case the lamp shines somewhat more dimly, but it shines. And regardless of how long you wait, it will never die out.

Again connect the battery instead of the transformer to the circuit and place the aluminum plate between the two iron electrodes such that the side that previously faced the cathode now faces the anode. Switch the power on and notice that the lamp shines initially, but as in the first case, it dies out in 10 to 15 minutes. Now feed the electrolytic cell with alternating current—the lamp will not shine at all.

What happened? Let's think. When direct current passed through the aluminum electrode (as in fig. 1), the following reactions took place on its surface:

$$(-)2H^+ + 2e^- \rightarrow H_2;$$
$$(+)4OH^- + 2e^- \rightarrow O_2 + 2H_2O;$$
$$4Al + 3O_2 \rightarrow 2Al_2O_3.$$

In other words, the entire "positive" side of the electrode turned into a dielectric, the circuit was broken, and the lamp went out. At the same time, the "negative" side of the aluminum electrode preserved electric conductivity and conducted electric current, although only in one direction! When the electrolytic cell was connected to the transformer (alternating current), the current could flow only in one direction. This means that the plate, oxidized on one side only, played the role of a current rectifier, which turned the alternating current into a pulsating current. So the lamp shone continuously when our electrolytic cell was fed by the transformer.

After the aluminum electrode was turned around and the device was fed with direct current, this electrode was oxidized on both sides. Covered by aluminum oxide films, it turned into an insulator. The oxidized electrode could neither rectify the current nor conduct it. This was the reason the lamp did not shine when the aluminum plate that was oxidized on both sides was submersed in the electrolyte.

What happens if sodium bicarbonate is substituted for sodium chloride? It turns out that a similar effect is observed with the iron plates but not with the aluminum plate because in the latter case some other chemicals are produced during electrolysis, and the "semiconductor" film of Al_2O_3 is not formed. The ability of aluminum to produce insulating films on its surface during electrolysis is widely used to insulate aluminum tools.

Quantum articles about electrolysis:

N. Paravyan, "A Tell-tale Trail and a Chemical Clock," September/October 1995, 42–43.

Homogeneous equations

by L. Ryzhkov and Y. Ionin

EXAMPLES OF WHAT MATHematicians call *homogeneous equations* abound in high school mathematics. When considering systems of linear equations, we can call the following system homogeneous in x and y:

$$\begin{cases} ax + by = 0 \\ cx + dy = 0. \end{cases}$$

Again, when solving trigonometric equations, we can call equations like

$$\sin x - \cos x = 0$$

and

$$\sin^2 x - 5 \sin x \cos x + \cos^2 x = 0$$

homogeneous in $\sin x$ and $\cos x$.

What property unifies all these examples, and how should we understand the term *homogeneous equation*?

Definition. A polynomial function of two variables $f(u, v)$ is called a "homogeneous polynomial of degree n" if all of its monomials are of degree n. For example,

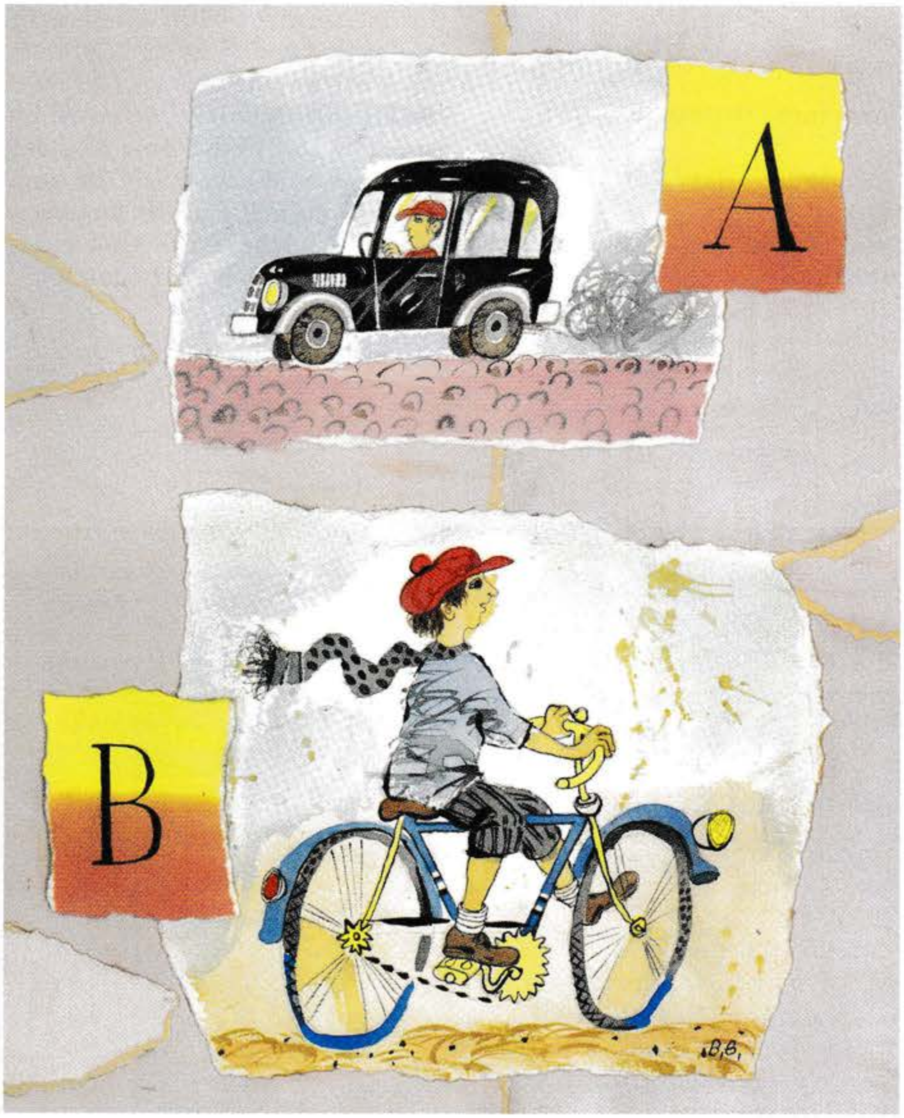
$$f(u, v) = 2u^2 - 7uv + 9u^2$$

is a homogeneous polynomial of the second degree, and

$$f(u, v) = u^3 - 15u^2v + 5v^3$$

is a homogeneous polynomial of the third degree.

Definition. The equation $f(u, v) = 0$ is called "homogeneous equation of degree k" if $f(u, v)$ is a homogeneous polynomial of degree k.



Art by Vasily Vlasov

Note that the notion of homogeneity can be generalized to encompass polynomials of more than two variables. For example, we say that the equation

$$x^3 + 3x^2y + 3xyz + z^3 = 0$$

is a homogeneous equation of the third degree with respect to the variables x, y, and z.

Consider a homogeneous equation of degree n in two unknowns x and y:

$$f(x, y) = a_0x^n + a_1x^{n-1}y + a_2x^{n-2}y^2 + \ldots + a_ny^n = 0 \quad (1)$$

We suppose that $a_0 \neq 0$.

The case when $a_0 = 0$ reduces to a case of lower degree. Indeed, if we factor out the lowest possible degree of y, a homogeneous polynomial with a nonzero leading term appears as the other factor. For instance,

$$x^4y^2 + 3xy^5 - 7y^6 = 0$$

turns into

$$y^2(x^4 + 3xy^3 - 7y^4) = 0.$$

Now it is enough to consider the equations

$$y^2 = 0$$

and

$$x^4 + 3xy^3 - 7y^4 = 0.$$

Note that the pair $x = 0$, $y = 0$ is a solution of equation (1), and the pair $x = x_0$, $y = 0$ is not if $a_0 \neq 0$ and $x_0 \neq 0$. If we divide both sides of equation (1) by y^n, we obtain an equation of degree n in one unknown, $t = x/y$:

$$a_0t^n + a_1t^{n-1} + a_2t^{n-2} + \ldots + a_n = 0 \quad (2)$$

We can use this property of homogeneous equations to solve certain problems.

Let's find the set of points on the plane whose coordinates satisfy equation (1). First of all, this set includes the origin. Then, if $a_0 = 0$, we get the solution $y = 0$, which defines the x-axis. And the roots $t_1, t_2, \ldots, t_k$ of equation (2) define the lines $x = t_1y$, $x = t_2y, \ldots,$ $x = t_ky$, passing through the origin. (In figure 1 we take, for example, $t_1 = 1$, $t_2 = -1$, $t_3 = 2$).

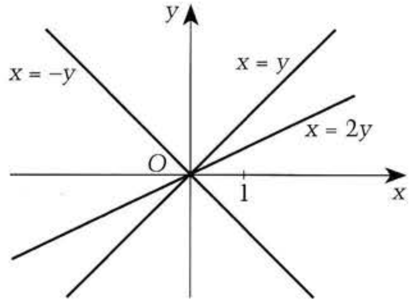


Figure 1

Problem 1. A car drives with constant velocity from town A to town B. At the same moment a bicycle leaves B for A. Three minutes after they have met, the car turns back abruptly, pursues the bicycle and, having caught up with it, turns again and arrives at B. If the car had turned back one minute after they had met and if the biker had increased his speed by 15/7 at the same moment, the car would have spent the same time getting to B. Find the quotient of the car's and the bicycle's velocities.

Problem 2. Solve the equation

$$20\left(\frac{x-2}{x+1}\right)^2 - 5\left(\frac{x+2}{x-1}\right)^2 + 48\frac{x^2-4}{x^2-1} = 0.$$

Problem 3. Solve the equation

$$2(x^2 + x + 1)^2 - 7(x-1)^2 = 13(x^3 - 1).$$

Problem 4. Solve the equation

$$x^2 + 2x + 15 = 2x\sqrt{2x+15}. \quad (3)$$

Problem 5. Solve the equation

$$2\sqrt[3]{x+1} - \sqrt[3]{x-1} = \sqrt[6]{x^2-1}.$$

Problem 6. Solve the system

$$\begin{cases} 3x^2 - 2xy - y^2 = 0, \\ x^2 + 5y = 6. \end{cases}$$

Problem 7. Solve the system

$$\begin{cases} 3x^2 - \dfrac{25}{12}xy + 3y^2 = 50, \\ x^2 + y^2 = 25. \end{cases} \quad (4)$$

Problem 8. Solve the system

$$\begin{cases} 2x^2 - 3xy + y^2 = 3, \\ x^2 + 2xy - 2y^2 = 6. \end{cases}$$

Let's now consider some examples of trigonometric equations that can be reduced to homogeneous equations.

Problem 9. Solve the equation

$$2\sin x + 3\cos x = 0.$$

Problem 10. Solve the equation

$$\sin^2 x - 3\sin x \cos x + 2\cos^2 x = 0.$$

There are many ways to solve the equation $a \sin x + b \cos x = c$. For example, it can be deduced to a homogeneous equation with respect to $\sin(x/2)$ and $\cos(x/2)$ by means of the substitution

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2},$$
$$\cos x = \cos^2\frac{x}{2} - \sin^2\frac{x}{2},$$
$$c = c\left(\cos^2\frac{x}{2} + \sin^2\frac{x}{2}\right).$$

Problem 11. Solve the equation

$$3\sin x + 5\cos x = -3 \quad (6)$$

The possibility of increasing by 2 the degree of monomials of the type $u = \sin x$, $v = \cos x$ by multiplying them by $1 \equiv \sin^2 x + \cos^2 x$ enriches the technique of using homogeneity to solve trigonometric equations.

Problem 12. Solve the equation

$$4\sin^3 x = \sin x + \cos x.$$

Sometimes we can reduce an exponential equation to a homogeneous one by an appropriate substitution.

Problem 13. Solve the equation

$$4^x = 2 \cdot 14^x + 3 \cdot 49^x.$$

Exercises

1. Solve the equation

$$6\sqrt[3]{x-3}+\sqrt[2]{x-2}=5\sqrt[6]{(x-2)(x-3)}.$$

2. Solve the equation

$$5^{\log_2\left(x^2\right)}-3^{2\log_4\left(x^2/2\right)}=\sqrt{3^{\log_{\sqrt{2}}\left(2x^2\right)}-5^{\log_2\left(x^2\right)-1}}.$$

3. Solve the equation

$$\left(3-2\sqrt{2}\right)^{x^2-6x+9}+\left(3+2\sqrt{2}\right)^{x^2-6x+9}=6.$$

4. Solve the equation

$$\sin 2x-\tan\frac{\pi}{6}\cos 2x=1.$$

5. Solve the equation

$$2\sin 4x-3\sin^2 2x=1.$$

6. Solve the equation

$$3\cos^2 x-\sin^2 x-\sin 2x=0.$$

7. Solve the system

$$\begin{cases}x^2+y\sqrt{xy}=420,\\ y^2+x\sqrt{xy}=280.\end{cases}$$

8. Solve the system

$$\begin{cases}\dfrac{x^2+y^2}{xy}=\dfrac{5}{2},\\ x^2-y^2=3.\end{cases}$$

9. Solve the equation

$$7\cdot 3^{x+1}-5^{x+2}=3^{x+4}-5^{x+3.}$$

10. An empty tank is being filled with the help of two pumps: one pumping pure water and the other a solution of an acid of fixed concentration. When the tank is full, the concentration of the acid in it is 5%. If the water pump had been turned off when the tank was half full, then the final concentration of the acid would have been 10%. Determine which pipe works faster and the ratio of the pumps' rates.

11. Solve the equation

$$\frac{(34-x)\sqrt[3]{x+1}-(x+1)\sqrt[3]{34-x}}{\sqrt[3]{34-x}-\sqrt[3]{x+1}}.$$

12. The system

$$\begin{cases}a_1x^2+b_1xy+c_1y^2=d_1,\\ a_2x^2+b_2xy+c_2y^2=d_2\end{cases}$$

has two solutions:

$$x_1=2,\ y_1=3,\ x_2=-5,\ y_2=-7.$$

Can you say whether or not there are some other solutions to the system? If so, what are these extra solutions?

13. Solve the equation

$$2^{2x+2}-6^x-2\cdot 3^{2x+2}=0.$$

14. Solve the equation

$$\left(\sqrt{\sqrt{x^2-8x+7}+\sqrt{x^2-8x-9}}\right)^x+\left(\sqrt{\sqrt{x^2-8x+7}-\sqrt{x^2-8x-9}}\right)^x=2^{x+1}.$$

15. Solve the system

$$\begin{cases}2^{x-y}-3\cdot 2^x+2^{x+y+2}=2,\\ 2^{x-y+1}-2^x-5\cdot 2^{x+y}=1.\end{cases}$$

16. Solve the equation

$$2\left(x^2+2\right)=5\sqrt{x^2+1}.$$

ANSWERS, HINTS & SOLUTIONS ON PAGE 56

Bulletin Board

Cyberwinners, old and new

Due to some crossed wires at CyberTeaser headquarters, the winners of the March/April contest were not posted here in the previous issue. Here are the first 10 contestants who correctly determined the extent of Boris's ice cream gluttony (brainteaser B228):

Liro Goldenberg (Holon, Israel)
Suhas Nayak (New South Wales, Australia)
David Friend (Winchester, England)
Hana Bizek (Argonne, Illinois)
Aleksei Lukashkin (Brighton, East Sussex, England)
Bob Cordwell (Ellicott City, Maryland)
Jaak Sarv (Tallinn, Estonia)
Bruno Konder (Rio de Janeiro, Brazil)
Alex Wissner-Gross (New Hyde Park, New York)
Theo Koupelis (Wausau, Wisconsin)

The next CyberTeaser (brainteaser B231 in this issue) proved to be trickier than it seemed, judging from the near misses and not-so-near misses we received. But we applaud everyone who put a human face on our clock puzzle by sending in an answer.

Here are the first 10 clockwatchers who found the degree of separation:

May T. Lim (Diliman, Philippines)
Clarissa Lee (Perak, Malaysia)
Bruno Konder (Rio de Janeiro, Brazil)
Jaak Sarv (Tallinn, Estonia)
Theo Koupelis (Wausau, Wisconsin)
Aleksei Lukashkin (Brighton, East Sussex, England)
Frank Lipinski (Batavia, Illinois)
Worawat Meevasana (Mercersburg, Pennsylvania)
Alex Wissner-Gross (New Hyde Park, New York)
Leo Borovskiy (Brooklyn, New York)

All our winners will receive a *Quantum* button and a copy of this issue. (The B228 winners received a bonus copy of the last issue as well.) And, as usual, everyone who sent in a correct solution was eligible to win a copy of our brainteaser collection *Quantum Quandaries.*

The next CyberTeaser awaits. Go to **www.nsta.org/quantum** and follow the links (no big challenge—consider it a little warm-up).

Infusium 23 Awards

A high school senior in a Texas border town is studying household and toxic waste. A Massachusetts teen with a pacemaker aspires to be a molecular biologist specializing in the effects of cardiovascular surgery. A 14-year-old Wisconsin student says she will be the first person to walk on Mars.

These three outstanding young women are representative of the 23 winners chosen last year from a field of 5,000 entries in the first annual Infusium 23 Women in Science Awards. This program recognizes high school females who plan to pursue science as a career. Applications for the second annual Infusium 23 Awards are now being accepted.

"Science and technology will increasingly provide the widest range of desirable jobs in the twenty-first century." says Johnna Doyle, Manager of Research and Development at Infusium and chairperson of the awards program. "Infusium 23 (a haircare line steeped in sciences) established this awards program to encourage young women to look at science beyond a specific science project or an extracurricular program that might earn them a merit badge, and to embrace career choices in science that have traditionally been categorized as male."

Infusium 23 will award 23 grants of $1,000 each, earmarked for female high school students to use to attend college, take summer classes, participate in extracurricular activities, or pursue some other educational opportunity. This award is open to all female students in high schools during the 1997–98 term. Applicants must submit an essay that includes

- what career in science she wants to pursue and why,
- extracurricular activities demonstrating career commitment, and
- the person or persons who have influenced her career choice and why.

There is no application form for this program. Completed statements or essays must be submitted along with the applicant's name, home address, phone number, age, and grade as of April 15, 1998, along with the name of her school, to Infusium 23 Women in Science Awards, 40 West 57th Street, 23rd Floor, New York, NY 10019. Applications must be postmarked by June 30, 1998.

Duracell/NSTA winners

Seven junior and senior high school students developed six inventions to claim top honors in the 16th Annual Duracell/NSTA Scholarship Competition. Over $100,000 in savings bonds is being awarded to more than 100 high school inventors, and this year's competition is recognizing over 1,500 U.S. students for their innovative ideas.

The seven first- and second-place winners, accompanied by their sponsoring teachers and parents, were honored in Las Vegas at the National Science Teachers Association's annual convention.

Kenneth Louie of Fort Salonga, New York, is this year's first-place winner of a $20,000 savings bond among 10th- to 12th-grade entries. His optical microphone, called **LightTalks**, works in areas prone to electromagnetic interference. Among 7th to 9th graders, the first-place $20,000 savings bond was awarded to a pair of 9th graders, Brandy Curry of Ottsville, Pennsylvania, and Jason Lamontagne of Kintnersville, Pennsylvania, for **Safety Seat Belt**, a portable and versatile device that ensures parents that their children's seatbelts are fastened.

Second place $10,000 savings bonds were awarded to 9th grader Whitney Blake from Manhasset, New York, for **The Child Safety Door Alarm**; 8th grader Benjamin Kendall of State College, Pennsylvania, for **Super Solder**; 12th grader Richard Barton from Gaithersburg, Maryland, for **Lapotron;** and 11th grader Jason Rolfe from Ridgewood, New Jersey, for **The Programmable Logic Emulator.**

In its 16th year, the Duracell/NSTA Scholarship Competition is also awarding 10 third-place $1,000 savings bonds, 24 fourth-place $500 bonds, and 60 fifth-place $200 bonds. All 1,542 students who entered the 1998 competition will receive a gift and a certificate of participation. Many of the 100 finalist entries will be displayed locally at conventions and science exhibits throughout the country.

"These results show the amazing creativity and scientific prowess of our nation's youth. The Duracell Competition proves that students are motivated when they are challenged to meet practical problems. This year's top winners also show that high school students can actively participate in significant R&D work," says Arthur Eisenkraft, competition judging chair.

A battery of winners

The top inventions cover a wide range of interests and expertise and demonstrate that our nation's youth can conduct scientific research, produce useful safety devices, and develop handy inventions for home, recreational, and industrial uses.

LightTalks, according to its inventor Louie, is a new concept. "Two years ago I learned from a physician that microphones could not be used in MRI (magnetic resonance imaging) rooms because they have too much interference. So this prompted me to find a way to develop one," the Kings Park High School senior says. Louie used a reflective mylar membrane, optical fibers and a white-light to produce the microphone that works not only in MRI areas but also in other locations with high electromagnetic interference, such as power stations or explosive environments. It is powered by four D-size Duracell batteries. Louie has applied to the U.S. Air Force Academy where he hopes to be a freshman in the fall. His sponsoring teacher is Jane Schoch.

In describing Safety Seat Belt, the first place team of Curry and Lamontagne say, "Our device will not only prevent children from getting seriously injured, it will also allow drivers to pay full attention to the road." Powered by two AA-size Duracell batteries, Safety Seat Belt sounds an alarm when a child is not seat-belted. The system can be shut off, but its switch is protected by a childproof cap. Curry and Lamontagne came up with the safety idea during their Palisades High School advanced learning class, taught by their sponsoring teacher, Pat Peterson.

Second-place winner Blake says she was motivated to develop The Child Safety Door Alarm to ensure that her young cousins would not fall into a family swimming pool. "It also works on a gate or a door and it can be left on at all times because it doesn't sound when a taller person passes," she says. Blake attends Manhasset High School in Manhasset, New York, and is sponsored by science research teacher, Peter Guastella. Her invention runs on two 9 V Duracell batteries.

Kendall got his winning idea as he watched his father solder some wires. "I thought it would be nice to have a way to heat wires and apply solder at the same time using only one hand," he says, "so that the other hand is free." His invention, Super Solder, is a "Why-didn't-I-think-of-that!" idea that features a small motor powered by two AA-size Duracell batteries attached to the side of the soldering gun.

The competition

To enter the Duracell/NSTA Scholarship Competition, 7th- through 12th-grade students design and build devices that are educational, useful, or entertaining, and powered by one or more Duracell batteries. Judging is based on creativity, practicality and energy efficiency of the invention, and clarity of the written description. Submissions are due at NSTA each January.

Sponsored by Duracell Inc. and administered by the National Science Teachers Association, the Duracell/NSTA Scholarship Competition has awarded over $750,000 in scholarships, savings bonds, and cash awards to over 850 students since 1983. Student inventors retain all rights to their devices.

For more information on the competition, call (888) 255-4242 or use our fax on demand service at (888) 400-6782 and, when prompted, ask for document 511. You can also visit the Duracell homepage at *www.nsta.org/programs/97/duracell.shtml.*

Massachusetts mathematics

PROMYS, the Program in Mathematics for Young Scientists, offers a lively mathematical environment in which ambitious high school students explore the creative world of mathematics. Through their efforts to solve challenging number theory problems, participants practice the art of mathematical discovery—numerical exploration, formulation and critique of conjectures, and techniques of proof and generalization.

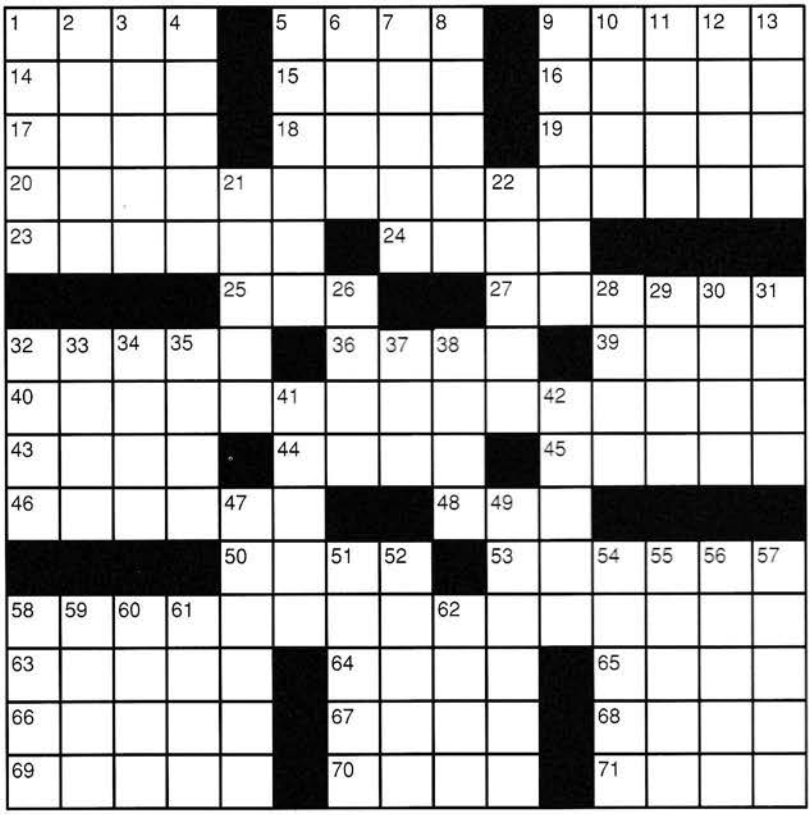
More experienced participants may also study hyperbolic geometry, mathematics of algorithms, and modular forms. Problem sets are accompanied by daily lectures given by research mathematicians with extensive experience in Professor Arnold Ross's longstanding Summer Mathematics Program at Ohio State University. In addition, a highly competent staff of 15 college-aged counselors lives in the dormitories and is always available to discuss mathematics with students. Each participant belongs to a problem-solving group that meets with a professional mathematician three times per week. Special lectures by outside speakers offer a broad view of mathematics and its role in the sciences.

PROMYS is a residential program designed for 60 ambitious students entering grades 10 through 12 held in Boston University from June 28 to August 8, 1998. Admission decisions are based on an applicant's solutions to a set of challenging problems included with the application packet; teacher recommendations; high school transcripts; and essays explaining their interest in the program. The approximate cost of room and board is $1400. Books may cost an additional $100. Tuition is $1500. Financial aid is available, and PROMYS is dedicated to the principle that no student will be unable to attend because of financial need.

PROMYS is directed by Professor Glenn Stevens. For applications, write to PROMYS, Department of Mathematics, Boston University, 111 Cummington Street, Boston, MA 02215; e-mail promys@math.bu.edu; or call (617) 353-2563. Applications will be accepted until June 15, 1998.

Criss × cross science

by David R. Martin



ACROSS

1 53,210 (in base 16)
5 River near Orford
9 Russian edict
14 Claim on property
15 Secular
16 ___ and raves
17 45,021 (in base 16)
18 Crazy
19 Aberdeen ___
20 Like some chemical reactions
23 Lustrous
24 503
25 Carbohydrate: suff.
27 Specialized secretors
32 Politician Stewart Lee ___
36 Netting
39 10: pref.
40 Like some nerves
43 Swedish botanist ___ Afzelius (1750–1837)
44 Placed
45 Desert fertile spot
46 African
48 ___ Paulo, Brazil
50 Jewish month
53 Slanted
58 $(NH_4)HCO_3$
63 Indicating NH_2
64 Reverberation
65 Tennis's Arthur
66 Unit of volume
67 Ursidae member
68 Plant starter
69 Group of cousins
70 Comedian Johnson
71 Goes astray

DOWN

1 Organism category
2 Quantity of alcohol
3 912,878 (in base 16)
4 1965 Physiol. Nobelist ___ Lwoff
5 Metal mixtures
6 Asian country
7 Cubed
8 Gram-negative aerobic bacteria
9 RNA base
10 German philosopher
11 Vessel: comb. form
12 Alphabet run
13 Exist in France
21 Compounds containing –CH=C(OH)–
22 ___ curve (lemniscate of Gerono)
26 Anarchist ___ Goldman
28 44,522 (in base 16)
29 Gains
30 602
31 Soft-walled cavities
32 Arrow poison
33 Pedestal part
34 Ancient Syria
35 Dalai ___
37 Upon: pref.
38 ___-serif (font)
41 Organic compound
42 ___'s law (stress proportional to strain)
47 City in Pakistan
49 To land
51 Unicellular protozoan
52 Competitor
54 Delete
55 More sage
56 Former hypothetical medium of space
57 Legal documents
58 Magnesium silicate
59 German painter ___ Nolde
60 Ramachandra's wife
61 52,717 (in base 16)
62 Talk

SOLUTION IN THE NEXT ISSUE

SOLUTION TO THE MARCH/APRIL PUZZLE

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

ANSWERS, HINTS & SOLUTIONS

Math

M231

Let's prove that the answer is $2 \cdot 1998 - 2 = 3994$. First we can point out a set that satisfies the conditions of our problem in which 3994 is the greatest number:

$$1998 - 1 = 1997, 1998, 1999, \ldots , 3994.$$

(There are 1998 terms in this progression, and the sum of any two of these numbers is greater than 3994.)

Next we show that it is impossible to find a set satisfying the conditions of the problem in which the greatest number is less than 3994. If there were such a set, then all the numbers that belonged to it would vary from 1 to 3993. Let $A \le 3993$ be the greatest of them. Then we can distribute all the numbers from 1 to A into the pairs

$$(1, A - 1), (2, A - 2), \ldots$$

(the sum of the numbers in each pair is A). There would be no more than $A/2 < 3994/2 = 1997$ such pairs. And there are 1997 different numbers that are less than A in the considered set. Thus there must be at least two numbers from one pair, and their sum will be equal to A, which contradicts the conditions of the problem.

M232

Consider parallelogram $BB'C'C$, which is equal to parallelogram $ABCD$ (fig. 1). Take point M' inside this parallelogram, such that $M'B = MA$ and $M'C = MD$. Then MM' is parallel to AB. In fact, triangles $BM'C$ and AMD are congruent because they have three equal sides, and thus $ABM'M$ is a parallelogram. Besides

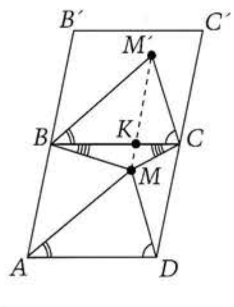


Figure 1

this, $\angle M'BC = \angle MAD = 40°$ and $\angle M'CB = \angle MAD = 70°$. Therefore

$$\angle M'BM = \angle M'BC + \angle MBC = 40° + 20° = 60°,$$

and

$$\angle M'CM = \angle M'CB + \angle MCB = 70° + 50° = 120°.$$

So $\angle M'BM + \angle M'CM = 180°$, and the quadrilateral $M'BMC$ is inscribed in a circle.

Now we see that the angles $\angle MBC$ and $\angle MM'C$ are equal since they subtend the same arc in a circle. Finally, if K is the point where MM' meets BC, then in the triangle $KM'C$ we know two angles: $\angle KM'C = 20°$ and $\angle KCM' = 70°$. So, $\angle M'KC = 180° - 70° - 20° = 90°$. However, because MM' is parallel to AB, $\angle BAD = \angle M'KC$, and all the angles of the parallelogram are right.

M233

We will solve this problem by setting up a geometric model. In our model, each fraction p/q, whether or not it is in lowest terms, corresponds to the point (q, p) on the coordinate plane. The reader can verify the following:

(1) Equivalent fractions lie on the same line through the origin. If the value of the fraction is k, this line has the equation $y/x = k$. Thus each rational number corresponds to a line through the origin.

(2) The fraction representing the rational number k in lowest terms is the integer point on the line $y/x = k$ that is closest to the origin.

(3) The fractions we are interested in (where q and p are positive, $q < 100$, and p/q is close to the fraction 5/8) correspond to points in the square $\{(x, y)$ such that $0 < x < 100, 0 < y < 100\}$ (that is, we don't have to worry about negative values, and we don't have to worry about fractions greater than 1).

(4) The fractions that are closest in value to 5/8 are those whose corresponding lines make the smallest angles with the line $y/x = 5/8$ (because we want the differences in the slopes to be as small as possible).

Now the point (5, 8) lies on the line $y/x = 5/8$, or $8y - 5x = 0$. We can solve our problem by considering all the lines parallel to this one; that is, all the lines whose equations are $8y - 5x = r$, for some number r. If such a line contains a point with integer coordinates (p, q), then r is certainly an integer. Furthermore, the closer r is to 0, the closer the line $8y - 5x = r$ is to the line $8y - 5x = 0$.

Now each point (p, q) is associated with two lines: the line $8y - 5x = r$, which is parallel to $y/x = 8/5$ (the point's "parallel"), and the line $y/x = p/q$, which goes through the origin (its "ray"). It is not hard to see that the fractions we want must be as far out as possible, along a parallel with the smallest r (in absolute value) as possible. Therefore, we should look on the lines where $r = \pm 1$.

Let us look at $r = +1$ first. We want integer points on the line $8y - 5x = 1$, and it is not difficult to see that the smaller angles will be made by rays of points that are further out on this parallel. By inspection, the point (3, 2) is on this line, so the points (3 + 8k, 2 + 5k) will be on this line as well (this is a standard technique of number theory), and we need the largest possible k. Since $y < 99$, we quickly find that the point we want is (99, 62), and the corresponding fraction is 62/99.

Similarly, if $r = -1$, the corresponding fraction is 58/93. The reader can check, using our geometric model, that in fact these points are the ones required by the problem.

M234

Let's transform this formula

$$\begin{aligned}
&\sin x \cos^2 y + \sin y \cos^2 x \\
&= \sin x\,(1 - \sin^2 y) + \sin y\,(1 - \sin^2 x) \\
&= \sin x + \sin y \\
&\quad - (\sin x \sin^2 y + \sin y \sin^2 x) \\
&= (\sin x + \sin y)(1 - \sin x \sin y) = 0.
\end{aligned}$$

The first parenthesis gives the equations $\sin x = -\sin y$ and $x = -y + 2\pi k$. This set consists of lines (fig. 2). The

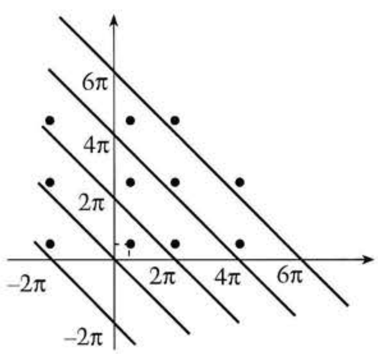


Figure 2

second parenthesis means that $\sin x = \sin y = 1$. Thus $x = \pi/2 + 2\pi m$ and $y = \pi/2 + 2\pi n$. This is the grid shown in fig. 2.

M235

First, we note that the area of the shadow depends only on the angle between the axis of the cylinder and the plane. Second, we note that the shadow consists of three parts: two shadows of semicircles (top and bottom of the cylinder) and the shadow of the axial section of the cylinder (fig. 3). Now, if α is the angle between the axis of the cylinder and the plane, then the angle between the plane of projection and the axial section is equal to α and the angle between the plane of projection and

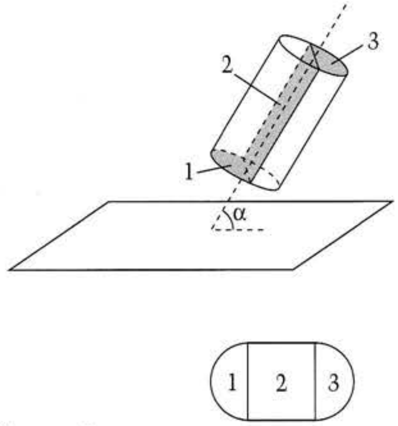


Figure 3

plane in which the top (bottom) circle lies is 90° – α. So, the area[1] of the shadow is

$$2\left(\frac{\pi r^2}{2}\right)\cos\left(90^\circ - \alpha\right) + h2r\cos\alpha$$
$$= \pi r^2 \sin\alpha + 2hr\cos\alpha.$$

Now it is not difficult to show that the maximal value of the expression $a \sin x + b \cos x$ is $\sqrt{a^2 + b^2}$. We concluded that the greatest possible area of the shadow is

$$\sqrt{\left(\pi r^2\right)^2 + (2hr)^2}.$$

To find the maximum value of $a \sin x + b \cos x$, note that there is an angle ϕ such that

$$\cos\phi = \frac{a}{\sqrt{a^2 + b^2}}$$

and

$$\sin\phi = \frac{b}{\sqrt{a^2 + b^2}}$$

(since the sum of the squares of these expressions is 1). Thus we can write $a \sin x + b \cos x$ as

$$\sqrt{a^2 + b^2}\,(\sin x \cos\phi + \cos x \sin\phi)$$
$$= \sqrt{a^2 + b^2}\,\sin(x + \phi).$$

[1]If the angle between two planes is α, and if the area of a figure on one of the planes is K, it is well known that the area of the orthogonal projection of the figure onto the other plane is $K \cos\alpha$.

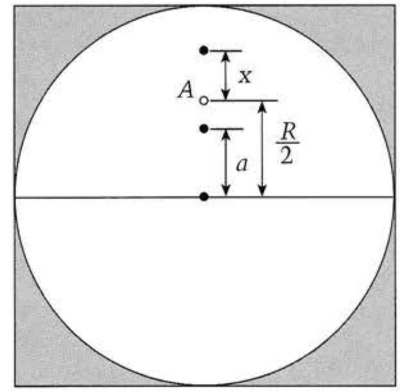


Figure 4

It follows that the maximum value of this expression is $\sqrt{a^2 + b^2}$, and it is achieved, for example, when $x + \phi = 90°$.

Physics

P231

The center of mass A of a square's half is located at the distance $R/2$ from its diameter (fig. 4). Let's denote by x the vertical distance from the corner's center of mass to point A. The ratio of the masses of the parts of any uniform sheet equals the ratio of their areas. Taking torques about point A (for a quarter of the square) yields

$$\left(R^2 - \frac{\pi R^2}{4}\right)x = \frac{\pi R^2}{4}\left(\frac{R}{2} - a\right),$$

from which we get

$$x = \frac{\pi\left(\dfrac{R}{2} - a\right)}{4 - \pi} = \frac{R\left(\pi - \dfrac{8}{3}\right)}{2(4 - \pi)} = 0.277R.$$

The distance from the center of

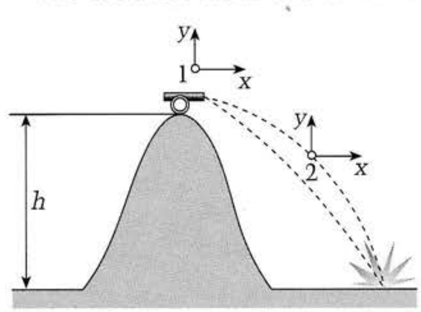


Figure 5

mass of the corner piece to the nearest corner of the square is about $0.316R$.

P232

Let's assume that both shells were fired by the same gun and consider their motion in the (laboratory) reference system *1* fixed to the gun (fig. 5). Up to the instant the second shot was fired, the first shell traveled $x_0 = v_0 t_0$ horizontally and $y_0 = g t_0^2/2$ vertically. At this instant its velocity was

$$v_1 = \sqrt{v_x^2 + v_y^2} = \sqrt{v_0^2 + (g t_0)^2},$$

and it was directed at an angle of $\alpha = \arctan(g t_0/v_0)$ to the horizon. Assume that the second shell was fired with velocity v_1 and at the angle α below the horizon. In this case both shells will move separated by the constant distance $s_0 = \sqrt{x_0^2 + y_0^2}$.

Now let's fix the dynamic system 2 to this hypothetical flight of the second shell and consider the flight of the actual shell in it. The second shell will overtake the first one in this dynamic system if it approaches with relative velocity $\mathbf{v}'$ along the straight line connecting the shells. Let's find its value and direction.

The approaching velocity cannot be too small because the flight time must be less than

$$t = \sqrt{\frac{2h}{g}}.$$

Thus,

$$v' \geq \frac{s_0}{t} = s_0 \sqrt{\frac{g}{2h}}.$$

The vector of relative velocity $\mathbf{v}'$ should form an angle α' with the horizon so that

$$\tan\alpha' = \frac{y_0}{x_0} = \frac{g t_0}{2 v_0}.$$

Now we return to the laboratory system *1*. Here the projections of the velocity $\mathbf{v}$ should be equal to

$$v_x = v_0 + v'\cos\alpha',$$
$$v_y = -gt - v'\cos\alpha',$$

where

$$v' \geq s_0 \sqrt{\frac{g}{2h}}.$$

The minimal value of velocity

$$v' = s_0 \sqrt{\frac{g}{2h}}$$

corresponds to the shells meeting near the Earth's surface, and the corresponding angle is

$$\alpha_0 = \arctan\left(\frac{v_y}{v_x}\right)\bigg|_{t=t_0} = 0.019 \text{ rad} = 1°5'.$$

which is read downward from the horizontal. In this case, the initial speed of the second must be

$$v = \sqrt{v_x^2 + v_y^2}\bigg|_{t=t_0} = 535.1 \text{ m/s}.$$

P233

Since the matter does not leak through the tunnel's wall, the mass flow (the mass of gas and drops passing through the tunnel's cross-section, including the tunnel's inlet and outlet) is

$$(\rho + \rho')uS = \left(\rho_0 + \rho_0'\right)u_0 S = \text{const.}$$

We took into account that the gas with density ρ and the drops with density ρ' move with the same speed. By the conditions of the problem, $\rho_0' = \rho_0$.

According to Bernoulli's equation (or Newton's second law of motion),

$$\left(\rho + \rho'\right)uS\Delta u = -\Delta P S.$$

Let's evaluate the pressure change ΔP along the tunnel. As the drops evaporate, the gas density increases, its temperature drops, and its speed decreases. (These qualitative descriptions will be confirmed at the end of the solution.) Therefore, the maximum speed change is $|\Delta u|_{\max} = u_0$ (it corresponds to complete arrest of the flow). The obtained equations result in

$$\left(\rho_0 + \rho_0'\right)u_0 \Delta u = -\Delta P,$$

from which we get

$$|\Delta P|_{\max} = \left(\rho_0 + \rho_0'\right)u_0^2 \cong 200 \text{ Pa}.$$

Thus, the largest possible change in pressure is three orders of magnitude less than atmospheric pressure. Therefore, we can assume the pressure to be constant along the stream. In such a case, the temperature of the evaporating drops can also be considered constant. Indeed, the temperature of boiling water at atmospheric pressure is constant, and the same is true for liquid nitrogen, whose boiling point (77 K) is far less than that of water. The ideal gas equation and constancy of pressure yield the reverse proportionality of density and temperature of the carrier gas:

$$\frac{\rho}{\rho_0} = \frac{T_0}{T}.$$

Each unit of mass transfers the kinetic energy of the chaotic molecular motion and the potential energy of their interaction. Assuming the carrier gas to be ideal, we can write its specific energy (energy per unit mass) as $c_p T$. Remember that in an ideal gas the energy of mutual interaction of the molecules is zero. The subscript "p" reflects the constant pressure in the tunnel. The same expression $c_p T'$ describes the energy of the gas evaporated from a drop with a temperature T'. In the condensed state (in the liquid drop) the energy equals $c_p T' - L$, where L is the latent heat of evaporation, which includes the potential energy of the interacting molecules.

The total energy of matter passing through any section of the tunnel per unit time (that is, the flow of total energy) does not vary:

$$c_p T \rho u S + (c_p T' - L)\rho' u S = \text{const.}$$

In this equation we neglect the kinetic energy of macroscopic movement of the mixture $(\rho + \rho')(u^2/2)uS$

because it is far less than the thermal energy:

$$u^2 << c_p T$$

because

$$10^2 << 10^3 \cdot 300 = 3 \cdot 10^5,$$

where

$$c_p \cong 10^3\ \text{J/(kg} \cdot \text{K)}.$$

We can express the density of the dispersed drops via the drop's radius:

$$\frac{\rho'}{\rho_0'} = \left(\frac{r}{r_0}\right)^3.$$

Now we can write all the dynamical parameters of the gas as functions of r/r_0:

$$\frac{u}{u_0} = \frac{1-\varepsilon_0 l}{1-\varepsilon_0 l\left(\frac{r}{r_0}\right)^3},$$

$$\frac{\rho}{\rho_0} = \frac{T_0}{T} = \frac{1+\varepsilon_0}{1+\varepsilon_0\left(\frac{r}{r_0}\right)^3}\frac{u_0}{u},$$

where

$$l = \frac{L}{c_p T_0} - \frac{T_0'}{T_0}, \text{ and } \varepsilon_0 = \frac{\rho_0'}{\rho_0}.$$

By the conditions of the problem, $\varepsilon_0 = 1$, $l \cong 0.4$. After complete evaporation, $r/r_0 = 0$. Finally, we have

$$\frac{u}{u_0} = 1 - l \cong 0.6.$$

Therefore,

$$u \cong 6 \text{ m/s}$$

and

$$\frac{\rho}{\rho_0} = \frac{T_0}{T} = \frac{2}{1-l} \cong 3.$$

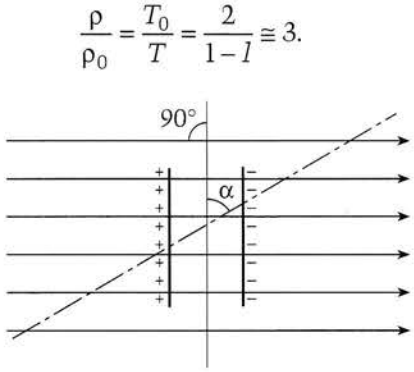


Figure 6

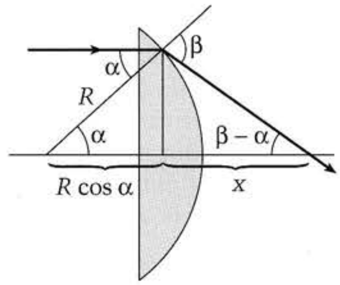


Figure 7

Therefore,

$$\rho \cong 3 \text{ kg/m}^3$$

and

$$T \cong 100 \text{ K}.$$

P234

Let's denote by U_0 the energy of the external uniform field in the absence of the capacitor, and by

$$U_C = \frac{\varepsilon_0 E_C^2 V}{2}$$

the energy of the charged capacitor ($\mathbf{E}_C$ is the electric field generated by the charges on the capacitor's plates, V is the volume of the capacitor). Before the capacitor was placed into the external field, the total energy of the system was

$$U = U_0 + U_C.$$

After the capacitor was placed in the position shown in fig. 6, the energy of the system became

$$U_1 = \frac{U_0 - \varepsilon_0 E_0^2 V}{2} + \frac{\varepsilon_0(\mathbf{E}_0 + \mathbf{E}_C)^2 V}{2},$$

where $\mathbf{E}_0$ is external electric field. The respective work W_1 equals the change in the system's energy:

$$W_1 = U_1 - U = \varepsilon_0 E_0 E_C V.$$

The rotation of the capacitor through angle α resulted in a change in the system's energy to the value

$$U_2 = \frac{U_0 - \varepsilon_0 E_0^2 V}{2} + \frac{\varepsilon_0 |\mathbf{E}_0 + \mathbf{E}_C|^2 V}{2}$$
$$= \frac{U_0 - \varepsilon_0 E_0^2 V}{2} + \frac{\varepsilon_0 E_0^2 V}{2} + \frac{\varepsilon_0 E_C^2 V}{2} + \varepsilon_0 E_0 E_C V \cos\alpha.$$

The work W_2 performed during the rotation is given by

$$W_2 = U_2 - U_1 = \varepsilon_0 E_0 E_C V(\cos\alpha - 1).$$

The work ratio we are looking for is

$$\frac{W_2}{W_1} = \frac{\varepsilon_0 E_0 E_C V(\cos\alpha - 1)}{\varepsilon_0 E_0 E_C V} = \cos\alpha - 1.$$

This solution is approximate because we neglected the redistribution of charges on the capacitor's plates during its rotation through angle α.

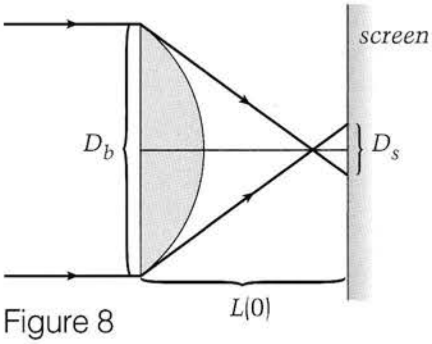


Figure 8

P235

Let's direct a parallel beam of light on the plane surface of the lens along its principal axis. We'll consider the refraction of an arbitrary ray of this beam and determine the distance from the center of the spherical surface to the point where the ray crosses the principal axis after refraction (fig. 7):

$$x = \frac{R\sin\alpha}{\tan(\beta-\alpha)}$$
$$= R\sin\alpha\frac{1+\tan\alpha\tan\beta}{\tan\beta - \tan\alpha},$$
$$\sin\beta = n\sin\alpha,$$
$$L(\alpha) = R\cos\alpha + x$$
$$= \frac{R}{\cos\alpha - \sqrt{\frac{1}{n^2} - \sin^2\alpha}}.$$

This expression shows that $L(\alpha)$ is a monotonic function of angle α: It decreases when the angle grows. The diaphragm restricts the angle to very small values—that is, to zero:

$$L(0) = F = \frac{R}{1-\frac{1}{n}} = 26.7 \text{ cm}.$$

The maximum angle α_{max} corresponds to the case when the beam hits the lens's rim. Geometry tells us that

$$\cos \alpha_{max} = 1 - d/R = 0.98,$$

and

$$L(\alpha_{max}) = 25.8 \text{ cm}.$$

Similarity of triangles (fig. 8) results in

$$D_s = D_b \frac{L(0) - L(\alpha_{max})}{L(\alpha_{max})} = \frac{2R\sin\alpha_{max}\left(L(0) - L(\alpha_{max})\right)}{L(\alpha_{max})} = 0.13 \text{ cm}.$$

Kaleidoscope

1. Large dust particles quickly settle down on the Earth, yet small dust is suspended in the atmosphere due to the chaotic motion of air molecules. There is no atmosphere on the Moon, so both large and small dust particles settle down on its surface almost simultaneously.

2. The partial pressure of water vapor will be enhanced near the humid soil. According to Dalton's law, the partial pressures of nitrogen (and oxygen) should be somewhat less over the humid soil than over the dry soil.

3. In accordance with the ideal gas law, any pair of three parameters (pressure, volume, and temperature) completely determine the state of gas.

4. The density decreases.

5. The less massive (and thus more mobile) hydrogen molecules more readily penetrate the partition, and thus increase the pressure in the air compartment. However, as the air crosses the partition, the pressures in the compartment equalize.

6. The readings of both manometers will be somewhat larger due to the additional weight of the gaseous column. The first manometer will read a lower pressure than the second one because the first column is shorter.

7. Weightlessness does not stop the chaotic motion of molecules in the spacecraft's "atmosphere."

8. No, it does not. The molecules, which move upward and hit the ceiling, are decelerated by the force of gravity. Thus, the collisions with the ceiling are not as "vigorous" as those with the floor.

9. No, it does not. Pressure is determined by the mean kinetic energy of the molecules, and this value doesn't depend on the wall's surface, provided the gas is in thermal equilibrium with it.

10. They are not equal. It is the mean kinetic energy of translational motion of the molecules that is equal in these gases. However, since nitrogen is a diatomic gas, the total kinetic energy of nitrogen includes the energy of molecular rotation, and consequently, it is larger than the total kinetic energy of neon.

11. No. The decrease in the kinetic energy of molecules near the cold wall is compensated by an increase in their concentration, and vice versa.

12. Since the gas does no work, its internal energy does not vary, and therefore, its temperature remains the same.

13. At the instant the vessel is dropped, the gas's density is larger at the bottom of the vessel than at the top. In the state of free fall, the molecules are distributed uniformly throughout the volume of the vessel. However, their total kinetic energy does not vary, so the gas's temperature remains the same.

14. The total kinetic energy of the gas motion as a whole transforms into the internal energy of the gas and the vessel, which means that the temperature of the gas will increase. Therefore, pressure also increases.

15. No, it doesn't. The air temperature is determined not by the velocity of the wind, but by the chaotic motion of its molecules that occurs independently of and in addition to the motion of the gas as a whole.

Microexperiment

Our thermal perception is not influenced by the total internal energy of the air (by the way, during the heating of the room the total internal energy does not vary—why?). However, we are very sensitive to temperature, which is determined by the mean kinetic energy of individual molecules. This value does increase during the heating of the room.

Brainteasers

B231

The short (hour) hand of a clock turns at a rate of 360°/12 hours = 30° per hour, or (1/2)° per minute. The long (minute) hand turns at the rate of 360°/60 minutes = 6° per minute. Thus, the angle between the short hand and the direction of 7 o'clock is 38 · (1/2)° = 19°. The angle between the long hand and the same direction is 38 · 6° – 210° = 18°. Therefore, the answer is 1°.

B232

Yes, it is possible. For instance, suppose there are 13 teams in the tournament. Say one team won five matches and lost seven, and all other matches ended in draws. Then the team would take first place (it would earn 5 · 3 = 15 points, and the other teams would at best have 11 · 1 + 3 = 14 points). If the points had been given as before, the winner would have had 2 · 5 = 10 points, and the rest of the teams would have had at least 11 · 1 = 11 points.

B233

See figure 9. We divide the grid into four 4 × 4 square sections and locate the one with the missing box. The remaining three 4 × 4 sections form one of our three pieces. Then we divide the fourth section into 2 × 2 square sections and locate the one with the missing box. The remaining

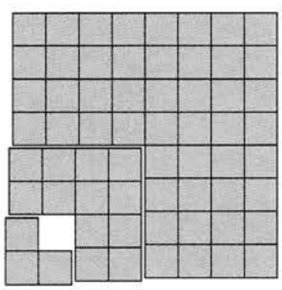


Figure 9

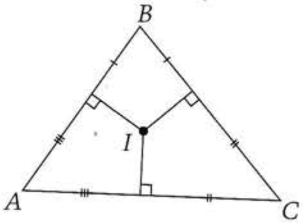


Figure 10

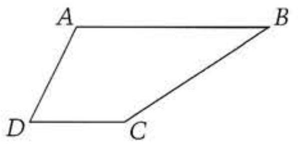


Figure 11

three 2×2 sections form the second of our three pieces, and the last piece is a 2×2 square with one box deleted.

With these three pieces we can form an 8×8 grid and place the missing square in any of the 64 possible locations by "spinning" each piece accordingly.

B234

See figure 10. Point I is the center of the inscribed circle. (This is just one of the possible solutions.)

B235

The body's weight decreased due to an increase in buoyancy.

At the Blackboard I

1. See figure 11. Since $\angle B$ and $\angle D$ are acute and $\sin \angle B = \cos \angle D$, then $m\angle B + m\angle D = 90$. But $m\angle A = 180 - m\angle D$, so

$$m\angle A - m\angle B$$
$$= (180 - m\angle D) - (90 - m\angle D) = 90.$$

2. We have $T = 90$ and $K = 6$. Since triangle APL is similar to triangle MNC,

$$\frac{AP}{PL} = \frac{MN}{NC},$$

giving

$$(AP)(NC) = (PL)(MN) = (PN)^2 = K^2 = 36.$$

Note that $AP \cdot NC$ is invariant.

3. We have $T = 36$ and $K = 6$. Letting a be the first term we have

$$\frac{(2a + K - 1)K}{2} = 75.$$

Therefore, K and $(2a + K - 1)$ are factors of 150. Possible factor pairs for (K, a) are $(2, 37)$, $(3, 24)$, $(5, 13)$, $(6, 10)$, and $(10, 3)$. We have $K = 6$, so $a = 10$.

Creence Lin's Relay

1. Figure 12 demonstrates that the answer is 6.

2. The length of the diagonal of the square is $6n$, so the square's area is $18n^2$. If C is the circumference, then $18n^2 = Cn$, so

$$C = 18n = 18 \times 6 = 108.$$

3. Let a be the current age of the aged tofu and r be the current age of the regular tofu. Then the time referred to in the problem was $a - r$ years (or months or days) ago, so the regular tofu's age was $r - (a - r) = 2r - a$. Then $a = 5(2r - a)$, which leads to $a = 5r/3$. Then $r + (5r/3) = 8$, which leads to $r = 3$.

1996 Super Relay

1. $N = (1/2)(2N)(N - 2)$, so $N^2 - 3N = 0$, and thus $N = 3$.

2. The slope m of the line equals

$$\frac{T^2 - 1}{1 - T} = -(T + 1).$$

Since $T = 3$, $m = -4$.

3. The x-coordinate of the vertex equals

$$\frac{-T^2}{2T} = \frac{-T}{2}.$$

Therefore, $C = -(-4)/2 = 2$.

4. $(2 + Ti)^2 = (4 - T^2) + 4Ti$. Thus $a + b = 4 + 4T - T^2$, which for $T = 2$ equals 8.

Braid #	Left	Center	Right
0	R	W	B
1	W	R	B
2	W	B	R
3	B	W	R
4	B	R	W
5	R	B	W
6	R	W	B

Figure 12

5. Modulo 10, $199^T \cdot 299^{T-1} = 9^{2T-1} = 9^{2T+1} = 81^T \cdot 9 = 9$. Thus the last digit is 9 and the value of T is irrelevant.

6. $(T - 1)^3 = K^3$, so $K = T - 1 = 9 - 1 = 8$.

7. The sum of the exterior angles is 360, the sum of the interior angles is $(T - 2)180$, so

$$(T - 2)180 - 360 = 180(T - 4).$$

For $T = 8$, the difference $= 720$.

8. $K = 720/60 = 12$. $A = K^2/\sqrt{3}$ so $A\sqrt{3}/9 = K^2/9 = 144/9 = 16$.

9. $\sin T \cos 286° - \cos T \sin(-106°) = -\sin T \cos 106° + \cos T \sin 106° = \sin(106° - T)$. Since $T = 16$, $\theta = 90$. We have $\cos 90 + i \sin 90 = i$.

10. $$\left(\frac{|i|}{i}\right)^2 = \frac{1}{i^2} = -1.$$

Note: For real numbers other than 0, the answer is 1; for pure imaginary numbers the answer is -1. Unfortunately for the clever ARML participant, both answers lead to nice results down the line.

11.

$$10(2x) - 5(3x - T) = 70 - 2(T + x)$$
$$5x + 5T = 70 - 2T - 2x$$
$$7x = 70 - 7T \to x = 10 - T.$$

Since $T = -1$, $x = 11$.

12.

$$\frac{\sin 9T}{\cos 9T} = \frac{\cos\theta}{\sin\theta}$$
$$\sin\theta \sin 9T = \cos\theta \cos 9T$$
$$\cos(\theta + 9T) = 0$$
$$\theta + 9T = 90 + 180K$$
$$\theta = 9(10 - T).$$

Since $T = 11$, $\theta = -9$.

13. If $y - 9 = T(x + 2)$, then for $y = 0$, $x = -9/T - 2$. Since $T = -9$, $x = 1 - 2 = -1$.

14. Multiplying the top equation by T and subtracting the bottom equation yields

$$(T + 4)y = T^2 + 6T + 8 = (T + 4)(T + 2).$$

Thus $y = T + 2$. So $x + T + 2 = T + 7$, giving $x = 5$. Then $x - y = 3 - T$. Since $T = -1$, $x - y = 4$.

15. Let n be the number of rows in the square. If n is even, the number of tiles in the diagonals is $2n$, if n is odd, the number is $2n - 1$. Since $2T^2 + 5$ is odd, we have

$$2n-1=2T^2+5 \to n=T^2+3.$$

Since $T=4$, $n=19$, and the number of tiles is 361.

At the Blackboard II

Problem 1. Let x be the car's velocity and let y be the biker's velocity, both in miles per minute. The time it takes the car to go from A to where it first meets the biker is the same in both cases and so is the time from the moment of this meeting till the car's arrival at B. Therefore the times that pass from the moment of first meeting till the moment when the car crosses the first meeting place for the last time are equal as well. We are going to calculate these times.

I. After the first meeting the car moves toward B for 3 minutes. So, it will take the car 3 minutes to return to the place where the meeting occurred. The biker will by this moment be $6y$ miles from this place. The car will be catching up with the bicycle with a velocity of $(x-y)$ miles/minute, so it will catch up with the bicycle $6y/(x-y)$ minutes later. The return to the first meeting place will also take $6y/(x-y)$ minutes, and the total time will equal

$$3+3+2\left(\frac{6y}{x-y}\right)=6+\frac{12y}{x-y} \text{ min.}$$

II. Analogously, we obtain

$$1+1+\frac{2\frac{15}{7}y}{x-\frac{15}{7}y}+\frac{2\frac{15}{7}y}{x-\frac{15}{7}y}=\left(2+\frac{60y}{7x-15y}\right) \text{min.}$$

Equating these expressions, we get

$$6+\frac{12y}{x-y}=2+\frac{60y}{7x-15y},$$

from which we obtain

$$7x^2-16xy-15y^2=0.$$

This is a homogeneous equation of the second degree in the two unknowns x and y. Putting $t=x/y$ (the required quotient), and solving the equation $7t^2-16t-15=0$, we find $t=3$ (the negative root is extraneous).

In this problem we had only one equation in two independent variables. Such occurrences are very rare. Usually some additional relations are imposed upon variables in a problem.

Problem 2. Introducing the new unknown quantities

$$u=\frac{x-2}{x+1}, v=\frac{x+2}{x-1},$$

we obtain the equation

$$20u^2-5v^2+48uv=0,$$

which is homogeneous and of the second degree in u and v. Now we divide this equation by v^2 (there are no solutions when $u=v=0$) and put $t=u/v$. Solving the equation $20t^2+48t-5=0$, we find $t_1=-5/2$ and $t_2=1/10$.

In the first case we get

$$\frac{u}{v}=-\frac{5}{2},$$
$$\frac{x-2}{x+1}:\frac{x+2}{x-1}=-\frac{5}{2},$$
$$\frac{x^2-3x+2}{x^2+3x+2}=-\frac{5}{2}.$$

This equation has no real roots.

In the second case

$$\frac{u}{v}=\frac{1}{10},$$
$$\frac{x^2-3x+2}{x^2+3x+2}=\frac{1}{10},$$
$$3x^2-11x+6=0,$$
$$x_1=3,$$
$$x_2=2/3.$$

Problem 3. Let $u=x-1$ and let $v=x^2+x+1$. The equation turns into $2v^2-7u^2=13uv$. We see that there are no solutions if $u=v=0$. Thus, we divide both parts of the equation by v^2. Let's introduce a new variable $t=u/v$. We get

$$7t^2+13t-2=0.$$

The roots of this equation are $t_1=1/7$, $t_2=-2$.

Taking $t=t_1$, we obtain

$$\frac{x-1}{x^2+x+1}=\frac{1}{7},$$
$$x^2-6x+8=0,$$
$$x_1=2,$$
$$x_2=4.$$

If $t=t_2$:

$$\frac{x-1}{x^2+x+1}=-2,$$
$$2x^2+3x+1=0,$$
$$x_3=-1,$$
$$x_4=-1/2.$$

Answer: $x_1=2$, $x_2=4$, $x_3=-1$, $x_4=-1/2$.

Problem 4. Introducing the additional variable $u=\sqrt{2x+15}$, we obtain the homogeneous equation $x^2+u^2=2xu$. We can either solve it as we did in problem 3, or notice that it can be written in the form $(x-u)^2=0$, from which we get $x=u$. From the equality $x=\sqrt{2x+15}$, we obtain $x^2-2x-15=0$, $x_1=-3$, $x_2=5$. The value $x=-3$ does not satisfy equation (3), but $x=5$ does.

Problem 5. If we were to let $u=x+1$ and $v=x-1$, we would get an equation that we might call "homogeneous of degree 1/3." To avoid these fractional powers, we let

(a) $u=\sqrt[6]{x+1}$, $v=\sqrt[6]{x-1}$, if $x\ge 1$.

(b) $u=\sqrt[6]{-x-1}$, $v=\sqrt[6]{-x+1}$, if $x\le -1$.

In the first case we arrive at the equation $2u^2-uv-v^2=0$. Dividing it by v^2, we find

$$\sqrt[6]{\frac{x+1}{x-1}}=1$$

(the second root of the equation $2t^2-t-1=0$ must be rejected);

$$\frac{x+1}{x-1}=1, x+1=x-1,$$

and thus the original equation has

no roots in the interval $x \geq 1$.

In the second case (when $x \leq -1$) we obtain $-2u^2 - uv + v^2 = 0$, from which we get

$$\sqrt[6]{\frac{x+1}{x-1}} = \frac{1}{2}, x = -\frac{65}{63}.$$

Since $-65/63 \leq -1$, this is a root of the original equation.

Problem 6. The first equation of the system is homogeneous of degree 2. If we divide both its parts by y^2 (since $x = y = 0$ is not a solution of the system) and put $t = x/y$, we arrive at the quadratic equation

$$3t^2 - 2t - 1 = 0,$$

from which we get $t_1 = 1$, $t_2 = -1/3$. Thus we have either $y = x$ or $y = -3x$. First substituting $y = x$ and then $y = -3x$ in the second equation, we find four solutions:

$$x_1 = y_1 = 1,$$
$$x_2 = y_2 = -6,$$
$$x_{3,4} = \frac{15 \mp \sqrt{249}}{2},$$
$$x_{3,4} = \frac{-45 \pm \sqrt{249}}{2}.$$

Problem 7. The essential point of this system is that all its non-constant terms are of the second degree. To get rid of the constant term, we multiply the second equation by 2 and subtract it from the first. Consider the system consisting of the resulting equation and the second equation of the original system:

$$\begin{cases} x^2 - \dfrac{25}{12}xy + y^2 = 0, \\ x^2 + y^2 = 25. \end{cases} \qquad (5)$$

System (5) is equivalent to system (4).

The first equation of system (5) gives $t_1 = 3/4$ and $t_2 = 4/3$, where $t = x/y$. Substituting $x = (3/4)y$ and $x = (4/3)y$ in the second equation, we find four solutions:

$$x_{1,2} = \pm 3, y_{1,2} = \pm 4, x_{3,4} = \pm 4, y_{3,4} = \pm 3.$$

The first equation of system (5) defines the lines $x = (3/4)y$ and $x = (4/3)y$, and the second equation of system (5) is an equation for a circle of radius 5 and center at the origin of coordinates. The solutions of this system are the coordinates of the points where they meet.

Problem 8. Multiplying the first equation of the system by 2 and subtracting the second equation from it, we obtain the system

$$\begin{cases} 3x^2 - 8xy + 4y^2 = 0, \\ x^2 + 2xy - 2y^2 = 6, \end{cases}$$

where the first equation is homogeneous. Its solutions are $x = 2y$ and $x = (2/3)y$. Substituting them in the second equation, we find two solutions: $x_{1,2} = \pm 2$, $y_{1,2} = \pm 1$.

Problem 9. This equation is homogeneous of the first degree with respect to $\sin x$ and $\cos x$. We don't lose any roots if we divide both its parts by $\cos x$. We have

$$2\tan x + 3 = 0,$$

$$\tan x = -3/2,$$

and

$$x = -\arctan(3/2) + \pi k$$

(for any integer k).

The equation

$$a \sin^2 x + b \sin x \cos x + c \cos^2 x = 0$$

can be transformed into

$$a \tan^2 x + b \tan x + c = 0.$$

Problem 10. We do not lose any roots if we divide the equation by $\cos^2 x$. We get the equation

$$\tan^2 x + 3 \tan x + 2 = 0,$$

from which we get $\tan x_1 = 1$, $x_1 = \pi/4 + \pi k$; $\tan x_2 = 2$, $x_2 = \arctan 2 + \pi m$ (for any integer k, m).

Problem 11. Using the hint in the text, we have

$$\sin^2\frac{x}{2} - 3\sin\frac{x}{2}\cos\frac{x}{2} - 4\cos^2\frac{x}{2} = 0,$$

$$\tan^2\frac{x}{2} - 3\tan\frac{x}{2} - 4 = 0,$$

$$\tan\frac{x_1}{2} = 4, x_1 = 2\arctan 4 + 2\pi k,$$

$$\tan\frac{x_2}{2} = -1, x_2 = -\frac{\pi}{2} + 2\pi m$$

(for any integer k, m).

We can also solve this equation by relating it to the formula for $\sin(A + B)$. We can do this by multiplying and dividing by $\sqrt{34}$:

$$3\sin x + 5\cos x$$
$$= \sqrt{34}\left(\frac{3}{\sqrt{34}}\sin x + \frac{5}{\sqrt{34}}\cos x\right)$$
$$= \sqrt{34}\sin\left(x + \arcsin\frac{5}{34}\right),$$

$$\sin\left(x + \arcsin\frac{5}{34}\right) = -\frac{3}{\sqrt{34}},$$

$$x = (-1)^{m+1}\arcsin\frac{3}{\sqrt{34}}$$
$$-\arcsin\frac{5}{34} + \pi m$$

(for any integer m).

Problem 12. If we multiply the righthand side of this equation by $1 \equiv \sin^2 x + \cos^2 x$ and then divide both parts by $\cos^3 x$, we'll obtain the equation

$$3 \tan^3 x - \tan^2 x - \tan x - 1 = 0.$$

We can guess (and check) that $\tan x = 1$ is a root of this equation. So, if we factor the polynomial in $\tan x$, one of the factors will be $\tan x - 1$. These considerations allow us to write:

$$(\tan x - 1)(3 \tan^2 x + 2 \tan x + 1) = 0,$$

which leads to

$$x = \pi/4 + \pi k \text{ (for any integer } k).$$

Problem 13. Putting $2^x = u$ and $7^x = v$, we get a homogeneous equation of degree 2:

$$u^2 - 2uv - 3v^2 = 0.$$

Solving, we find

(a) $u/v = -1$, $(2/7)^x = -1$ (there are no roots in this case);

(b) $u/v = 3$, $(2/7)^x = 3$, $x = \log_{2/7}3$.

Bingo

Problem 1. There are 10 numbers that can go in the first position. Since the numbers on a card must be distinct, there are nine remaining possibilities for the second position, for a total of $10 \cdot 9 = 90$. However, the cards $\{1, 2\}$ and $\{2, 1\}$ are considered to be the same card, since the first card wins exactly when the second card wins. So we must divide by 2 to get an answer of 45.

Problem 2. By the same argument as above, the answer is $[n(n-1)]/2$.

Problem 3. The cards $\{1, 3\}$ and $\{2, 3\}$ are the only cards that contain the number 3, so the answer is 2. Remember that $\{3,1\}$ is the same card as $\{1, 3\}$.

Problem 4. Since there are n balls and we don't allow duplicate numbers on a card, $n-1$ numbers can go with 1 to complete a card. So the answer is $n-1$. Similarly, there are $n-1$ cards that contain 2.

For $3 \le i \le n$, only 2 cards contain i: $\{1, i\}$ and $\{2, i\}$.

Problem 5. There are $n-1$ cards containing the number 1. Similarly, $n-1$ cards contain the number 2. Every card contains either a 1 or a 2, so we've counted every card. The card $\{1, 2\}$ was counted twice, so the total is $(n-1)+(n-1)-1 = 2n-3$.

Problem 6. There are n possible balls that can be drawn first and $n-1$ balls that can be drawn second. Since order doesn't count, the first two balls can be drawn in $[n(n-1)]/2$ possible ways. Since only one combination leads to a win for the card $\{1, 2\}$, this card's winning probability is $2/[n(n-1)]$. To compare this with $1/(2n-3)$, we subtract

$$\frac{1}{2n-3}-\frac{2}{n(n-1)} = \frac{(n^2-n)-(4n-6)}{n(n-1)(2n-3)} = \frac{(n-2)(n-3)}{n(n-1)(2n-3)},$$

which is positive if $n \ge 4$. This shows that the card $\{1, 2\}$ is doing worse than it would if the game were fair. Note that when n is large, this difference is approximately $1/(2n-3)$, which means that the winning chance of the card $\{1, 2\}$ is negligible compared to what would be fair.

Problem 7. By a symmetry argument, every card other than $\{1, 2\}$ has the same winning probability. This is the main idea of the section "Using Symmetry."

Problem 8. The sum of the winning probabilities for each card must total 1 because there is always a winner and because for a tie the prize is shared. We know that there are $2n-3$ cards, and the card $\{1, 2\}$ has a winning probability of $2/n(n-1)$. There are $2n-4$ cards remaining. Since each of these cards has the same winning probability, let x be this probability. Then,

$$(2n-4)x+\frac{2}{n(n-1)}=1.$$

Solving for x yields

$$x=\frac{n+1}{2n(n-1)}.$$

Subtracting $1/(2n-3)$ from this, we see that these cards are doing slightly better than what would be fair.

Problem 9. There are n balls, so $n-1$ numbers can go with 1 to complete a card.

Problem 10. The ball 3 appears on the cards $\{1, 3\}$ and $\{3, 4\}$. A similar argument works for ball 4. Every other number $i \ne 1, 3, 4$ appears only on the card $\{1, i\}$.

Problem 11. Adding the card $\{3, 4\}$ to the collection of $n-1$ cards containing a 1 gives a total of n cards.

Problem 12. Every card other than $\{3, 4\}$ contains a 1.

Problem 13. Every card either contains a 1 or is the card $\{3, 4\}$. Once a 1 and some other number i is drawn, then the card $\{1, i\}$ wins. Once both 3 and 4 are drawn, then the card $\{3, 4\}$ wins.

Problem 14. There are six permutations since $3 \cdot 2 \cdot 1 = 6$. Because each of these three balls is equally likely to be drawn last among these three balls, the probability that 1 is drawn last is 1/3.

Problem 15. The card $\{1, 2\}$ has a better winning probability than $\{1, 3\}$, because $\{1, 3\}$ has a number in common with the "big winner" card. By a symmetry argument, $\{1, 2\}$ and $\{1, 5\}$ have the same winning probability, as does any pair of cards not including $\{1, 3\}$, $\{1, 4\}$, or $\{3, 4\}$. Similarly, $\{1, 3\}$ and $\{1, 4\}$ have the same winning probability.

Problem 16. The winning probability of the "big winner" card $\{3, 4\}$ is 1/3. The winning probability of the cards $\{1, 3\}$ and $\{1, 4\}$ is

$$\frac{1}{n}-\frac{n-3}{2n(n-1)}.$$

The winning probability of the cards $\{1, i\}$ for $i = 2, 5, 6, 7, \ldots, n$ is

$$\frac{1}{n}-\frac{n-4}{3n(n-1)}.$$

To compute this, write out the ways that each card can win alone, win in a 2-way tie, win in a 3-way tie, ..., win in a j-way tie. Then add, remembering that tied cards share the prize. Also, recognize the appearance of triangular numbers in the case of $\{1, 3\}$ and $\{1, 4\}$, and the appearance of pyramidal numbers in the case of $\{1, i\}$ for $i = 2, 5, 6, 7, \ldots, n$. Note that every card other than the "big winner" is doing worse than what would be fair.

Problem 17. Figure 13 gives the solution. In this example, there are n balls and $2n-2$ players. The winning probability of the "big loser" card $\{1, 2\}$ is $2/[n(n-1)]$, and the winning probability of the "big winner" card $\{3, 4\}$ is 1/6. Computing the winning probabilities of the other cards is difficult.

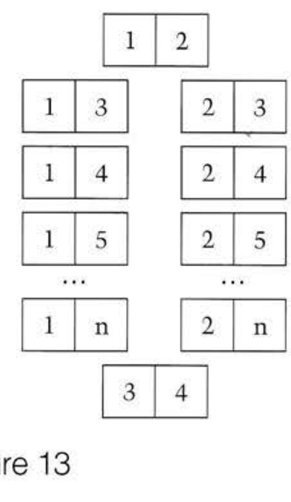


Figure 13

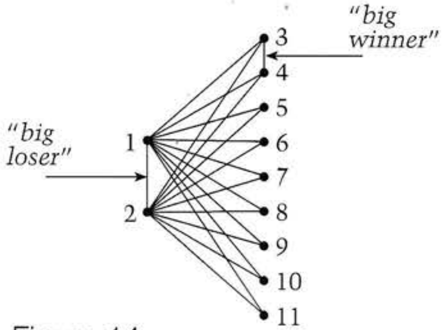


Figure 14

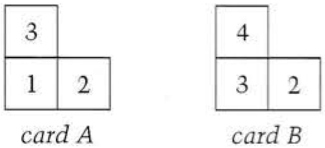


Figure 15

The next section of the article considers graphs of these games. The graph for this Bingo game is shown in figure 14.

Problem 18. Every number appears on exactly two cards.

Problem 19. The four cards {1, 2}, {2, 3}, {3, 4}, and {1, 4} each have a winning probability of 16/105. The three cards {5, 6}, {6, 7}, and {5, 7} each have a winning probability of 16/105.

Problem 20. Just check that the edges match. For example, {1, 3} in the graph on the left is matched with {1, 2} on the right.

Problem 21. To show that {1, 2} and {2, 3} have the same winning probability, consider the congruence of the graph with itself that sends (1, 2, 3, 4, 5) to (2, 3, 4, 5, 1). A similar argument shows that any other pair of edges has the same winning probability.

Problem 22. By a symmetry argument, the six cards with entries in (1, 2, 3, 4) each have the exact same winning probability of 17/252. Similarly, the 10 cards with entries in (5, 6, 7, 8, 9) have the same winning probability of 5/84. To do the computation, note that each card has a 1/36 chance of winning when the first two balls are drawn. Then, some card with entries in (1, 2, 3, 4) wins when three balls are drawn with probability

$$\frac{5}{9}\cdot\frac{4}{8}\cdot\frac{3}{7}\cdot 2=\frac{5}{21}.$$

By symmetry, the contribution due to each card is

$$\frac{1}{6}\cdot\frac{5}{21}=\frac{5}{126}\text{ and }\frac{1}{36}+\frac{5}{126}=\frac{12}{252}.$$

A similar argument works for the 10 cards with entries in (5, 6, 7, 8, 9).

Problem 23. The corner number 1 is involved in two winning combinations, and the two edge numbers 2 and 3 are involved in only one winning combination. This is not as symmetric as 2×1 Bingo, where each number on a card is involved in only one winning combination.

Problem 24. The center square is involved in four winning combinations, but it is a free space. The 8 numbers on the main diagonals are involved in three winning combinations, one horizontal, one vertical, and one diagonal. The remaining 16 numbers are involved in two winning combinations, one horizontal and one vertical.

Problem 25. The are five possibilities for the corner number, which is involved in two winning combinations. There are $4 \cdot 3 = 12$ remaining ordered pairs possible for the two edge numbers, but their order doesn't matter, so there are only six possibilities. The answer is $(5 \cdot 4 \cdot 3)/2 = 30$. Similarly, for n balls, the answer is $[n(n-1)(n-2)]/2$.

Problem 26. Enumerating all 24 orders in which the four balls can be drawn, we see that card *A* wins 12 times out of 24, card *B* wins 10 times out of 24, and there is a tie 2 times out of 24. Therefore, card *A* has the advantage. An alternative justification uses the asymmetry of the L-shaped Bingo card because card *A*'s corner number, 1, doesn't appear on card *B* at all, whereas card *B*'s corner number, 3, appears on card *A*.

Problem 27. Note that the set of winning combinations for each player is the same in either model.

Problem 28. The pair of cards *A* and *B* is the pair in figure 15. The other two pairs are equivalent to the pair in figure 15, using an appropriate renumbering of the balls. When all three cards are played together, a symmetry argument shows that they all have the same winning chance.

Problem 29. We list a few of the differences. A normal 5×5 Bingo card has 12 winning combinations, some having 4 numbers and some having 5 numbers, due to the free space. There are restrictions on which numbers appear in which columns. There are only 75 balls, rather than an arbitrary number n. A Bingo

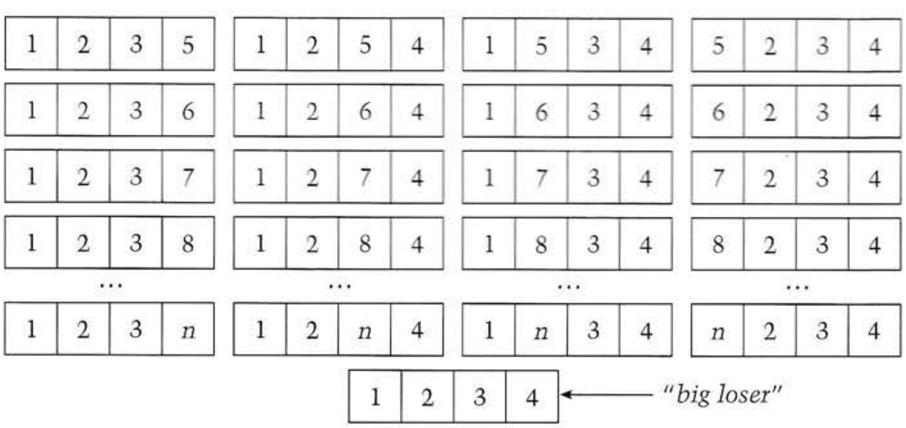


Figure 16

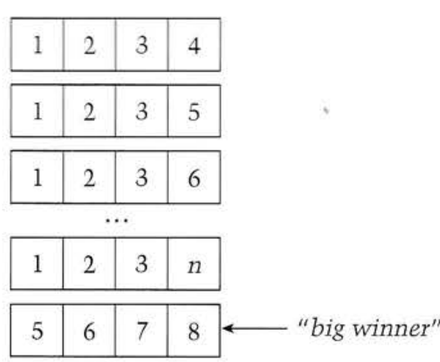


Figure 17

1	16	31	$45+i$	61
2	$15+i$	32	47	62
3	18	*free*	48	$60+i$
4	19	$30+i$	49	64
i	20	35	50	65

i	16	31	46	61
2	17	32	$45+i$	62
3	$15+i$	*free*	48	63
4	19	34	49	$60+i$
5	20	$30+i$	50	65

1	16	31	46	$60+i$
2	17	$30+i$	47	62
i	18	*free*	48	63
4	19	34	$45+i$	64
5	$15+i$	35	50	65

1	16	$30+i$	46	61
i	17	32	47	62
3	18	*free*	$45+i$	63
4	$15+i$	34	49	64
5	20	35	50	$60+i$

1	16	31	46	61
2	17	32	47	62
3	18	*free*	48	63
4	19	34	49	64
5	20	35	50	65

"big loser"

Figure 18

parlor does not return all of the entry fees as prizes, which significantly cuts into any potential advantage. People may not want to wait for us to examine every possible card and compute the optimal card. Ties are not shared. The person who calls "Bingo!" first wins.

Problem 30. The card {1, 5} is better than {1, 3}. In some sense, {1, 3} is more crowded than {1, 5}.

Problem 31. With three or five players, the game is unfair. With four or six players, the game is fair. To see this, draw the graphs. Remember that a player who can improve his or her winning probability by changing cards will switch. Is there a player that is always at an advantage?

Problem 32. The player at a disadvantage will switch cards, which causes the other player to want to switch, so there is an infinite cycle. To study what happens as players switch cards, we must have a more precise definition of the problem. We assume that players are asked in turn if they want to switch. A player follows a greedy strategy, switching to the card that is best based on what the other players currently have. If two cards are equally good, the player chooses according to some fixed rule. Now it is possible to study this problem for various collections of Bingo cards. Note that the problem is even harder if we allow players to form coalitions or to enter and leave the game.

Problem 33. We think that these two examples are the most biased 2×1 Bingo games possible, but we have no proof. What are the most biased examples for other variants of Bingo?

Problem 34. Try using symmetry. Are there any asymmetric, fair 2×1 Bingo games (that is, fair 2×1 Bingo games that are proven to be fair by some means other than symmetry)? What does the previous question mean?

Problem 35. Try to work specific examples, and then generalize them.

Problem 36. It might be useful to experiment with the help of a computer.

Problem 37. Duplicate cards or duplicate winning combinations allow for an arbitrarily unfair game. For this reason, when we're considering the fairness of a Bingo game, we should make sure that none of the cards being considered contain a duplicate winning combination. In L-shaped Bingo and in normal 5×5 Bingo, cards may not be exact duplicates, but share one or more winning combinations. In 5×5 Bingo, we also have to worry about a winning combination on one card being a proper subset of a winning combination on another card because of the free space.

Problem 38. To get a graph for an L-shaped Bingo game, we would have to use two edges to represent each card. Working with different colors for each card might be useful. In normal 5×5 Bingo, there are more obstacles since a winning combination must connect 4 or 5 numbers.

Problem 39. Figure 16 contains every 4×1 Bingo card containing at least three of {1, 2, 3, 4}. Figure 17 contains every 4×1 Bingo card containing 1, 2, and 3, along with the card {5, 6, 7, 8}.

These examples might be the most unfair. What would be an example with both a "big loser" card and a "big winner" card? What are the corresponding examples for 5×1 Bingo?

1	16	31	$45+i$	61
2	$15+i$	32	46	62
3	17	*free*	47	$60+i$
4	18	$30+i$	48	63
i	19	33	49	64

6	20	35	50	65
7	21	36	51	66
8	22	*free*	52	67
9	23	37	53	68
10	24	38	54	69

"big winner"

Figure 19

1	16	31	46	61
2	17	32	47	62
3	18	*free*	48	63
4	19	33	49	64
5	20	34	50	65

card A

2	16	31	46	61
6	17	35	47	66
3	18	*free*	48	63
4	21	33	49	64
5	20	34	51	65

card B

Figure 20

Problem 40. In figure 18, i varies from 6 to 15. In figure 19, i varies from 5 to 15.

Figure 20 uses the fact that the corner number is involved in three winning combinations, whereas the edge numbers are involved in two winning combinations. This is analogous to the example given in figure 15 for L-shaped Bingo. The main idea is that we embed an L-shaped Bingo card in the upper-left corner.

It is probably possible to construct examples that are more unfair than these.

Problem 41. One method for creating an approximately fair Bingo game would be to make sure that the numbers on the cards are uniformly distributed, but that would not be enough to create a perfectly fair Bingo game. To set up an unfair Bingo game, you could create a "big winner" card and give it to someone who you wanted to win.

Problem 42. Define a winning combination to be a set of numbers and define a card to be a set of winning combinations. No winning combination on one card may be equal to or a proper subset of a winning combination on another card, for the reasons given in the solution to problem 37.

Gradus

Problem 2. We have

$$\alpha^2 + 2\alpha\beta + \beta^2 = (\alpha + \beta)^2 = 3^2 = 9.$$

Problem 3. Method I: We have

$$\alpha^2 + \beta^2 = (\alpha + \beta)^2 - 2\alpha\beta = (-3)^2 - 2(-5) = 19.$$

Method II: Since α and β are roots of the given equation, we know that

$$\alpha^2 - 3\alpha - 5 = 0,$$
$$\beta^2 - 3\beta - 5 = 0.$$

Adding, we find that

$$\alpha^2 + \beta^2 - 3(\alpha + \beta) - 10 = 0,$$

or

$$\alpha^2 + \beta^2 - 3(3) - 10 = 0$$

(since $\alpha + \beta = 3$) or

$$\alpha^2 + \beta^2 = 9 + 10 = 19.$$

Problem 4. Using either method I or method II from problem 3, we find that $\alpha^2 + \beta^2 = p^2 - 2q$.

Problem 5. We have

$$\frac{1}{\alpha} + \frac{1}{\beta} = \frac{\alpha+\beta}{\alpha\beta} = \frac{p}{q}.$$

Problem 6. Method I: Expand $(\alpha + \beta)^3$, and use the techniques of problem 3, method I.

Method II: We already know the value of $\alpha^2 + \beta^2$ in terms of p and q: It is $p^2 - 2q$. We can use this, and method II of problem 3, to "bootstrap" our computation. For any value of x such that $x^2 - px + q = 0$, we know that $x^3 - px^2 + qx = 0$ as well. Since α and β satisfy the original equation, we have

$$\alpha^3 - p\alpha^2 + q\alpha = 0,$$
$$\beta^3 - p\beta^2 + q\beta = 0.$$

Adding, we find that

$$(\alpha^3 + \beta^3) - p(\alpha^2 + \beta^2) + q(\alpha + \beta) = 0,$$

or

$$(\alpha^3 + \beta^3) = p(\alpha^2 + \beta^2) - q(\alpha + \beta) = p(p^2 - 2q) - pq = p^3 - 3pq.$$

Problem 7: Adding the two fractions, we find that their value is simply -1 (and is independent either of α and β or of any equation they satisfy).

Problem 9. Following problem 3, method I, we find that

$$(\alpha + \beta + \gamma)^2 = \alpha^2 + \beta^2 + \gamma^2 + 2(\alpha\beta + \beta\gamma + \gamma\alpha),$$

or

$$p^2 = \alpha^2 + \beta^2 + \gamma^2 + 2q.$$

It follows that

$$\alpha^2 + \beta^2 + \gamma^2 = p^2 - 2q.$$

Compare this with the corresponding expression for the sum of the squares of the roots of a quadratic equation.

Problem 10. We follow problem 3, method II:

$$\alpha^3 - p\alpha^2 + q\alpha - r = 0,$$
$$\beta^3 - p\beta^2 + q\beta - r = 0,$$
$$\gamma^3 - p\gamma^2 + q\gamma - r = 0.$$

Adding, we have

$$\begin{aligned}\alpha^3 + \beta^3 + \gamma^3 &= p(\alpha^2 + \beta^2 + \gamma^2) - q(\alpha + \beta + \gamma) + 3r\\ &= p(p^2 - 2q) - pq + 3r\\ &= p^3 - 3pq + 3r.\end{aligned}$$

Problem 11. We can "bootstrap" this to find the sums of higher powers of α, β, and γ. (In fact, we have been beaten to this idea by Sir Isaac Newton, who is credited with a general formula for the sums of powers of roots of a polynomial equation.)

Suppose α, β, and γ are roots of

$$x^3 - px^2 + qx - r = 0.$$

Then $x^4 - px^3 + qx^2 - rx = 0$ as well, so

$$\alpha^4 - p\alpha^3 + q\alpha^2 - r\alpha = 0,$$
$$\beta^4 - p\beta^3 + q\beta^2 - r\beta = 0,$$
$$\gamma^4 - p\gamma^3 + q\gamma^2 - r\gamma = 0.$$

Adding, we obtain

$$\alpha^4 + \beta^4 + \gamma^4 - p(\alpha^3 + \beta^3 + \gamma^3) + q(\alpha^2 + \beta^2 + \gamma^2) - r(\alpha + \beta + \gamma) = 0$$

or

$$\alpha^4 + \beta^4 + \gamma^4 - p(p^3 - 3pq + 3r) + q(p^2 - 2q) - rp = 0.$$

This gives our required expression:

$$\alpha^4 + \beta^4 + \gamma^4 = p^4 - 4p^2q + 4pr + 2q^2.$$

Problem 12. The given expression equals

$$\frac{\alpha^2\gamma^2 + \beta^2\gamma^2 + \gamma^2\beta^2}{\alpha^2\beta^2\gamma^2}.$$

The denominator is clearly r^2. The numerator looks like it is related to q^2. Indeed, we have

$$q^2 = (\alpha^2\beta^2 + \beta^2\gamma^2 + \alpha^2\gamma^2) + 2\alpha\beta\gamma(\alpha + \beta + \gamma).$$

Using this information, we quickly find that our required expression equals

$$\frac{q^2 - 2pr}{r^2}.$$

Problem 13. The given expression

equals

$$\frac{\alpha^2\beta^2+\beta^2\gamma^2+\alpha^2\gamma^2}{\alpha\beta\gamma}.$$

Using the computations from problem 12, we find that this is

$$\frac{q^2-2pr}{r}.$$

Problem 14. This system of equations is not difficult to solve, for example by substitution. But we can make the solution easier still by relating it to the previous discussion. We can think of forming a quadratic equation with roots α and β. It will be $x^2 - 5x + 6 = 0$. But this equation solves easily by factoring, and its roots are 2 and 3. Hence $\alpha = 2$ and $\beta = 3$, or $\alpha = 3$ and $\beta = 2$. Note that the solution turns on the fact that the equations are symmetric in α and β.

Problem 15. We can proceed as in example 14. The equation satisfied by α and β is $x^2 - 8x - 7 = 0$, and this time it doesn't factor over the integers, so we use the quadratic formula to find that $\alpha = 4 + \sqrt{23}$, $\beta = 4 - \sqrt{23}$ (or vice versa).

Problem 16. Let the numbers α, β, γ be solutions to the equation

$$x^3 + 9x^2 + 19x + 11 = 0.$$

By inspection, $x = -1$ is a solution, so $(x + 1)$ is a factor. Division reveals that the other factor is $x^2 + 8x + 11$. Setting this factor equal to 0, we find that the other roots are $-4 \pm \sqrt{5}$. But which value is α? Again, the symmetry of the original equations tells us that α could be any of the numbers -1, $-4 + \sqrt{5}$, $-4 - \sqrt{5}$, if β and γ are the other two. There are six solutions in all.

Problem 17. We would like to write down a cubic equation whose roots are α, β, and γ. But we need the value of $q = \alpha\beta + \beta\gamma + \alpha\gamma$. We can use the method of problem 9. If the equation is

$$x^2 - px^2 + qx - r = 0,$$

then $p = 5$, $r = -28$, and

$$\alpha^2 + \beta^2 + \gamma^2 = p^2 - 2q = 29.$$

It follows that $q = -2$.

Now we know that the equation we seek is

$$x^3 - 5x^2 - 2x + 24 = 0.$$

As in problem 16, we can find one root by inspection, then the others by division. We find that the values of α, β, and γ are -2, 3, 4 (in any order).

Problem 18. Let

$$P(x) = x^5 + px^4 + qx^3 + rx^2 + sx + t$$

be the polynomial with lead coefficient 1 whose roots are a, b, c, d, and e. Then

$$p = -(a + b + c + d + e)$$

and

$$2q = (a + b + c + d + e)^2 - (a^2 + b^2 + c^2 + d^2 + e^2),$$

so p and q are both divisible by n. Writing algebraically that a, b, c, d, and e satisfy $P(x) = 0$ and adding the resulting equations (as in problem 3, method II) yields

$$\begin{aligned}&(a^5 + b^5 + c^5 + d^5 + e^5)\\ &\quad + p(a^4 + b^4 + c^4 + d^4 + e^4)\\ &\quad + q(a^3 + b^3 + c^3 + d^3 + e^3)\\ &\quad + r(a^2 + b^2 + c^2 + d^2 + e^2)\\ &\quad + s(a + b + c + d + e) + 5t = 0.\end{aligned}$$

Since p, q, $a^2 + b^2 + c^2 + d^2 + e^2$, and $a + b + c + d + e$ are all divisible by n, it follows that $(a^5 + b^5 + c^5 + d^5 + e^5) + 5t$ is also divisible by n. But $t = -abcde$, hence the conclusion.

Where did we use the fact that n is odd?

Problem 19: Since $r_1^2 + 2r_1 + 3 = 0$ and $r_2^2 + 2r_2 + 3 = 0$, the required expression reduces to

$$\frac{2r_1+2}{3r_1+1}+\frac{2r_2+2}{3r_2+1}=\frac{12r_1r_2+8(r_1+r_2)+4}{9r_1r_2+3(r_1+r_2)+1}.$$

Since $r_1 + r_2 = -2$ and $r_1r_2 = 3$, the required value is

$$\frac{36-16+4}{27-6+1}=\frac{12}{11}.$$

Index of Advertisers

COWCULATIONS

Come, bossy

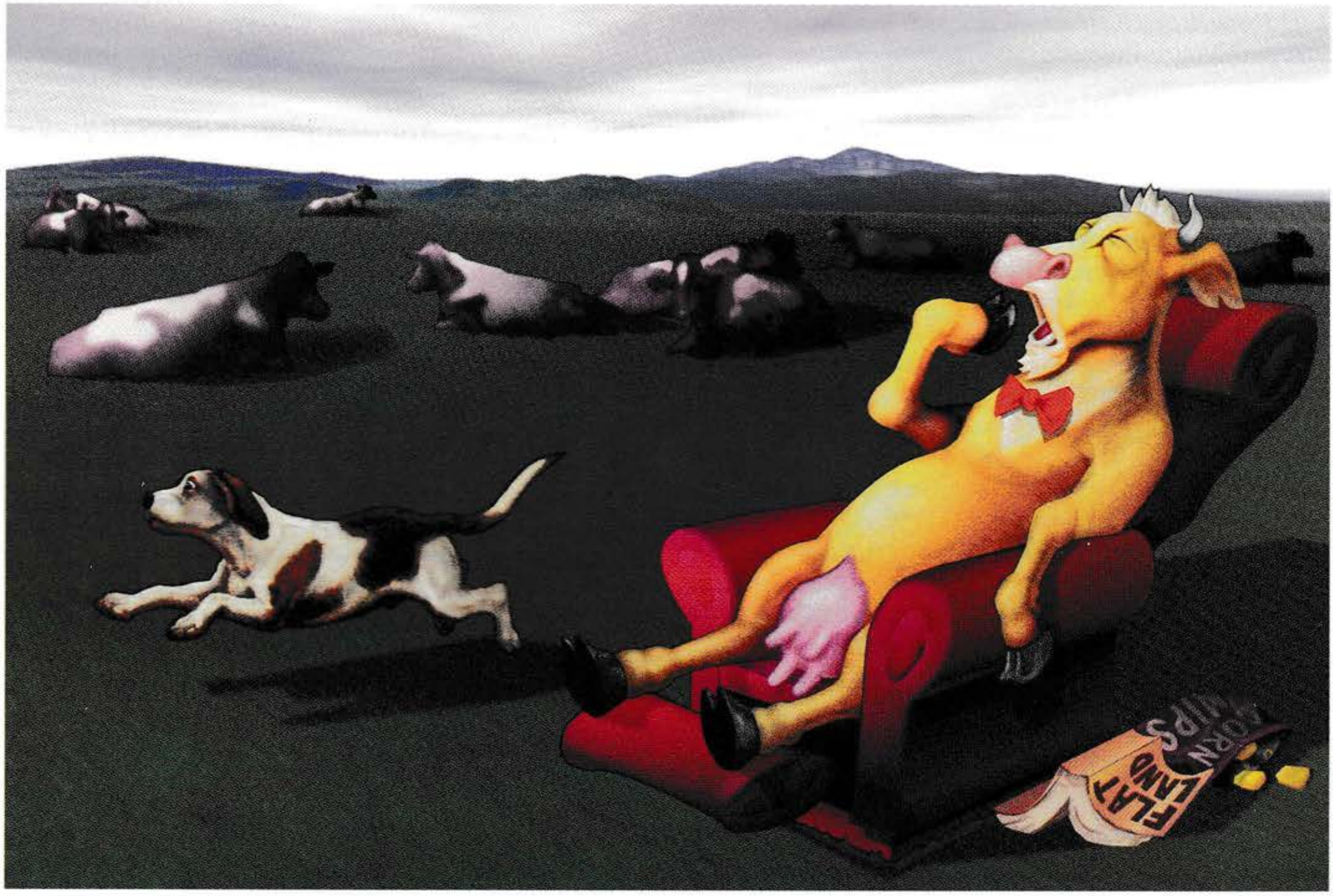
Rounding up the herd

by Dr. Mu

WELCOME BACK TO COWCULATIONS, THE column devoted to problems best solved with a computer algorithm. At farmer Paul's dairy, milking is done twice a day: 5 o'clock in the morning and 5 o'clock in the evening, 365 (or 366) days a year. After 20 years, farmer Paul knows the routine so well he can do it in his sleep, and occasionally has.

Everyone on the farm pitches in to do his or her part. Paula gets up early to make coffee and breakfast and get the kids' school lunches packed. The kids, Mike and Jean, do their chores, feeding the cows and washing the udders. Farmer Paul does the milking, with the help of milking machines. It is left to Bud, the family dog, to round up the cows and bring them in for milking.

Bud has been with the family ever since he walked down the road and onto the farm looking for something to eat. Some town folks had gotten tired of feeding their Christmas present along about March and had taken Bud on a one-way ride into the country. Ever since, Bud has been earning his keep bringing home the cows. As farmer Paul frequently says, "I'm not running a retirement home here. Everyone must earn his keep."



Art by Mark Brenneman

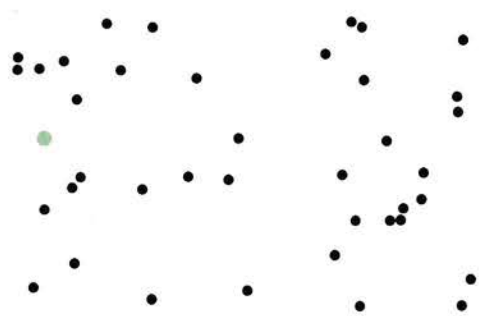
When Bud first started going after cows, he was very disorganized. He'd run back and forth in an almost random pattern, moving around the herd from cow to cow. What was a typical path like?

Let's put our 40 cows out in the pasture at random with *Mathematica.*

```
cow := {Random[], Random[]}
cows = Table[cow, {40}];
```

Next, we'll make points out of the cows and show a graph. (We'll single me out for special treatment and color me green.)

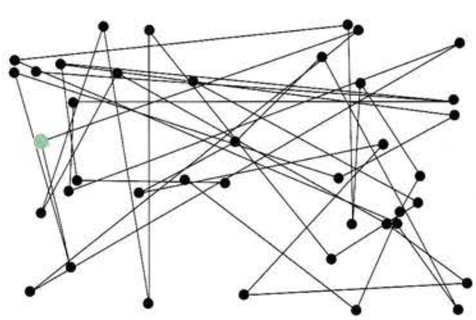
```
drmu = First[cows];
herd = {{PointSize[.02], Point/@ cows},
  {PointSize[.03], RGBColor[0, 1, 0],
    Point [drmu]}};
Show[Graphics[herd]]
```



Now I'll show you how disorganized Bud was in the beginning, keeping track of his path length with a special pathlength function.

```
route = Line[Join[cows, {First[cows]}]];
pathlength[x_] := (Apply[Plus, Map[√#.# &,
  x - RotateRight[x]]])
Print["Random path length = ",
  pathlength[cows]];
Show[Graphics[{route, herd}]]
```

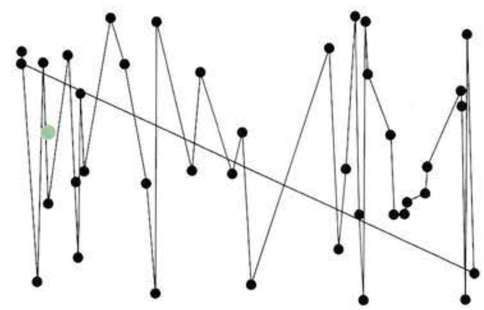
```
Random path length = 22.8515
```



After a month or so, Bud devised another plan. He decided to move among the cows from west to east and then back to the beginning. This cut his path length considerably.

```
cows = Sort[cows];
route = Line[Join[cows, {First[cows]}]];
Print["West to East path length = ",
  pathlength[cows]];
Show[Graphics[{route, herd}]]
```

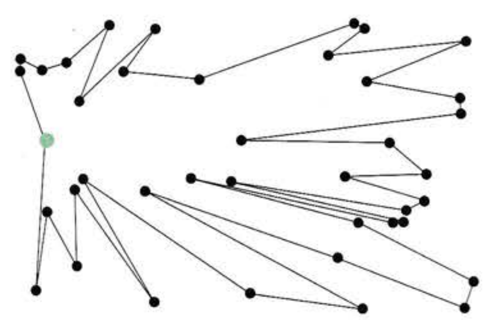
```
West to East path length = 16.9239
```



Still not satisfied, Bud went to work on a way to list the cows based on which ones he saw as he turned in a full circle chasing his tail. He decided to do his 360-degree scan from the position of Dr. Mu. The result is a simple closed path that cuts down the path length even further.

```
angle[a_, b_] := Apply[ArcTan, (b - a)]
cows = Complement[cows, {drmu}];
cows = Sort[cows, (angle[drmu, #1] <=
  angle[drmu, #2])&];
cows = Join[cows, {drmu}];
route = Line[Join[cows, {First[cows]}]];
Print["Simple closed path length = ",
  pathlength[cows]];
Show[Graphics[{route, herd}]]
```

```
Simple closed path length = 9.97219
```

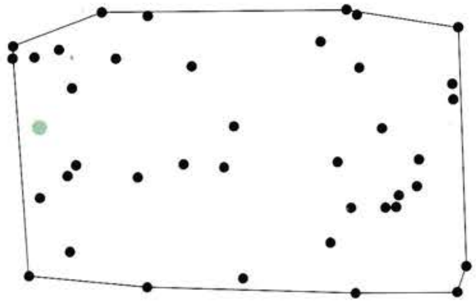


Early last year, Bud became even wiser. He discovered, quite by accident, a remarkable fact about cows. On a whim he made a quick trip around the outside of the herd. Much to his surprise, once the cows on the perimeter got up and started moving toward the barn, the others followed in line. Thus he realized it's not necessary to visit all the cows to bring 'em home, just to beat a path around them. This suggests a computing problem, which, you guessed it, is your next "Challenge Outta Wisconsin."

COW 10

Write a program that finds the shortest path around any herd and cowculates its length.

The shortest path around the herd has length = 3.4436.



Bring 'em home, that's your task.
Surround the herd, make it fast.
With the distance, be a miser.
Do it best, like old "Bud wiser."
—Dr. Mu

Path length

Mathematica is an extremely mathematical programming language. In mathematics, sets are objects upon which operations are performed. In *Mathematica*, lists are sets of objects upon which operations are defined. Algorithms, or procedures, in mathematics are easily and naturally transformed into *Mathematica* expressions or functions. As an example of this natural transformation from mathematics to *Mathematica*, let's construct, from scratch, the pathlength function.

A path in the plane is an ordered set of ordered pairs of real numbers.

```
path = {{x1, y1}, {x2, y2}, {x3, y3}, {x4, y4}};
```

Our goal is to write a function that will take any path object and return its path length, which is the length of the lines between neighboring points, where the first and last points are considered neighbors. We are interested in the closed path length.

We can match up consecutive points very easily by rotating the path:

```
RotateRight[path]
```

$\{\{x_4, y_4\}, \{x_1, y_1\}, \{x_2, y_2\}, \{x_3, y_3\}\}$

and subtracting from the original path.

```
path - RotateRight[path]
```

$\{\{x_1 - x_4, y_1 - y_4\}, \{-x_1 + x_2, -y_1 + y_2\}, \{-x_2 + x_3, -y_2 + y_3\}, \{-x_3 + x_4, -y_3 + y_4\}\}$

Now we'd need to square the x and y coordinates of all the pairs, add them together, and take the square root to get the distance between neighboring points. This is done by mapping a function that does this precisely onto the pairs. But first, notice what the dot product followed by the square root $\sqrt{A.A}$ does to a point A in the plane.

$$\sqrt{\{x_1, y_1\}.\{x_1, y_1\}}$$

$$\sqrt{x_1^2 + y_1^2}$$

This is exactly what we want to apply to each pair above, which is done with a mapping as follows:

```
temp = [Map√#.# &, path - RotateRight[path]]
```

$$\sqrt{(-x_1+x_2)^2+(-y_1+y_2)^2} + \sqrt{(-x_2+x_3)^2+(-y_2+y_3)^2} + \sqrt{(-x_1+x_4)^2+(-y_1+y_4)^2} + \sqrt{(-x_3+x_4)^2+(-y_3+y_4)^2}$$

Finally, we add up the distances by applying the Plus operation:

```
Apply[[Plus, temp]]
```

$$\sqrt{(-x_1+x_2)^2+(-y_1+y_2)^2} + \sqrt{(-x_2+x_3)^2+(-y_2+y_3)^2} + \sqrt{(-x_1+x_4)^2+(-y_1+y_4)^2} + \sqrt{(-x_3+x_4)^2+(-y_3+y_4)^2}$$

This sequence of mathematical transformations is now captured in a new function that we name pathlength and define exactly as described above, but now in one line.

```
pathlength[x_] := (Apply[Plus, Map[√#.# &,
  x - RotateRight[x]]])
```

What you have just seen is an example of creating a custom function in *Mathematica*. Notice that we never once dealt with elements of any list one at a time. Also, there are no Do, For, or While loops. Operating on whole lists is a signature of functional programming. By starting with the original path list, we transformed, via fundamental *Mathematica* functions, a path list onto a pathlength. Functional transformations are the essence of *Mathematica* programming. For those interested in learning more about programming in *Mathematica*, I recommend *An Introduction to Programming with* Mathematica, by Gaylord, Kamin, and Wellin, published by TELOS (http://www.telospub.com).

And finally ...

Send your solution to me at drmu@cs.uwp.edu. Past solutions are available at http://usaco.uwp.edu/cowculations. If competitive computer programming is your path around the herd, stop by the USA Computing Olympiad web site at http://usaco.uwp.edu. The 1998 USA National Championship has just concluded, and the top 15 students have won an all-expense-paid trip to the University of Wisconsin–Parkside this summer. There they will compete for one of four positions on the team that will represent the United States at the 10th International Olympiad in Informatics to be held in Setúbal, Portugal, September 5–12, 1998.