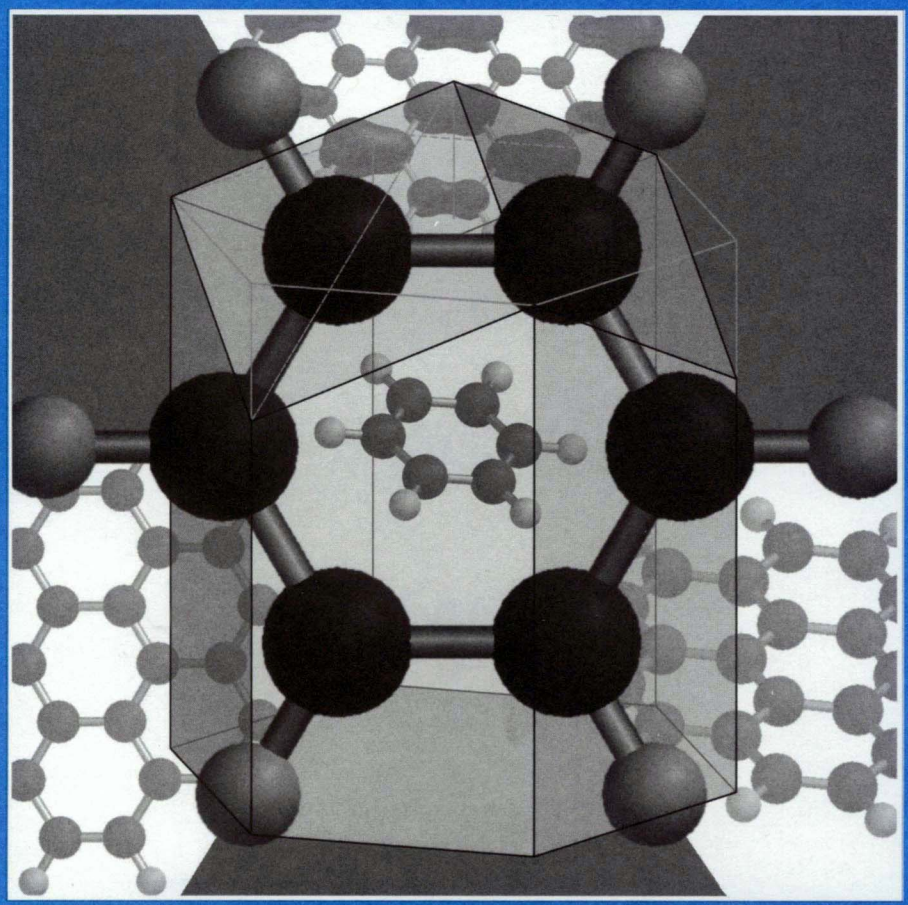
Vol. 74, No. 3, June 2001

MATHEMATICS MAGAZINE



- Elusive Optimality in the Box Problem
- The Anxious Gambler's Ruin
- Counting Perfect Matchings in Hexagonal Systems Associated with Benzenoids
- Using Less Calculus in Teaching Calculus

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

Mathematics Magazine aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for the *Magazine*. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Manuscripts on history are especially welcome, as are those showing relationships among various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 74, pp. 75–76, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Submit new manuscripts to Frank A. Farris, Editor, Mathematics Magazine, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053-0373. Manuscripts should be laser printed, with wide line spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should mail three copies and keep one copy. In addition, authors should supply the full five-symbol 2000 Mathematics Subject Classification number, as described in *Mathematical Reviews*. Copies of figures should be supplied on separate sheets, both with and without lettering added.

Cover image, *Nature's Hexagons*, by Jason Challas, who lectures on Computer Art at Santa Clara University. Thanks to John Thoburn of the Santa Clara University Department of Chemistry.

AUTHORS

Nelson Blachman got his B.S. in physics from Case School of Applied Science (now part of Case Western Reserve University) in 1943 and then went to work at Harvard University's Underwater Sound Laboratory. In 1947 he got his Ph.D. in engineering sciences and applied physics from Harvard University, subsequently working at the Brookhaven National Laboratory, the US Office of Naval Research in Washington and in London, and, for forty years, the GTE Government Systems Corp. Most of his work has dealt theoretically with the effects of noise on communication systems, but he's also investigated various other applications of probability theory.

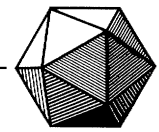
D. Marc Kilgour is Professor of Mathematics at Wilfrid Laurier University, Director of the Laurier Centre for Military Strategic and Disarmament Studies, and Adjunct Professor of Systems Design Engineering (University of Waterloo). He studied under Anatol Rapoport for his Ph.D. in mathematics (University of Toronto, 1973). His doctoral thesis was on truels, and his specialization in game theory may account for his interests in applications such as arms control, deterrence, environmental management, bargaining, arbitration, voting, fair division, and computerized advice to decision-makers in strategic conflict. His interest in the Box Problem is easier to explain—he has always liked puzzles.

Joseph Bak received his B.A., M.A., and Ph.D. degrees from Yeshiva University. Since 1970 he has been teaching at The City College of New York. His primary area of research is approximation theory, and he also has a special interest in complex analysis, in which he co-authored a text with Donald J. Newman. His growing interest in probability, and in the ideas of this article in particular, arose in connection with undergraduate and graduate education courses that he has taught for the past several years at City College.

Fred J. Rispoli received his Ph.D. from SUNY Stony Brook in 1990 under the direction of Michel Balinski. He has taught at Dowling College since 1986, and is interested in applications of graph theory and combinatorial properties of polytopes. Motivated by the NSF initiative to integrate mathematics across the curriculum, he became interested in hexagonal systems and benzenoids after several discussions with chemistry professors. These discussions and the need to calculate determinants also led to his interest in developing the use of spreadsheets in mathematics and science courses.

Radoslav M. Dimitrić obtained his B.S. and M.Sci. degrees from The University of Belgrade in Serbia and his Ph.D. from Tulane University in New Orleans. He held positions at universities in Ireland and England before settling in the U.S. where his career has been revolving almost exclusively around the universities in Northern California. In addition to his research in algebra (look for his monograph "Slender Modules, Slender Rings" to be published by Cambridge University Press) Dimitrić is passionate about the issues of teaching mathematics; hence the present paper that is an amalgam with another of his interests: history of mathematics.

Vol. 74, No. 3, June 2001



MATHEMATICS MAGAZINE

MATHEMATICS MAGAZINE (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.
The annual subscription price for *MATHEMATICS MAGAZINE* to an individual member of the Association is $16 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $64. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 40% dues discount for the first two years of membership.) The nonmember/library subscription price is $68 per year.
Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Advertising correspondence should be addressed to Dave Riska (*driska@maa.org*), Advertising Manager, the Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

Periodicals postage paid at Washington, D.C. and additional mailing offices.

Postmaster: Send address changes to Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036-1385.

Printed in the United States of America

Elusive Optimality in the Box Problem

NELSON M. BLACHMAN
33 Linda Avenue, Apt. 2208
Oakland, CA 94611-4819
blachman@gte.net

D. MARC KILGOUR
Wilfrid Laurier University
Waterloo, ON N2L 3C5 CANADA
mkilgour@wlu.ca

Introduction

Imagine that on a game show you are presented with two identical boxes: B_s, which contains an amount of money $\$S > 0$, and B_l, which contains $\$L = \$2S$. You pick one box—say B_x—which might be either B_s or B_l. Now you must decide whether to keep B_x and the $\$X$ it contains, or to exchange it for B_y and $\$Y$. You do not know the values of S or L, but before you make your decision you may peak inside B_x to learn the value of X.

According to the following argument, what you should do is clear:

> Because X is equally likely to be S or L, you are equally likely to have $2X$ or $X/2$ after trading. Consequently, your expected gain from trading boxes is $\frac{1}{2}(2X) + \frac{1}{2}(X/2) = \frac{5}{4}X > X$. On average, therefore, trading results in a 25% improvement.

In other words, you optimize your expected net gain by always trading. Moreover, the value of X does not affect the decision, so there is no need to look inside B_x.

Most people find this conclusion paradoxical, as trading merely adds one ineffective step to the random selection of a box. Because X and Y take the same values with the same probabilities (they are "stochastically equal"), their averages must be identical. Another argument is that trading B_x for B_y is equally likely to result in a gain of $L - S = S$ or a loss of the same amount; so the expected change in your wealth must be zero. Using a *geometric* mean, the "average" of doubling and halving would be the identity. But expected values seem appropriate here, and they are weighted *arithmetic* means.

This puzzle, which we call the Box Problem, dates back at least to 1953 [**9**]. It has also been named the "wallet game" [**8**], the "exchange paradox" [**12, 15**], the "two-envelope paradox" [**5, 10**], and the "Ali–Baba paradox" [**1, 14**].

An obvious criticism of the "always trade" argument is that the expected value calculation is the same no matter what the value of X. Intuitively, we feel some values of X are so large that $X = L$ is much more likely than $X = S$. To address this criticism, it is natural to use a Bayesian analysis, in which new information (like the value of X) can change your probabilities, and therefore your decision.

Bayesian analyses of the Box Problem did not appear until the past decade but now there are several [**4, 6, 10, 11, 12**]. These show that always trading is not optimal for *every* prior distribution of S, but—surprisingly—they have identified *some* prior distributions for which always trading *is* optimal, apparently. (Christensen and Utts [**6**]

claimed that this phenomenon could not occur when the underlying distribution is continuous, but corrected their error later [**3**].)

Thus, the 25% expected gain argument is only the beginning of the puzzle. Prior distributions exist for which always trading is optimal, so the Bayesian perspective provides no grounds for flatly rejecting the 25% argument. These "anomalous distributions," and the "always trade" conclusion they support, thus constitute a deeper paradox—which this article aims to explore.

A strategy for the Box Problem is a rule to determine whether to trade or not, given each possible amount that you could discover in B_x. We begin by introducing some strategies and discuss how to evaluate and compare them. Without using conditional probabilities, we determine optimal strategies whenever they exist, obtaining results that illuminate the anomalous cases while remaining consistent with Bayesian analysis of the "ordinary" cases. Then we exhibit the results of thirteen different strategies in a million-round simulation of the Box Problem under an anomalous distribution, showing how the benefits of theoretically superior strategies can be extremely elusive in practice. We conclude with some further comments about "solving" the Box Problem.

Strategies in the box problem

A strategy is a rule telling you to trade boxes, or not, based on the value of X. To specify a strategy, we must know what values of X are possible; to evaluate the benefit of a strategy, we must have probability distributions for X and Y. To keep the analysis simple without losing anything essential, we will suppose that the only possible amounts that could be found in the boxes are powers of 2.

Assume that the distribution of S is given by a doubly-infinite sequence of nonnegative numbers, $\ldots, p_{-2}, p_{-1}, p_0, p_1, p_2, \ldots$, satisfying $\sum_k p_k = 1$, where

$$p_k = \Pr[S = 2^k]$$

for each $k \in \mathbb{Z}$. (Hereinafter, summations run from $-\infty$ to ∞ unless otherwise specified.) Note that p_k might equal 0 for many values of k, perhaps when $k < 0$ because small fractions of a dollar are hard to put into boxes, or when $k \gg 1$ because very large amounts of money are not available.

We can now determine the distributions and expected values of X, Y, L, and S in terms of $\{p_k\}$. First, the expected value of S is the summation of the product of every possible value of S times its probability, i.e.,

$$\mathbf{E}[S] = \sum_k 2^k p_k. \tag{1}$$

Of course, $\mathbf{E}[S] = \infty$ is possible (i.e., the summation in (1) may not converge), in which case we say that $\mathbf{E}[S]$ is not finite [**7**, p. 221], or that S has "a heavy tail."

Because $L = 2S$, it follows that $\Pr[L = 2^k] = p_{k-1}$ and $\mathbf{E}[L] = 2\mathbf{E}[S]$. To find the distribution of X, we note that X can equal 2^k in two mutually exclusive ways: either $S = 2^{k-1}$ and $X = L$, or $S = 2^k$ and $X = S$. The first has probability $\frac{1}{2}p_{k-1}$, and the second $\frac{1}{2}p_k$. Hence,

$$\Pr[X = 2^k] = \frac{1}{2}p_{k-1} + \frac{1}{2}p_k, \tag{2}$$

so the expected value of X equals

$$\mathbf{E}[X] = \sum_k 2^{k-1}[p_{k-1} + p_k] = \sum_k 2^{k-1} p_{k-1} + \sum_k 2^{k-1} p_k$$
$$= \sum_k 2^k p_k + \sum_k 2^{k-1} p_k = \mathbf{E}[S] + \frac{1}{2}\mathbf{E}[S] = \frac{3}{2}\mathbf{E}[S]. \tag{3}$$

Note that $\mathbf{E}[Y] = \frac{3}{2}\mathbf{E}[S]$, because $X + Y = L + S = 3S$. (Alternatively, $\mathbf{E}[Y]$ could be calculated directly.) Clearly $\mathbf{E}[L]$ and $\mathbf{E}[X] = \mathbf{E}[Y]$ are all finite if and only if $\mathbf{E}[S]$ is finite.

A strategy for the Box Problem is a rule telling you to trade boxes, or not, as a function of the value of X. To be as general as possible, we allow a strategy to give you a probabilistic instruction ("Trade with probability q"). Thus, a strategy σ is a doubly-infinite sequence

$$\sigma = \{\ldots, q_{-2}, q_{-1}, q_0, q_1, q_2, \ldots\},$$

where q_k is the probability of trading boxes upon finding that $X = 2^k$. (We'll write $q_k(\sigma)$ instead of q_k when the name of the strategy is not clear from the context.)

If you use strategy σ, we will say that the box you end up with is $B_{z(\sigma)}$, containing the amount $Z(\sigma)$. For example, two very simple strategies are the always-trade strategy, σ_{always}, which has $q_k = 1$ for all k and $Z(\sigma_{\text{always}}) = Y$, and the never-trade strategy, σ_{never}, which has $q_k = 0$ for all k and $Z(\sigma_{\text{never}}) = X$.

A general expression for the distribution of $Z(\sigma)$ will be useful. There are three ways for $Z(\sigma)$ to equal 2^k: either $X = 2^k$ and you don't trade, or $X = S = 2^{k-1}$ and you do trade, or $X = L = 2^{k+1}$ and you do trade. The probabilities of these three mutually exclusive events are, respectively, $\Pr[X = 2^k](1 - q_k)$, $\frac{1}{2}p_{k-1}q_{k-1}$, and $\frac{1}{2}p_k q_{k+1}$. Thus, using (2),

$$\Pr[Z(\sigma) = 2^k] = (1 - q_k)\Pr[X = 2^k] + \frac{1}{2}p_{k-1}q_{k-1} + \frac{1}{2}p_k q_{k+1}$$
$$= \Pr[X = 2^k] + \frac{1}{2}(q_{k-1} - q_k)p_{k-1} - \frac{1}{2}(q_k - q_{k+1})p_k \tag{4}$$

for each $k \in \mathbb{Z}$. Of course, $S \le Z(\sigma) \le L$, so $\mathbf{E}[S] \le \mathbf{E}[Z(\sigma)] \le \mathbf{E}[L] = 2\mathbf{E}[S]$. It follows that $\mathbf{E}[Z(\sigma)]$ is finite if and only if $\mathbf{E}[S]$ is finite. In this case, (4) implies that

$$\mathbf{E}[Z(\sigma)] = \sum_k 2^k \Pr[Z(\sigma) = 2^k]$$
$$= \mathbf{E}[X] + \sum_k 2^{k-1}[(q_{k-1} - q_k)p_{k-1} - (q_k - q_{k+1})p_k]. \tag{5}$$

We now define $G(\sigma)$, the gain from strategy σ, as $G(\sigma) = Z(\sigma) - X$. The value of $G(\sigma)$ is the amount by which σ is better than doing nothing. It is easy to verify that

$$\Pr[G(\sigma) = 2^k] = \frac{1}{2}p_k q_k \qquad \Pr[G(\sigma) = -2^{k-1}] = \frac{1}{2}p_{k-1}q_k \tag{6}$$

for each $k \in \mathbb{Z}$, so that the expected value of $G(\sigma)$ is

$$\mathbf{E}[G(\sigma)] = \sum_k 2^{k-1} p_k q_k - \sum_k 2^{k-2} p_{k-1} q_k. \tag{7}$$

If $\mathbf{E}[S]$ is finite, then $\mathbf{E}[G(\sigma)] = \mathbf{E}[Z(\sigma)] - \mathbf{E}[X]$ and (5) and (7) are equivalent, as is easily shown. The benefit of (7) is that it may be meaningful even when $\mathbf{E}[S]$ is infinite, in which case $\mathbf{E}[Z(\sigma)]$ and $\mathbf{E}[X]$ are infinite also.

Now we turn to some more interesting strategies. The simplest useful strategy for the Box Problem is the *threshold* strategy of trading if and only if the observed value of X is below some fixed amount. Intuition suggests that threshold strategies are promising because they specify trading when the amount in hand, X, is small, but not when it is large. For $K \in \mathbb{Z}$, we define σ_K as $q_k = 1$ if $k \leq K$ and $q_k = 0$ if $k > K$. Substitution in (4) shows that

$$\Pr[Z(\sigma_K) = 2^k] - \Pr[X = 2^k] = \begin{cases} -\frac{1}{2}p_K & \text{if } k = K \\ \frac{1}{2}p_K & \text{if } k = K+1 \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Thus the net effect of σ_K is to transfer the probability $\frac{1}{2}p_K$ from 2^K to 2^{K+1}. Without regard to whether $\mathbf{E}[S]$ is finite, we can use (7) to find the expected gain for σ_K relative to X,

$$\mathbf{E}[G(\sigma_K)] = 2^{K-1}p_K. \tag{9}$$

Thus the threshold strategy σ_K is never worse than doing nothing, and is strictly better (according to the expectation-of-gain criterion) whenever $p_K > 0$.

A *subset strategy* is defined by a subset $U \subseteq \mathbb{Z}$ such that trading occurs if and only if $\log_2 X \in U$, i.e., if and only if the amount in B_x equals 2^k for some $k \in U$. The simplest subset strategies are *point strategies*, which call for trading if and only if $X = 2^K$ for some specific value of K, i.e., $U = \{K\}$. Of course, threshold strategies are also subset strategies; the subset corresponding to σ_K is $U = \{k \in \mathbb{Z} : k \leq K\}$.

Two more strategies in the subset class will be used as illustrations below. Define σ_{even} by $q_k = 1$ if k is even and $q_k = 0$ if k is odd. Define σ_{odd} by $q_k = 1$ if k is odd and $q_k = 0$ if k is even. These two strategies are "complementary" in the sense that their respective q_ks sum to 1 for every k, reflecting that the subsets defining σ_{even} and σ_{odd} are complements.

As already noted, (5) can be solved for $\mathbf{E}[G(\sigma)] = \mathbf{E}[Z(\sigma)] - \mathbf{E}[X]$ when $\mathbf{E}[S] < \infty$. When $1 - q_k$ is substituted for q_k in every term of (5), the value of $\mathbf{E}[G(\sigma)]$ obtained in this way changes sign. Thus, the expected gains from complementary strategies, if both finite, are equal and opposite in sign. Hence (assuming $\mathbf{E}[S] < \infty$) $\mathbf{E}[G(\sigma_{\text{even}})] + \mathbf{E}[G(\sigma_{\text{odd}})] = 0$ and $\mathbf{E}[G(\sigma_{\text{never}})] + \mathbf{E}[G(\sigma_{\text{always}})] = 0$.

A strategy not in the subset class was suggested independently by Ross [**13**] and Bruss [**5**]. Based on a probabilistic threshold, it includes trade probabilities q_k that are neither zero nor one. The idea is to compare the observed amount, X, with the value of a random variable, T, and trade if and only if $X \leq T$. An equivalent way to specify this *random-threshold* strategy is via the trade probabilities $q_k = \Pr[T \geq 2^k]$. Of course, $\lim_{k\to\infty} q_k = 0$. Ross [**13**] showed that no matter what the distribution of S, a random threshold strategy σ satisfies $\Pr[Z(\sigma) = L] > \Pr[Z(\sigma) = S]$ provided $\{q_k\}$ is a strictly decreasing sequence, which can be arranged by choosing T suitably.

To illustrate random threshold strategies, we will use *offset geometric* strategies, $G(\alpha, K)$, where $0 < \alpha < 1$ and $K \in \mathbb{Z}$, $K \geq 0$. To define $G(\alpha, K)$, set $q_k = 1$ for $k < K$, and $q_k = \alpha^{k-K}$ for $k \geq K$. (The trade probability, q_k, is the complement of the cumulative distribution function of a geometric distribution with parameter α, "offset" K units to the right.) Note that either $q_k = 1$ or $q_k < q_{k-1}$.

We are now ready to compare strategies, and to classify distributions of S as "ordinary" or "anomalous."

Improving your strategy

Our objective is to find "good" strategies for the Box Problem in the anomalous case when the apparently optimal strategy is "always trade." To specify this case, we must determine when optimal strategies exist, and identify them.

We wish to find all strategies $\sigma^* = \{\ldots, q_{-1}^*, q_0^*, q_1^*, \ldots\}$ that maximize $\mathbf{E}[G(\sigma)]$. Assume that the underlying distribution $\{\ldots, p_{-1}, p_0, p_1, \ldots\}$ is such that $\mathbf{E}[S] < \infty$, so that $\mathbf{E}[G(\sigma)] = \mathbf{E}[Z(\sigma)] - \mathbf{E}[X]$. In this case, the summations in (7) (or (5)) can be regrouped to produce

$$\mathbf{E}[G(\sigma)] = \sum_k 2^{k-1} p_k q_k - \sum_k 2^{k-2} p_{k-1} q_k = \sum_k 2^{k-2} q_k [2p_k - p_{k-1}]. \quad (10)$$

Since the probability q_k must lie between 0 and 1, it is clear from (10) that $\mathbf{E}[G(\sigma)]$ is maximized by any strategy satisfying

$$q_k^* = \begin{cases} 1 & \text{if } p_k > \frac{1}{2} p_{k-1} \\ \text{arbitrary} & \text{if } p_k = \frac{1}{2} p_{k-1} \\ 0 & \text{if } p_k < \frac{1}{2} p_{k-1}. \end{cases} \quad (11)$$

Equation (11) is equivalent to the *Exchange Condition for Discrete Distributions* (named by Brams and Kilgour [**4**], but also discovered, at least in special cases, by several others): When you find $\$2^k$ in B_x, you may trade boxes if and only if

$$p_k \geq \frac{p_{k-1}}{2}, \quad (12)$$

and you must trade boxes if the inequality in (12) is strict.

An optimal strategy σ^* satisfying (11) is a Bayesian optimum in that it is "local," taking into account all the information you have (i.e. the realized value of k) at the time you make your decision. This is a consequence of the term-by-term maximization of (10). For the same reason "global" effects, such as convergence or divergence of expectations, are irrelevant to the characterization of σ^* in (11). Thus, (11) seems to make sense even when it shouldn't—when the underlying distribution fails the condition $\mathbf{E}[S] < \infty$, so that the argument leading to (11) cannot be justified.

In summary, Exchange Condition (11) gives all optimal strategies, σ^*, for any Box Problem $\{p_k\}$ for which $\mathbf{E}[S] < \infty$, i.e., for which the summation in (1) is convergent. But should this condition fail, we have no reason to believe that a strategy satisfying (11) is better than any other. The series in (5) and (7) are at best conditionally convergent in this case so, as discussed in detail by Norton [**12**], the manipulation of (7) to obtain (10) is unjustified.

Yet (11) seems to leave a loophole for the dubious "always trade" strategy! Define $\{p_k\}$ to be an *anomalous distribution* if it satisfies $p_k \geq \frac{1}{2} p_{k-1}$ for all $k \in \mathbb{Z}$. By (11), no strategy can be better than "always trade" ($q_k^* = 1$ for all k) if the distribution is anomalous. And the really bad news is that there are plenty of anomalous distributions (see below).

Have we come all this way to return to the conclusion implied by the seemingly paradoxical 25% expected gain calculation? No. The reason is that no anomalous distribution could possibly satisfy the condition $\mathbf{E}[S] < \infty$. To see this, note first that each anomalous distribution is characterized by a value of H, with $-\infty \leq H < \infty$, such that $p_k = 0$ for $k < H$ and $p_k > 0$ for $k \geq H$. (Let $H = \inf\{k : p_k > 0\}$, which must exist since p_k must be positive for at least one value of k. It is then immediate that $p_k > 0$ for any $k > H$.) Now select any $K > -\infty$ such that $K \geq H$. Because

$p_k \geq \frac{1}{2}p_{k-1}$ for $k = K+1, K+2, \ldots, K+h$, it follows that $p_{K+h} \geq 2^{-h}p_K$. Then substitution in (1) yields

$$\mathbf{E}[S] \geq \sum_{h=0}^{\infty} 2^{K+h} p_{K+h} \geq \sum_{h=0}^{\infty} 2^K p_K = \infty,$$

because $p_K > 0$. Therefore, we should not be surprised if (11) leads to nonsensical conclusions for anomalous distributions, because all such distributions fail a condition that is crucial to the derivation of (11).

We turn to the problem of comparing strategies when $\{p_k\}$ is anomalous. We now know that $\mathbf{E}[Z(\sigma)] = \infty$ for every strategy σ, because $\mathbf{E}[S] = \infty$. Our comparisons will rely on expectation of gain and on stochastic dominance.

A random variable W_1 *stochastically dominates* a random variable W_2 if and only if

(a) for every w, $\Pr[W_1 \leq w] \leq \Pr[W_2 \leq w]$, and

(b) for some w_0, $\Pr[W_1 \leq w_0] < \Pr[W_2 \leq w_0]$.

We say that strategy σ_1 *stochastically dominates* strategy σ_2 if and only if $Z(\sigma_1)$ stochastically dominates $Z(\sigma_2)$; if so, there is a strong argument that you are better off choosing σ_1 instead of σ_2. If equality holds in (a) for all w, W_1 and W_2 are *stochastically equal*; strategies σ_1 and σ_2 are stochastically equal if $Z(\sigma_1)$ is stochastically equal to $Z(\sigma_2)$.

For example, the best-case outcome L stochastically dominates X, and X stochastically dominates the worst-case outcome S. (To verify part (a) of the definition, note that, for any integer K, $\Pr[S \leq 2^K] = \sum_{k=-\infty}^{K} p_k$, $\Pr[X \leq 2^K] = \sum_{k=-\infty}^{K-1} p_k + \frac{1}{2}p_K$, and $\Pr[L \leq 2^K] = \sum_{k=-\infty}^{K-1} p_k$. To verify part (b), note that these three quantities are different for any $K \in \mathbb{Z}$ satisfying $p_K > 0$, and that such a K must exist.) Similarly, the strategies σ_{never} and σ_{always} are stochastically equal because $X = Z(\sigma_{\text{never}})$ and $Y = Z(\sigma_{\text{always}})$ are stochastically equal.

Recall that $G(\sigma) = Z(\sigma) - X$, where $X = Z(\sigma_{\text{never}})$. It follows from (8) that the threshold strategy σ_K stochastically dominates σ_{never} if and only if $p_K > 0$, since σ_K transfers probability $\frac{1}{2}p_K$ from 2^K to 2^{K+1}. (If $p_K = 0$, the two are stochastically equal.) Thus, stochastic dominance tells us that any threshold strategy is better than nothing, provided the exact threshold occurs with positive probability.

Unfortunately, comparison of threshold strategies using stochastic dominance is not very useful. Suppose that K and M are integers such that $K > M$. From (4) and (2), it follows that $\Pr[Z(\sigma_K) = 2^k] = \Pr[X = 2^k] = \frac{1}{2}[p_{k-1} + p_k]$ for all k, except that $\Pr[Z(\sigma_K) = 2^K] = \frac{1}{2}p_{K-1}$ and $\Pr[Z(\sigma_K) = 2^{K+1}] = \frac{1}{2}[2p_K + p_{K+1}]$; similarly for $\Pr[Z(\sigma_M)]$. Therefore

$$\Pr[Z(\sigma_M) \leq 2^k] - \Pr[Z(\sigma_K) \leq 2^k] = \begin{cases} -\frac{1}{2}p_M & \text{if } k = M \\ \frac{1}{2}p_K & \text{if } k = K \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

It follows that σ_K and σ_M are stochastically equal if $p_K = p_M = 0$, σ_K stochastically dominates σ_M if $p_M = 0$ and $p_K > 0$, and σ_M stochastically dominates σ_K if $p_M > 0$ and $p_K = 0$. If $p_K > 0$ and $p_M > 0$, there are no stochastic dominance relations between σ_K and σ_M. Thus stochastic dominance tells us that a threshold strategy σ_M is a poor choice if $p_M = 0$; any threshold strategy σ_K, where $p_K > 0$, would be better. But this result does not help in the context of distributions for which $p_k > 0$ for all large enough k. Thus, for anomalous distributions, no threshold strategy (with sufficiently large threshold) stochastically dominates any other.

Stochastic dominance can be used to compare random threshold strategies. Suppose $q_k(\sigma) \leq q_k(\sigma')$ for all $k \in \mathbb{Z}$, with strict inequality for at least one value of k where $p_k > 0$. Then it follows that σ stochastically dominates σ'. Among offset geometric strategies, for example, it is easy to show that, for each α and for $K \in \mathbb{Z}$ large enough, $G(\alpha, K)$ is stochastically dominated by $G(\alpha, K+1)$, which is stochastically dominated by $G(\alpha, K+2)$, etc.

For anomalous distributions, however, stochastic dominance has not helped very much. Threshold strategies with large thresholds are incomparable. Random threshold strategies can always be improved by shifting to the right. So we try another method of comparison, based on expectation.

Expectation is a natural way to evaluate policies in any probabilistic situation. If you intend to play the box game $n \gg 1$ times using strategy σ and if the Law of Large Numbers (the "law of averages") applies, then the total you will receive is likely to be relatively close to $n \times \mathbf{E}[Z(\sigma)]$, which justifies the choice of a strategy σ that maximizes $\mathbf{E}[Z(\sigma)]$. Expectations also have the advantage of assigning numerical values, so every strategy is evaluated on the same scale.

But the disadvantage of expectation when $\mathbf{E}[S] = \infty$ is, as already noted, that $\mathbf{E}[Z(\sigma)] = \infty$ for every strategy σ, so there are no grounds for comparison of strategies. For example, L always stochastically dominates X and Y, which always dominate S, even though when $\mathbf{E}[S] = \infty$ all of L, X, Y, and S have the "same" expected value.

We can, however, use expectation of gain for comparison. Recall that $G(\sigma) = Z(\sigma) - X$, and that (7) may make it possible to calculate $\mathbf{E}[G(\sigma)]$ even when $\mathbf{E}[S] = \infty$. It is reasonable to prefer σ_1 to σ_2 when $G(\sigma_1) - G(\sigma_2)$ tends to be positive. Moreover, $\mathbf{E}[G(\sigma_1) - G(\sigma_2)] = \mathbf{E}[G(\sigma_1)] - \mathbf{E}[G(\sigma_2)]$, so if (7) can be used to show that $\mathbf{E}[G(\sigma_1)] > \mathbf{E}[G(\sigma_2)]$, then we are justified in preferring σ_1 to σ_2.

We use the expectation-of-gain criterion to try to sort out the threshold strategies. Assume that M and K are integers and that $M < K$. Using (9), it is easy to show that

$$\mathbf{E}[G(\sigma_K) - G(\sigma_M)] = \frac{1}{2}(p_K 2^K - p_M 2^M). \tag{14}$$

An alternative derivation can be based on the calculations leading to (13).

There are distributions $\{p_k\}$ with $\mathbf{E}[S] = \infty$ for which (14) identifies a unique threshold strategy that is preferred to all others on the basis of the expectation-of-gain criterion. (One such distribution is given by $p_k = 0$ for $k < 0$, $p_1 = \frac{3}{4}$, and $p_k = 2^{-k-1}$ for $k \geq 2$. The threshold strategy recommended by (14) is σ_1.) But for anomalous distributions, we are no further ahead because, as is easy to verify, $p_K 2^K \geq p_M 2^M$ is true for every K and M such that $K > M$, where $K \geq H$ and $K > -\infty$. (Recall that $H = \inf\{k : p_k > 0\} \geq -\infty$.)

The optimal selection of a threshold strategy for an anomalous distribution is therefore a real conundrum. Equation (14) demonstrates that on the expectation-of-gain criterion, a higher threshold is always better than a lower one. In other words, σ_K gets better and better as K increases. But note that $\sigma_K \to \sigma_{\text{never}}$ as $K \to -\infty$, and that $\sigma_K \to \sigma_{\text{always}}$ as $K \to \infty$. The implication is that even a low threshold is better than never trading, higher thresholds are better than lower ones, and the best possible strategy is to increase the threshold to the extreme, approaching in the limit the strategy of always trading—which on this argument must be "optimal." But these conclusions are impossible to reconcile with what we already know, that σ_{always} is no better and no worse than σ_{never}. We seem to have an infinite sequence of uphill steps that, as in an M. C. Escher drawing, returns us to the same level. To understand the situation better, and to show that promising strategies can have elusive benefits, we will use simulations.

Some simulation results

We now illustrate some of the strange things that can occur in the Box Problem when the underlying distribution is anomalous, i.e., when (12) holds for all $k \in \mathbb{Z}$. Brams and Kilgour [**4**] provide a useful collection of such distributions, including the geometric distribution with ratio $r = \frac{2}{3}$:

$$p_k = \begin{cases} 0 & \text{for } k < 0 \\ \frac{1}{3}(\frac{2}{3})^k & \text{for } k \geq 0 \end{cases} \tag{15}$$

The random variable S with this distribution belongs to the family introduced by Norton [**12**]; it was also considered by Linzer [**10**] *inter alios*. In fact, any geometric distribution with $r > \frac{1}{2}$ would serve our purposes.

The distribution (15) is anomalous because $p_k > \frac{1}{2}p_{k-1}$ for $k = 1, 2, \ldots$, so $\mathbf{E}[S] = \infty$. Still, what should you do if you believe that S is distributed according to this distribution when the host asks you whether you want to trade? You know that (11) appears to suggest always trading, but you also know that (11) is not applicable.

Let $K \in \mathbb{Z}$ satisfy $0 \leq K < \infty$, and consider the threshold strategy σ_K. From (9), it follows that

$$\mathbf{E}[G(\sigma_K)] = \mathbf{E}[Z(\sigma_K) - X] = \frac{1}{2}p_K 2^K = \frac{2^{2K-1}}{3^{K+1}}. \tag{16}$$

Note that the right side of (16) is an increasing function of K—it is of the same order as $(\frac{4}{3})^K$. In other words, it makes the "promise" we have seen before: choose a higher and higher threshold and you will do better and better.

We now use simulation to show how difficult it is to pin down this gain in practice. Table I summarizes the outcomes of 1,000,000 replications of the Box Problem with S distributed according to (15). Each entry in the body of the table is the average gain over 100,000 replications using the indicated strategy. The final entry of each row is the average of the preceding entries, and is thus the average over all 1,000,000 replications. The use of *Mathematica* [**2**] to obtain the numbers in Table I required about ten minutes of computing time on a 200-MHz PC.

TABLE I: Average outcomes for various strategies in the Box Problem when S is distributed as in (15). Each entry is the average outcome for the indicated strategy over 100,000 replications. The last column is the average over all 1,000,000 replications.

σ_{never} (X)	\$345,829	\$5,167	\$26,252	\$15,290	\$9,734	\$8,634	\$14,176	\$24,879	\$359,159	\$9,310	\$81,843
σ_{always} (Y)	173,621	7,556	46,602	18,656	6,704	14,070	25,286	14,107	700,763	9,360	101,672
σ_4	345,829	5,167	26,252	15,290	9,735	8,634	14,177	24,879	359,159	9,311	81,843
σ_9	345,831	5,169	26,254	15,292	9,736	8,636	14,178	24,881	359,161	9,312	81,845
σ_{14}	345,837	5,175	26,259	15,298	9,745	8,644	14,184	24,890	359,168	9,321	81,852
σ_{19}	345,870	5,196	26,307	15,347	9,787	8,670	14,208	24,929	359,205	9,386	81,890
σ_{24}	345,923	5,375	26,637	15,468	9,724	8,534	14,549	24,845	359,347	9,192	81,959
σ_{29}	345,419	7,556	25,127	18,656	6,704	14,070	14,549	24,845	362,534	9,360	82,882
σ_{34}	345,419	7,556	46,602	18,656	6,704	14,070	25,286	14,107	357,166	9,360	84,493
σ_{odd}	173,649	5,199	47,435	16,633	8,650	10,007	14,093	14,200	702,715	9,410	100,199
σ_{even}	345,800	7,524	25,419	17,312	7,788	12,697	25,369	24,786	357,207	9,260	83,316
$G(\frac{5}{6}, 0)$	345,825	5,144	26,254	15,292	9,645	8,634	14,179	24,868	359,184	9,298	81,832
$G(\frac{5}{6}, 25)$	173,621	7,556	46,602	18,320	8,718	11,386	14,549	14,443	700,763	10,031	100,599
Minimum (S)	173,150	4,241	24,285	11,315	5,479	7,568	13,154	12,995	353,307	6,223	61,172
Maximum (L)	346,300	8,482	48,569	22,630	10,959	15,136	26,308	25,991	706,615	12,447	122,344

The first row of Table I represents strategy σ_{never} and the second σ_{always}. Because $Z(\sigma_{\text{never}}) = X$ and $Z(\sigma_{\text{always}}) = Y$, these rows record the average contents of B_x and B_y, respectively. Similarly, the last two rows of Table I give S, the smaller of the amounts in the two boxes, and L, the larger of the two amounts.

Table I has several remarkable features. First, stochastic dominance and stochastic equality don't seem to mean much. The average gains for two stochastically equal strategies, σ_{never} and σ_{always}, are far from equal. Stochastically dominant strategies do not necessarily average better than the strategies they dominate (all of the threshold strategies average more than σ_{never}, but less than σ_{always}). Some comparisons are as expected: the offset geometric strategy $G(\frac{5}{6}, 25)$ averages better than $G(\frac{5}{6}, 0)$, and L averages better than X and Y, which average better than S—but this is hardly surprising, as it is always true that $L = \max(X, Y) > \min(X, Y) = S$.

The key feature of Table I is the importance of a few very large values, which simply dominate everything else. Almost all of the variability in the first column, for example, can be explained only by occurrence of one large value, $S = 2^{34}$, in box B_y. This single event swamps everything else in the column; strategies that did not trade the huge value of $X = L$ averaged about \$345,000, whereas those that traded it away averaged about \$173,000.

The single largest value in the simulation appeared in the ninth column, where $S = 2^{35}$ occurred in box B_x. In this column, the average for a strategy depends almost entirely on whether this huge value of X was traded for the even more enormous Y. In this case, strategies that traded averaged about \$700,000, and strategies that did not averaged about \$360,000.

What's more, nothing affects the overall averages (final column) like the single enormous value of S in the ninth column. The three strategies that traded this particular X for Y average a little more than \$100,000 over the 1,000,000 replications; all of the others, a little more than \$80,000. Thus, most of the variability in the final column of the table tells us only about success or failure in this one (in a million) event.

Still, despite the masking effect of rare extreme values of S, Table I contains some evidence that strategies we have identified as preferable tend to do a little better. For example, note the steady increase in the final column from σ_4 to σ_9 to ... to σ_{34}. Of course, all of these strategies failed to trade for the very large Y in the ninth column, and would have done much better had they done so. What's more, the next strategy in the sequence, σ_{39}, would have produced the same result as σ_{always}, for the simple reason that never in the simulation did any value of X or Y exceed 2^{36}.

Table I also shows simulation results for four more strategies, two of the offset geometric random threshold type, $G(\frac{5}{6}, 0)$ and $G(\frac{5}{6}, 25)$, plus the strategies σ_{odd} (trade only when X is an odd power of 2) and σ_{even} (trade only when it is an even power). As usual, strategies that traded for the huge value of Y appearing somewhere in the ninth group of 100,000 replications did much better. Close examination shows, for example, that even though the second random threshold strategy stochastically dominates the first, the empirical evidence is not overwhelming. The dominant strategy does better in seven of the ten groups, but its higher overall average can be entirely attributed to its decision on the extreme event in the ninth group.

Finally, Table I makes clear that patience is essential to enjoy the benefits of better strategies. In each column, the benefit of σ_9 over σ_4 is very stable, at \$1 or \$2, consistent with the theoretical average improvement of \$1.693 from (14). On the other hand, the average gain from σ_{34} exceeded the average gain from σ_{29} over 1,000,000 replications, but the two averages were equal in five of the groups of 100,000; σ_{29} was better in three of the remaining five groups, and σ_{34} was better in only two. But the clearest indication of the almost chaotic results of the simulation may be that, even

with 1,000,000 replications, the best average was achieved by σ_{always}, a strategy that could hardly be less sophisticated.

Conclusion

The Box Problem is more general than it may seem. It does not depend, for example, on the ratio $L/S = 2$; whenever $0 < L/S \neq 1$, the same paradoxes arise. If you do not look inside B_x, or have no way to utilize this information, then trading increases your expected value by the factor

$$\frac{1}{2}\left(\frac{L}{S} + \frac{S}{L}\right) = 1 + \frac{(L-S)^2}{2LS}, \tag{17}$$

which always exceeds 1, so you will always be tempted to trade boxes. Even if $L > 0$ and $S > 0$ are random quantities (whether or not they are independent of each other), then L/S will have no fixed ratio, but the expected value of X/Y will exceed 1 because of (17), and again you will be tempted to trade. And we have already shown that "always trade" is a useless strategy.

A possible escape from the Box Problem is to measure values by using utility rather than dollars. Utility is a numerical measure of worth that can capture relevant aspects of the problem, such as the declining worth of an additional dollar as your wealth increases. But, as Brams and Kilgour [**4**] argue, the temptation to trade still arises if one box contains twice the utility of the other. If expected utility is finite (corresponding to a light-tailed distribution of S), then an optimal strategy can be found by modifying the derivation of (11). Let $u_k = u(2^k)$ denote your utility for $\$2^k$. Then substituting u_k for 2^k as in (5) produces

$$\mathbf{E}[u(Z)] = \mathbf{E}[u(X)] + \sum \frac{1}{2} q_k[(u_{k+1} - u_k)p_k - (u_k - u_{k-1})p_{k-1}].$$

To maximize the utility of the outcome, you may trade $X = 2^k$ if and only if $(u_{k+1} - u_k)p_k \geq (u_k - u_{k-1})p_{k-1}$, and you must trade $X = 2^k$ if the inequality is strict. This is the analogue for utilities of the Exchange Condition, (12).

But what if the distribution has a heavy tail? One idea is to find a "nearby" problem with a light tail. Returning for simplicity to the problem without utilities, begin by deciding how many rounds, say n, that you intend to play. During n rounds, you can expect that S will fall only once inside the top $\frac{100}{n}\%$ of its probability distribution, i.e., the part beyond $S = 2^{\kappa(n)}$, where $\kappa(n)$ is the smallest integer for which $\sum_{\kappa(n)}^{\infty} p_k \leq 1/n$. Hence, your total outcome from the n rounds will hardly be affected by the shape of the tail of the distribution of S beyond $\$2^{\kappa(n)}$. In particular, it is unlikely to matter much whether the tail is finite, infinite but light, or—the actual case—infinite and heavy. Create a light-tailed distribution by re-assigning the probability beyond $S = 2^{\kappa(n)}$, and calculate an optimal strategy using (11).

While the light-tailed distribution you construct is likely to be close to the original heavy-tailed distribution, it must be noted that the events removed are exactly those unlikely extremes that swamp the optimality calculation. Of course, we recommend the use of utility whenever your value for very great gains is less than proportionate.

Another approach is to count time as a cost in your utility function. Selecting a strategy with a higher anticipated outcome might be less appealing if you must wait for a long time to obtain the benefit. Counting time as a cost may avoid many problems associated with heavy-tailed distributions.

The Box Problem captures a fundamental difficulty with the concept of optimality. It is sometimes impossible to provide reasonable advice to a decision-maker. Indeed, we have identified a situation when "optimality" seems not to exist, and when the strategy recommended by a straightforward "local" (Bayesian) analysis is simply useless. A decision-maker who relies on certain promising strategies is almost sure to find their benefits elusive. We have shown that this elusiveness arises from the "heavy tail" of a distribution, where events seem so remote that they can be of no practical significance. But given such a distribution, the *elusiveness problem* seems to us unavoidable.

Note added in proof: We have shown that good strategies are elusive when $L = 2S$ and S has an anomalous distribution. Both conditions seem essential: when X and Y are statistically independent and identically distributed (iid) in accordance with (15), a carefully-chosen threshold strategy achieves an average outcome Z very close to the average of $\max(X, Y)$, provided the number of rounds is large. A future paper will provide details, and further elucidate good switching strategies.

Acknowledgment. We thank the following individuals for valuable comments, useful materials, and helpful discussions: David Blackwell, Steven J. Brams, Lorne Campbell, Henry Cejtin, Herman Chernoff, David desJardins, Elliot Linzer, Samuel Merrill III, John Norton, Richard Potthoff, Lloyd S. Shapley, and Rolf Turner.

REFERENCES

1. M. Bickis, The real paradox of Ali–Baba, *SSC Liaison* **12**, No. 2 (1998), 20–25.
2. N. R. Blachman and C. P. Williams, *Mathematica: A Practical Approach* (Second Edition), Prentice–Hall, Englewood Cliffs, NJ, 1999.
3. N. M. Blachman, R. Christensen, and J.M. Utts, Comment on Christensen and Utts (1992), *The Amer. Statist.* **50** (1996), 98–99.
4. S. J. Brams and D. M. Kilgour, The box problem: to switch or not to switch, this MAGAZINE **68** (1995), 27–34.
5. F. T. Bruss, The fallacy of the two envelopes problem, *Math. Scientist* **21** (1996), 112–119.
6. R. Christensen and J. M. Utts, Bayesian resolution of the exchange paradox, *The Amer. Statist.* **47** (1992), 274–276. For correction of important errors see [**3**].
7. R. Feller, *An Introduction to Probability Theory and Its Applications*, Volume 1 (Third Edition), J. W. Wiley and Sons, New York, 1968.
8. M. Gardner, *Gotcha! Paradoxes to Puzzles*, W. H. Freeman, New York, 1982.
9. M. Kraitchik, *Mathematical Recreations*, 2nd ed., Dover, New York, 1953.
10. E. Linzer, The two-envelope paradox, *The Amer. Math. Monthly* **101** (1994), 417–419.
11. W. Mixon, S.J. Brams, and D.M. Kilgour, Letter to the Editor and response, this MAGAZINE **68** (1995), 322–323.
12. J. D. Norton, When the sum of our expectations fails us: the exchange paradox, *Pacific Philosophical Quarterly* **79** (1998), 34–58.
13. S. M. Ross, Comment on Christensen and Utts (1992), *The Amer. Statist.* **48** (1994), 267.
14. R. Turner et al., Ali–Baba paradox, *SSC Liaison* **11**, No. 4 (October, 1997), 34 and **12**, 1 (February, 1998), 34–39.
15. S. L. Zabell, Symmetry and its discontent, in *Causation, Chance, and Credence*, B. Skyrms and W. L. Harper (eds), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

The Anxious Gambler's Ruin

JOSEPH BAK
City College of New York
New York, NY 10031

Introduction

Suppose a person with A dollars decides to risk it in the hope of increasing it to B dollars. This could represent the final attempt of a gambler who has lost almost all of his money in Las Vegas, and now seeks to win just enough for transportation home. He finds a game where the probability of winning is p and where a bet of A dollars potentially pays A dollars; of course, if he loses, he loses that money.

In his last-ditch effort, he continues to stake the largest possible amount toward reaching, but not exceeding his goal. Thus, if he currently has A dollars and $A \leq B/2$, he risks all of it; if he is more than half-way to his goal, he bets the difference, $B - A$, since winning that amount will bring his purse exactly to B. He stops gambling only when he has either lost all of his money or reached his goal of B.

Our gambler might be described as desperate, but the adjective *anxious* seems more appropriate. Aside from the connotation of nervousness about the outcome, it also suggests the desire to come to a conclusion as soon as possible. This paper deals with both of these aspects of the game: the gambler's chance of reaching his goal (which we will call *success*), and the expected number of games until he has either achieved success or lost all of his money (which we will call *failure*). Obviously, this analysis also applies to perfectly calm a person willing to risk a fixed amount of disposable income for a chance at a specific sum of money, but we continue with the image of the anxious gambler.

This is a variation of the classic problem known as "the gambler's ruin." In that case, a gambler with A dollars continually bets \$1 against an opponent with $B - A$ dollars until one of them is "ruined". If the gambler with an original amount of A has probability p of winning each game, then his probability of *not* being ruined (his chance of achieving B dollars before falling to 0) is given by:

$$\mathcal{P}(A, B, p) = \begin{cases} \dfrac{(q/p)^A - 1}{(q/p)^B - 1} & \text{if } p \neq \frac{1}{2},\ q = 1 - p \\ \\ A/B & \text{if } p = \frac{1}{2}. \end{cases} \tag{1}$$

See, for instance, Feller [**7**].

Obviously, our anxious gambler problem can be similarly cast as a contest between two individuals, one with an initial fortune of A, the other with an initial fortune of $B - A$, culminating when one of them is ruined. The only difference, of course, is that the anxious gambler bets a varying amount on each game, in his eagerness to meet, but not exceed, his goal.

To distinguish between the different but related variables under discussion, p and q will always denote the gambler's probability of winning and losing, respectively, on each bet. The script letters $\mathcal{P}(A, B, p)$ and $\mathcal{D}(A, B, p)$ represent the probability of success and the expected number of games (or duration) in the classic case, with

uniform bets of \$1 each. Roman letters $P(A, B, p)$ and $D(A, B, p)$ denote the corresponding probability and duration for the anxious gambler.

First, we present an extremely simple proof that $P(A, B, 1/2) = A/B$. We then generalize the result and show how the key idea of the proof was used by de Moivre to obtain formula (1) for $\mathcal{P}(A, B, p)$, even when $p \neq 1/2$. Another section deals with the transition from the classic gambler's ruin to the anxious gambler's ruin by considering the effects of increasing the stakes in the classic case to a fixed amount of \$2 (or more) per bet.

We then address general values of p, and find explicit formulae for $P(A, B, p)$ and $D(A, B, p)$. We also cite a theorem of Dubins and Savage on an extremal property of $P(A, B, p)$, deriving a corresponding result for $D(A, B, p)$. The last two sections offer some ramifications of these ideas, relating them to certain aspects of sporting events and lotteries, and examining a more theoretical possibility: playing against an infinitely rich adversary.

From $P(A, B, \frac{1}{2})$ to $\mathcal{P}(A, B, p)$ and $\mathcal{D}(A, B, p)$

Recall that $P(A, B, 1/2)$ is the probability that the anxious gambler eventually reaches the goal of B, given that he starts with A and uses the anxious strategy in a game with even odds of winning.

PROPOSITION 1. *For all positive integers A, B such that $A < B$, $P(A, B, 1/2) = A/B$.*

Proof. Let $X_0 = A$, and let X_i represent the gambler's winnings on bet i, $i \geq 1$, so that $S_n = X_0 + X_1 + X_2 + \cdots + X_n$ represents his holdings after n bets. A simple induction argument shows that each X_i has one positive value and one matching negative value with equal probabilities. Thus, $\mathbf{E}[X_i] = 0, i \geq 1$, and for all positive integers n,

$$\mathbf{E}[S_n] = X_0 = A.$$

The probability that the game will continue indefinitely is 0 since the probability that the game will continue beyond n games is at most $1/2^n$. Hence, with probability 1, S_n converges to a limit function S, which, according to the gambler's strategy, has only two possible values: B (with probability $P(A, B, 1/2)$), or 0, with the complementary probability. Thus, on the one hand,

$$\mathbf{E}[S] = \lim \mathbf{E}[S_n] = A$$

while, on the other hand,

$$\mathbf{E}[S] = B\ P\left(A, B, \frac{1}{2}\right)$$

from the definition of expected value. A comparison of the two expressions for $\mathbf{E}[S]$ shows that $P(A, B, 1/2) = A/B$. (A formal proof that $\mathbf{E}[S] = \lim \mathbf{E}[S_n]$ can be given using the Bounded Convergence Theorem, which can be found in Billingsley [**1**, p.338].) ■

The proof of Proposition 1 can be applied with minor modification to obtain the same result in the classic problem when $p = 1/2$. In fact, the proposition can be extended to a variety of cases, including fixed bets of any size, or any other type of "fair" bet, such as the type obtained when "true odds" are offered (e.g., a 35:1 payoff in a

game with a 1/36 chance of winning). The only requirements for the type of play are that:

C1. there are two fixed positive numbers m and q_o such that, in each game, the probability of losing at least m is at least q_o , or

C1′. there are two fixed positive numbers m and p_o such that, in each game, the probability of winning at least m is at least p_o (either condition will guarantee that the probability of continuing indefinitely is zero),

C2. the only possible end results are 0 or B,

C3. each game is fair, i.e. $\mathbf{E}[X_i] = 0$ for all $i \geq 1$.

Thus we have

PROPOSITION 2. (PROPOSITION 1 GENERALIZED) *The probability that a gambler with initial value A will succeed in achieving his goal of B is A/B, whenever the individual bets and the overall strategy satisfy conditions C1–C3.*

According to condition C3, the sequence of random variables $\{S_n\} = \{X_0 + X_1 + X_2 + \cdots + X_n\}$ is a *martingale*. (See Doob's article [**3**], *What is a martingale?* for a good introduction.) Thus Proposition 1 can be viewed as an example of martingale theory. It is interesting to note, however, how the defining property, condition C3, was used by de Moivre to solve the classic gambler's ruin problem almost 200 years before the development of martingale theory.

Solutions to the classic gambler's ruin problem date back as far as 1654. Although no proof of the general case was published until 1711, Edwards shows how the ideas expressed in a series of letters between Pascal and Fermat indicated their knowledge of the general result. Edwards even offers a likely reconstruction of their proofs, based on their approaches to similar problems, and on hints derived from their correspondence [**5**, 1982] [**6**, 1983].

None of the earliest proofs indicated the simplicity of the proof for $p = 1/2$. In the first published proof, however, de Moivre [**2**, 1711] actually used the method of the generalized proposition to solve the classic problem *in all cases, i.e., even if* $p \neq 1/2$. His ingenious approach consisted of changing the values of the coins used for betting, assigning them values in such a way as to guarantee that condition C3 is satisfied. To that end, he imagined that the player with A coins has them all in a pile, but rather than all having the same unit value, the bottom coin is given the value q/p, the one above that is given the value $(q/p)^2$, etc. with the top coin having a value of $(q/p)^A$. His opponent's $B - A$ coins are likewise piled up and given the values $(q/p)^{A+1}$ for the top coin, $(q/p)^{A+2}$ for the one underneath, down to $(q/p)^B$ for his bottom coin. Moreover, the transfer of a coin after any game is always done by placing the loser's top coin on top of the winner's pile. Thus, $\mathbf{E}(X_i)$, the expected value for the first player in any game i, is given by a combination of terms of the form $p(q/p)^{(j+1)} - q(q/p)^j$, all of which equal 0. Replacing A by the new initial fortune of the first player, replacing B by the sum of the two players' initial fortunes, and arguing as in the proposition, de Moivre obtained

$$\mathcal{P}(A, B, p) = \left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^A\right] \Bigg/ \left[\left(\frac{q}{p}\right) + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^B\right]$$

which matches formula (1) for all values of p.

Later, de Moivre attacked the problem of finding the duration in the classic problem. Reassigning the value 1 to each coin, he noted that the expected gain for the gambler in each game is $p - q$. The overall expected gain is $\left[\mathcal{P}(A, B, p)\right](B - A) -$

$\left[1 - \mathcal{P}(A, B, p)\right] A$. Thus, using the fact that the product of the expected number of games with the expected gain per game should equal the overall expected gain, de Moivre concluded that the expected number of games is given by

$$\mathcal{D}(A, B, p) = \left[A - B\,\mathcal{P}(A, B, p)\right] / (q - p) \qquad \text{if } p \neq 1/2. \tag{2}$$

See Thatcher's account [**12**] from 1957. De Moivre's argument cannot be applied if $p = 1/2$. In that case, the result can be obtained by solving a difference equation and ›y Feller [**7**, p. 349] as

$$\mathcal{D}\left(A, B, \frac{1}{2}\right) = A(B - A). \tag{3}$$

Bolder play: increasing the stakes in the classic gambler's ruin

The relation between the gambler's approach in the classic problem and the approach of our anxious gambler can be seen as follows. Suppose the classic gambler were to raise the stakes to \$2 a game, or to any larger amount S. (Technically, of course, this is only possible if both A and B are divisible by S. If that is not the case, one could use a mixed strategy, switching back to \$1 per game when the fortune falls below S or above $B - S$. In our discussion, however, we will simply assume that S is a common divisor of A and B.) If $p = 1/2$, according to our general proposition, the increased stakes would have no effect on the gambler's probability of success. On the other hand, if $p < 1/2$, the probability of success increases.

Feller [**7**, p.346] notes that increasing the stakes to S is equivalent to changing A and B into A/S and B/S, respectively, and he proves that this leads to an increased probability of success if $S = 2$. In a recent note, Isaac [**8**, p. 406] gives a very neat proof of the general result, showing that the corresponding probability of failure is a decreasing function of S, for all positive S. Thus,

$$\mathcal{P}(A/S, B/S, p) > \mathcal{P}(A, B, p) \quad \text{for} \quad p < \frac{1}{2}, \quad S > 1. \tag{4}$$

Since the gambler's probability of winning is the same as his opponent's probability of losing, and since his opponent is playing the same game with A replaced by $B - A$ and p replaced by q, it follows that $P(A, B, p) = 1 - P((B - A), B, q)$. Along with inequality (4), this yields

$$\mathcal{P}(A/S, B/S, p) < \mathcal{P}(A, B, p) \quad \text{for} \quad p > \frac{1}{2}, \quad S > 1. \tag{5}$$

Dubins and Savage offer no formal proof of the above inequalities, but they note [**4**, p. 83] that, for $p < 1/2$, the strong law of large numbers would "stimulate an interest in large bets". Indeed, as the reference to the strong law of large numbers implies, inequalities (4) and (5) go hand in hand with a decrease in the duration of play. The proposition below amplifies this idea.

PROPOSITION 3. *For all values of p between* 0 *and* 1, *and all* $S \geq 2$ *(which divide A and B),* $\mathcal{D}(A/S, B/S, p) < \mathcal{D}(A, B, p)/S$.

Proof. According to de Moivre's formula (2):

$$\frac{\mathcal{D}(A, B, p)}{\mathcal{D}(A/S, B/S, p)} = \frac{S\left[A - B\,\mathcal{P}(A, B, p)\right]}{A - B\,\mathcal{P}(A/S, B/S, p)}.$$

If $p < 1/2$, both $\mathcal{P}(A, B, p)$ and $\mathcal{P}(A/S, B/S, p)$ are less than A/B, so that both expressions on the right side of the equation above are positive. It follows that the inequality in our proposition is equivalent to inequality (4). Similarly, if $p > 1/2$, the proposition is equivalent to inequality (5). Finally, if $p = 1/2$, we can apply formula (3) to obtain a more explicit result:

$$\mathcal{D}\left(\frac{A}{S}, \frac{B}{S}, \frac{1}{2}\right) = \mathcal{D}\left(A, B, \frac{1}{2}\right)/S^2. \qquad \blacksquare$$

According to Proposition 3, $\mathcal{D}(2, 6, p) \geq 2\mathcal{D}(1, 3, p)$ for all p, and (as noted above) $\mathcal{D}(2, 6, 1/2) = 4\mathcal{D}(1, 3, 1/2)$. Some examples of these values and the corresponding values of $\mathcal{P}$ are given in FIGURE 1 for p between .1 and .9.

p	$\mathcal{P}(1, 3, p)$	$\mathcal{P}(2, 6, p)$	$\mathcal{D}(1, 3, p)$	$\mathcal{D}(2, 6, p)$	$\frac{\mathcal{D}(2,6,p)}{\mathcal{D}(1,3,p)}$
0.1	0.0110	0.0002	1.2088	2.4989	2.0672
0.2	0.0476	0.0037	1.4286	3.2967	2.3077
0.3	0.1139	0.0277	1.6456	4.5843	2.7859
0.4	0.2105	0.1203	1.8421	6.3910	3.4694
0.5	0.3333	0.3333	2	8	4
0.6	0.4737	0.6090	2.1053	8.2707	3.9286
0.7	0.6203	0.8214	2.1519	7.3213	3.4022
0.8	0.7619	0.9377	2.1429	6.0440	2.8205
0.9	0.8901	0.9877	2.0879	4.9074	2.3504

Figure 1 All numbers with 4 decimal places are approximate

$P(A, B, p)$ and $D(A, B, p)$ in relation to $\mathcal{P}(A, B, p)$ and $\mathcal{D}(A, B, p)$

To obtain the general formula for $P(A, B, p)$, we again consider the random variables $S_i = X_0 + \cdots + X_i$, which represent the gambler's fortune after i games, and let $R_i = S_i/B$ represent the corresponding ratio of his fortune to his ultimate goal, for $i \geq 0$. (Thus $R_0 = A/B$.) If R_i is less than $1/2$, R_{i+1} will be either 0 (with probability q) or $2R_i$ (with probability p), since S_{i+1} will be either 0 or $2S_i$ in the respective cases. Similarly, if R_i is greater than or equal to $1/2$, R_{i+1} will be either 1 (with probability p) or $2R_i - 1$ (with probability q), since S_{i+1} will either B or $S_i - (B - S_i) = 2S_i - B$. (This recursive formula is found, in a somewhat more abstract setting, in Dubins [**4**, p. 85].) The four cases above allow us to categorize the i^{th} bet ($i \geq 1$) in two ways:

i) The i^{th} bet will be the last (concluding with $S_i = 0$, with probability q, if $R_{i-1} \leq 1/2$; and concluding with $S_i = B$, with probability p, if $R_{i-1} \geq 1/2$).

ii) There will be an $(i + 1)st$ bet. In that case, $R_{i-1} \neq 1/2$, and if R_{i-1} has binary representation $= .b_1b_2b_3\ldots$, R_i will equal the one-position shift $= .b_2b_3b_4\ldots$. This follows since $R_{i-1} < 1/2$ implies $b_1 = 0$ and $2R_{i-1} = .b_2b_3b_4\ldots$, while $R_{i-1} > 1/2$ implies $b_1 = 1$ and $.b_2b_3b_4\ldots = 2R_{i-1} - 1$. Proceeding inductively, then, it follows that the only possible value for R_i, other than 0 or 1, has a binary representation equal to an i-position shift of the binary representation for R_0. As an example, the chart below depicts the possible values of R_i, beginning with $R_0 = A/B = 11/32 = .01011$ (base 2). Obviously, if $R_i = 0$ or 1, all subsequent R_j, $j > i$, have the same value. For simplicity, these "inherited" values of 0 or 1 have been omitted.

To determine $P(A, B, p)$, let W_i denote the event that the gambler achieves his goal of B dollars on the i^{th} bet. Note that W_i is the intersection of the events $R_i = 1$ and $R_j \neq 0$ or $1, 0 < j < i$. In addition, R_i can equal 1 only if $R_{i-1} \geq 1/2$. Thus the first binary digit of R_{i-1}, which equals the i^{th} binary digit of R_0 must be 1. If $R_0 = A/B = \sum \left(\frac{1}{2}\right)^{n_k}$, where the n_ks are positive integers, it follows that $\Pr(W_i) > 0$ if and only if i is equal to n_k, for some k. To assure that $R_j \neq 0$ or 1, $0 < j < i$, while $R_i = 1$, each of the bets $1, 2, \ldots n_k$ must end in a win with the exception of games n_j, $j < k$, which must result in a loss. Thus $\Pr(W_i) = p^{n_k - k + 1} q^{k-1}$. Combining these results yields

THEOREM 1. *For all p, $0 < p < 1$,*

$$P(A, B, p) = \sum \Pr(W_i) = \sum p^{n_k - k + 1} q^{k-1}, \tag{6}$$

where the increasing sequence $\{n_k\}$ represents the positions of the 1s in the binary representation of A/B.

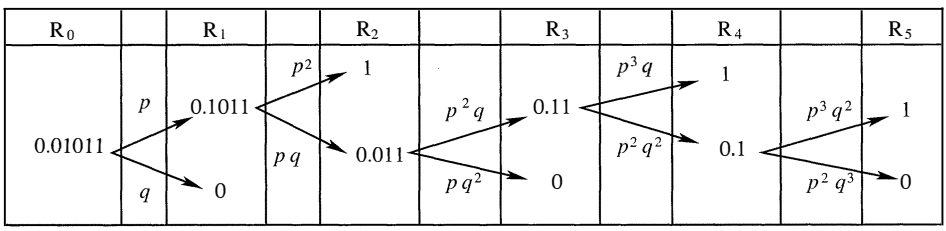


Figure 2

For example, in FIGURE 2 above, $P(A, B, p) = p^2 + p^3 q + p^3 q^2$.

Theorem 1 shows that $P(A, B, p)$ is actually a function of the ratio $r = A/B$, and the probability p. In fact, for fixed p, $P(A, B, p) = P(r, p)$ is a continuous function of the ratio r. If two ratios r_1 and r_2 are sufficiently close, their binary representations will agree in the first m digits, and according to (6), the difference in their associated values of P cannot exceed $\sum_{n_k > m} p^{n_k - k + 1} q^{k-1}$. This, in turn, is easily seen to be less than q^m if $p \leq 1/2$, and p^m for $p \geq 1/2$. Thus the difference approaches 0 as m approaches infinity. Similarly, it is easy to show that P is an increasing function of r for fixed p, and an increasing function of p for fixed r. Note that if $p = q = 1/2$, formula (6) becomes

$$P\left(A, B, \frac{1}{2}\right) = \sum \left(\frac{1}{2}\right)^{n_k} = A/B.$$

The fact that an increase in the stakes *in the classic case* increases the probability of success (with $p < 1/2$) suggests the following theorem on the optimality of the anxious gambler approach:

THEOREM 2. (DUBINS AND SAVAGE) *If $p < 1/2$, $P(A, B, p)$ is at least as large as the probability offered by any strategy subject to the single restriction that the possible payoffs in each game consist of a loss equal to the amount staked, with probability q, or a gain of that same amount, with probability p.*

The proof, involving some general results about optimal strategies and Markov processes, is given by Dubins and Savage [**4**, pp. 87–89]. Thus, $P(A, B, p)$ not only equals or exceeds $\mathcal{P}(A, B, p)$, but is at least as large as any mixed strategy of the type

described above. In fact, it is the ultimate mixed strategy, involving the largest possible reasonable bet at each stage!

We now turn our attention to $D(A, B, p)$, the expected number of games, or duration, in the anxious gambler approach.

THEOREM 3. *If A/B is equal to the terminating binary fraction $\sum_{j=1}^{k} \left(\frac{1}{2}\right)^{n_j}$, then*

$$D(A, B, p) = \frac{1}{q} + \left(\frac{1}{p} - \frac{1}{q}\right) \sum_{j=1}^{k-1} \left(\frac{q}{p}\right)^{j-1} p^{n_j} - \left(\frac{q}{p}\right)^{k-2} p^{n_k - 1}. \tag{7}$$

If A/B has an infinite binary representation of the above form, then

$$D(A, B, p) = \frac{1}{q} + \left(\frac{1}{p} - \frac{1}{q}\right) \sum_{j=1}^{\infty} \left(\frac{q}{p}\right)^{j-1} p^{n_j}. \tag{8}$$

Note: Since $D(A, B, p)$, like $P(A, B, p)$, depends only on the ratio A/B (and p), we can always take A, B relatively prime. In that case, formula (7) applies if $B = 2^{n_k}$ for some integer n_k, and (8) is the appropriate formula in all other cases.

Proof. Suppose $A/B = \sum \left(\frac{1}{2}\right)^{n_j}$, where the sum runs from 1 to k if $B = 2^{n_k}$, and from 1 to infinity otherwise. Let $N(A, B, p)$ denote the number of games until the gambler reaches a conclusion, and for all positive integers i, let $d(i)$ denote the probability that $N(A, B, p) = i$, with $D(i)$ equal to the probability that $N(A, B, p)$ is greater than or equal to i.

By definition, $D(A, B, p) = \sum i d(i)$. We will find it more convenient, however, to obtain $D(A, B, p)$ as the equivalent sum of the series: $\sum D(i)$. To that end, recall the notation R_i which was introduced in the derivation of formula (6), and note that the number of games will be at least n if and only if, for all $j < n - 1$,

i) $R_j \neq 1/2$ and
ii) game $j + 1$ results in a win if $R_j < 1/2$, and
iii) game $j + 1$ results in a loss if $R_j > 1/2$.

Since the binary representation for R_j is simply the binary representation for A/B starting with the $(j + 1)^{st}$ digit, we can derive (7) by sectioning $\sum D(i)$ into parts, of which we show the first, second, and last:

$$\sum_{i=1}^{n_1} D(i) = 1 + p + p^2 + \cdots + p^{n_1 - 1}$$

$$\sum_{i=1+n_1}^{n_2} D(i) = p^{n_1 - 1} q \left(1 + p + p^2 + \cdots + p^{n_2 - n_1 - 1}\right)$$

$$\sum_{i=1+n_{k-1}}^{n_k} D(i) = p^{n_{k-1} - k + 1} q^{k-1} \left(1 + p + p^2 + \cdots + p^{n_k - n_{k-1} - 1}\right)$$

(Note that in this case we need not consider any further terms in the series since, with $B = 2^{n_k}$, $D(i) = 0$ for all $i > n_k$.) The partial sums above can be simplified as

$$(1 - p^{n_1})/q + \frac{1}{p}(p^{n_1} - p^{n_2}) + \frac{1}{p}\frac{q}{p}(p^{n_2} - p^{n_3}) + \cdots$$
$$+ \frac{1}{p}\frac{q}{p}^{k-2}(p^{n_{k-1}} - p^{n_k}),$$

and combining terms with like powers of p yields formula (7). Formula (8) follows by letting k approach infinity and observing that, since $n_k \geq k$, the final expression in (7) is bounded by pq^{k-2} and approaches 0 as k approaches infinity. ■

Notes:

1) As noted above, $D(A, B, p)$, like $P(A, B, p)$, is actually a function of the ratio $r = A/B$, and of p. Unlike, P, however, $D(A, B, p) = D(r, p)$ is a discontinuous function of r. In fact, it has a removable discontinuity at every dyadic rational, and is continuous at all other points.

   To prove the continuity, note that if r is sufficiently close to $r_0 = A_0/B_0$, which has an *infinite* binary representation, the binary representation for r is equal to that of r_0 in the first M digits. Hence, $|D(r, p) - D(r_0, p)|$ is bounded by the tail of the (convergent) series in (8) and is arbitrarily small for sufficiently large M.

   On the other hand, assume $B_0 = 2^{n_k}$ so that $D(r_0, p)$ is given by (7). Then, if $r > r_0$ is sufficiently close to r_0, its binary representation will equal that of r_0, as well as one or more additional binary digits, all in positions arbitrarily far from n_k. Thus, revisiting the original expressions which gave rise to (8), we see that $D(r, p) - D(r_0, p)$ is equal to $(1/p)\,(q/p)^{k-1}\left(p^{n_k} - p^{n_k+M}\right)$ and similar terms involving powers of p beyond $n_k + M$, so that $\lim_{r \to r_0+} [D(r, p) - D(r_0, p)]$ equals $(1/p)\,(q/p)^{k-1}\,p^{n_k}$.

   If $r < r_0$ is sufficiently close to r_0, it has the same binary digits as r_0 with the last digit replaced by a sufficiently large consecutive string of ones starting in the next position. Again, considering the term lost in the expression for $D(r_0, p)$ and the alternative terms introduced shows that $\lim_{r \to r_0-} [D(r, p) - D(r_0, p)]$ is also equal to $(1/p)\,(q/p)^{k-1}\,p^{n_k}$. Thus, $D(r, p)$ has a limit as r approaches r_0 and has a removable singularity at $r = r_0 = A_0/B_0$ with a "gap" equal to $(1/p)(q/p)^{k-1}p^{n_k}$, where $B_0 = 2^{n_k}$ and k is the number of ones in the binary expansion of A.

2) If $p = 1/2$, all the terms of the series in (7) and (8) equal 0. Hence, as long as A, B are relatively prime $D(A/B, 1/2)$ is independent of A. In fact, $D\,(A/B, 1/2)$ equals $2 - (1/2)^{k-1}$ if $B = 2^k$, and is equal to 2 otherwise. Thus, with $p = 1/2$, the countably many discontinuities can be viewed as the markings on an infinitely detailed ruler. Imagine a wooden ruler of width 2 cm, with a 1 cm mark (at the top of the ruler) above the center of the ruler, a 1/2 cm mark above the points representing the fractions 1/4 and 3/4 and so on. Then the length of unmarked wood above each point A/B represents the value of $D\,(A/B, 1/2)$ at that point. See FIGURE 3 below.

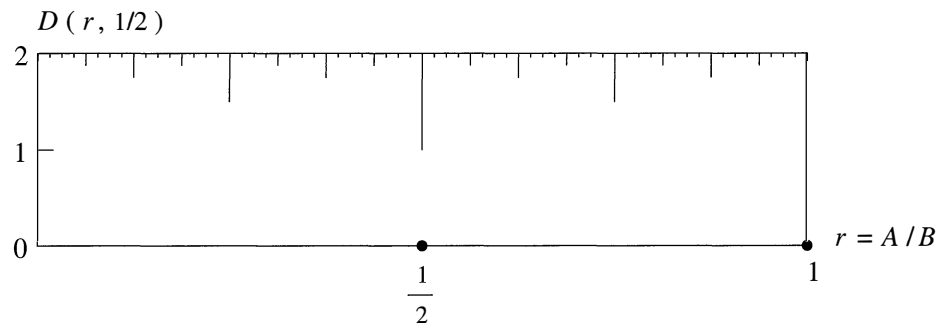


Figure 3

3) For fixed p, $D(r, p)$ is bounded. If

$$p \le \frac{1}{2}, \quad \sup_r D(r, p) = \frac{1}{q} + \left(1 - \frac{p}{q}\right) \sum q^{j-1} = \frac{1}{p}.$$

If $p \ge 1/2$, $\sup_r D(r, p) = 1/q$. This follows from (8) or from the identity $D(A/B, p) = D((B - A)/B, q)$.

To further examine the relation between $D(A, B, p)$ and $\mathcal{D}(A, B, p)$, we first establish the following results about the classic gambler problem:

LEMMA 1. *For all values of p and all positive integers A, $\mathcal{P}(A, A + 1, p) \le 2Ap/(A + 1)$.*

Proof. The proof is by induction on A. If $A = 1$, $\mathcal{P}(1, 2, p) = p$ and the result is proven. To complete the proof, let $P_k = \mathcal{P}(k, k + 1, p)$. Assume $P_k \le 2kp/(k + 1)$. The identity $P_{k+1} = p + q\ P_k\ P_{k+1}$ shows, then, that $P_{k+1} = p/(1 - q\ P_k) \le (k + 1)\ p/(k + 1 - 2pqk) \le 2(k + 1)\ p/(k + 2)$, since $pq \le 1/4$, and the proof is complete. ∎

LEMMA 2. *For all positive integers $A < B$, and all $p \le 1/2$, $\mathcal{D}(A, B, p) \ge A$.*

Proof. According to (3), $\mathcal{D}(A, B, 1/2) = A(B - A) \ge A$. If $p < 1/2$, we note the obvious inequality $\mathcal{D}(A, B, p) \ge \mathcal{D}(A, A + 1, p) = (A - (A + 1)\mathcal{P}(A, A + 1, p))/(q - p)$, according to (2). Hence, according to Lemma 1, $\mathcal{D}(A, B, p) \ge A(1 - 2p)/(q - p) = A$. ∎

The identity $\mathcal{D}(A, B, p) = \mathcal{D}(B - A, B, q)$ shows that the corresponding inequality for $p \ge 1/2$ is $\mathcal{D}(A, B, p) \ge B - A$.

Note that Lemma 2 is, in a sense, unimprovable since $\mathcal{D}(A, A + 1, 1/2) = A$ and for any fixed A and B, $\lim_{p \to 0} \mathcal{D}(A, B, p) = A$.

Finally, to complete our sequence of results highlighting the connection between the chance of success and the expected number of games, we establish the following complement to Theorem 2:

THEOREM 4. *For all integers $A < B$, and $0 < p < 1$,*

$$D(A, B, p) \le \mathcal{D}(A, B, p).$$

As an immediate corollary, $D(A, B, p) = D(A/S, B/S, p) \le \mathcal{D}(A/S, B/S, p)$ for all common divisors, S, of A and B. Thus, the duration in the anxious gambler approach is less than the corresponding duration using fixed bets of any size.

Proof. If $B = 2$ or $B = 3$, the classic approach is identical with that of the anxious gambler. Thus we need only consider $B \ge 4$. If $p = 1/2$, $D(A/B, 1/2) \le 2$ (see note 2 following Theorem 3). On the other hand, $\mathcal{D}(A, B, 1/2) = A(B - A) > 2$ since $B \ge 4$. Thus we can assume $p \ne 1/2$, and by the symmetry of both D and $\mathcal{D}$ about $p = 1/2$, we need only consider $p < 1/2$. We complete the proof by considering four cases, where the inequality is fairly tight in the first case:

a) Assume $A/B \le 1/4$. Then $D(A/B, p) \le \lim D(A/B, p)$ as $A/B \to 1/4$, which is equal to $D(.0011111111\ldots, p) = 1 + p + p^2 + p^2q + \cdots = 1 + 2p$. On the other hand, since $A \le B/4$, $\mathcal{D}(A, B, p) \ge \mathcal{D}(1, 4, p)$. This is clear since B is at least 3 more than A, so that exchanging A and B for 1 and 4, respectively, can only decrease the value of $\mathcal{D}$. Note then that with $A = 1$ and $B = 4$, the number of games will equal 1 with probability q and will equal at least 3 with probability p. Thus $\mathcal{D}(1, 4, p) \ge q + 3p = 1 + 2p$. (As an illustration of the closeness of these estimates for small p, $\mathcal{D}(1/40, .1)$ is roughly 1.11 while $\mathcal{D}(1, 40, .1)$ is about 1.25).

b) If

$$\frac{1}{4} < \frac{A}{B} \le \frac{1}{2}, \mathcal{D}\left(\frac{A}{B}, p\right) \le \lim D\left(\frac{A}{B}, p\right) \text{ as } \frac{A}{B} \to \frac{1}{2}$$
$$= D(.0111111\ldots, p) < 2,$$

while $B \ge 4$ implies $A > 1$ and $\mathcal{D}(A, B, p) \ge A \ge 2$.

c) If $A/B > 1/2$, $\mathcal{D}(A, B, p) \ge A > B/2$. To find an upper bound for $D(A/B, p)$ note that if $B \le 2^n$, then $A/B \le 1 - 1/2^n$ and $D(A/B, p) \le \lim D(A/B, p)$ as $A/B \to 1 - 1/2^n = D(.11111101111\ldots, p)$ where the 0 appears in the $(n-1)$st position. Hence $D(A/B, p) \le n$, and $D(A, B, p) \le \lceil \log_2 B \rceil$, where the latter symbol represents the smallest integer $\ge \log_2 B$. It follows that $D(A/B, p) \le \mathcal{D}(A, B, p)$ as long as $\lceil \log_2 B \rceil \le B/2$. It is easy to see that this inequality is valid for all $B \ge 4$ with the exception of $B = 5$. Since $A > B/2$, the proof will be completed by considering the two special cases, $A = 3$ and $A = 4$:

d) $D(3/5, p) = D(.00110011\ldots, p) = 1 + p + p^2 + p^2q + p^2q^2(1 + p + p^2 + p^2q) + \cdots < (16/15)(1 + p + p^2 + p^2q)$, since $p^2q^2 < 1/16$ and $D(3/5, p) < 3$. A similar argument shows that $D(4/5, p) = D(.11001100\ldots, p) < 4$. On the other hand, $\mathcal{D}(3, 5, p) \ge 3$ and $\mathcal{D}(4, 5, p) \ge 4$, by Lemma 2. ■

The figure below depicts all values of the four variables discussed for $p = .3$ and $B = 20$, as well as the difference between $P(A/B, p)$ and $\mathcal{P}(A, B, p)$, and an indication of the "gap" at the points of discontinuity of $D(A/B, p)$. Note that in this chart, the largest difference between P and $\mathcal{P}$ corresponds to the value of A, namely 17, which also has the largest associated value of $\mathcal{D}$. As expected, the correlation between these two variables is very high, although the exact correspondence of their maxima does not hold for all values of B and p.

A	$P(A/20, .3)$	$\mathcal{P}(A, 20, .3)$	$P - \mathcal{P}$	$D(A/20, .3)$	"Gap"	$\mathcal{D}(A, 20, .3)$
1	0.0043	0.0000	0.0043	1.4368		2.5000
2	0.0144	0.0000	0.0144	1.4560		5.0000
3	0.0300	0.0000	0.0300	1.4858		7.5000
4	0.0480	0.0000	0.0480	1.5200		9.9999
5	0.09	0.0000	0.0900	1.3	0.3	12.4999
6	0.1001	0.0000	0.1001	1.6192		14.9996
7	0.1236	0.0000	0.1236	1.6640		17.4992
8	0.1601	0.0000	0.1601	1.7334		19.9981
9	0.2020	0.0001	0.2019	1.8134		22.4955
10	0.3	0.0002	0.2998	1	1	24.9896
11	0.3101	0.0005	0.3096	2.0192		27.4756
12	0.3336	0.0011	0.3325	2.0640		29.9431
13	0.3701	0.0027	0.3674	2.1334		32.3672
14	0.4120	0.0062	0.4058	2.2134		34.6902
15	0.51	0.0145	0.4955	1.7	0.7	36.7771
16	0.5336	0.0337	0.4999	2.4448		38.3132
17	0.5884	0.0787	0.5097	2.5494		38.5641
18	0.6735	0.1837	0.4898	2.7114		35.8163
19	0.7714	0.4286	0.3428	2.8979		26.0714

Figure 4 All numbers with 4 decimal places are approximate

Implications for sporting events and lotteries

The equivalence of the inequalities in the proof of Proposition 3 showed that an increased probability of success, with $p < 1/2$, goes hand in hand with a reduction in the expected number of games. Theorems 2 and 4 reinforced this idea. This basic idea is a bit of conventional wisdom understood by all sports fans (and commonly expressed as "anything can happen in a short series"). That is, the probability of a weaker team winning a playoff series increases as the number of games in the series decreases.

While this element of unpredictability lends additional interest to shorter series, the various leagues make adjustments to decrease the likelihood that a clearly inferior team will win a sequence of playoff series and emerge as the champion. In some cases, either the lowest ranked teams must play an additional round, or the highest ranked teams draw a "bye", playing one round less. In many sports, in the early, shorter rounds, the lowest ranked teams must compete against the highest ranked, thereby reducing the probability of a win by the underdog, in spite of the shortness of the series.

Lotteries offer an interesting example of increasing the probability of reaching a goal in an unfair game by minimizing the duration of the game. Suppose, for example, that a person with \$10 decides to take a stab at earning \$10,000 by continually playing a game such as blackjack where his probability of winning each game might be .4. Then, *betting \$10 at a time*, the probability that he will reach his goal, $\mathcal{P}(1, 1000, .4)$, would be less that 10^{-175}. If he follows the anxious gambler strategy, his probability of success, $P(10, 10000, .4)$ would go up to just over .0001. According to Theorem 2, this is the optimum probability as long as each bet involves a possible win and a possible loss of equal value.

His probability of success would be almost five times as high, however, if he were given a 1/2000 chance with a payoff of 999:1. This phenomenon, in spite of the obvious decrease in the "fairness" of the game (from an expected loss of \$.20 per dollar waged to an expected loss of more than \$.50 per dollar waged), may be attributed to the fact that the duration of the game has now been reduced to its ultimate low value of 1. It also demonstrates an intuitive sense among the many people who buy lottery tickets or play slot machines, even when alternative games with a higher expected value are available. The truth is that a lottery ticket may actually offer the best available chance of winning a million dollars. Sadly, a recent survey [**10**] showed that many Americans also believe that winning a lottery or sweepstakes offered them the *overall* best chance to obtain half a million dollars or more in their lifetime. Less than half agreed with the assertion that "saving and investing some of their income was the most reliable route to wealth."

An infinitely rich adversary

In the classic case of gambling with fixed stakes, it is possible to consider playing against an infinitely rich adversary. While this adversary cannot be ruined, we can interpret $\mathcal{P}(A, \infty, p)$ as the probability that the gambler with initial fortune A will also never be ruined (and, in fact, will get infinitely rich). Formula (1) shows that $\lim_{B\to\infty} \mathcal{P}(A, B, p) = 0$ if $p \leq 1/2$. However, if $p > 1/2$, $\lim_{B\to\infty} \mathcal{P}(A, B, p) = 1 - (q/p)^A$ so that there is a positive probability of survival, and a corresponding infinite duration: $\lim_{B\to\infty} \mathcal{D}(A, B, p) = \infty$.

On the other hand, using the anxious gambler approach versus an infinitely rich opponent would be obviously catastrophic. We have already seen that increasing the stakes reduces the probability of success when $p > 1/2$. In this case, the probability of success would shrink to 0, since the probability of surviving through n games would

equal p^n. In fact, $N(A, \infty, p)$, the number of games until ruin, would have a geometric distribution with probability q, and $D(A, \infty, p)$ would equal the finite value $1/q$ (see Note 3 following Theorem 3). Hence we would have the paradoxically unfortunate situation where the expected value after any finite number of games would be positive *and increasing*, while the ultimate expected value would be 0.

REFERENCES

1. P. Billingsley, *Probability and Measure*, 3rd ed., Wiley, New York, 1995.
2. A. de Moivre, De Mensura Sortis, *Phil. Trans. R. Soc.* **27** (1711), 213–264. Translated in *Internat. Statist. Rev.*, **52** (1984), no. 3, 237–262.
3. J. L. Dobb, What is a martingale?, *Amer. Math. Monthly* **78** (1971), 451–463.
4. L. E. Dubins and L. J. Savage, *How to Gamble If You Must: Inequalities for Stochastic Processes*, McGraw-Hill, New York, 1965.
5. A. W. F. Edwards, Pascal and the problem of points, *Internat. Statist. Rev.* **50** (1982), 259–266.
6. A. W. F. Edwards, Pascal's problem: The 'Gambler's Ruin', *Internat. Statist. Rev.* **51** (1983), 73–79.
7. W. Feller, *An Introduction to Probability Theory and Its Application*, 3rd ed., Wiley, New York (1968).
8. R. Isaac, Bold play is best: a simple proof, this MAGAZINE **72** (1999), 405–407.
9. W. D. Kaigh, An attrition problem of Gambler's Ruin, this MAGAZINE **52** (1979), 22–25.
10. Jonathan Karl and Associated Press, Many see lottery, not saving, as way to wealth, `www.cnn.com/US/9910/28/savings/`, Associated Press, 1999.
11. E. Shoesmith, Huygen's solution to the Gambler's Ruin problem, *Historia Math.* **13** (1986), 157–164.
12. A. R. Thatcher, A note on the early solutions of the problem of the duration of play, *Biometrika* **44** (1957), 515–518.

50 Years Ago in the MAGAZINE

In Volume 24, No. 1 (January–February, 1951), there appeared an article "On approximating the roots of an equation by iteration," by Jerome Hines. Mr. Hines explained ways of accelerating Picard iteration, including a nice exposition of Newton's Method. His biographical sketch included the following:

> Jerome Hines, well known singer with the Metropolitan Opera Company, wrote his paper, appearing in this issue, while an undergraduate at the University of California at Los Angeles. While in college he majored in both chemistry and mathematics.... Mr. Hines won the Metropolitan $1000 Caruso award and has been with the Metropolitan since 1946–47. He has more than 30 operatic roles in his repertoire, including that of Swallow which he created at the Metropolitan premiere of "Peter Grimes." Despite the crowded life of a Metropolitan star Mr. Hines manages to continue his studies in mathematics, in which he became especially interested while in college....

Counting Perfect Matchings in Hexagonal Systems Associated with Benzenoids

FRED J. RISPOLI
Dowling College
Oakdale, NY 11769

Introduction

Many hydrocarbons involve hexagonal rings; benzene, consisting of a single hexagon of carbons with hydrogens attached, is the original example. It turns out that the number of perfect matchings in certain associated graphs is relevant to the chemistry of these hydrocarbons. Furthermore, standard methods of undergraduate linear algebra and discrete mathematics can be used to count the matchings, and familiar counting numbers show up. In this article we present this easily accessible but seemingly little known connection between chemistry and mathematics.

According to the chemistry folklore, the German chemist August Kekulé (1829–1896) discovered the molecular structure of benzene after he dreamed of a snake swallowing its own tail. Apparently, the dream led to his conjecture that benzene consists of 6 carbon atoms, each linked to 1 hydrogen atom via a carbon-hydrogen bond, and that the carbon atoms are linked to each other via a cycle of length 6 consisting of alternating single and double carbon-carbon bonds. FIGURE 1 illustrates a molecular model of benzene. Kekulé's discovery initiated the study of special types of graphs used to model benzene-like molecules called benzenoids.

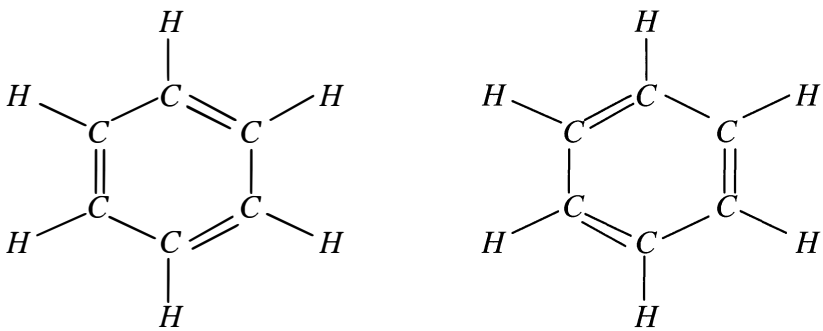


Figure 1 The molecular structures associated with benzene

A graph G is called 2-connected if it is connected and at least 2 vertices must be removed to make G disconnected. A *hexagonal system* is a 2-connected planar graph such that each bounded face can be drawn (and will be drawn) as a regular hexagon (see, e.g., FIGURES 2-5). Notice that this condition forces the valence (i.e., degree) of each vertex in a hexagonal system to be either 3 or 2; vertices of valence 2 can appear only on the unbounded face. Moreover, each pair of adjacent hexagons has exactly one edge in common.

Given a graph G, a *perfect matching* M in G is a subgraph containing all the vertices of G such that every vertex has degree 1; the number of perfect matchings in G is denoted by $\Phi(G)$. In essence, a perfect matching is a pairing of two subsets (say,

white and black) of vertices. For example, in FIGURE 2 the thick edges in H_2 illustrate a perfect matching. The reader should confirm that $\Phi(H_1) = 2$, $\Phi(H_2) = 3$, and $\Phi(H_3) = 0$. (Hint: For H_3, consider the parity of black and white vertices.)

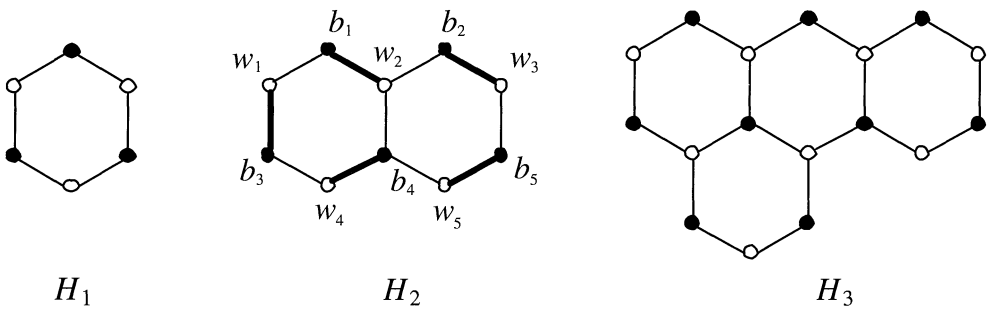


Figure 2 Hexagonal systems. $\Phi(H_1) = 2$, $\Phi(H_2) = 3$, and $\Phi(H_3) = 0$

A *benzenoid* is a special type of hydrocarbon molecule. Given a molecular model of a benzenoid, its corresponding hexagonal system H is obtained by removing the edges representing carbon-hydrogen bonds and letting the remaining edges of H represent either single or double carbon-carbon bonds. The graph H_1 in FIGURE 2 is the hexagonal system corresponding to benzene. It turns out that all hexagonal systems that arise from benzenoids admit perfect matchings, and that each perfect matching is a possible location for all the double carbon-carbon bonds. Conversely, experimental chemistry tells us that a benzenoid may be synthesized for each hexagonal system containing a perfect matching. Therefore, chemists are interested in knowing, for a given hexagonal system H, whether $\Phi(H) = 0$. In addition, chemical properties of a benzenoid such as stability and energy levels depend on the number of perfect matchings in its corresponding hexagonal system, so chemists seek efficient methods to calculate $\Phi(H)$. (For more details, see [**2**], [**5**], and [**7**].)

For an arbitrary graph G with n vertices, determining whether $\Phi(G) = 0$ can be solved using Edmond's matching algorithm, which requires $O(n^3)$ operations (see, e.g., [**3**]). For hexagonal systems, more efficient algorithms are known (see, e.g., [**4**] and [**6**]). However, computing $\Phi(G)$ is known to be an NP-hard problem (that is, there is no algorithm to compute $\Phi(G)$ involving $O(n^k)$ operations, where k is a fixed constant). This is so even if G is *bipartite* (that is, a graph with vertex set $V = V_1 \cup V_2$, such that $V_1 \cap V_2 = \emptyset$ and all edges join vertices in V_1 to vertices in V_2. (For more on matching algorithms and computing Φ, see [**3**].)

For some special classes of graphs, such as planar graphs, Φ is more easily determined. Here we describe how to determine Φ for hexagonal systems in particular, by calculating the determinant of a certain adjacency matrix. We also obtain simple explicit formulas for Φ for some special classes of hexagonal systems.

Counting perfect matchings

Consider the hexagonal systems given in FIGURE 2. All three graphs are bipartite, since we can color the vertices black and white such that no two adjacent vertices have the same color. In fact, every hexagonal system is bipartite (this can be proved by induction on the number of hexagons), so every hexagonal system may be represented by a matrix, defined as follows. Let H be a hexagonal system and let E denote the set of edges in H. Let $B \cup W$ be the set of vertices of H, where $B = \{b_1, \ldots, b_n\}$,

$W = \{w_1, \ldots, w_m\}$, and all edges in H join vertices in B to vertices in W. The *biadjacency matrix*, written $A(H) = [a_{ij}]$, is defined by $a_{ij} = 1$ if the edge $\{b_i, w_j\} \in E$, and $a_{ij} = 0$ if $\{b_i, w_j\} \notin E$.

We'll assume henceforth that B and W contain the same number of vertices (i.e., $|B| = |W|$) since this is a necessary condition for the existence of a perfect matching. In this case, the biadjacency matrix is square. For example, the biadjacency matrix for the hexagonal system H_2 in FIGURE 2 is

$$\begin{array}{c} \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{array} \begin{array}{c} \begin{array}{ccccc} w_1 & w_2 & w_3 & w_4 & w_5 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array}$$

Recognizing a hexagonal system H as a bipartite graph with edges from $B = \{b_1, \ldots, b_n\}$ to $W = \{w_1, \ldots, w_n\}$ sets up a correspondence between perfect matchings in H and permutations of $\{1, 2, \ldots, n\}$. For a given hexagonal system H and a perfect matching M in H, with associated permutation σ in the symmetric group S_n, we define the *sign* of M to be $+1$ if σ is an even permutation and -1 if σ is odd.

For instance, the perfect matching shown in H_2 of FIGURE 2 corresponds to the permutation $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{smallmatrix}\right)$. Thus, for every hexagonal system H with $|B| = |W|$ there is a one-to-one correspondence between the nonzero terms in the expansion of the determinant of $A(H)$ and the perfect matchings in H. Indeed, more is true:

THEOREM 1. *For a hexagonal system H, $\Phi(H) = |\det A(H)|$.*

Notice that since each entry a_{ij} in $A(H)$ is either 0 or 1, the nonzero terms in the expansion of $\det A(H)$ are all either 1 or -1. Thus the theorem holds if we know that all the nonzero summands in $\det A(H)$ have the same sign. Lemmas 1 and 2 imply this, and so prove Theorem 1.

In what follows we let $|G|$ denote the number of edges in a graph G. We also recall *Euler's formula* for a connected planar graph G: If G has v vertices, e edges, and f faces (including the unbounded face), then $v - e + f = 2$. (For a proof, see, e.g., [**8**].)

LEMMA 1. *Let H be a hexagonal system and let M and M^* be perfect matchings in H. Then every cycle C in $M \cup M^*$ satisfies $|C| \equiv 2 \bmod 4$.*

Proof. Let C be a cycle in $M \cup M^*$. Let r be the number of hexagons inside C, v_{int} the number of vertices inside C, and e_{int} the number of edges inside C. Applying Euler's formula to C and its interior gives $(v_{\text{int}} + |C|) - (e_{\text{int}} + |C|) + (r + 1) = 2$, so $e_{\text{int}} = v_{\text{int}} + r - 1$. Since every hexagon has 6 edges and every edge in the interior of C is in exactly two hexagons, the number of edges in C and its interior is $e_{\text{int}} + |C| = 6r - e_{\text{int}}$. The last two equations imply that $|C| = 4r - 2v_{\text{int}} + 2$.

To complete the proof we need only show that v_{int} is even. Since the edges on C come alternately from M and M^*, no vertex on C can be matched to any vertex in the interior of C by an edge in M. Since H is a planar graph and M is a perfect matching, each vertex b in the interior of C lies in a unique edge $\{b, w\}$ in M, with w also in the interior of C. Thus v_{int} is even. ∎

LEMMA 2. *In a hexagonal system H, all perfect matchings have the same sign.*

Proof. (The example following this proof illustrates the idea.) Let M and M^* be any two perfect matchings in H. Without loss of generality, we may label the vertices in H so that M corresponds to the identity permutation, say σ. Then M^* corresponds to a permutation we denote by σ^*.

Now $M \cup M^*$ is a union of disjoint cycles and isolated edges. Let $C_1, \ldots, C_k$ denote the cycles in $M \cup M^*$. By Lemma 1, $|C_i| \equiv 2 \bmod 4$ for all i. Observe that each C_i corresponds to a cyclic permutation, say σ_i. Moreover, the length of each cyclic permutation σ_i is $|C_i|/2$, which is odd. Therefore each σ_i can be factored into an *even* number of transpositions. Since $\sigma^* = \sigma \cdot \sigma_1 \cdot \sigma_2 \cdots \sigma_k$, both σ and σ^* have the same sign. ■

Example. Consider the hexagonal system H given in FIGURE 3. Let M and M^* be the perfect matchings shown in the right-hand figure: the solid edges are in M, the dotted edges in M^*, and the thick edges in $M \cap M^*$. Then M corresponds to the identity, and M^* corresponds to $\sigma^* = \left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 2 & 3 & 1 & 4 & 7 & 11 & 8 & 6 & 9 & 10 \end{smallmatrix}\right)$. Notice that $M \cup M^*$ contains cycles of length 6 and 10, and that σ^* can be factored into cyclic permutations of length 3 and 5, given by $\sigma^* = \sigma_1 \cdot \sigma_2$, where (in cycle notation) $\sigma_1 = (1\ 5\ 4)$ and $\sigma_2 = (6\ 7\ 11\ 10\ 9)$.

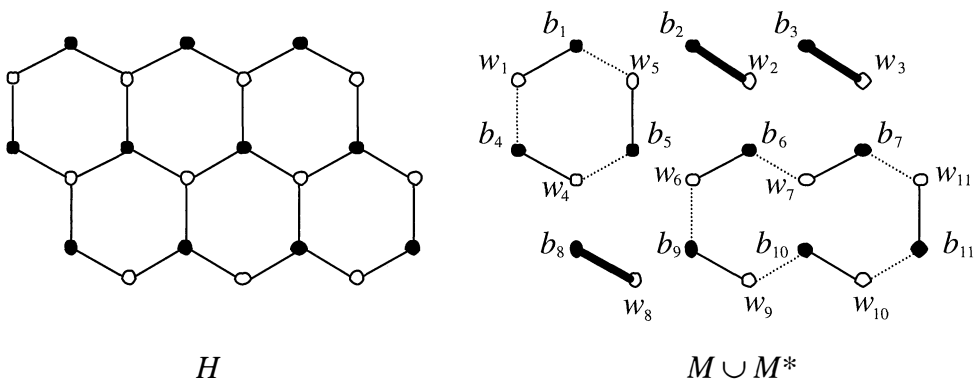


Figure 3 The union of perfect matchings in a hexagonal system

In summary, to compute $\Phi(H)$ for a hexagonal system H, label the vertices, obtain a biadjacency matrix $A(H)$, and calculate $|\det A(H)|$. For example, the matrix $A(H_2)$ for H_2 in FIGURE 2 has determinant 3, which verifies that $\Phi(H_2) = 3$. A larger example—with 980 perfect matchings—is shown in FIGURE 4. This counting method is due to Kasteleyn [**5**]; a comprehensive discussion also appears in [**3**].

Special classes of hexagonal systems Certain classes of hexagonal systems have special structures that further simplify computing Φ. Several examples appear in FIG-

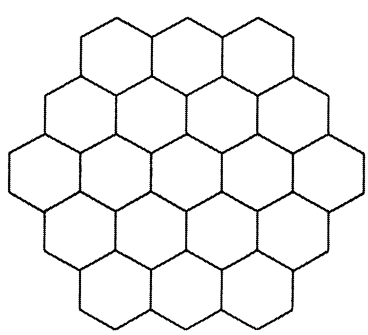


Figure 4 A hexagonal system with 980 perfect matchings

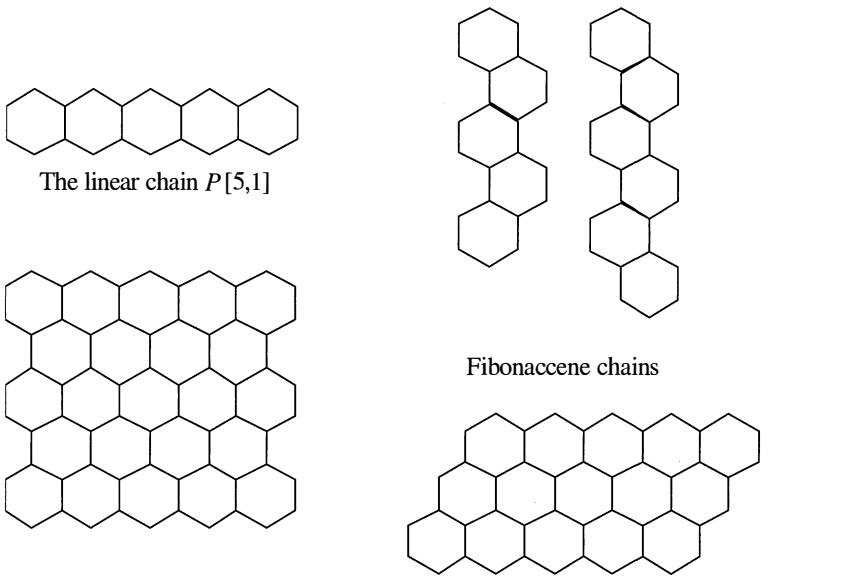


Figure 5 Special classes of hexagonal systems

URE 5. For convenience, we draw all hexagonal systems so that (as shown) each hexagon has east and west vertical edges; the remaining edges are northeast, northwest, southeast, and southwest.

In Theorems 2, 3, and 4, which follow, we give formulas for counting perfect matchings in the systems of FIGURE 5. Various chemists have discovered these formulas; for more details, see [**7**].

A hexagonal system is called a *fibonaccene chain* if it consists of a chain of hexagons $H_1, \ldots, H_p$, with H_1 on top, and only the following shared edges: For i even and $1 < i < p$, H_i shares its northwest edge with H_{i-1} and its southwest edge with H_{i+1}; H_1 and H_p share edges as indicated in FIGURE 5. The name is due to the fact that Φ satisfies a Fibonacci-style recurrence relation, as we now prove.

THEOREM 2. *Let H be a fibonaccene chain with h hexagons; let $a_h = \Phi(H)$. Then $a_0 = 1$, $a_1 = 2$, and $a_h = a_{h-1} + a_{h-2}$ for $h \geq 2$.*

Proof. Let H be the chain $H_1, \ldots, H_h$, and let M be a perfect matching in H. If M contains the *northwest* edge of H_1, it must also contain the southwest edge of H_1 and the east vertical edges of both H_1 and H_2. The remaining edges in M can be any perfect matching in the hexagonal system $H_3, \ldots, H_h$. Hence there are a_{h-2} perfect matchings in H that contain the northwest edge of H_1.

If, instead, M contains the *northeast* edge of H_1, it must also contain the west vertical edge of H_1, and the remaining edges of M can now be any perfect matching of $H_2, \ldots, H_h$. Hence there are a_{h-1} perfect matchings in H that contain the northeast edge of $H1$. Since every perfect matching of H contains either the northwest edge or the northeast edge of H_1, but not both, $a_h = a_{h-1} + a_{h-2}$. ■

FIGURE 5 shows fibonaccene chains with 5 and 6 hexagons. By Theorem 2, we may calculate Φ using the Fibonacci sequence 1, 2, 3, 5, 8, 13, 21; here $a_5 = 13$ and $a_6 = 21$.

A hexagonal system consisting of h hexagons such that all adjacent pairs of hexagons share exactly one vertical edge and no nonvertical edges is called a *linear chain* of length h. A *parallelogram hexagonal system*, denoted by $P[h, p]$, consists of p linear chains $L_1, \ldots, L_p$, each of length h, such that for $i = 1, \ldots, p-1$, all southern edges of hexagons in L_i, except for the southeast edge of the eastern-most hexagon in L_i, are also northern edges of hexagons in L_{i+1}, except for the northwest edge of the western-most hexagon in L_{i+1}. FIGURE 5 shows $P[5, 1]$ and $P[5, 3]$.

For n and r nonnegative integers, with $0 \leq r \leq n$, we write

$$C(n, r) = \frac{n!}{r!(n-r)!}.$$

THEOREM 3. *For a parallelogram hexagonal system $P[h, p]$, we have*

$$\Phi(P[h, p]) = C(h + p, p).$$

Proof. The proof is by induction on $k = h + p$. Clearly, the formula holds for $P[1, 1]$ (which corresponds to benzene). Notice that $P[1, h]$ and $P[h, 1]$ are linear chains of length h; it is easy to see that $\Phi(P[h, 1]) = h + 1$. Thus the formula holds for both $P[h, 1]$ and $P[1, h]$. Assume that the result holds for all parallelogram hexagonal systems $P[h, p]$ with $h + p = k$.

Consider $P[h, p + 1]$ (the case $P[h + 1, p]$ is similar). Let L be the northern-most linear chain of $P[h, p + 1]$, and let M be a perfect matching of $P[h, p + 1]$. Suppose that M contains the eastern-most *vertical* edge of L. Then M must also contain the northwest edge of every hexagon in L. The remaining edges in M can be any perfect matching of the hexagonal system $P[h, p + 1]$ with L removed—that is, $P[h, p]$. By the inductive assumption, there are $C(h + p, p)$ such perfect matchings.

Now suppose that M contains the eastern-most *northeast* edge of L. Then M must also contain the southeast edge of the eastern-most hexagon in every linear chain (i.e., row) in $P[h, p + 1]$. The remaining edges of M can now be any perfect matching of the hexagonal system $P[h, p + 1]$ with the eastern-most hexagon removed from every row—that is, $P[h - 1, p + 1]$. By the inductive assumption, there are $C(h + p, p + 1)$ such perfect matchings.

Every perfect matching in $P[h, p + 1]$ contains either the eastern-most vertical edge of L or the eastern-most northeast edge of L, but not both. Thus $\Phi(P[h, p + 1]) = C(h + p, p) + C(h + p, p + 1)$, so by Pascal's identity we have $\Phi(P[h, p + 1]) = C(h + p + 1, p + 1)$. ∎

A *rectangular hexagonal system*, denoted by $R[h, p]$, consists of p linear chains $L_1, \ldots, L_p$ of length h, together with $p - 1$ linear chains $\bar{L}_1, \ldots, \bar{L}_{p-1}$ of length $h - 1$, such that for $i = 1, \ldots, p - 1$, all northern edges of $\bar{L}_i$ are southern edges of L_i, and all southern edges of $\bar{L}_i$ are northern edges of L_{i+1}. For example, $R[5, 3]$ appears in FIGURE 5. A proof of the following theorem is left as an exercise.

THEOREM 4. *For a rectangular hexagonal system $R[h, p]$, we have $\Phi(R[h, p]) = (h + 1)^p$.*

By Theorem 3, $\Phi(P[5, 3]) = C(8, 3) = 56$; by Theorem 4, $\Phi(R[5, 3]) = 6^3 = 216$.

It is natural to wonder: Are the special classes discussed here common in nature? The answer is yes. For example, *naphthacene* is the benzenoid whose hexagonal system is the linear chain $R[4, 1]$; it is used to help derive the antibiotic *aureomycin*. *Chrysene*, the benzenoid whose hexagonal system is a fibonaccene chain with 4 hexagons, is present in coal heated at high temperatures. *Benzo[a]pyrene*, a compound known to be

present in tobacco smoke, can be obtained by combining pyrene (which corresponds to $P[2, 2]$) with benzene. For further discussion about some remarkable chemical properties of benzenoids (also known as aromatic compounds), and connections between chemistry and hexagonal systems, see [**1**] or [**7**].

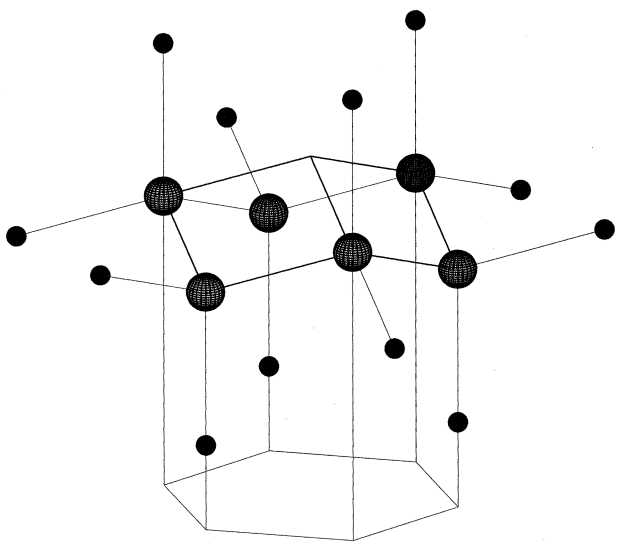
Acknowledgment. We thank the referee for many useful suggestions.

REFERENCES

1. J. Aihara, Why aromatic compounds are stable, *Scientific American*, March 1992, 62–28.
2. S. Cyvin and I. Gutman, *Kekulé Structures in Benzenoid Hydrocarbons*, Springer-Verlag, New York, NY, 1988.
3. L. Lovasz and M. D. Plummer, *Matching Theory*, North–Holland, Amsterdam, The Netherlands, 1986.
4. P. Hansen and M. Zheng, A linear algorithm for perfect matching in hexagonal systems, *Discrete Math.* **122** (1993), 179–196.
5. P.W. Kasteleyn, Graph theory and crystal physics, in F. Harary, editor, *Graph Theory and Theoretical Physics*, Academic Press, New York, NY, 1967.
6. H. Sachs, Perfect matchings in hexagonal systems, *Combinatorica* **4** (1984), 89–99.
7. N. Trinajstić, *Chemical Graph Theory*, 2nd ed., CRC Press, Boca Raton, FL, 1992.
8. R. Wilson and J. Watkins, *Graphs: An Introductory Approach*, Wiley, New York, NY, 1990.

Cyclohexane and Honeycomb Cells

The benzene rings of Rispoli's article are planar. Six carbon atoms can also join to form a skew hexagon, as seen in the cyclohexane molecule. Here, the bonds between carbons are single bonds, and each carbon has two hydrogen atoms attached. The cyclohexane molecule has been superimposed on the ideal shape of a honeycomb cell, FIGURE 2A from the next article. Is it a coincidence that these two skew hexagons from nature have exactly the same shape?



Thanks are due to John Thoburn of the Santa Clara University Department of Chemistry.

Using Less Calculus in Teaching Calculus: An Historical Approach[1]

R. M. DIMITRIĆ
Berkeley Basic Research Institute
1176 University Ave.
Berkeley, CA 94702

I would like to compare through some historical examples the use of non-calculus (mostly geometric or algebraic) and calculus methods in teaching students mathematical content. In the present college and precollege school system the latter methods are overdone at the expense of the former, in cases where both can be used to explain material or solve problems. Examples that follow are aimed at students of varying levels of sophistication. These range from facts about e to the ideal shape of honeybee cells.

Introduction

It is well known [**5**] that the original Euclid's *Elements* contained few geometrical drawings, abstract in their nature. Neat explanatory drawings and constructions were added by Euclid's translators and commentators at times when doing mathematics by geometric means flourished. There are indications that Isaac Newton did not like the prevalent geometric method of his time (as much as he utilized it) [**7**] and that this dislike played a role in his shared discovery of calculus. Gradually, geometric reasoning and visualization were forgotten and analytic methods became king in texts of Lagrange, Russell [**19**], and other members of the French, German and English mathematics schools.

Calculus nowadays has the same role category theory will have in the future: it is used mainly as a unifying and generalizing tool that can tie together seemingly separate problems by resolving them via the same (calculus) methods. Both of these tools (and here we refer to them as tools in teaching) are often overdone and abused to the point where simple problems are treated with the heavy machinery that is inappropriate for a particular context. This practice has negative didactic consequences for it obfuscates intuition and reduces learning to rote mechanical performance of calculus rules. I think that a careful combination of ("purely") non-calculus (geometric/algebraic) and calculus methods is most conducive to learning new material in analysis and demonstrating its real powers. History of mathematics offers a good source of examples that can be used to compare geometric and calculus methods.

Powers and radicals

It is useful to teach easy properties of powers and radicals, such as the following: If $a > 0$, then $\sqrt[n]{a} > 1$ for $a > 1$ and $\sqrt[n]{a} < 1$ for $a < 1$; in addition, $\lim_{n\to\infty} \sqrt[n]{a} = 1$. Now, what about a more difficult question, namely to find $\lim_{n\to\infty} \sqrt[n]{n}$? A hint might come from playing with the calculator (there is in fact a sheer pleasure of calculating the values of $\sqrt[n]{n}$): $\sqrt{2} \approx 1.4142$, $\sqrt[3]{3} \approx 1.4422$, $\sqrt[4]{4}$ (I pause here to break the routine

[1]The author gave a talk on this subject at the AMS–MAA Meeting in Baltimore, January 1998, in a special section on the role of history in teaching mathematics.

and ask the students for the answer, without a calculator—they should note the delightful fact that the fourth root of four equals $\sqrt{2} \approx 1.4142$), $\sqrt[5]{5} \approx 1.3797$, $\sqrt[6]{6} \approx 1.3480$, $\sqrt[7]{7} \approx 1.3205$, and so on. (I branch out by asking them to find all natural numbers c, d, such that the c-th root of c is same as the d-th root of d; in fact, 2 and 4 are the only such pair, which can be inferred from the subsequent discussion.) Some patterns in these roots would emerge, namely that $\sqrt[n]{n}$ first increases, then decreases (where the turn happens between 2 and 4) and if one goes far enough, one might see that the sequence converges to 1, which, of course, can be rigorously justified. The notion of powers with fractional exponents should have been already discussed and this is a good time to talk about exponents that are not necessarily rational. I leave students to ponder on the lingering questions such as

$$\text{Which is greater: } \sqrt[c]{c} \text{ or } \sqrt[d]{d}, \text{ for } c, d > 0? \tag{1}$$

and whether indeed $\lim_{n\to\infty} \sqrt[n]{n} = 1$. In fact, these question are answered by exploring the number e, as follows.

The Euler number e

> If thou lend money to any of my people that is poor by thee, thou shalt not be to him as an usurer, neither shalt thou lay upon him usury. (Exodus 22:25, King James Version; see also Leviticus 25:36 and Deuteronomy 23:19,20)

Discussions on powers should also lead to an introduction of e, a number not as famous as π at the elementary level (although $\pi = \ln \sqrt[i]{-1}$). An enticing way to introduce e is through the mixed-up envelopes problem of Nikolaus Bernoulli I (1687–1759) (see [**18**, p. 46], where reference is made to De Montmort: *Essai d'analyse sur les jeux de hasard*, Paris, 1713): suppose n letters are written to n different persons, whose addresses are written on n different envelopes. What is the probability that all n letters will be put in wrong envelopes? (Brawner [**4**] and Margolius [**15**] present nice discussions of this and similar problems in the MAGAZINE.) It turns out that this probability is $1/e$ when $n \to \infty$. This reciprocal is in fact close to the "base" that John Napier (1550–1617) first used when he anticipated what later became the familiar logarithms. His contemporary Jobst Bürgi (1552–1632) had similar ideas, except his base was close to e. This approach is somewhat convoluted and I prefer to ask this question after I bring e into the picture as follows.

By way of a non-rigorous (but inspiring) presentation, I also avoid introducing e the way Euler (1707–1783) introduced it (via the infinite binomial series [**9**]), and use instead the computation of interest: \$ 1 is deposited at 100% annual interest compounded n times per year. It is plausible that the "bankers" have arrived at these calculations before everybody else, minus the problematically high interest rate. Whence the amount has grown to $(1 + 1/n)^n$ at the end of year one, or to $(1 + r/n)^{nx}$, after x years, if the annual rate is r.

For nonstop compounding we let n grow unchecked and want to see how much money is available at the end of a year. At the rigorous level, this sequence is shown to be convergent, and at the relaxed level calculators come in handy to dispel the popular student opinion that the former sequence converges to 1, rather than to the new mysterious number e (as shown by Daniel Bernoulli (1700–1782)). The money accumulated cannot be \$ 1 for another reason: surely the interest added something to the principal. Note that if the money is left to mature for time x and at the same interest rate, with simple interest, then the money grows to $1 + x$.

It is worth noting that with continuous compounding at the rate r during a period x the amount accumulated is the same as with continuous compounding at the rate x and the time r.[2] This may look intuitively clear to some, and to the others the fact that both of the following two quantities converge to e^{rx} should be convincing enough to that end (the standard replacement r/n, $x/n = 1/k$, establishes this quickly, for large n):

$$\left(1+\frac{r}{n}\right)^{xn} \approx \left(1+\frac{x}{n}\right)^{rn}. \tag{2}$$

By this heuristic, the exponential function $y = e^x$ is introduced. My question is always: "What is more profitable: to have computation of interest done at the year end at 100% interest, or at 50% interest every 6 months, with the interest from the first half of the year added to the principal for computation of interest at the end of the second half?" The answer is unanimously that the latter method will give more interest (I do not know whether the answer would be as unanimous in less monetarily conscious cultures). And so much the better for continuous compounding, we arrive at an intuitive understanding of the following nice inequality

$$e^x \geq 1 + x \tag{3}$$

(the lost art of inequalities!). The inequality is clearly true for $x \leq -1$, but also for $x \in (-1, 0)$; the latter can be seen from an inequality $e^t \geq et$, $t \in (0, 1)$ that may be intuitive to some (the money e^t that \$ 1 accumulates at the rate of 100% after a fractional time t with continuous compounding is greater than the t-"prorated" amount et of the final amount e). Otherwise, it is justified as in the sequel.

In fact, it is instrumental to have Jakob Bernoulli's (1654–1705) inequality set beforehand:

$$(1+a)^n \geq 1 + na, \text{ for every } a \geq -1, \text{ and natural } n$$

([**2**], [**3**, p.380]; see also [**1**, Lectio VII, §XIII, p. 224]). Proving this inequality by induction is straightforward, but one can resort to comparison of compounding interest (once a year, with gain in interest equal a, after n years) and simple interest, for an equally elegant intuitive proof. The case when a is negative can be interpreted via depreciation of property. Incidentally, this inequality can be used in a rigorous proof that the sequence for e converges.

Now we can make a calculus inference: drawing the graphs of $(1 + x)^n$ and $1 + nx$ we note that the two curves have the point (0, 1) in common, and the Bernoulli inequality does not allow any more common points to the right of -1, hence the line is the tangent to the curve at this point. The calculus fact we derive is that the slope of the tangent at (0,1) to the curve $(1 + x)^n$ is n (FIGURE 1b).

Using Bernoulli's inequality, we have $(1 + \frac{x}{n})^n \geq 1 + x$, for every natural n and every real $x \geq -n$. Thus, passing to the limit we again get our inequality $e^x \geq 1 + x$, *for all* x, and we conclude that the slope of the tangent of e^x at (0,1) is $1 = e^0$—see the graphs in FIGURE 1a. Moreover we can also claim that $e^h \approx 1 + h$, for small h, for in this case the time to maturity is as small as we wish and the simple interest calculation is close to a continuous one (for this is the meaning of the tangent to e^x at (0, 1)). This is good to know, for if we are looking at a general tangent to e^x at some point (x_0, e^{x_0}) with slope m, then for x close to x_0, we have $e^x \approx m(x - x_0) + e^{x_0}$ (tangent approximated by a secant); algebra leads to $e^{x_0}\frac{e^h-1}{h} \approx m$, where we set $h = x - x_0$.

[2]The interchangeability of time and interest is a deep secret around which the world revolves and the old proverb should actually be rephrased: "Time is interest!"

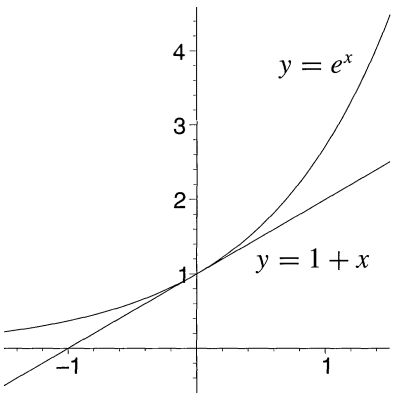


Figure 1a

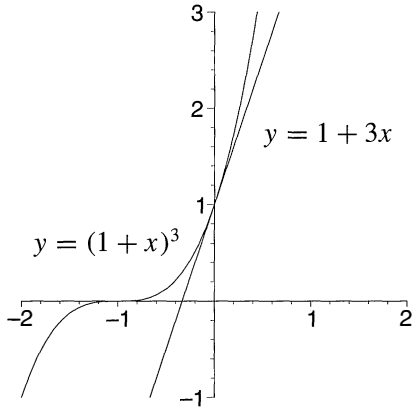


Figure 1b

The consequence is that the slope at a point (x, e^x) is same as e^x, which is another crucial calculus fact, obtained yet again by algebraic reasoning. Maor [**14**] has more information on the events related to the development of the story of e.

When I have clarified the inequality (3), I ask whether the same inequality would hold for other bases greater than 1. At this initial stage when students' grasp of exponential functions and calculus is meager, they are tempted to think that the same inequality holds regardless of the base, especially if they rely on loose graphs to judge (to use educational jargon: this is the teacher's *pedagogical content knowledge (PCK)*—the teacher's ability to predict (incorrect) responses of the students). The proper inequality is

$$a^x \geq x \ln a + 1, \tag{4}$$

which we can establish using once more our fundamental inequality: $a^x = e^{x \ln a} \geq 1 + x \ln a$. The tangent at 0 has slope $\ln a$. There is yet another way to look into this inequality, namely as *dual* to (3). If $f(x) \geq g(x)$, for invertible functions f, g, then their inverses satisfy $f^{-1}(x) \leq g^{-1}(x)$; this is how we obtain $\ln x \leq x - 1$ $(x > 0)$ and the consequence is again (4).

Our inequality (3), or (4), can be cast now in a different shape by replacing x by $x/e - 1$ (or by $x/a - 1$) in the inequality $e^x \geq 1 + x$ (respectively (4)) and simplifying to get

$$e^{x/e} \geq x, \tag{5}$$

or

$$a^{x/a} \geq (x - a) \ln a + a, \quad \text{for all } x; \tag{6}$$

voilà, we have the answer to our question (1): If $e \leq a < x$, then $\sqrt[a]{a} > \sqrt[x]{x}$, or in power notation, $a^x > x^a$ and similarly, if $a < x \leq e$, then $\sqrt[a]{a} < \sqrt[x]{x}$ (in the former case $(x - a) \ln a + a \geq x - a + a$), or $a^x < x^a$, and clearly $\sqrt[e]{e}$ is the largest of them all. One should take some steps to justify why we make the above substitution and this is an opportunity to point out the beauty of how symmetry works: in light of the symmetry of a and x in inequality (4), it pays to look into the case when that symmetry is perfect—when $x = a$ and thus to measure how far we are from that perfection for other choices of x and a; compare x and a and see how much their quotient differs from 1. This is the time to talk also about some mathematical ingenuity—use of foresight that requires more than an immediate one-step procedure.

Somewhat inconspicuously and, to the student's untrained eye, unrelated, one might ask the following, to set the stage for later use: Given $a > 0$, find positive numbers y, z subject to $a = yz$, so that $P = z^y$ is maximized.

It is very instructive to show the students how a great mathematician Jacob Steiner (1796–1863) arrives at the same questions and answers outlined above. In [**21**], he discusses the following problem.

Steiner's problem on the number e

For what positive x is the x-th root of x the greatest?

Steiner lived considerably later than Newton (1643–1727) and Leibnitz (1646–1716), and thus he knew calculus. He begins [**21**] by noting that a rectangle with prescribed circumference has the greatest area if the rectangle is a square—a fact that must have been known even to pre-Greek mathematicians. In other words, if $x + y = a$, then the product xy is maximal if and only if $x = y$.

Steiner goes on to say that the same is true when a fixed quantity $a > 0$ is broken up into 3 or n summands: the product of the summands is the largest, when they are all equal. This small generalization is not often seen in textbooks and should be introduced, for it has a pedagogical value in that it can be geometrically interpreted, but can also be used to illustrate basic notions when discussing partial derivatives and extrema in several variables (knowing this geometric-algebraic fact makes the solution via partial derivatives not only easy, but explainable). Again it is important to ask the students to work out this expanded problem both geometrically or algebraically (if possible).[3] This is a perfect space to take advantage of the harmonic–geometric–arithmetic mean inequalities ($H \leq G \leq A$) or to introduce them, if this has not been done before.

For the more advanced students I have the following *dual* statements for them to understand, prove, elaborate, etc.; I use it to *expand* and *relate* and talk about the notion of *duality* and what it means in the following particular statements.

a) If r_i are rational numbers and $x_i > 0$ with $s_1x_1 + \ldots + s_nx_n = a$, for some numbers s_i, a, then the product $P = x_1^{r_1} \cdot \cdots \cdot x_n^{r_n}$ attains its maximum when

$$\frac{s_1x_1}{r_1} = \cdots = \frac{s_nx_n}{r_n}.$$

b) If r_i are rational numbers and $x_i > 0$ with $P = x_1^{r_1} \cdot \ldots \cdot x_n^{r_n}$, for some P, then the sum $s_1x_1 + \ldots + s_nx_n = a$ is minimal when the same equalities hold:

$$\frac{r_1}{s_1x_1} = \cdots = \frac{r_n}{s_nx_n}.$$

Claim a) can be proved by first assuming that r_i are natural numbers and using the geometric-arithmetic inequality applied to the quantities $(s_ix_i/r_i)^{r_i}$; the case $r_i = p_i/q_i$ is then straightforward. The proof of b) is shorter, for one can refer to the parts of the proof for a). These inequalities can be used extensively to find extremal values of various functions, without using calculus techniques.

Steiner generalizes his first remark further to something related to the above statements a) and b): to find out when the product $P = \prod x_i$ is maximal, if $\sum x_i = a$, for

[3]Here again one can introduce a method not used previously for solving the problem, using the arithmetic–geometric mean inequality, which in turn can be proved geometrically, as well as using calculus methods: Start with $a/n = (x_1 + \ldots + x_n)/n \geq \sqrt[n]{x_1 \ldots x_n}$), where the equality, if and only if all the x's are equal, producing the maximum of the product.

a fixed $a > 0$ (e.g. $a = e$). The index set need not be the integers; this is where the generalization is not a straightforward one from the square case. The intrinsic value of this one page paper without proofs lies in his truly great *interpretation*, for Steiner allows for index sets, other than the positive integers. The solution however is just like for the square case: each part of the partition equals e and there are a/e of those parts (that is the "number" of the summands is a/e so that in effect the product is the power $e^{a/e}$). One first shows that the parts have to be equal: for if there are two unequal parts, then reduce the problem to the case of two summands to get a contradiction. Thus if we have $a > 0$ and $a = yz$, we can maximize $P = z^y$, or $P = z^{a/z}$.

Finally, after learning more about derivatives we can learn more about the function $f(x) = \sqrt[x]{x}$, $x > 0$. We find the local extrema, if they exist (whether they do should be discussed beforehand, with a more advanced slant). One may work this out using logarithmic differentiation: $\ln f = \frac{1}{x} \ln x$ and $df = f \frac{1}{x^2}(1 - \ln x)\, dx$ and one can show that the maximum occurs for $x = e$ and then $\sqrt[e]{e} \approx 1.4447$. The graph is now sketched more or less routinely, using whatever appropriate techniques one teaches at this time.

This is done after thorough preparation, without which we would indulge in a bare exercise of routine logarithmic differentiation that would not reveal much of the essence. Discussions of the kind I presented here should be introduced much sooner than the subject of logarithmic differentiation, perhaps sooner than after having taught students much calculus at all. Both geometric-algebraic and derivative methods have their own merits: geometric-algebraic methods are less routine, and teach students several things about the exponential function in the process, such as its derivative; plus that beautiful picture (Figure 1a,b) .... It is worth working out examples of this kind when geometric methods turn out to be too complicated, and calculus proves invaluable and superior. This happens often (or sometimes?) in the case of curves of higher degrees, outside of the scope of conic sections. Care should be exercised, for non-calculus methods work well with many curves other than conic sections. One can for instance find extremal values of a wide array of functions utilizing the inequalities we mentioned here; try, for instance, $f(x) = x^2(2 - x^2)$—rewrite it and then use the harmonic–geometric–arithmetic mean inequalities.

Réaumur's honeybee cell problem as solved by Boscovich

Close the top of a regular hexagonal prism with a roof made of three congruent rhombi in such a way as to get a prescribed volume of honey with a minimal expenditure of wax.

The ancient thinkers were considerably puzzled by cell constructions the bees make. Pappus of Alexandria (ca. 300 A.D.) attributes the hexagonal shape of the hive-bee cells to reasons of economy [**16**]. Of the three regular polygons that tile the plane, the hexagon encloses the largest area for the prescribed perimeter. The density of a regular tessellation is defined to be the reciprocal of the ratio of the area of the polygon to the area of its inscribed circle. Hilbert and Cohn-Vossen [**10**] show that the best (largest density) plane tessellation is hexagonal, with the density ≈ 0.9069. Kepler (1571–1630) elaborates more fully on the shape of bee cells in [**11**].

Réaumur (1683–1757) in his monumental treatise [**17**] on insects clarifies that such an ideal shape of a honeybee cell is indeed rare and that the real bee cell is more often a crude approximation of the ideal geometrical shape. These qualifications were resounded by Darwin (1809–1882) who cites additional sources in [**6**], testifying to economizing patterns of bees in cell making as the means of natural selection (see the section "Cell-making instinct of the hive-bee," in chapter VIII (titled "Instinct")

of [**6**])—apparently wax-economizing behavior and the unique shape of bee cells are obtained in time through gradual selection in that the bees which economize more, are more likely to survive. Close scrutiny of Darwin's related passages show that he is wavering on the subject, for instance in not being able to decide whether it is economy of wax or labor that matters. It is not fully clear that the bees have the same sense of economizing as the humans. I refer to [**22**] for variations on this optimization problem.

Réaumur challenged several mathematicians of the time with the problem and a German mathematician Koenig gave a calculus solution [**12**], which however had some mistakes in it (see the sequel). We give here two solutions by R. Boscovich (1711–1787), in reverse order from his own (the solutions are somewhat modified by using his ideas from both of the solutions [**20**]). We will explore the honeybee cell problem in more detail elsewhere [**8**]. It is worth mentioning that the cells often "stand" on their roofs (the upside-down version of our Figure 2a) and that a honeycomb consists of two interlocking layers of adjacent hexagonal cells, their roofs fitting without interstices in the middle, their bases roughly forming two parallel surfaces (planes) tessellated by hexagons—the easily identifiable part of a honeycomb.

I give this problem to more advanced students, as a project with several aspects to deal with. To begin with, making a drawing using principles of descriptive geometry may be very useful and instructive.

In this idealized honeybee cell, the base is a regular hexagon $ABCDEF$ (Figure 2a), but the top is not closed by a parallel and congruent regular hexagon $GHIKLM$ as in a regular prism with rectangular sides. Instead the vertical sides are congruent trapezoids whose edges form a nonplanar hexagon $NHOKPM$.

If we raise a perpendicular QR from the center Q of $GHIKLM$ and find the points N, O, P such that $QR = GN = IO = LP$, then, by way of symmetry, the resulting rhombi $RMNH$, $RHOK$, $RKPM$ are congruent. $MQHG$ is a rhombus (Figure 2b), hence the pyramids $RMQH$ and $NMGH$ are of the same volume, since they have equal bases MQH and MGH as well as heights RQ and NG respectively. The same reasoning applies to other rhombi. Thus, no matter what shape the rhombi are, all the cells will have the same volume as if the top were closed by the regular hexagon. We now need to minimize the amount of wax used to build such a cell.

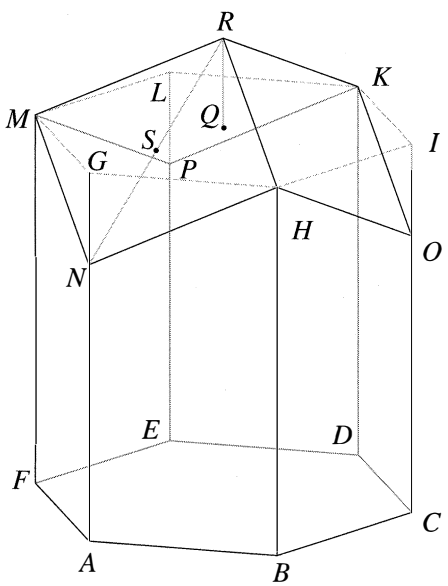


Figure 2a

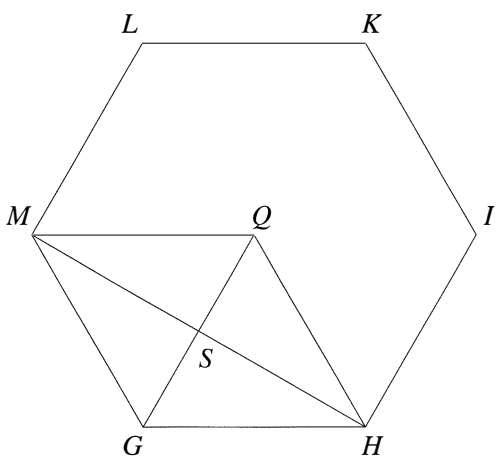


Figure 2b

Boscovich's second solution The variable part of the total area is

$$\text{Area} = 6\ \text{Area}(ANHB) + 3\ \text{Area}(RMNH).$$

If we denote the lengths $AB = GH = a$, $AG = b$, $GN = x$, then $\text{Area}(ANHB) = ab - ax/2$, and

$$\text{Area}(RMNH) = MH \cdot SN = a\sqrt{3}\sqrt{GN^2 + SG^2} = a\sqrt{3}\sqrt{x^2 + a^2/4}.$$

Thus we need to minimize the function $A(x) = 6ab - 3ax + (3\sqrt{3}a/2)\sqrt{4x^2 + a^2}$ and the minimum is easily found (after differentiation) to occur for

$$a^2 = 8x^2, \qquad (*)$$

(thus, we need to assume that $a < 2\sqrt{2}b$; in reality, bees make deep cells, hence this condition is satisfied).

Unknown to Boscovich, Newton's student Maclaurin had a very similar calculus solution [**13**] simultaneously, or before Boscovich. It is a good challenge to arrive at the minimizing quantities using inequalities, not derivatives. What however escaped both Boscovich and Maclaurin[4] is the following algebraic way to minimize the above quantity: In fact we need to minimize $C(x) = \sqrt{3}\sqrt{x^2 + a^2/4} - x$.

If we denote $y^2 = x^2 + a^2/4$, then this is the constraint for minimizing $C(x, y) = \sqrt{3}y - x$ (which is the same as minimizing C^2). Although, x and y are not perfectly symmetrical participants, we can emphasize their dual roles by introducing $D(x, y) = \sqrt{3}x - y$ (by now, duality would have been our old friend). We have now easily $C^2 - D^2 = 2(y^2 - x^2) = a^2/2$ and thus $C^2 = a^2/2 + D^2$; hence, C is minimized for $D = 0$ and that leads to $(*)$.

Because of equality $(*)$ we get $RN^2 = (2SN)^2 = 12x^2 = (3/2)a^2$, and since $MH^2 = 3a^2$, we get the equality

$$MH^2 = 2RN^2, \qquad (**)$$

which determines the shape of the rhombi. The angles of the rhombi are easily determined: $\tan \angle RNH = SH/SN = MH/RN = \sqrt{2}$, thus the angles are $\angle MNH \approx 109°28'16.39''$ and $\angle NHR \approx 70°31'43.61''$.[5] The volume is $V = (3\sqrt{3}/2)a^2b$, whereas the minimal (closed) cell area is $\text{Area} = (3(\sqrt{3} + \sqrt{2})/2)a^2 + 6ab$.

Here Boscovich found errors in computations (or measurements, which Maraldi claimed to have performed) of his predecessors who worked on the same problem. In addition, Réaumur claimed that Koenig proved that by making such a pyramidal roof, rather than the flat one, the bees save in wax by as much as the amount needed to build the flat roof. Boscovich said that this was wrong and it is a good exercise for the students to find the correct "savings."

We now show an interesting fact, namely that the rhombic angles are equal to the corresponding non-right angles of the side trapezoids. Using $(**)$ and the double angle formula for tan we get $\tan \angle MNH = -2\sqrt{2} = -\tan \angle RHN$. On the other hand,

[4]But then, it may not have escaped them—they may have wanted to prove the point of calculus' usefulness. Maclaurin was a natural preacher in the new discipline of calculus and Boscovich was showing that he had adopted the new invention. Unlike most commentators, I think that Newton (and hence his students) was not comfortable with geometry and this, among other factors, may have facilitated a quest for something else. On the other hand, Boscovich adored the older geometric method and it resulted in a number of ingenious geometric solutions to various problems, while he at the same time may have had somewhat shorter calculus solutions.

[5]Figure 2a is a cell view that distorts the angles; this view was chosen to reduce the "clutter in the attic."

$\tan \angle NHB = a/x = 2\sqrt{2}$, by $(*)$, hence $\angle RHN = \angle NHB$; thus all the other corresponding angles are equal.

Boscovich however goes further to show something no one involved noticed, namely that the *spatial* angles between the faces are all equal (except the right angles at the base), and equal to 120°: We have just established that all the plane angles at vertex N are equal, thus the angles formed by the respective planes must be equal too. The angle between the planes $FANM$ and $NABH$ is the same as $\angle FAB = 120°$. Now the same arguments may be repeated for the vertices O, P, R, in place of N, since the trihedra at those vertices are congruent to the one at N. The same reasoning (and the same angles) applies to another group of vertices H, K, M that have four quadrilaterals abutting at them.

This wonderful reasoning is then carried over to the construction of honeycomb. Boscovich thinks that bees have special instruments that they can use only in such a way as to produce the prescribed angles to connect two planes. These instruments are never so perfect, just like human limbs; thus the cells will deviate from the perfect form. Let us add that in our industrial age, the cells are often started (by a manufacturer) in a form of a planar network of perfect hexagons and then offered to bees to finish the construction. Boscovich is nonetheless awed by the intention of the Divine Creator of nature who gave these little animals the tools and instinct that dictate the shape of the cells with the greatest saving of wax.

Boscovich's first solution Let us look now into Boscovich's geometric solution (in fact his first solution) to the problem of the shape of the honeybee cell. He is pushing geometric considerations to their limits, and literally at that, for he utilizes limiting processes, without explicitly using our modern terminology in this respect. In this solution he gets the rhombi shape relation $(**)$, which can be shown to be equivalent to $(*)$, obtained earlier via derivatives. He says that the minimizing position $MNHR$ will be such that any roof plate passing through MH (say $MN'HR'$, Figure 3a), and close enough to the minimizing position, will assume equal areas in moving from a position just before, to some position just after the minimizing position.

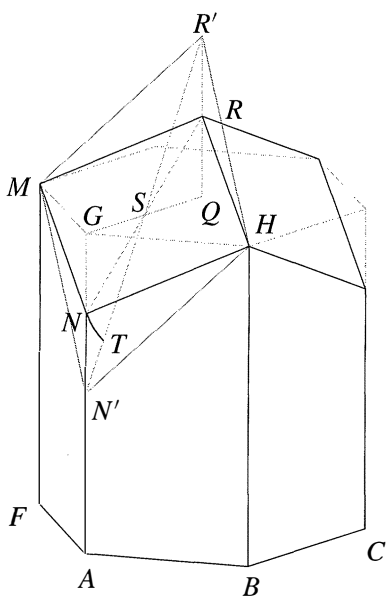


Figure 3a

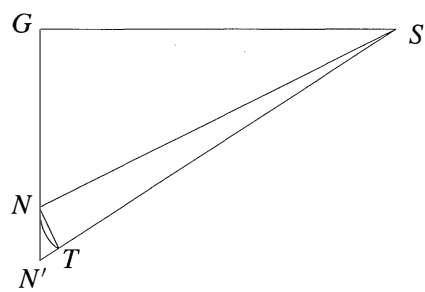


Figure 3b

The extra area that $MN'HR'$ has over $MNHR$ equals the sum of the areas of triangles NHN' and NMN'—by which the areas of $ANHB$ and $ANMF$ were reduced:

(1) $\text{Area}(MN'HR') - \text{Area}(MNHR) = \text{Area}(NHN') + \text{Area}(NMN')$.

The diagonals of the rhombus $MNHR$ are perpendicular and bisect each other and we can see then that GQ and RN pass through the mid-point S. Using the data we have and Figure 2a,b we get

(2) $GH = 2GS$

and

(3) $GH^2 : GS^2 : SH^2 = 4 : 1 : 3$.

Within the plane NSN', use an arc with center S and radius SN to get a point T on SN'. We then have

(4) $\text{Area}(MN'HR') - \text{Area}(MNHR) = MH \cdot N'T$

and

(5) $\text{Area}(NHN') + \text{Area}(NMN') = NN' \cdot GH$,

as triangles with the same base NN' and altitude GH. From (1), (4), and (5) we get $MH \cdot N'T = NN' \cdot GH$, or $NN' : N'T = MH : GH$, and by (2), we have

(6) $NN' : N'T = SH : SG$.

Here is where Boscovich applies limiting reasoning (N' is close to N): The triangle NTN' can be considered to be right, with the right angle at T (Figure 3b), and angles GNS and $NN'S$ approximately equal, and we can consider the triangles SNG and NTN' to be similar, thus we have a relation $NN' : N'T = SN : NG$ so, by (6), we have

(7) $SN : NG = SH : SG$; hence, by (3), $SN^2 : NG^2 = 3 : 1$,

and then

(8) $SG^2 : NG^2 = 2 : 1$, since $SG^2 = SN^2 - NG^2$.

From (7) and (8) we now have $2SN^2 = SH^2$ and, since $NS = (1/2)NR$, $SH = (1/2)MH$, we get $2NR^2 = MH^2$ and this is the relation $(**)$. We can now reconstruct the rest as in the other proof.

Boscovich prefers this geometric solution which appears to him as simpler and more elegant, as it often happens ("*Et quidem saepe accidit, potissium in hujusmondi problematis admodum simplicibus, ut Geometria simpliciores, & elegantiores determinationes exhibeat, quam calculus.*"). He also rightly says that the geometric approach gets some relationships straightforwardly, unlike the solution with the derivative. Going back to my introductory remarks, this inevitably happens when we apply unifying tools (such as calculus and category theory)—they give us automatic snappy solutions, but a more detailed and perhaps a deeper insight is lost, while we gain in a global picture. In any case, both of these solutions are a good example of a happy mixture of geometric and analytic methods working together. They both emulate this mix, one using Newton's (derivatives), another Leibnitz's (infinitesimals) flavor of differential calculus. Most often I present the method that uses derivatives to minimize, however with a more refined audience I inevitably reach for the purely geometric option.

The honeycomb discussion I touched upon here contains more mathematics than meets the eye, and is far from exhausted; I will devote a more thorough study to it in the future [**8**].

Acknowledgment. I thank the referees for a number of comments, suggestions and corrections that resulted in a better presentation of the paper. I hope that the paper still contains imperfections, without which it could never aspire to be perfect.

REFERENCES

1. Barrow, Isaac, *The Mathematical Works*, Ed. W. Whewell, xx+320 pp. +220 figures. Cambridge, UK, 1860, reprinted G. Olms Verlag, 1973.
2. Jakob Bernoulli, Positiones Arithmeticae de seriebus infinitas, earumque summa finita, Basileae, 1689, *Opera* **1**, p. 375–402.
3. Jakob Bernoulli, *Opera*, 2 volumes, VIII+48+1139 pp., Ed. H. Cramer and F. Philibert, Geneva, 1744.
4. James N. Brawner, Dinner, dancing and tennis anyone?, this MAGAZINE, **73** (2000), 29–35.
5. Pierre Cartier and Marjorie Senechal, The continuing silence of Bourbaki—an interview with Pierre Cartier, June 18, 1997, *The Math. Intelligencer*, **20** (1998), No. 1, 22–28.
6. Charles Robert Darwin, *On the Origin of Species by Means of Natural Selection*, John Murray, London, 1859.
7. Radoslav Dimitrić, Sir Isaac Newton, *Math. Intelligencer*, **13**(1991), No. 1, 61–65.
8. Radoslav Dimitrić, *Honeycombs*, A book in preparation, ca. 2001.
9. Leonard Euler, Introductio in Analysin Infinitorum, vol. 1. *Opera Omnia*, Ser. 1, vol. **8**, Leipzig, 1922. English edition: Introduction to the Analysis of the Infinite, Springer Verlag, New York, 1990.
10. D. Hilbert, and S. Cohn-Vossen, *Geometry and the Imagination*, Chelsea, New York, 1952.
11. Ioannis Kepleris, *Strena Seu De Niue Sexangula*, Francofurti ad Moenum apud Godefridum Tampach, 1611. English edition: The six-cornered snowflake, Clarendon Press, Oxford, 1966.
12. Johann Samuel Koenig, Lettre de Monsieur Koenig à Monsieur A.B., écrite de Paris à Berne le 29 novembre 1739 sur la construction des alvéoles des abeilles, avec quelques particularités littéraires, *Journal Helvétique*, April 1740, 353–363.
13. Colin Maclaurin, Of the bases of the cells wherein the Bees deposite their honey. Part of a letter from Mr. Maclaurin, Professor of Mathematics at Edinburgh, and F.R.S. to Martin Folkes, Esq; pr. R.S., *Philosophical Transactions*, **42** (1742/43), 565–571.
14. Elie Maor, *e. The Story of a Number*, Princeton University Press, Princeton, 1994.
15. Barbara H. Margolius, Avoiding your spouse at a bridge party, this MAGAZINE, **74** (2001), 33–41.
16. Pappus of Alexandria, (ΠΑΠΠΟΥ ΑΛΕΞΑΝΔΡΕΩΣ ΣΥΝΑΓΩΓΗ) *Pappi Alexandrini Collectionis Reliquiae*, Liber V. A Greek-Latin edition with commentaries by Fridericus Hultsch. Berolini, apud Weidmannos, 1876. (See also an excerpt in English translation in: Thomas Heath, *A History of Greek Mathematics*, vol. II, Dover Publications, Inc., New York, 1981: On the sagacity of bees in building their cells, *ca.* 340? A.D.)
17. René Antoine Ferchault de Réaumur, *Mémoires pour Servir à l'Histoire des Insectes* (1734–42) IIIrd chapter, fifth volume, 1740.
18. W. W. Rouse Ball, *Mathematical Recreations & Essays*, The MacMillan Company, New York, 1947.
19. Bertrand Russell, *An Essay on the Foundations of Geometry*, Cambridge: At the University Press, 1897.
20. Benedicto Stay, *Philosophiae Recentioris a Benedicto Stay*, Versibus Traditae Libri X ..., Romae 1755. Addendum pp. 498–504 of book II is by Boscovich, Ruggiero, *De apium cellulis*.
21. Jacob Steiner, Über das grösste Product der theile oder summanden jeder Zahl, *J. Reine Angew. Math.* = *"Crelle's Journal"* **40** (1850), No. 3, 208.
22. L. Fejes Tóth, What the bees know and what they do not know, *Bull. Amer. Math. Soc.* **70** (1964), 468–481.

NOTES

Strategies for Rolling the Efron Dice

CHRISTOPHER M. RUMP
Department of Industrial Engineering
University at Buffalo
State University of New York
Buffalo, NY 14260-2050
crump@eng.buffalo.edu

Consider a two-player competitive game involving the rolling of six-sided dice. The object is to see which player rolls a larger number. However, these dice are not the usual type with sides numbered 1 through 6. Rather, this game is played with a set of four dice designed by Bradley Efron, a statistician at Stanford University. These four dice are "unfolded" in Figure 1.

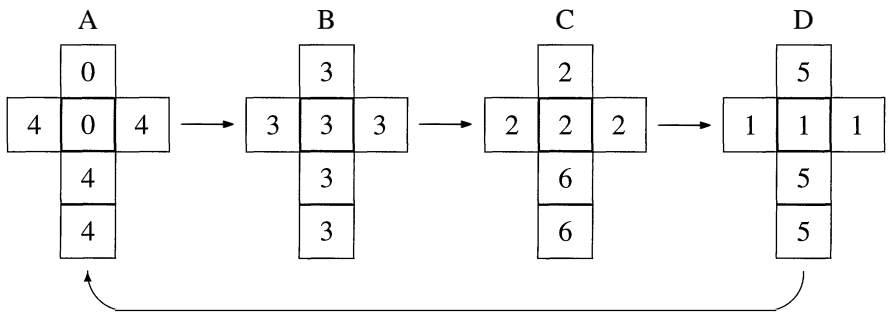


Figure 1 Efron dice

What is peculiar about these dice is that they are probabilistically non-transitive [**1, 3**]. This is due to the fact that die A is twice as likely to beat die B, die B is twice as likely to beat die C, die C is twice as likely to beat die D, and, paradoxically, die D is twice as likely to beat die A! Thus, in a "gentleman's game," as Ross Honsberger calls it [**2**], we graciously let our opponent choose a die to roll, so that we then can pick a die that beats it two times out of three. Clearly, a "gentleman's game" is not attractive for the person who chooses first.

Suppose now, instead, that the players, each equipped with a personal set of four Efron dice, simultaneously choose a die to roll without revealing their selections until the dice are rolled. In repeated plays of this game, a player's choice of a die to roll is not so clear. What is clear is that a deterministic strategy, that is, one that involves a completely predictable sequence such as always choosing a particular die, can be soundly beaten. Thus, players must keep their opponents guessing by choosing a mixed strategy that randomly picks among the four dice.

One might suspect that an optimal strategy would be to randomly choose among the four dice with equal (uniform) probability. After all, this type of strategy is optimal for the classic non-transitive game of rock-scissors-paper (rock beats scissors, scissors beats paper, paper beats rock) [**6**]. However, this strategy is not optimal for the Efron dice game.

To see why, we first consider the probability that each die will beat any other on a play of this game. In doing so, we create a matrix of probabilities, where each row of the matrix corresponds to the choice of a die by the the row player, let's call her Rose, and each column corresponds to a choice by the column player, Colin. Each probability then expresses the probability that Rose beats Colin for that particular combination of dice. It is not terribly difficult to show that this matrix is given by

$$\begin{array}{cc} & \text{Colin} \\ & \begin{array}{cccc} \text{A} & \text{B} & \text{C} & \text{D} \end{array} \\ \text{Rose} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \\ \text{D} \end{array} & \left[\begin{array}{cccc} 1/2 & 2/3 & 4/9 & 1/3 \\ 1/3 & 1/2 & 2/3 & 1/2 \\ 5/9 & 1/3 & 1/2 & 2/3 \\ 2/3 & 1/2 & 1/3 & 1/2 \end{array}\right] \end{array}$$

Since each probability in this "payoff" matrix expresses the probability that Rose beats Colin, which is one less the probability that Colin beats Rose, this constitutes a two-person, one-sum game. By symmetry of their possible die choices, both Rose and Colin are equally likely to win this game. Thus, the game is fair, and its value is the probability that Rose wins, namely $1/2$.

Note that when both Rose and Colin roll die B, a tie value results with no winner. We have chosen to arbitrarily assign a probability value of $1/2$ in this situation (the (B, B) diagonal element in the matrix), with the interpretation that the players will each choose a (possibly) new die to roll, and the probability that Rose will win on this new roll (in a fair game) is $1/2$.

We can find the optimal strategy for either player (by symmetry their optimal strategies are the same) in the same way as for the more common zero-sum games by transforming the game into a linear programming problem [**6**]. Let $\boldsymbol{x}$ denote the strategy vector for the row player Rose. This is a probability vector, where the ith element, x_i, denotes the probability that Rose will choose to roll the die corresponding to row i. Multiplying this vector by the jth column of the payoff matrix gives the expected probability of Rose winning when Colin chooses to roll the die corresponding to column j. Classic game theorist that she is, Rose employs a *maximin* criterion that attempts to maximize the minimum (worst-case) expected probability of winning. She therefore seeks to maximize the lower bound on her expected probability of winning over all possible actions from Colin via the linear program

$$\text{Maximize } v,$$

subject to

$$\begin{aligned} \frac{1}{2}x_A + \frac{1}{3}x_B + \frac{5}{9}x_C + \frac{2}{3}x_D &\geq v \\ \frac{2}{3}x_A + \frac{1}{2}x_B + \frac{1}{3}x_C + \frac{1}{2}x_D &\geq v \\ \frac{4}{9}x_A + \frac{2}{3}x_B + \frac{1}{2}x_C + \frac{1}{3}x_D &\geq v \\ \frac{1}{3}x_A + \frac{1}{2}x_B + \frac{2}{3}x_C + \frac{1}{2}x_D &\geq v \\ x_A + x_B + x_C + x_D &= 1 \end{aligned}$$

and $x_i \geq 0$, for $i = A, B, C, D$.

Using the substitution $x_D = 1 - x_A - x_B - x_C$, this problem reduces to

$$\text{Maximize } v,$$

subject to

$$\begin{aligned}
-\frac{1}{6}x_A - \frac{1}{3}x_B - \frac{1}{9}x_C + \frac{2}{3} &\geq v \\
\frac{1}{6}x_A \qquad\quad - \frac{1}{6}x_C + \frac{1}{2} &\geq v \\
\frac{1}{9}x_A + \frac{1}{3}x_B + \frac{1}{6}x_C + \frac{1}{3} &\geq v \\
-\frac{1}{6}x_A \qquad\quad + \frac{1}{6}x_C + \frac{1}{2} &\geq v \\
x_A + \quad x_B + \quad x_C \qquad &\leq 1
\end{aligned}$$

and $x_i \geq 0$, for $i = A, B, C$.

Since, by symmetry, the value of this fair game is $v^* = 1/2$, the second and fourth constraints at optimality imply that $\pm\frac{1}{6}(x_A^* - x_C^*) \geq 0$, i.e., $x_A^* = x_C^*$. Hence, Rose should choose dice A and C with equal probability. This fact reduces the remaining constraints to

$$\begin{aligned}
5x_A^* + 6x_B^* &= 3 \\
2x_A^* + \quad x_B^* &\leq 1,
\end{aligned}$$

where the equation follows from combining the first and third original constraints. These remaining constraints then imply the existence of multiple optimal solutions given by

$$\begin{aligned}
x_A^* = x_C^* &\in \left(0, \frac{3}{7}\right) \\
x_B^* &= \frac{1}{2} - \frac{5}{6}x_A^* \\
x_D^* &= 1 - 2x_A^* - x_B^* = \frac{1}{2} - \frac{7}{6}x_A^*.
\end{aligned}$$

This optimal solution set is plotted (as a function of the first row probability x_A^*) in FIGURE 2.

It is a bit surprising that among the infinitely many optimal solutions, the uniform solution $x_i = 1/4$, $i = A, \ldots, D$, is not among them. A fairly close alternative is the solution $\boldsymbol{x}^* = (6/24, 7/24, 6/24, 5/24)$. The most interesting solutions are the boundary solutions $\boldsymbol{x}^* = (0, 1/2, 0, 1/2)$ and $\boldsymbol{x}^* = (3/7, 1/7, 3/7, 0)$. The first of these solutions suggests that it is optimal to play by throwing only dice B and D with equal likelihood, even if your opponent chooses to play optimally with all four dice. The other boundary solution suggests playing with only the first three dice, leaving out die D.

Two dice Now, consider the case where each player rolls two Efron dice, with the goal of rolling the highest total. If a player is allowed to roll the same die twice, the

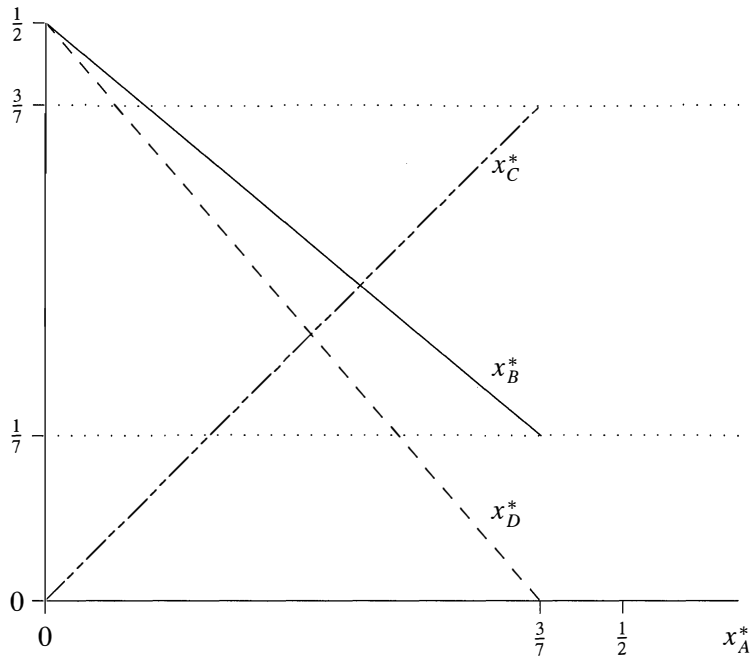


Figure 2 Optimal strategy solutions

probability matrix becomes

	AA	AB	AC	AD	BB	BC	BD	CC	CD	DD
AA	1/2	16/27	4/9	10/27	4/9	8/27	2/5	16/49	14/27	4/9
AB	11/27	1/2	16/27	1/2	2/3	4/9	1/3	8/27	2/5	7/12
AC	5/9	11/27	1/2	17/27	1/2	16/27	1/2	4/9	10/27	1/2
AD	17/27	1/2	10/27	1/2	1/3	2/5	7/12	14/27	4/9	3/8
BB	5/9	1/3	1/2	2/3	1/2	2/3	1/2	4/9	1/3	1/2
BC	19/27	5/9	11/27	3/5	1/3	1/2	2/3	16/27	1/2	5/12
BD	3/5	2/3	1/2	5/12	1/2	1/3	1/2	2/5	7/12	1/2
CC	33/49	19/27	5/9	13/27	5/9	11/27	3/5	1/2	17/27	5/9
CD	13/27	3/5	17/27	5/9	2/3	1/2	5/12	10/27	1/2	5/8
DD	5/9	5/12	1/2	5/8	1/2	7/12	1/2	4/9	3/8	1/2

Note that when both players choose BB a tie always results. Again, we assign this a value of $1/2$ as before. In other cases where a tie may result, we have assumed that the players roll these same dice again until a victor emerges.

Choices AA and BD would never be used since they are dominated by choice CC. Also, choice AB is dominated by CD, and choice AD is dominated by BC. Therefore, we can remove these four dominated strategies from consideration. Solving the remaining matrix game, we find the unique optimal solution $x_{BB}^* = 5/17$, $x_{BC}^* = 3/17$, and $x_{CC}^* = 9/17$. Thus, the optimal mixed strategy mixes between dice B and C and never uses dice A and D. Most of the time the player uses just one die (either die B or C) and rolls it twice.

If we disallow the option of rolling a single die twice, and require each player to roll two different Efron dice, we simply further eliminate choices BB, CC, and DD. This leaves only choices AC, BC, and CD to choose from. Thus, die C is always used and the choice becomes one of choosing one of the remaining dice A, B, or D. Choosing one of these dice to roll is equivalent to the original probability matrix with choice C removed, viz.

$$\begin{array}{cc} & \text{Colin} \\ & \begin{array}{ccc} \text{A} & \text{B} & \text{D} \end{array} \\ \text{Rose} \begin{array}{c} \text{A} \\ \text{B} \\ \text{D} \end{array} & \left[\begin{array}{ccc} 1/2 & 2/3 & 1/3 \\ 1/3 & 1/2 & 1/2 \\ 2/3 & 1/2 & 1/2 \end{array}\right] \end{array}$$

Clearly, option B is dominated by option D. Removing option B yields a 2×2 matrix between options A and D, in which A is dominated by D as well. Thus, option D constitutes a dominant strategy or saddle point for this fair game. Hence, when choosing two unique dice from the set of four Efron dice to roll, each player should choose dice C and D.

REFERENCES

1. M. Gardner, Mathematical games: the paradox of the nontransitive dice and the elusive principle of indifference, *Scientific American* **223** (1970), 110–114.
2. R. Honsberger, Some surprises in probability, in *Mathematical Plums*, R. Honsberger, ed., The Mathematical Association of America, 1979.
3. R.P. Savage, The paradox of nontransitive dice, *Amer. Math. Monthly*, **101**:5 (1994), 429–436.
4. K.J. Smith, *Finite Mathematics: A Discrete Approach*, Scott, Foresman and Co., Glenview, IL 1975.
5. R.L. Tenney and C.C. Foster, Non-transitive dominance, this MAGAZINE **49**:3 (1976), 115–120.
6. W. Winston, *Operations Research: Applications and Algorithms*, 3rd ed., Duxbury Press, Belmont, CA, 1994.

Probabilities of Consecutive Integers in Lotto

STANLEY P. GUDDER
JAMES N. HAGLER
University of Denver
Denver, CO 80208

Introduction In a typical lotto game (such as the Colorado Lottery [**3**]), ping-pong balls numbered 1 through 42 are placed in a clear plastic drum and thoroughly mixed. Six balls are then randomly selected; if your previously chosen six numbers agree with those on the selected balls (in any order) then you win the jackpot. The lottery is repeated twice a week, and the payoff depends on how many times it is played until there is a winner. If more than one player wins, the proceeds (millions of dollars) are divided among the winners. The chance of winning with a single pick is, of course,

$$1/\binom{42}{6} = 1/5245786,$$

so such a win is quite unlikely.

A sample of winning combinations shows that, very frequently, there are at least two consecutive numbers among the six. For example, on December 30, 1998, the winning set in the Colorado Lottery was {2, 3, 14, 17, 19, 22}, which contains the consecutive numbers 2 and 3. To satisfy our curiosity we examined the 312 different winning combinations over the years 1996–1998 (all data were gathered from the Colorado Lottery's website [**3**]) and we found that about 53% (164 out of 312) of the time this combination contains at least two consecutive numbers. Considering the spread between 1 and 42 it was surprising to us that this percentage was so high. Also, these

data showed that about 39% of the time (121 times) there were exactly two consecutive numbers. This means that in about 14% of the cases there were more than two consecutive numbers. For example, there could be two separated pairs of consecutive numbers such as {13, 23, 24, 27, 28, 39} (December 7, 1996) or a triple of consecutive numbers, such as {10, 11, 12, 19, 23, 41} (January 10, 1998).

Since not all lotteries have the same rules, we frame computations in the general setting of an (n, m)-lotto game, where the winning set is an m-element subset of $\{1, \ldots, n\}$. First we compute probabilities of consecutive integers theoretically, using a simple formula (Proposition 1), from which related results on the clustering of numbers on a winning ticket are easily derived. Then we observe that 6 is the smallest m for which the probability of at least two consecutive integers in a winning set in $(42, m)$-lotto exceeds $1/2$; thus $m = 6$ is the $1/2$-*threshold* for $n = 42$. Examining threshold values for other ns suggests that they grow with order $\sqrt{n}$; standard approximations from calculus let us conjecture a closed form. In Theorem 5 we show that this conjecture is "almost correct."

Our work is similar in spirit to a nice extension of the birthday problem by Abramson and Moser [**1**], who compute the probability that no two birthdays among n people lie within k days of each other. The interesting book by Henze and Riedwyl [**5**] presents strategies for avoiding sharing prizes if you win the lottery, and much statistical analysis of real lotto data. Our work also overlaps that of Henze [**4**], who studies the distribution of spaces between numbers on winning lotto tickets.

Lotto probabilities We define an (n, m)-*lotto game* as one in which the winning set is an m-element subset of $\{1, \ldots, n\}$. An m-element set A in $\{1, \ldots, n\}$ is called an m-set and when we write $A = \{a_1, a_2, \ldots, a_m\}$, we always assume that $a_1 < a_2 < \cdots < a_m$. If A contains at least two (respectively no) consecutive integers, then A is called a *consecutive* (respectively *nonconsecutive*) m-set. Let $K(n, m)$ and $N(n, m)$ denote the number of consecutive and nonconsecutive m-sets. The probability $P(n, m)$ that a randomly chosen m-set is consecutive is

$$P(n, m) = K(n, m) \Big/ \binom{n}{m}.$$

For example, in the Colorado lotto game, a $(42, 6)$-lotto, we want to find

$$P(42, 6) = K(42, 6) \Big/ \binom{42}{6};$$

thus, finding $P(n, m)$ amounts to finding $K(n, m)$. This is not as simple as it may look and we invite readers to try their hands at this before proceeding to the solution. The result (Proposition 1) is known (see, for instance, [**2**, pp. 30–1] or [**6**, Example 4c, pp. 7–8]). It is also the starting point for Henze [**4**], who studies spacing in lotto combinations.

As often happens in such combinatorial problems, it is easier to find $N(n, m)$ and then use the relation $K(n, m) = \binom{n}{m} - N(n, m)$. Of course, if $m > (n + 1)/2$ then $N(n, m) = 0$, so we assume henceforth that $m \leq (n + 1)/2$. The cases $m = 0$ and $m = 1$ are also trivial, so we assume that $m \geq 2$. Finally, we assume that $\binom{r}{j} = 0$ when $j > r$ or $r < 0$.

PROPOSITION 1.

$$N(n, m) = \binom{n - m + 1}{m}.$$

Proof. We display a bijection between the set of m-sets in $\{1, \ldots, n-m+1\}$, which has cardinality $\binom{n-m+1}{m}$, and the set of nonconsecutive m-sets in $\{1, \ldots, n\}$. Indeed, if $\{a_1, a_2, \ldots, a_m\}$ is an m-set in $\{1, \ldots, n-m+1\}$, then $\{a_1, a_2+1, \ldots, a_m+m-1\}$ is a nonconsecutive m-set in $\{1, \ldots, n\}$; this correspondence is clearly bijective. ∎

By Proposition 1,

$$K(n, m) = \binom{n}{m} - \binom{n-m+1}{m}. \tag{1}$$

In (42, 6)-lotto, for example, we have

$$P(42, 6) = 1 - \binom{37}{6} \Big/ \binom{42}{6} \approx 0.557.$$

The idea of the preceding proof can be generalized to gaps of two or more. Let us say that an m-set A in $\{1, \ldots, n\}$ has *gaps of at least* k if $a_{j+1} - a_j \geq k+1$ for $j = 1, \ldots, m-1$. Then the proof of Proposition 1 (or see [**4**]) can be generalized easily to show the following result:

PROPOSITION 2. *The number of m-sets in $\{1, \ldots, n\}$ in which there is a gap of at least k between every pair of integers is*

$$\binom{n-km+k}{m}.$$

Proposition 1 provides a starting point for computing probabilities of events naturally associated with clustering of winning numbers in lotto. Let us say that an m-set A in $\{1, \ldots, n\}$ has *exactly two consecutive integers* if for some $k < m$, $a_k + 1 = a_{k+1}$ and $a_j + 1 < a_{j+1}$ for $j \neq k$. The next result lets us count these sets.

PROPOSITION 3. *The number of m-sets in $\{1, \ldots, n\}$ with exactly two consecutive integers is*

$$(m-1)\binom{n-m+1}{m-1}.$$

Proof. Each nonconsecutive $(m-1)$-set $A = \{a_1, a_2, \ldots, a_{m-1}\}$ in $\{1, \ldots, n-1\}$ generates the following $m-1$ m-sets, each with exactly two consecutive integers:

$$\{\mathbf{a}_1, \mathbf{a}_1+\mathbf{1}, a_2+1, \ldots, a_{m-1}+1\}$$
$$\{a_1, \mathbf{a}_2, \mathbf{a}_2+\mathbf{1}, a_3+1, \ldots, a_{m-1}+1\}$$
$$\vdots$$
$$\{a_1, a_2, \ldots, \mathbf{a}_{m-1}, \mathbf{a}_{m-1}+\mathbf{1}\}.$$

Observe that each m-set with exactly two consecutive integers can be written in precisely one of these ways. Since each A generates $m-1$ of these sets, their total number is

$$(m-1)N(n-1, m-1) = (m-1)\binom{(n-1)-(m-1)+1}{m-1}$$
$$= (m-1)\binom{n-m+1}{m-1}.$$ ∎

In (42, 6)-lotto, for example, the probability of exactly two consecutive integers in the winning lotto combination is

$$5\binom{37}{5}\Big/\binom{42}{6} \approx 0.415.$$

(In the Colorado lottery, this phenomenon occurred 121 out of 312 times, for a frequency of 0.388.)

The definition and the proposition extend easily to m-sets in $\{1, \ldots, n\}$ with exactly r consecutive integers. We invite the reader to prove the next result by mimicking the proof of Proposition 3.

PROPOSITION 4.

(a) The number of m-sets in $\{1, \ldots, n\}$ with exactly r consecutive integers is

$$(m - r + 1)\binom{n - m + 1}{m - r + 1}.$$

(b) The number of m-sets in $\{1, \ldots, n\}$ with exactly 2 separated pairs of consecutive integers is

$$\binom{m - 2}{2}\binom{n - m + 1}{m - 2}.$$

In (42, 6)-lotto, for example, the probability of exactly 3 consecutive integers in the winning set is $4\binom{37}{4}/\binom{42}{6} \approx 0.050$, which agrees well with the observed Colorado frequency of 0.055 (17 out of 312). The probability of exactly two separated pairs is $\binom{4}{2}\binom{37}{4}/\binom{42}{6} \approx 0.076$, while the observed frequency was 0.058 (18 out of 312). We urge readers to compute these and other probabilities (e.g., 3 separated pairs, a separated triple and a pair, 4 consecutive integers, etc.) for their own lotteries, and to compare predictions with actual data.

Threshold numbers The *birthday problem* (see, e.g., [**6**, Example 5j, pp. 40–1]) asks how many people are necessary to have a probability of at least 1/2 of two matching birthdays. The same kind of analysis can be applied to a general (n, m)-lotto game, where we seek, for fixed n, the smallest m which gives the probability of more than 1/2 that the winning set is consecutive. Perhaps surprisingly, there is an "almost closed" form for m in terms of n that can be derived using only calculus.

We saw above that the probability of at least two consecutive integers in (42, 6)-lotto exceeds 1/2. What if only 5 numbers were selected? From (1) we see that $P(42, 5) = 1 - \binom{38}{5}/\binom{42}{5} \approx 0.41 < 1/2$, so we call 6 the 1/2-*threshold number* for $n = 42$. In general, we'll call the smallest m such that $P(n, m) > 1/2$ the *threshold number* for n; we denote it by $T(n)$. The following table shows several threshold numbers:

TABLE 1: Threshold Numbers

n	5	10	12	42	49	100	116	1000	10^4	10^5	10^6
$T(n)$	3	3	4	6	7	9	10	27	84	264	833

Notice that $T(n)$ appears to have order $\sqrt{n}$. But is there a closed-form expression for $T(n)$? We shall now try to conjecture one.

We seek the smallest m such that $K(n,m)/\binom{n}{m} > 1/2$; by (1), this is equivalent to

$$\binom{n-m+1}{m} \Big/ \binom{n}{m} < \frac{1}{2}. \tag{2}$$

Writing out the binomial coefficients using factorials, we see that (2) is equivalent to

$$\frac{n-m}{n} \cdot \frac{n-m-1}{n-1} \cdot \cdots \cdot \frac{n-m-(m-2)}{n-(m-2)} < \frac{1}{2}. \tag{3}$$

The left side of (3) has $m-1$ factors, which decrease from left to right. Thus (3) certainly holds if

$$\frac{n-m}{n} \le 2^{-1/(m-1)}, \tag{4}$$

which is equivalent to

$$n\left(2^{1/(m-1)} - 1\right) \le m\, 2^{1/(m-1)}. \tag{5}$$

Since $2^{1/(m-1)} \approx 1 + \frac{\ln 2}{m-1}$, we can rewrite (5) *approximately* as

$$\frac{n}{m-1} \ln 2 \le m\left(1 + \frac{\ln 2}{m-1}\right). \tag{6}$$

From (6) we have $n \ln 2 \le m^2 - m(1 - \ln 2) \le m^2$. Hence, $\sqrt{n \ln 2} \le m$ and since m is an integer we have $\lceil \sqrt{n \ln 2} \rceil \le m$, where $\lceil \cdot \rceil$ denotes the ceiling function.

Based on the preceding heuristic argument, we might conjecture that the smallest possible integer m is $S(n) := \lceil \sqrt{n \ln 2} \rceil$. This conjecture is indeed correct for many values of n. For example, it works for all n in Table 1 other than $n = 5, 12, 49, 116$. For these values, we have

$$S(5) = 2 < 3 = T(5); \qquad S(12) = 3 < 4 = T(12);$$
$$S(49) = 6 < 7 = T(49); \qquad S(116) = 9 < 10 = T(116).$$

Thus it seems that our original conjecture is off by at most 1. In Theorem 5 we state and prove our corrected conjecture:

THEOREM 5. *The threshold function $T(n)$ satisfies*

$$\left| T(n) - \left\lceil \sqrt{n \ln 2} \right\rceil \right| \le 1.$$

Remark. Theorem 5 can be sharpened to show that $T(n)$ is either $\lceil \sqrt{n \ln 2} \rceil$ or $\lceil \sqrt{n \ln 2} \rceil + 1$. We omit the proof, which is considerably more detailed than that following.

Proof. Let $a = \frac{1}{2} + \sqrt{\frac{1}{4} + n \ln 2}$. (It is convenient to consider a because it is close to $\sqrt{n \ln 2}$ and it satisfies the simple quadratic equation $a^2 - a = n \ln 2$.) In particular,

$$\frac{\ln 2}{a-1} = \frac{a}{n}. \tag{7}$$

We first show that (4) holds for $m = \lceil a \rceil$. Using (7), we get

$$\frac{n-a}{n} \le 2^{-1/(a-1)} \iff 2^{1/(a-1)}\left(1-\frac{a}{n}\right) \le 1 \iff \exp\left(\frac{\ln 2}{a-1}\right)\left(1-\frac{a}{n}\right) \le 1$$

$$\iff \exp\left(\frac{a}{n}\right)\left(1-\frac{a}{n}\right) \le 1.$$

To see that the last inequality holds, let $f(x) = e^x(1-x)$. Then $f(0) = 1$ and $f'(x) < 0$ if $x > 0$. Since $0 < \frac{a}{n} < 1$, we have $f\left(\frac{a}{n}\right) < f(0) = 1$, so

$$\frac{n-a}{n} < 2^{-1/(a-1)}, \tag{8}$$

as claimed. Letting $m = \lceil a \rceil > a$, we conclude from (8) that (4) holds. Hence (2) holds, so

$$T(n) \le \left\lceil \frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2} \right\rceil.$$

Having obtained an upper bound for $T(n)$, we next derive a lower bound. Letting $b = a - 1 = -\frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2}$, we have $b^2 + b = n\ln 2$, so

$$\frac{\ln 2}{b-1} = \frac{b(b+1)}{n(b-1)}. \tag{9}$$

Using (9), the estimate $e^x > 1 + x$, and the fact that $b^2 < n + 1$, we find after calculation that

$$\exp\left(\frac{\ln 2}{b-1}\right)\left(1 - \frac{b}{n-b+2}\right) > 1,$$

which implies

$$\frac{n-2b+2}{n-b+2} > 2^{-1/(b-1)}. \tag{10}$$

It follows from (10) that if $m = \lfloor b \rfloor < b$ (where $\lfloor \cdot \rfloor$ denotes the floor function), then

$$\frac{n-2m+2}{n-m+2} > 2^{-1/(m-1)}. \tag{11}$$

Now the left side of (11) is the *smallest* factor on the left side of (3), so

$$\binom{n-m+1}{m} \Big/ \binom{n}{m} > 1/2.$$

Thus $\left\lfloor -\frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2} \right\rfloor < T(n)$, and we conclude that

$$\left\lceil -\frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2} \right\rceil \le T(n) \le \left\lceil \frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2} \right\rceil. \tag{12}$$

Now a routine calculation shows that

$$\left\lceil -\frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2} \right\rceil \le \left\lceil \sqrt{n\ln 2} \right\rceil \le \left\lceil \frac{1}{2} + \sqrt{\frac{1}{4} + n\ln 2} \right\rceil. \tag{13}$$

Since the upper and lower bounds differ by 1, the theorem follows from (12) and (13).

■

We could also consider threshold numbers $T_\alpha(n)$ for probabilities α other than $1/2$. Our first conjecture would then be replaced by the guess that $T_\alpha(n) \approx \sqrt{n \ln\left(\frac{1}{1-\alpha}\right)}$. For example, if $\alpha = 1 - e^{-1} \approx 0.6321$, then our conjecture would have the simple form $T_\alpha(n) = \lceil \sqrt{n} \rceil$.

Our proof of Theorem 5 is easy to modify to show that when $0 < \alpha < 1 - e^{-1}$, then

$$\left| T_\alpha(n) - \left\lceil \sqrt{n \ln\left(\frac{1}{1-\alpha}\right)} \right\rceil \right| \leq 1. \tag{14}$$

Our numerical experiments suggest that (14) holds (for sufficiently large n) also when $1 - e^{-1} \leq \alpha < 1$, but we do not yet have a proof.

Acknowledgment. The authors thank the referees for helpful advice and for pointing out the references [**1**] and [**5**]. The latter reference led us to [**4**].

REFERENCES

1. M. Abramson and W. Moser, More birthday surprises, *Amer. Math. Monthly* **77** (1970), 856–8.
2. C. Berge, *Principles of Combinatorics*, Academic Press, New York, NY, and London, UK, 1971.
3. Website of the Colorado Lottery: http://www.coloradolottery.com/.
4. N. Henze, The distribution of spaces on lottery tickets, *The Fibonacci Quarterly* **33** (1995), 426–431.
5. N. Henze and Hans Riedwyl, *How to Win More: Strategies for Increasing a Lotto Win*, A.K. Peters, Natick, MA, 1998.
6. S. Ross, *A First Course in Probability (5th Ed.)*, Prentice Hall, Upper Saddle River, NJ, 1998.

Pythagorean Boxes

RAYMOND A. BEAUREGARD
E. R. SURYANARAYAN
University of Rhode Island
Kingston, RI 02881

Introduction A Pythagorean rectangle is one with integer sides and integer diagonals. Much has been written about these in the context of Pythagorean triangles which are represented by Pythagorean Triples (PTs), that is, integer triples (a, b, c) satisfying $a^2 + b^2 = c^2$ (See, for instance [**12**]). A Pythagorean box is a box whose edges and inside diagonals are integers. These are represented by ordered quadruples (x, y, z, w) whose components are integers satisfying the equation

$$x^2 + y^2 + z^2 = w^2, \tag{1}$$

where $w > 0$. We refer to such quadruples as PBs.

Our purpose is to look at the geometric and algebraic properties of PBs with one eye on the many nice properties that PTs exhibit. For example, every PB corresponds to the rational point $(x/w, y/w, z/w)$ on the unit sphere; this is analogous to the correspondence of PTs with rational points on the unit circle. It is well known that every pair of positive integers n, m determines a PT, namely $(n^2 - m^2, 2nm, n^2 + m^2)$. By analogy we show how every pair of positive rationals determines a PB. The set of PTs can be made into a group in at least two different ways [**2, 4**]. We show how the set of PBs is a group extension of each of these. In fact, the set of PBs and the set of PTs are made

into fields, the former an extension of the latter of degree 2. We analyze Pythagorean boxes with a square base in terms of a Pell equation, and close by discussing perfect Pythagorean boxes.

We will find it convenient to allow x, y, and z to have negative values (thus including each octant of the unit sphere). We allow degeneracy, letting x, y, or z be 0 so as to include PTs among the PBs. For reasons that will become clear shortly y and z are not allowed to be simultaneously zero; thus $y^2 + z^2 \neq 0$.

New PBs from old Motivated by the ways in which PTs can be composed to produce new ones, we define the following operation on PBs:

$$(x_1, y_1, z_1, w_1) * (x_2, y_2, z_2, w_2) = \tag{2}$$
$$(x_1w_2 + w_1x_2, y_1z_2 + z_1y_2, z_1z_2 - y_1y_2, x_1x_2 + w_1w_2).$$

It is not difficult to show that the quadruple defined in (2) is a PB. One may use Cramer's Rule, or argue directly, to see that the two (new) middle components are not simultaneously zero. To illustrate the binary operation (2), we have $(3, 2, 6, 7) * (3, 6, 22, 23) = (90, 80, 120, 170)$. Thus, a product of *primitive* PBs (where the components are relatively prime) can be far from primitive.

This generalizes some previously known operations on PTs. When $x = 0$, equation (1) defines the PT (y, z, w) which can be associated with the rational point $(y/w, z/w)$ on the unit circle. In this case, the operation in (2) is similar to the one described by Eckert [**4**] (where additive notation is used). On the other hand, if $y = 0$, then $z \neq 0$, and (1) defines the PT (x, z, w) associated with the rational point $(w/z, x/z)$ on the unit hyperbola $(w/z)^2 - (x/z)^2 = 1$. In this case, equation (2) reduces to the binary operation described by the authors in [**2**].

The operation above might seem more natural if we associate the PB defined by (1) with the matrix

$$\begin{bmatrix} w & x & 0 & 0 \\ x & w & 0 & 0 \\ 0 & 0 & z & y \\ 0 & 0 & -y & z \end{bmatrix} \tag{3}$$

having integer entries and equal (2×2 main-diagonal) block determinants. Then our rule of composition (2) corresponds to matrix multiplication. Since the set of such matrices is a multiplicative semigroup, the set

$$\mathcal{B} = \{B : B = (x, y, z, w) \text{ satisfies (1)}, w > 0, y^2 + z^2 \neq 0\}$$

of PBs is also a semigroup. Notice that the identity matrix corresponds to the PB $(0, 0, 1, 1)$ which is then the identity element for $\mathcal{B}$.

The determinant of a matrix like (3) is nonzero and hence it is invertible; its inverse corresponds to the quadruple

$$\frac{1}{(y^2 + z^2)}(-x, -y, z, w), \tag{4}$$

which represents a *rational box*. We define the related PB, $B^{\#} = (-x, -y, z, w)$ to be the *quasi-inverse* of B. Note that $B * B^{\#} = (0, 0, y^2 + z^2, y^2 + z^2)$. Quasi-inverses are useful in establishing that the semigroup $\mathcal{B}$ is cancellative.

Let us define a relation $\sim$ on $\mathcal{B}$ by $(x_1, y_1, z_1, w_1) \sim (x_2, y_2, z_2, w_2)$ if there are positive integers n and m such that $(nx_1, ny_1, nz_1, nw_1) = (mx_2, my_2, mz_2, mw_2)$. It

is not difficult to show that $\sim$ is an equivalence relation; it has the effect of identifying all PBs of the form (nx, ny, nz, nw), as n varies over the positive integers. We write $[x, y, z, w]$ for the equivalence class containing (x, y, z, w), and use the symbol $\mathcal{R}$ for the set of these equivalence classes. In fact, $\sim$ is a *congruence* on $\mathcal{B}$ in the sense that if $A_1 \sim A_2$ and $B_1 \sim B_2$ then $A_1 * B_1 \sim A_2 * B_2$. Thus $\mathcal{R}$ becomes a group under the induced operation. The mapping

$$[x, y, z, w] \to \frac{1}{\sqrt{(y^2+z^2)}} \begin{bmatrix} w & x & 0 & 0 \\ x & w & 0 & 0 \\ 0 & 0 & z & y \\ 0 & 0 & -y & z \end{bmatrix} \tag{5}$$

is an isomorphism of $\mathcal{R}$ onto the group of these unimodular (that is, determinant 1) matrices.

Parameters Let (x, y, z, w) be a PB. Recalling that the square of an odd integer reduces to 1 (mod 8), we see that at least two of x, y, z in (1) must be even. Henceforth we take y and z to be even. Notice that if (x, y, z, w) is nondegenerate ($xyz \neq 0$) and primitive then x and w must be odd.

For any three integers a, b, c with $c > 0$ and a and b not both 0, a simple computation shows that the quadruple (x, y, z, w) defined by

$$x = (a^2 + b^2 - c^2)/c, \ \ y = 2a, \ \ z = 2b, \ \ w = (a^2 + b^2 + c^2)/c \tag{6}$$

is a PB provided that c divides $a^2 + b^2$ (see [**11**, p. 68]). A calculation shows that if $c_1 c_2 = a^2 + b^2 \neq c^2$ then we obtain the same PB using a, b, c_1 as we do using a, b, c_2 except for the algebraic sign of the first component. For example, (1, 3, 2) yields the PB (3, 2, 6, 7) and (1, 3, 5) yields (−3, 2, 6, 7). Conversely if (x, y, z, w) is a given PB we may find such parameters a, b, c from the equations

$$a = y/2, \ \ b = z/2, \ \ c = (w - x)/2, \tag{7}$$

and (6) holds. Thus the correspondence

$$(x, y, z, w) \to 2(a, b, c)$$

is one to one. This correspondence carries through to matrices via

$$\begin{bmatrix} w & x & 0 & 0 \\ x & w & 0 & 0 \\ 0 & 0 & z & y \\ 0 & 0 & -y & z \end{bmatrix} \to \begin{bmatrix} w-x & 0 & 0 \\ 0 & z & y \\ 0 & -y & z \end{bmatrix} = 2 \begin{bmatrix} c & 0 & 0 \\ 0 & b & a \\ 0 & -a & b \end{bmatrix}. \tag{8}$$

The composite mapping

$$[x, y, z, w] \to \frac{1}{\sqrt{(y^2+z^2)}} \begin{bmatrix} w & x & 0 & 0 \\ x & w & 0 & 0 \\ 0 & 0 & z & y \\ 0 & 0 & -y & z \end{bmatrix} \to \frac{1}{c} \begin{bmatrix} b & a \\ -a & b \end{bmatrix} \tag{9}$$

is an isomorphism of the group $\mathcal{R}$ with the group of 2×2 matrices

$$\left\{ A : A = \begin{bmatrix} r & s \\ -s & r \end{bmatrix}, r, s \in \mathbb{Q}, r^2 + s^2 \neq 0 \right\}. \tag{10}$$

For example, let $r = 2/3$ and $s = 4/5$. Using (6) and (9), the matrix

$$M = \begin{bmatrix} 2/3 & 4/5 \\ -4/5 & 2/3 \end{bmatrix} = \frac{1}{15}\begin{bmatrix} 10 & 12 \\ -12 & 10 \end{bmatrix}$$

gives rise to the quadruple $(19/15, 24, 20, 469/15)$, which represents a rational box; the corresponding PB is $(19, 360, 300, 469)$, whose parameters $a = 150$, $b = 180$, $c = 225$ yield matrix M.

Subgroups of PTs Consider $[x, y, z, w]$ with $xyz = 0$, and look back at (6). When $x = 0$, then $a^2 + b^2 = c^2$, so these parameters form a PT (a, b, c) arising from $[0, a, b, c]$, and (9) gives an isomorphism of this subgroup (corresponding to the rational points on the unit circle) with the subgroup of the group in (10) consisting of unimodular matrices. On the other hand, when $y = 0$ then $z \neq 0$ and (9) reduces to

$$[x,\ 0,\ z,\ w] \to \begin{bmatrix} (w+x)/z & 0 \\ 0 & (w+x)/z \end{bmatrix}, \tag{11}$$

since $a = 0$ and $b/c = z/(w - x) = (w + x)/z$. This gives an isomorphism of this subgroup (corresponding to the rational points on the unit hyperbola), with the group of nonzero rationals. The fraction in (11), when written in lowest terms, yields the PT parameters for (x, z, w) as described by the authors in [**2**]. However, the group of PTs there is isomorphic to the subgroup of positive rationals due to the fact that only the middle components are allowed to be negative in that discussion.

The subset of $\mathcal{R}$ with $z = 0$ is of no algebraic interest since it is not closed under our binary operation.

The field of PBs; the subfield of PTs The matrices described by (10), together with the zero matrix, form a field $\mathcal{F}$ isomorphic to the field of Gaussian rationals. If we make an exception and include $[1, 0, 0, 1]$ in the set $\mathcal{R}$ of equivalence classes of PBs, we can extend the composite mapping (9), and pair $[1, 0, 0, 1]$ with the 2×2 zero matrix. In this way, the thus-enlarged $\mathcal{R}$ becomes a field isomorphic to $\mathcal{F}$, the addition in $\mathcal{R}$ induced by that in $\mathcal{F}$. The zero element of $\mathcal{R}$ is $[1, 0, 0, 1]$, and the additive inverse of $[x, y, z, w]$ is $[x, -y, -z, w]$. One must be careful to interpret $n[x, y, z, w]$ correctly for positive integers n: it is *not* $[nx, ny, nz, nw]$ (which is the same as $[x, y, z, w]$). A computation shows that

$$n[x, y, z, w] = [(n^2 + 1)x + (n^2 - 1)w, 2ny, 2nz, (n^2 - 1)x + (n^2 + 1)w].$$

Clearly the general formula for addition in $\mathcal{R}$ is complicated; addition in particular cases is best carried out using parameters. For example,

$$[3, 2, 6, 7] + [4, 2, 4, 6] \to \begin{bmatrix} 3/2 & 1/2 \\ -1/2 & 3/2 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} =$$

$$\begin{bmatrix} 7/2 & 3/2 \\ -3/2 & 7/2 \end{bmatrix} \to [27, 6, 14, 31].$$

The subset $\mathcal{P}$ consisting of (equivalence classes of) PTs of $\mathcal{R}$ with $y = 0$ is a subfield of $\mathcal{R}$ which is isomorphic to the field of rationals; indeed the mapping (11), when extended to map $[1, 0, 0, 1]$ to the zero matrix, is a field isomorphism. Thus $\mathcal{R}$ is a 2-dimensional field extension of $\mathcal{P}$. Multiplication of PTs is illustrated in [**2**]. As an illustration of addition of PTs, the reader may check (using (11) and (7)) that

$$[3, 0, 4, 5] + [5, 0, 12, 13] = [45, 0, 28, 53].$$

Interestingly, the subset of $\mathcal{R}$ with $x = 0$ is not closed under addition.

The nature of components In this section we consider nondegenerate PBs. It is known [**11**, p. 407] that a positive integer w is the inside diagonal of a Pythagorean Box if and only if it is not of the form $w = 2^i$ or $w = 2^i \times 5$ (where i is a nonnegative integer). If, as we do for the remainder of this section, we restrict ourselves further to primitive PBs, then only odd integers $w \neq 1, 5$ can so serve (even values of w are easily ruled out because two other components are even). In contrast, every positive integer is one of the first three components (that is, an edge) of a PB. If $n > 1$ is an odd integer then $(n, \frac{n^2-1}{2}, \frac{n^4+2n^2-3}{8}, \frac{n^4+2n^2+5}{8})$ is a PB; for $n = 1$, we exhibit $(1, 2, 2, 3)$. For even integers we see that $(h^2 + k^2 - 1, 2h, 2k, h^2 + k^2 + 1)$ is a PB for any integers h and k. In fact, taking $h = k$ we see that every positive even integer is the edge of a square side of some Pythagorean Box. The same cannot be said of odd integers since at least two of the first three components of a PB must be even.

Let us pursue PBs of the form (x, y, y, w) with a square $y \times y$ base and altitude x. Thus

$$x^2 = w^2 - 2y^2. \tag{12}$$

When $x = 1$, (12) is a Pell equation with a well-known infinite set of primitive solutions [**1**]:

$$S = \{(w, y) = (3, 2), (17, 12), (99, 70), (577, 408), \ldots\}$$

Thus there are an infinite number of boxes with a square base and altitude 1. The first few such PBs are $(x, y, z, w) = (1, 2, 2, 3)$, $(1, 12, 12, 17)$, $(1, 70, 70, 99)$, $(1, 408, 408, 577)$. What other altitudes x are possible? The infinite solution set S together with the least positive solution of (12) can be used to generate an infinite solution set for (12), assuming that x is an integer for which (12) is solvable. This is done using Brahmagupta's identity [**1**, p. 320], a form of which asserts that if $t_1 = w_1^2 - 2y_1^2$ and $t_2 = w_2^2 - 2y_2^2$, then $t_1 t_2 = w_3^2 - 2y_3^2$, where $w_3 = w_1 w_2 + 2y_1 y_2$ and $y_3 = w_1 y_2 + w_2 y_1$. For example, the solutions in S are reproduced in this way, combining two (w, y) pairs to form a third.

Equation (12) is solvable exactly when $x = \pm 1$ or is a product of primes $\equiv \pm 1 \pmod 8$. To see why this is true let (x, y, w) be a primitive solution of (12) and let p be an odd prime divisor of x. Then w and p are relatively prime, so y is invertible (mod p) and we may let t be an integer such that $w \equiv ty \pmod p$. Substituting into (12), we see that $x^2 \equiv (t^2 - 2)y^2$ and so $(t^2 - 2) \equiv 0 \pmod p$. This shows that 2 is a quadratic residue of p. But then we must have $p \equiv \pm 1 \pmod 8$ (see [**1**, p. 130]). For the converse, it is known [**10**, p. 210] that if p is a prime $\equiv \pm 1 \pmod 8$ then (12) with p in place of x^2 has a primitive solution. Using Brahmagupta's identity, it follows immediately that (12) is solvable when x is a product of such primes.

For each of these allowable altitudes, x, the set of primitive solutions of (12) is infinite. For example, there are an infinite number of PBs with a square base and altitude 7, or 17, or $7 \times 17 = 119$, etc.

Perfect Pythagorean Boxes Are there any perfect Pythagorean Boxes in the sense that all of the sides are Pythagorean rectangles? This famous problem is treated extensively by Richard Guy [**5**, p. 173], where the discussion is based on a paper by Leech [**8**]. There is an interesting equivalent, but very different, formulation of this problem by Luca [**9**]. It remains an open question. Computer searches by Korec [**6, 7**] have established 10^6 as a lower bound for any edge and 10^9 as a lower bound for the largest edge for such boxes.

Two perpendicular sides of a Pythagorean Box can have an integer diagonal as illustrated with the PB (153, 104, 672, 697), which is the smallest known example; both

(153, 104, 185) and (104, 672, 680) are PTs. This is a particular case of the following. In general, it can be shown that if

$$x = p^2q^2 - r^2s^2, \qquad y = 2pqrs, \qquad z = p^2r^2 - q^2s^2, \tag{13}$$

then $x^2 + y^2$ and $y^2 + z^2$ are squares and

$$x^2 + y^2 + z^2 = (p^4 + s^4)(q^4 + r^4). \tag{14}$$

Expression (14) is a square if $p^4 + s^4 = q^4 + r^4$, and there are an infinite number of (primitive) solutions to this equation (see [**3**, p. 644], [**11**, p. 55], and [**13**]). The right-hand side of (14) can be a square without the factors being equal, as shown by our example where $(p, q, r, s) = (13, 1, 2, 2)$. An example that is not of the form (13) is given by (117, 520, 756, 925). Unlike the previous example, the two even components do not give rise to a PT. If we have whetted your appetite, further discussions are available ([**5**] and [**8**]).

REFERENCES

1. A. Adler and J. Coury, *The Theory of Numbers*, Jones & Bartlett Pub. Co., Boston, 1995.
2. R.A. Beauregard and E.R. Suryanarayan, Pythagorean triples: the hyperbolic view, *College Math. J.* **27** (1996), 170–181.
3. L.E. Dickson, *History of Theory of Numbers*, Vol. II, Chelsea, New York, 1952.
4. E. Eckert, The group of primitive Pythagorean triangles, this MAGAZINE **54** (1984), 22–27.
5. R.K. Guy, *Unsolved Problems in Number Theory*, Vol. I, Springer-Verlag, New York, 1994.
6. I. Korec, Nonexistence of a small perfect rational cuboid, II, *Acta Math. Univ. Comenian.*, **44/45** (1984), 39–48.
7. I. Korec, Lower bounds for perfect rational cuboids, *Math. Slovaca* **42** (1992), no. 5, 565–582.
8. John Leech, The rational cuboid revisited, *Amer. Math. Monthly* **84** (1977), 534–533; corrections (Jean Lagrange) **85** (1978), 473.
9. Florian Luca, Perfect cuboids and perfect square triangles, this MAGAZINE **73** (2000), 400–401.
10. T. Nagell, *Introduction to Number Theory*, John Wiley & Sons, New York, NY (1951).
11. W. Sierpinski, *Elementary Theory of Numbers*, Vol 32, North-Holland Mathematical Library, Amsterdam, 1988.
12. Darko Veljan, The 2500-year-old Pythagorean theorem, this MAGAZINE, **73** (2000), 259–272.
13. A. J. Zajta, Solutions of the Diophantine equation $x^4 + y^4 = z^4 + t^4$, *Math. Comp.* **41** (1983), 635–659.

A Simple Fact About Eigenvectors That You Probably Don't Know

WARREN P. JOHNSON
University of Wisconsin
Madison, WI 53706

Suppose, in a course in elementary linear algebra, we are doing a first example of the calculation of eigenvalues and eigenvectors, say for the matrix $A = \left(\begin{smallmatrix} 5 & 2 \\ 3 & 6 \end{smallmatrix}\right)$. The eigenvalues are the solutions of $(5 - \lambda)(6 - \lambda) - 2 \cdot 3 = 0$, which are $\lambda_1 = 3$ and $\lambda_2 = 8$. To get the eigenvector corresponding to $\lambda_1 = 3$, we find a basis for the null space of $A - 3I$:

$$A - 3I = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \quad \Longrightarrow \quad \vec{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{is an eigenvector.}$$

To get the eigenvector corresponding to $\lambda_2 = 8$, we find a basis for the null space of $A - 8I$:

$$A - 8I = \begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} \quad \Longrightarrow \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{is an eigenvector.}$$

The first time I ever did an example like this in front of a class, several students remarked at this point (paraphrased): "Hey look, both columns of $A - 3I$ are the same as $\vec{v}_2$! And, well, that's not quite true for $A - 8I$, but at least both columns of that are multiples of $\vec{v}_1$. Does that always happen?" In other words, they were led to conjecture

THEOREM 1. *Suppose A is a 2×2 matrix with distinct eigenvalues λ_1 and λ_2, and corresponding eigenvectors $\vec{v}_1$ and $\vec{v}_2$ respectively. Then both columns of $A - \lambda_1 I$ are multiples of $\vec{v}_2$, and both columns of $A - \lambda_2 I$ are multiples of $\vec{v}_1$.*

Did you know that? I didn't. My initial reaction (which I had just enough sense to hide from my students) was: it's probably false, because if it were true, it would be in all the textbooks. In fact it doesn't seem to be in any of the textbooks; neither has anyone I've shown it to recognized it. I found one person who had observed the phenomenon, and one or two others proved the theorem on the spot. We shall deduce it from

LEMMA 1. *Suppose A is a matrix with at least two distinct eigenvalues. Let λ be one of them, and let $\vec{v}$ be an eigenvector whose eigenvalue is not λ. Then $\vec{v}$ is in the column space of $A - \lambda I$.*

This follows in turn from the trivial

LEMMA 2. *If $\vec{v}$ is an eigenvector of A with eigenvalue μ, then $\vec{v}$ is also an eigenvector of $A - \lambda I$ with eigenvalue $\mu - \lambda$.*

Proof of the Lemmas. We have

$$(A - \lambda I)\,\vec{v} = A\vec{v} - \lambda\vec{v} = \mu\vec{v} - \lambda\vec{v} = (\mu - \lambda)\,\vec{v},$$

which is Lemma 2. If $\mu \neq \lambda$, this says that a nonzero scalar multiple of $\vec{v}$ is a linear combination of the columns of $A - \lambda I$, so $\vec{v}$ is itself a linear combination of the columns of $A - \lambda I$, *i.e.*, $\vec{v}$ is in the column space of $A - \lambda I$; thus Lemma 1.

This unprepossessing lemma has some interesting consequences, as we shall see. Theorem 1 follows immediately, since the column spaces of $A - \lambda_1 I$ and $A - \lambda_2 I$ there are one-dimensional. Moreover, we have the following generalization:

THEOREM 2. *Suppose A is an $n \times n$ matrix with a set $S = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of n independent eigenvectors. If λ is an eigenvalue of A, then*

(i) The eigenvectors in S that correspond to λ are a basis for the null space of $A - \lambda I$.

(ii) The eigenvectors in S that do not correspond to λ are a basis for the column space of $A - \lambda I$.

Proof. The point of the theorem is (ii). Since the eigenvectors of A corresponding to λ are all the vectors in the null space of $A - \lambda I$, (i) holds by definition. To prove (ii), suppose that there are k eigenvectors in S that correspond to λ. Then we know that

(a) There are $n - k$ eigenvectors in S that do not correspond to λ, all of which, by Lemma 1, are in the column space of $A - \lambda I$.

(b) The dimension of the null space of $A - \lambda I$ is k, and hence the dimension of the column space of $A - \lambda I$ is $n - k$.

(ii) follows from (a) and (b), since whenever we have $n - k$ independent vectors in a vector space of dimension $n - k$, they must be a basis for it. ∎

Theorem 1 provides a simple check on eigenvector calculations for most 2×2 matrices. It could also be used to shorten them: if we had Theorem 1 when we were doing our first example, consideration of either $A - 3I$ or $A - 8I$ would give us both eigenvectors of A at once. This idea might be used with advantage when the eigenvalues and eigenvectors are not so nice numerically, *e.g.*, if they are complex. The eigenvalues of $\left(\begin{smallmatrix} 1 & -2 \\ 5 & 3 \end{smallmatrix}\right)$ are $\lambda = 2 \pm 3i$. To find the eigenvector for $2 - 3i$ we should ordinarily have to find the nullspace of

$$\begin{pmatrix} 1 - (2 - 3i) & -2 \\ 5 & 3 - (2 - 3i) \end{pmatrix} = \begin{pmatrix} -1 + 3i & -2 \\ 5 & 1 + 3i \end{pmatrix}.$$

It is not hard to imagine a student making a mistake doing this. We might observe instead that either column of the above matrix, say

$$\begin{pmatrix} -2 \\ 1 + 3i \end{pmatrix}, \quad \text{must be an eigenvector for } 2 + 3i,$$

and hence

$$\begin{pmatrix} -2 \\ 1 - 3i \end{pmatrix} \quad \text{must be an eigenvector for } 2 - 3i.$$

The probability that Theorem 2 can be used in a similar way for a larger matrix is small, but not zero. One such example, the details of which we leave to the reader, is

$$\begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 0 \\ 3 & 6 & 5 \end{pmatrix}.$$

One might say that the most interesting thing about Theorem 2 is that it is *not* interesting—the general case is, in a sense, less interesting than the 2×2 case. But I do think Theorem 2 brings together some of the theory of a basic linear algebra course in an appealing way.

Lemma 1 further enables an easy proof of a standard result on eigenvectors of symmetric matrices:

THEOREM 3. *Let A be a real symmetric $n \times n$ matrix. If $\vec{v}$ and $\vec{w}$ are eigenvectors of A with different eigenvalues, then $\vec{v} \perp \vec{w}$.*

Recall that a real symmetric matrix has real eigenvalues and eigenvectors, so no complex conjugation is needed for the orthogonality.

Proof. Let λ be the eigenvalue for $\vec{v}$; this means that $\vec{v}$ is in the null space of $A - \lambda I$. Moreover, since λ is not the eigenvalue for $\vec{w}$, $\vec{w}$ is in the column space of $A - \lambda I$. Since A is symmetric, $A - \lambda I$ is also symmetric, and therefore $\vec{w}$ is in the row space of $A - \lambda I$. But the row space and the null space are orthogonal complements, so we must have $\vec{v} \perp \vec{w}$. ∎

This argument is outlined in problem 18 in section 6.4 of Strang [**1**]. It seems to me at least as good as the usual, more computational proof, which one can also find there (or in any number of other linear algebra books).

Acknowledgment. I became aware of Theorems 1 and 2 in Fall 1996, at Beloit College, when I taught linear algebra for the first time. I would like to thank that class for helping me learn these facts (and others), and for their part in transforming linear algebra from a subject that I did not particularly like as a student into one that I enjoy teaching very much. I also want to thank Gil Strang for his comments and encouragement after reading a preliminary version of this note (and for [**1**], which is my favorite linear algebra textbook). Paul Terwilliger made a helpful remark as well.

REFERENCES

1. Gilbert Strang, *Introduction to Linear Algebra* , 2nd Edition, Wellesley-Cambridge Press, Wellesley, MA, 1998.

A Generalized General Associative Law

WILLIAM P. WARDLAW
U. S. Naval Academy
Annapolis, MD 21402

This note is primarily evangelical. I urge you to adopt, as your favorite proof of the general associative law, a proof that generalizes that law.

Most of us would agree that the general associative law is an interesting and important concept of abstract algebra. But a cursory sampling of algebra texts shows that the treatment of this law differs considerably from author to author; some treatments are hazy, and others are incomplete.

I favor the job done by Nathan Jacobson in [**11**]. He begins on page 20 by inductively defining the *left associated product* $\prod_1^m a_i$ by the formulas

$$\prod_1^1 a_i = a_1, \qquad \prod_1^{r+1} a_i = \left(\prod_1^r a_i\right) a_{r+1}. \tag{1}$$

Next, Jacobson uses the associative law

$$(ab)c = a(bc) \tag{2}$$

and induction on m to establish, as a lemma, that

$$\prod_1^n a_i \prod_1^m a_{i+n} = \prod_1^{n+m} a_i. \tag{3}$$

Finally, Jacobson uses induction to show that all products associated with $(a_1 \cdot a_2 \cdots a_n)$ are equal:

$$\begin{aligned}(a_1 \cdot a_2 \cdots a_n) = p = uv &= (a_1 \cdots a_m)(a_{m+1} \cdots a_n) \\ &= \left(\prod_1^m a_i\right)\left(\prod_1^{n-m} a_{i+m}\right) = \prod_1^n a_i.\end{aligned}$$

Hungerford [**10**, p. 28] gives the same proof as Jacobson, but neatly shortens it to six equalities.

Zassenhaus [**19**, p. 1] tersely asserts: "A product of arbitrarily many factors is determined solely by the order of its factors." He lets n be greater than three, and assumes

that for every $m < n$ the product $a_1 \cdot a_2 \cdot a_m$ designates, "unambigously, an element of the group." Then

$$\begin{aligned} P = a_1 \cdot a_2 \cdots a_n = P_1 P_2 &= (a_1 \cdot a_2 \cdots a_m)(a_{m+1} \cdot a_{m+2} \cdots a_n) \\ &= (a_1(a_2 \cdots a_m))\,(a_{m+1} \cdot a_{m+2} \cdots a_n) \\ &= a_1\,((a_2 \cdots a_m)(a_{m+1} \cdot a_{m+2} \cdots a_n)) = a_1(a_2 \cdots a_n) \end{aligned} \tag{4}$$

uniquely determines the product. (I recommend that you read the original, where Zassenhaus makes exactly those comments necessary to make the proof totally unambiguous.)

Barnes [**2**] gives a statement and proof of the general associative law on pages 12 and 13 which is almost exactly the same as given by Zassenhaus, except that Barnes uses the left associated product and Zassenhaus uses the right associated product as their standard products. Michael Artin [**1**] and Ledermann [**13**] give valid proofs similar to (4), but I find both a bit ambiguous or hazy.

W. R. Scott [**14**, pp. 3–4] allows no ambiguity in his statement and proof of the General Associative Law 1.1.3. His distinctive approach is well worth reading!

Van der Waerden [**15**, p. 16] stops short of the general associative law and only proves (3). Both Chevalley [**5**, p. 4] and Zariski and Samuel [**18**, pp. 2–3] prove that if $n_0, n_1, \ldots, n_r$ are integers such that $0 = n_0 < n_1 < \cdots < n_r = n$, then

$$\prod_{j=1}^{r} \left(\prod_{k=n_{j-1}+1}^{n_j} a_i \right) = \prod_{1}^{n} a_k, \tag{5}$$

which stops short of the general associative law.

A number of authors do not prove the general associative law. Among these are Birkhoff and MacLane [**3**], Erlich [**6**], Hall [**7**], Hu [**9**], Lang [**12**], Walker [**16**], who say it can be shown, and Herstein [**8**], who seems not to mention it at all. Walker [**16**, p. 32], cites Jacobson. Hall [**7**] relegates the proof to Exercise 1 on page 24.

With such a plethora of proofs and nonproofs of the general associative law, one could rightfully wonder what could be added to the subject. But there is a proof in [**17**], much like the treatment by Barnes in [**2**] and by Zassenhaus in [**19**], which actually generalizes the general associative law. To see this, allow me to introduce some of the notions from [**17**].

We define a *groupoid* to be an ordered pair $(G, *)$, where G is a set and $*$ is a (full) *binary operation* on G; that is, $*$ is a function $* : G \times G \to G$ with domain $G \times G$ and range a subset of G. (See [**2**, p. 24] or [**4**, p. 1].) We denote the groupoid with a boldface $\mathbf{G}$, and let $\mathbf{G}^2$ denote the set of all products in $\mathbf{G}$, that is, the range of the binary operation $*$. As is usual, we will henceforth denote $a * b$ by ab. A groupoid $\mathbf{G}$ is *n-associative* if the product of any n elements is independent of how they are associated, that is, if $a_1 a_2 \cdots a_n$ denotes unambiguously an element of $\mathbf{G}$ independent of the way the product is parenthesized. Now we can state the "Generalized General Associative Law":

THEOREM 1. *Let* $\mathbf{G}$ *be a groupoid that is n-associative for some* $n \geq 3$. *Then* $\mathbf{G}$ *is* $(n+1)$*-associative.*

Proof. Assume that $\mathbf{G}$ is n-associative. Consider a product $p \equiv (a_1 a_2 \cdots a_n a_{n+1})$ of $n+1$ elements of $\mathbf{G}$ with some association of factors. (We adopt the convention of using "$\equiv$" between products which have the same parenthesization, or to assign values with a fixed parenthesization to a single variable, and using "$=$" when n-associativity implies the equality of the products.) Then there must be a number k with $1 \leq k \leq n$ such that the elements a_k and a_{k+1} are grouped to-

gether to form the element $b_k \equiv a_k a_{k+1}$ of $\mathbf{G}$. If $i < k$, let $b_i \equiv a_i$, and let $b_i \equiv a_{i+1}$ if $i > k$. Thus we have $p \equiv (b_1 b_2 \cdots b_n) = (b_1 \cdots b_{n-1}) b_n$. If $k < n$, the latter is $p = (b_1 \cdots b_{n-1}) a_{n+1} \equiv (a_1 \ldots a_n) a_{n+1}$. But if $k = n$, $n - 1 \geq 2$ and $p = (b_1 \ldots b_{n-1}) b_n = ((b_1 b_2) \ldots) b_n \equiv ((a_1 a_2) \ldots)(a_n a_{n+1}) \equiv (c_1 \ldots)(c_{n-1} c_n)$ with $c_1 = a_1 a_2$ and $c_i = a_{i+1}$ for $i > 1$. Then n-associativity gives $p = (c_1 \cdots c_{n-1}) c_n \equiv ((a_1 a_2) \cdots a_n) a_{n+1} = (a_1 \cdots a_n) a_{n+1}$. Thus, in any case, p is the unambiguously defined element $(a_1 \cdots a_n) a_{n+1}$ of $\mathbf{G}$. Therefore, $\mathbf{G}$ is $(n + 1)$-associative. ∎

This proof is quite similar to the Zassenhaus proof, but with one significant difference. In order to apply induction, Zassenhaus started with an "external factorization"

$$P \equiv (a_1 \cdots a_m)(a_{m+1} \cdots a_n),$$

while we found it necessary to use an "internal factorization"

$$p \equiv (a_1 \cdots (a_k a_{k+1}) \cdots a_{n+1}) \tag{6}$$

to reduce a product of $n + 1$ elements to a product of n elements, to which n-associativity could be applied. Of course, each of these factorizations rests on the requirement that a binary operation can only act on exactly two elements of the groupoid.

There are probably as many variations of the proof that n-associativity implies $(n + 1)$-associativity as there are of the standard general associativity law. One can be less formal and avoid introducing the variables b_i and c_i, or one can more formally define an n-product and show that every n-product is equal to a standard n-product, perhaps the left associated product (3). For example, see [**17**, p. 588]. It seems certain that the internal factorization (6) must be part of the proof. It would be nice if the awkwardness of the $k = n$ case in (6) could be more gracefully handled.

We claim that Theorem 1 is a proper generalization of the general associative law. To see this, we must produce examples of *strictly n-associative groupoids*, that is, groupoids that are n-associative but are not $(n - 1)$-associative. The construction of examples is hampered if one is unaware of the following result:

THEOREM 2. *Let* $\mathbf{G}$ *be a groupoid such that* $\mathbf{G}^2 = \mathbf{G}$. *Then* $\mathbf{G}$ *being* $(n + 1)$-*associative implies that* $\mathbf{G}$ *is n-associative.*

Proof. Let $(a_1 a_2 \cdots a_n)$ be an arbitrary product of n factors. Writing $a_1 \equiv b_1 b_2$ and $b_{i+1} \equiv a_i$ for $i = 2, 3, \ldots, n$, we see that $(a_1 a_2 \cdots a_n) \equiv (b_1 b_2 \cdots b_{n+1}) = \prod_1^{n+1} b_i \equiv \prod_1^n a_i$ after applying $(n + 1)$-associativity to $(b_1 b_2 \cdots b_{n+1})$. ∎

Once one knows to avoid putting all elements of $\mathbf{G}$ into the multiplication table, examples of strictly n-associative groupoids are readily available. Here are several examples from [**17**]:

Example 1. Let $\mathbf{G}$ be the groupoid with multiplication table

$*$	1	2	3	4
1	2	3	4	4
2	4	4	4	4
3	4	4	4	4
4	4	4	4	4

$\mathbf{G}$ is 4-associative because the product of any four elements of $\mathbf{G}$ is 4. Since $(1 * 1) * 1 = 2 * 1 = 4 \neq 1 * (1 * 1) = 1 * 2 = 3$, $\mathbf{G}$ is not 3-associative. Hence, $\mathbf{G}$ is strictly 4-associative.

Example 2. Let **G** be the groupoid with multiplication table

$*$	1	2	3	4
1	2	4	4	4
2	4	3	4	4
3	4	4	4	4
4	4	4	4	4

G is 5-associative because the product of any five elements of **G** is 4. Since $(1 * 1) * (1 * 1) = 2 * 2 = 3 \neq ((1 * 1) * 1) * 1 = (2 * 1) * 1 = 4 * 1 = 4$, **G** is strictly 5-associative.

These examples readily generalize.

Example 3. Let $\mathbf{G} = (\{1, 2, \ldots, k\}, *)$ for $k \geq 4$ be the groupoid with binary operation $*$ given by

$$a * b = \begin{cases} b + 1 & \text{if } a = 1 \text{ and } b < k, \\ k & \text{otherwise.} \end{cases}$$

Then **G** is strictly k-associative.

Example 4. Let $\mathbf{G} = (\{1, 2, \ldots, k\}, *)$ for $k \geq 4$ be the groupoid with binary operation $*$ given by

$$a * b = \begin{cases} b + 1 & \text{if } a = b < k, \\ k & \text{otherwise.} \end{cases}$$

Then **G** is strictly $(2^{k-2} + 1)$-associative.

These examples show that Theorem 1 is a proper generalization of the general associative law. The interested reader is directed to [**17**], which examines the possible cardinalities of strictly n-associative groupoids.

REFERENCES

1. Michael Artin, *Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1991.
2. Wilfred Barnes, *Introduction to Abstract Algebra*, Heath, Boston, MA, 1963.
3. Birkhoff and MacLane, *A Survey of Modern Algebra*, revised edition, Macmillan, NY, 1961.
4. R. H. Bruck, *A Survey of Binary Systems*, third printing corrected, Springer, NY, 1971.
5. Claude Chevalley, *Fundamental Concepts of Algebra*, Academic Press, NY, 1956.
6. Gertrude Erlich, *Fundamental Concepts of Abstract Algebra*, PWS-Kent, Boston, MA, 1991.
7. Marshall Hall, Jr., *The Theory of Groups*, Macmillan, NY, 1959.
8. I. N. Herstein, *Topics in Algebra*, Blaisdell, Waltham, MA, 1964.
9. S. T. Hu, *Elements of Modern Algebra*, Holden-Day, San Francisco, 1965.
10. T. W. Hungerford, *Algebra*, Springer, NY, 1987.
11. Nathan Jacobson, *Lectures in Abstract Algebra*, vol. I, van Nostrand, NY, 1951.
12. Serge Lang, *Algebra*, Addison Wesley, Reading, MA, 1965.
13. Walter Ledermann, *Introduction to the Theory of Finite Groups*, Oliver and Boyd (Interscience) NY, 1961.
14. W. R. Scott, *Group Theory*, Dover, NY, 1987.
15. B. L. van der Waerden, *Modern Algebra*, vol. I, Ungar, NY, 1953.
16. E. A. Walker, *Introduction to Abstract Algebra*, Random House, NY, 1987.
17. W. P. Wardlaw, Finitely associative groupoids and algebras, *Houston Journal of Mathematics* **9**:4, 1983.
18. Zariski and Samuel, *Commutative Algebra*, van Nostrand, NY, 1958.
19. Hans Zassenhaus, *The Theory of Groups*, 2ed, Chelsea, NY, 1958.

An Application of the Marriage Lemma

ANDREW LENARD
Indiana University
Bloomington, IN 47405-7106

An important proposition in graph theory is a celebrated theorem of König and Hall, often referred to as the *Marriage Lemma* [**1**]. Let us briefly recall what this theorem is about and why it deserves this name. By a *graph* one means a set (for our purposes always finite) whose elements are referred to as *vertices*, together with a certain distinguished set of unordered pairs of vertices, referred to as *edges*.

A graph is called *bipartite* if the set of vertices is partitioned into two parts so that no edge connects two vertices in the same part. These two parts shall be denoted by G and B, having the application in mind that G is a set of girls and B a set of boys; the elements will be called g-vertices and b-vertices (girls, respectively boys). In this application an edge between vertices g and b has the meaning that boy b is acceptable to girl g as a future marriage partner. In general, for a given bipartite graph we say that G can be *matched* into B, if it is possible to select for each g-vertex *one* edge incident with it such that no two edges thus selected terminate on the same b-vertex. Thus, a matching in the application indicates that it is possible to arrange that each girl shall be married to a boy acceptable to her, without thereby any bigamy being created.

A matching for a bipartite graph may or may not be possible, but if it is, then for every subset F of G the total number of b-vertices that are connected to some g-vertices in F cannot be less than the number of g-vertices in F. This much is obvious. But what is far from obvious is that this condition is also sufficient; that constitutes the content of:

THEOREM 1. (MARRIAGE LEMMA.) *If in a bipartite graph the condition* $|F| \leq |\{b \in B : b \text{ connected by an edge to some } g \in F\}|$ *holds for all subsets* $F \subseteq G$, *then a matching of the set* G *of* g*-vertices into the set* B *of* b*-vertices is possible.*

Here and in the following we use the notation that $|\cdot|$ is the number of elements of the set indicated between the bars.

We shall not prove the Marriage Lemma here; proofs are found in many books on graph theory, for instance in *The Theory of Graphs and its Applications* by Claude Berge [**1**]. What we shall prove, however, is a stronger sufficient (though not necessary) condition for the existence of a matching with the sufficiency of this condition following from the Marriage Lemma itself.

THEOREM 2. *If in a bipartite graph there is a number* k *such that*

$$|\{\text{edges incident on vertex } g\}| \geq k \text{ for all } g \in G$$

and

$$|\{\text{edges incident on vertex } b\}| \leq k \text{ for all } b \in B,$$

then a matching of the set G *of* g*-vertices into the set* B *of* b*-vertices is possible.*

Proof. Let F be any subset of the set G of g-vertices in the graph, and let A be the set of b-vertices that are connected by an edge of the graph to some g-vertex in F. Let m be the total number of edges incident on vertices $g \in F$, so by hypothesis $m \geq k|F|$. But m is then also equal to the total number of edges incident on vertices $b \in A$, hence

$m \leq k|A|$ as well. The last two inequalities imply that $|F| \leq |A|$. Thus the condition in the Marriage Lemma holds, and so a desired matching is possible. ■

Theorem 2 is *strictly* weaker than Theorem 1, since a matching may also be possible if the hypotheses of Theorem 2 are not satisfied. The simplest example is provided by taking two girls g_1 and g_2, two boys b_1 and b_2, and letting g_1 be connected by edges to both boys b_1 and b_2, while g_2 is connected only to b_2. Here the boy b_2 is connected to two girls, whereas the girl g_2 only to one boy, so no number k of the kind required for the theorem exists. Yet a matching is obviously possible: Marry g_1 to b_1, g_2 to b_2!

The purpose of this article is to apply Theorem 2 in a problem that concerns sequences formed from a finite set of symbols ("letters") and using each of the letters once. For the sake of brevity, we shall call such sequences *arrangements*.

We begin with a simple illustrative example. Consider the following set of six arrangements formed from the four letters A, B, C, and D:

A B C D	A C D B	A D B C
B C A D	D B A C	C D B A

This list of six arrangements, chosen from the total number $4! = 24$ possible, has an interesting property: All 15 non-empty *subsets* of the set $\{A, B, C, D\}$, when each is suitably ordered, can be seen as forming an initial segment of at least one of them! (Remember that the concept of set, in contrast to that of sequence, does not involve the order of the elements; for instance, the subset $\{B, A, D\}$ must be ordered as $A\ D\ B$ in order to occur as the initial segment of the arrangement $A\ D\ B\ C$ (third on our list).)

It is easy to see that a list of arrangements of four letters that has this property cannot have a length less than six. For $\{A, B, C, D\}$ has 6 two-element subsets, and a little thought shows that each of these has to occur in a different arrangement on the list.

Consider now the corresponding problem for arrangements formed from a set of n letters. The above reasoning shows that if such a list has the property that each of the $2^n - 1$ nonempty subsets occurs as an initial segment of one of the arrangements on the list, then the list must have a length no less than the binomial coefficient that occurs along the central line of symmetry of the Pascal Triangle, namely

$$\nu_n = \binom{n}{m},$$

where $m = \lfloor \frac{n+1}{2} \rfloor$ (the notation $\lfloor x \rfloor$ stands for the integer part of x, that is, the largest integer not exceeding it). The reason is that there are $\binom{n}{k}$ subsets of size k, and each of these must form the initial segment of a different sequence on the list. The maximum of $\binom{n}{k}$ for various k occurs for $k = m$. The question still remains whether such a list of ν_n sequences exists. In other words, it is an obviously necessary condition for such a list of arrangements that its length be no shorter than ν_n; but is that also sufficient? We shall see that the answer is affirmative.

The author is conscious of a piece of wisdom he learned from his late friend George Minty. Minty once exclaimed, "Many of the most beautiful theorems of mathematics are of the form: Such and such a necessary condition is also sufficient. The necessity is frequently obvious or at least easy to see, but to establish the sufficiency is the real trick."

THEOREM 3. *Given a set of n letters, there exists a collection C of ν_n arrangements of these letters, with the property that each subset of the set of letters, when suitably ordered, occurs as the initial segment of some arrangement in the collection C.*

The first few values of ν_n are shown in the table below

n	1	2	3	4	5	6	7	8	$\cdots$
ν_n	1	2	3	6	10	20	35	70	$\cdots$

These numbers are generated recursively by the equations

$$\nu_{2n} = 2\,\nu_{2n-1}$$
$$\nu_{2n+1} = \frac{2n+1}{n+1}\,\nu_{2n}.$$

The asymptotic behavior of ν_n for large n is obtained from Stirling's Formula

$$n! \sim (2\pi n)^{1/2} n^n e^{-n}, \qquad (\text{as } n \to \infty),$$

where the $\sim$ symbol means that the limit of the ratio of the functions of n on the two sides of it tends to 1 in the indicated limit. Stirling's Formula is a celebrated result that should appear in all current calculus textbooks but, apparently, it does not, due perhaps to the relatively advanced analysis needed to identify the exact numerical proportionality factor $(2\pi)^{1/2}$. (For a simple proof without the exact factor see Feller [**2**, Section II.9]) From it, and the explicit expression

$$\nu_n = \frac{n!}{m!(n-m)!},$$

one obtains

$$\nu_n \sim \left(\frac{2}{\pi\, n}\right)^{1/2} 2^n, \qquad (\text{as } n \to \infty).$$

The interesting feature of this asymptotic formula is that when n is large then ν_n is small compared the total number 2^n of subsets of the set of n letters. So while the list of arrangements contemplated is relatively short, it accommodates the much more numerous collection of subsets, in the manner specified.

Before proving Theorem 3 we need to consider a combinatorial proposition proved by means of Theorem 2.

THEOREM 4. *Suppose $\mathcal{S}$ is a set with n elements. If $k \le \lfloor \frac{n+1}{2} \rfloor$ then from each k-element subset of $\mathcal{S}$, it is possible to remove one element in such a manner that, in the collection of the remaining subsets, each $k-1$-element subset of $\mathcal{S}$ occurs at least once.*

This is illustrated in the case $n = 4$ and $k = 2$, by removing from the six 2-element subsets of $\{A, B, C, D\}$, namely

$$\{A, \underline{B}\},\ \{A, \underline{C}\},\ \{A, \underline{D}\},\ \{B, \underline{C}\},\ \{\underline{B}, D\},\ \{C, \underline{D}\},$$

the respective elements indicated by underlining. In the remaining collection of 1-element subsets $\{A\}$ occurs three times, while each of $\{B\}$, $\{C\}$ and $\{D\}$ occurs once.

Proof of Theorem 4. The condition on k is evidently necessary, since only then is $\binom{n}{k-1} \le \binom{n}{k}$. We must show its sufficiency. Here is where Theorem 2 is helpful. Define a bipartite graph whose g-vertices are the $(k-1)$-element subsets of the given set and whose b-vertices are the k-element subsets. Two such subsets are, by definition, to be connected by an edge if and only if the smaller is a subset of the larger.

We remind the reader that here the vertices, normally visualized as *points*, are actually *sets*. This is a good illustration of the power of abstraction in mathematics, for in graph theory the nature of vertices is immaterial, only our specification of which pairs of them share an edge is relevant.

Each $(k-1)$-element subset is contained in precisely $n-(k-1)=n-k+1$ k-element subsets, in other words, each g-vertex is incident with $n-k+1$ edges. Similarly, each k-element subset gives rise to precisely k subsets of size $(k-1)$, obtained by deleting each of its elements. Thus the number of edges incident with one b-vertex is k. But since $k \leq \lfloor \frac{n+1}{2} \rfloor \leq \frac{n+1}{2}$ implies $k \leq n-k+1$, the condition of Theorem 2 is satisfied. Hence, a matching of the g-vertices into the b-vertices is possible. In the present case, this means an assignment to *each* $(k-1)$-element subset a k-element subset of which it itself is a subset, but in such a manner that the same k-element subset is not assigned to two distinct $(k-1)$-element subsets. From each of these k-element subsets, delete the element so as to form the assigned $(k-1)$-element subset. From the rest of the k-element subsets delete any one element arbitrarily, as you please. This shows that the deletion described in the theorem is indeed possible. ■

Now we are ready to prove Theorem 3, that is to say, to construct a list of ν_n permutations formed from the given n letters, with the required property.

It should be noted though, that perhaps "construct" is not exactly the right word. For use will be made of Theorem 4, which was proved by means of Theorem 2. And the latter depends on the Marriage Lemma, a statement that only asserts the *existence* of a matching, without really specifying how one finds it. Given any bipartite graph, though, in which a matching is known to exists, one can actually find one—if worst comes to worst—by trial and error, in as much there are altogether only a finite number of mappings (regardless of whether they are matchings or not) of the set of g-vertices into the set of b-vertices! This is of course a very inefficient procedure, and there are better ones. In fact, when one examines a suitable proof of the Marriage Lemma, it itself suggests a route to actually finding a matching. The present author learned this too from his late friend George Minty who used to love to talk about the Marriage Lemma, a particular favorite of his.

Proof of Theorem 3. We begin by writing down all subsets of m elements of the given set of n elements. Next, we remove one element from each of these subsets so that the remaining subsets, of $m-1$ elements each, exhaust all subsets of that size. This is possible on account of Theorem 4. The element of each subset, distinguished by having been removed in this step, is made to be the m^{th} member of the arrangements to be formed from elements of the subset in question. Repeating this process, from the remaining subsets of $m-1$ elements each, we again remove one element, in accordance with the requirement of Theorem 3, so that the remaining subsets, of $m-2$ elements each, again exhaust all subsets of that size; and then take the elements removed in this step to become the $(m-1)^{\text{st}}$ of the arrangement formed from elements of the subset in question.

The process continues the same way until the arrangements of m members each have been constructed. Clearly, the list so produced has the property that if $k \leq m$ then every subset of k elements is found as an initial segment of some of the listed arrangements. Having accomplished this, we go through the same process with the complementary subsets of $n-m$ elements each, but this time construct the required arrangements in the opposite order, that is to say, proceeding from the left to right (instead of right to left, as before). When this is done then we simply concatenate each arrangement obtained from a subset of the given set in the first process with the arrangement obtained from the complementary subset in the second process.

The complete list of sequences, of length n each, then has the required property.

■

To understand this procedure better it is useful to see a concrete example. We illustrate the process on the set $\{A, B, C, D, E\}$ of five letters. In each step of the process, the distinguished ("removed") element will be underlined. We begin then with the list of $\binom{5}{3} = 10$ subsets of $\lfloor \frac{5+1}{2} \rfloor = 3$ elements:

{A,B,C}	{A,B,D}	{A,B,E}	{A,C,D}	{A,C,E}
{A,D,E}	{B,C,D}	{B,C,E}	{B,D,E}	{C,D,E}

The reader is asked to check that the property required by Theorem 4 holds. After having removed the underlined letters (which will become the 3rd members of the sequences to be constructed), we deal in a similar way with remaining sets of 2 elements each:

{A,B}	{A,D}	{A,E}	{A,C}	{C,E}
{D,C}	{B,C}	{B,E}	{B,D}	{C,D}

The letters underlined here become 2nd members of the sequences; and of course the remaining letters will be 1st members. The list of the ten arrangements appears then as follows:

A B C	A D B	E A B	A C D	C E A
D E A	B C D	B E C	B D E	C D E

Again, please check that all subsets of $\{A, B, C, D, E\}$ of size no more than 3 elements occur as some initial segment of the arrangements displayed above—as they have to.

According to plan, next we work on the subsets of 2 elements each, complementary to the previous ones. This gives rise to the following:

{D,E}	{C,E}	{C,D}	{B, E}	{B,D}
{B,C}	{A,E}	{A,D}	{A,C}	{A,B}

With the distinguished element put into 1st place one has then the 10 arrangements of length 2:

E D	E C	D C	B E	D B
C B	E A	A D	A C	A B

Concatenating each of the arrangements of 3 letters, obtained previously, with the corresponding arrangement of the complementary 2 letters, as above, we obtain then the final result in the form of a list of 10 arrangements of 5 letters each:

A B C E D	A D B E C	E A B D C	A C D B E	C E A D B
D E A C B	B C D E A	B E C A D	B D E A C	C D E A B

This list has the property required by Theorem 3.

Acknowledgment. The author wishes to dedicate this article to the memory of his late colleague and good friend George Minty, at one time Professor of Mathematics in Indiana University. His infectious enthusiasm for any subject that caught his fancy and his willingness to explain all aspects of it, contributed mightily to the mathematical education of the present writer. Minty is not forgotten.

REFERENCES

1. Claude Berge, *The Theory of Graphs and its Applications*, John Wiley, New York, 1962.
2. William Feller, *An Introduction to Probability Theory and Its Applications*, 2nd Edition, Wiley, New York, 1957.

PROBLEMS

ELGIN H. JOHNSTON, *Editor*
Iowa State University

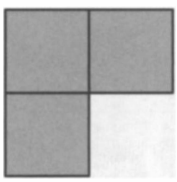
Assistant Editors: RAZVAN GELCA, Texas Tech University; ROBERT GREGORAC, Iowa State University; GERALD HEUER, Concordia College; PAUL ZEITZ, The University of San Francisco

Proposals

To be considered for publication, solutions should be received by November 1, 2001.

1623. *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, NY.*

Find the number of ways that k copies of the tromino



can be placed, with the orientation shown and without overlapping, on a $3 \times n$ rectangle.

1624. *Proposed by Murray S. Klamkin, The University of Alberta, Edmonton, AB, Canada.*

An ellipsoid is tangent to each of the six edges of a tetrahedron. Prove that the three segments joining the points of tangency of opposite edges are concurrent.

1625. *Proposed by Mihàly Bencze, Romania.*

Let $x_1, x_2, \ldots, x_n$ be positive real numbers and let $a_1, a_2, \ldots, a_n$ be positive integers. Prove that

$$\prod_{k=1}^{n}\left(1+x_k^{1/a_k}\right)^{a_k} \geq \left(1+\left(\prod_{k=1}^{n} x_k\right)^{1/\sum_{k=1}^{n} a_k}\right)^{\sum_{k=1}^{n} a_k}.$$

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution.

Solutions should be written in a style appropriate for this MAGAZINE. Each solution should begin on a separate sheet.

Solutions and new proposals should be mailed to Elgin Johnston, Problems Editor, Department of Mathematics, Iowa State University, Ames IA 50011, or mailed electronically (ideally as a LaTeX file) to ehjohnst@iastate.edu. All communications should include the readers name, full address, and an e-mail address and/or FAX number.

1626. *Proposed by Hojoo Lee, student, Kwangwoon University, Seoul, South Korea.*

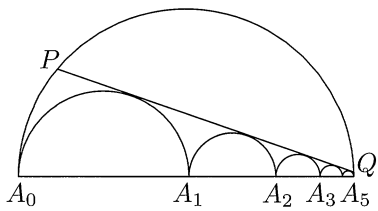
Let f, g, $h : \mathbb{R} \to \mathbb{R}$ be functions such that $f(g(0)) = g(f(0)) = h(f(0)) = 0$ and

$$f(x + g(y)) = g(h(f(x))) + y$$

for all $x, y \in \mathbb{R}$. Prove that $h = f$ and that $g(x + y) = g(x) + g(y)$ for all $x, y \in \mathbb{R}$.

1627. *Proposed by Jiro Fukuta, Shinsei-cho, Gifu-ken, Japan.*

Semicircle C has diameter A_0A_n. Semicircles $C_1, C_2, \ldots, C_n$ are drawn so that C_k has diameter $A_{k-1}A_k$ on A_0A_n. In addition, C_1 is internally tangent to C at A_0 and externally tangent to C_2 at A_1, C_n is internally tangent to C at A_n and externally tangent to C_{n-1} at A_{n-1}, for $2 \le k \le n-1$, C_k is externally tangent to C_{k-1} and C_{k+1} at A_{k-1} and A_k respectively, and each C_k, $1 \le k \le n$ is tangent to a chord PQ of C. The case $n = 5$ is illustrated in the accompanying figure.



(a) Let A_1A_1' and $A_{n-1}A_{n-1}'$ be perpendicular to A_0A_n at A_1 and A_{n-1}, respectively. Let circle X be externally tangent to C_2, internally tangent to C and tangent to A_1A_1' on the side opposite C_1, and let circle Y be externally tangent to C_{n-1}, internally tangent to C and tangent to $A_{n-1}A_{n-1}'$ on the side opposite C_n. Prove that circle X is congruent to circle Y.

(b) Suppose C_0 is a semicircle with diameter on A_0A_n and tangent to PQ. Let D and E be the endpoints of its diameter. Lines DD' and EE' are drawn perpendicular to A_0A_n. Let Z be the circle tangent to each of DD' and EE' and internally tangent to C. Show that Z is tangent to the circle with diameter A_1A_{n-1}.

Quickies

Answers to the Quickies are on page 246.

Q911. *Proposed by Murray S. Klamkin, The University of Alberta, Edmonton, AB, Canada.*

Two points P and Q are on opposite sides of a given plane in $\mathbb{R}^3$. Describe how to determine a point R in the plane so that $|PR - QR|$ is maximal.

Q912. *Proposed by David W. Carter, Draper Laboratory, Cambridge, MA.*

Let A, B, and C be distinct, non-collinear points in the plane. In which direction should we move B to obtain the largest rate of increase in $\angle ABC$?

Solutions

The Fall Is Steeper than the Rise **June 2000**

1599. *Proposed by Ice B. Risteski, Skopje, Macedonia.*

Given $\alpha > \beta > 0$ and $f(x) = x^\alpha(1-x)^\beta$. If $0 < a < b < 1$ and $f(a) = f(b)$, show that $f'(a) < -f'(b)$.

(I) *Solution by Reza Akhlaghi, Prestonsburg Community College, Prestonsburg, KY, and Fary Sami, Harford Community College, Bel Air, MD.*

Using

$$a^\alpha(1-a)^\beta = b^\alpha(1-b)^\beta, \tag{1}$$

we have

$$f'(a) + f'(b) = \frac{\beta}{b}a^\alpha(1-a)^\beta\left[\frac{\alpha}{\beta}\left(1+\frac{b}{a}\right) - \frac{b}{1-a}\left(1+\frac{1-a}{1-b}\right)\right]. \tag{2}$$

Let $\frac{\alpha}{\beta} = r > 1$ and $\frac{b}{a} = t > 1$. From (1) we then have $\frac{1-a}{1-b} = t^r$. Combining this equation with $b = at$ we can solve for a and b in terms of r and t to get

$$a = \frac{t^r - 1}{t^{r+1} - 1} \qquad \text{and} \qquad b = \frac{t(t^r-1)}{t^{r+1}-1}.$$

Substituting these expressions inside the square brackets in (2), then simplifying, we obtain

$$f'(a) + f'(b) = -\frac{\beta}{b}\frac{a^\alpha(1-a)^\beta}{t^{r-1}(t-1)}\left[t^{2r} - rt^{r+1} + rt^{r-1} - 1\right]. \tag{3}$$

Let $g(t) = t^{2r} - rt^{r+1} + rt^{r-1} - 1$, so

$$g'(t) = rt^{r-2}\left(2t^{r+1} - (r+1)t^2 + (r-1)\right) = rt^{r-2}h(t),$$

with $h(t) = 2t^{r+1} - (r+1)t^2 + (r-1)$. Because $h'(t) = 2(r+1)(t^r - t)$ is positive for $t > 1$, it follows that $h(t) > h(1) = 0$ for $t > 1$, and then that $g(t) > g(1) = 0$ for $t > 1$. Hence, from (3) we have $f'(a) + f'(b) < 0$, so $f'(a) < -f'(b)$.

(II) *Solution by the proposer.*

Let $\alpha/\beta = r$ and $b/a = t$. The equality $f(a) = f(b)$ then implies that $\frac{1-a}{1-b} = t^r$. Next note that the inequality $f'(a) < -f'(b)$ is equivalent to

$$\alpha\left(\frac{1}{a} + \frac{1}{b}\right) < \beta\left(\frac{1}{1-a} + \frac{1}{1-b}\right).$$

This can be rewritten as

$$\frac{\alpha}{\sqrt{ab}}\frac{t - \frac{1}{t}}{\sqrt{t} - \frac{1}{\sqrt{t}}} < \frac{\beta}{\sqrt{(1-a)(1-b)}}\frac{t^r - \frac{1}{t^r}}{\sqrt{t^r} - \frac{1}{\sqrt{t^r}}},$$

which reduces to

$$t - t^{-1} < \frac{1}{r}\left(t^r - t^{-r}\right). \tag{1}$$

Now let

$$g(x) = \frac{2\sinh(sx)}{x},$$

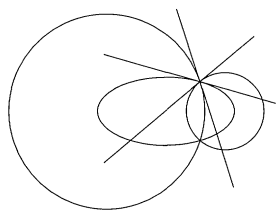
where $s = \ln t > 0$. It is easy to check that g is increasing on the interval $x > 0$. The inequality (1) then says $g(1) < g(r)$, which is true because $r > 1$.

Also solved by Jean Bogaert (Belgium), Con Amore Problem Group (Denmark), R. Daniel Hurwitz. There were two incorrect results submitted.

Tangent to Conic as Angle Bisector **June 2000**

1600. *Proposed by Juan-Bosco Romero Márquez, Universidad de Valladolid, Valladolid, Spain.*

Let $\mathcal{C}$ be either an ellipse or a hyperbola. For P a point on $\mathcal{C}$, prove that the tangent line to $\mathcal{C}$ at P bisects one of the angles formed by the tangents to the circles passing through P with centers at the foci of $\mathcal{C}$.



Solution by Michael Woltermann, Washington and Jefferson College, Washington, PA.

Let F_1 and F_2 be the foci of the conic. By the reflexive properties of the ellipse or hyperbola, the tangent at P makes equal angles with the focal radii F_1P and F_2P. It follows that the tangent and normal lines to the conic at P bisect the vertical angles formed by lines F_1P and F_2P, and thus also bisect the vertical angles formed by the perpendiculars to F_1P and F_2P at P. These perpendiculars are the tangents at P to the circles passing through P and centered at the foci of the conic.

Also solved by Reza Akhlaghi and Fary Sami, Henry J. Barten, Michel Bataille (France), J. C. Binz (Switzerland), Jean Bogaert (Belgium), Michael Brozinsky, Con Amore Problem Group (Denmark), Ragner Dybvik (Norway), Matt Foss, Ming-Lun Ho, Hans Kappus (Switzerland), Robert Mandl, José H. Nieto (Venezuela), Raul A. Simon (Chile), Peter Y. Woo, Li Zhou, and the proposer.

Regular Polygons on a Circle **June 2000**

1601. *Proposed by John Clough, State University of New York at Buffalo, Buffalo, NY; Jack Douthett, TVI Community College, Albuquerque, NM; and Roger Entringer, University of New Mexico, Albuquerque, NM.*

Let a and b be positive integers. Place a white points on a circle so that they form the vertices of a regular a-gon. Place b black points on the same circle so that they form the vertices of a regular b-gon and so that white and black points are distinct. Beginning with a black point whose clockwise distance from the nearest white point is a minimum, and proceeding clockwise, label the points with the integers 0 through $a + b - 1$. Prove that the black points have labels $\lfloor k(a+b)/b \rfloor$, $k = 0, 1, \ldots, b-1$, and that the white points have labels $\lceil k(a+b)/a \rceil - 1$, $k = 1, 2, \ldots, a$.

Solution by José H. Nieto, Universidad del Zulia, Maracaibo, Venezuela.

Let $\ell(P)$ denote the label assigned to point P. Let $B_0, B_1, \ldots, B_{b-1}$ be the black points, numbered clockwise, so that $\ell(B_0) = 0$. Let A_a be the white point from which the clockwise distance to B_0 is minimal, and number clockwise the other white points as $A_1, A_2, \ldots, A_{a-1}$. If we rotate clockwise the white points until A_a reaches B_0, then no white point passes over a black point. The label assigned to the kth black point must be $k + j$ where j is the number of rotated white points in arc $(B_0, B_k]$. Because the points A_r, B_s each are vertices of a regular polygon, it follows that j is the greatest integer such that $jc/a \le kc/b$, where c is the circumference of the circle. Hence $j = \lfloor ka/b \rfloor$ and $\ell(B_k) = k + \lfloor ka/b \rfloor = \lfloor k(a+b)/b \rfloor$. Next, observe that the minimal *counterclockwise* distance from a black point to a white point is attained from B_0 to A_a. Hence we may use the previous argument to calculate the labels for the white points, proceeding counterclockwise and interchanging the roles of of the black and the white points. More precisely, put $B'_i = A_{a-i}$ for $i = 0, 1, \ldots, a-1$ and $A'_j = B_{b-j}$ for $j = 1, 2, \ldots, b$, and let $\ell'(P')$ be the label assigned to P' in the counterclockwise numbering. Note that $\ell'(P) = a + b - 1 - \ell(P)$. We then have $\ell'(B'_k) = \lfloor k(a+b)/a \rfloor$, and hence,

$$\begin{aligned}\ell(A_k) = a + b - 1 - \ell'(B'_{a-k}) &= a + b - 1 - \lfloor (a-k)(a+b)/a \rfloor \\ &= -1 - \lfloor -k(a+b)/a \rfloor = \lceil k(a+b)/a \rceil - 1.\end{aligned}$$

Also solved by Hamza Ahmad and Nancy Colwell, Jean Bogaert (Belgium), Richard F. McCoart, Jr., and the proposers.

Points in a Prism **June 2000**

1602. *Proposed by Michael Golomb, Purdue University, West Lafayette, IN.*

Suppose S is a bounded set of points in $\mathbb{R}^n$, $n \ge 2$, such that every n-simplex whose vertices are points of S has volume at most V. Prove that there exists an n-prism of volume at most $2n(n-1)^{n-1}V$ that contains S.

(An n-prism is an n-dimensional solid bounded by two parallel hyperplanes and a finite number of hyperplanes, each of which contains a line parallel to a fixed line intersecting these two hyperplanes.)

Solution by the proposer.

By considering the closure of S, if necessary, we may assume that S is compact. We may also assume that S is not contained in an $(n-1)$-dimensional subset of $\mathbb{R}^n$ because such a set is included in a sufficiently large $(n-1)$-cube of n-volume 0.

We first consider the case $n = 2$. Let m be the maximal length of a segment connecting two points of S, and let ℓ be such a segment. At each end of ℓ draw a segment of length $4V/m$ perpendicular to ℓ and bisected by the endpoint of ℓ. These two segments are two sides of an m by $4V/m$ rectangle that contains S.

Now assume that $n \ge 3$. Let σ_0 be an $(n-1)$-simplex of maximum $(n-1)$-volume m among those whose vertices are points of S, and let H_0 be the hyperplane containing σ_0. Any n-simplex with base σ_0 and vertex $p \in S$ has volume not exceeding V if and only if $\operatorname{dist}(p, H_0) \le nV/m$. Therefore S is a subset of the region of $\mathbb{R}^n$ between H_+ and H_-, where H_+ and H_- are hyperplanes parallel to H_0 at a distance nV/m on either side of H_0.

Let v_i, $i = 1, 2, \ldots, n$ denote the vertices of σ_0, and let f_i be the $(n-2)$-simplicial face of σ_0 that is opposite vertex v_i. An $(n-1)$-simplex with base f_i and vertex $p \in S \cap H_0$ has $(n-1)$-volume not exceeding m only if p lies in the half-space of H_0 that contains f_i and is bounded by the $(n-2)$-dimensional plane h_i that is parallel to f_i and passes through v_i. Thus $S \cap H_0$ is a subset of the $(n-1)$-simplex $\overline{\sigma}_0$ whose faces

lie in the planes $h_1, h_2, \ldots, h_n$. In the lemma below we prove that

$$\text{vol}(\overline{\sigma}_0) = (n-1)^{n-1}\,\text{vol}(\sigma_0) = (n-1)^{n-1}m,$$

where the volumes here are $(n-1)$-volumes.

Let H_i, $i = 1, 2, \ldots, n$ be the hyperplane that contains h_i and is orthogonal to H_0. We now define the prism π that contains S. Its lateral faces lie in the hyperplanes $H_1, H_2, \ldots, H_n$, and its base faces lie in the hyperplanes H_+ and H_-. Its mid-cross section, orthogonal to its axis, is the $(n-1)$-simplex $\overline{\sigma}_0$. The length of the axis is $2nV/m$ and the $(n-1)$-volume of its base is $(n-1)^{n-1}m$. It follows that the n-volume of the prism is $2n(n-1)^{n-1}V$, as desired. We must show that $S \subset \pi$.

Suppose that q is a point of $\mathbb{R}^n$ that is not in π. Then either q is not in the region of $\mathbb{R}^n$ between H_+ and H_-, or q is on the side of some H_j that does not contain $\overline{\sigma}_0$. In the former case, $q \notin S$, as shown above. In the latter case, let q_0 be the orthogonal projection of q onto H_0. The $(n-1)$-simplex σ_j with base f_j and vertex q has $(n-1)$-volume at least as large as that of the $(n-1)$-simplex σ_{j_0} with base f_j and vertex q_0. But since q_0 is on the side of h_j which does not include $\overline{\sigma}_0$, we have

$$\text{vol}(\sigma_{j_0}) > \text{vol}(\sigma_0) = m,$$

where the volumes here are $(n-1)$-volumes. It follows that the $(n-1)$-volume of σ_j exceeds m, so $q \notin S$.

It remains to prove the following lemma.

LEMMA. Let $\overline{\sigma}$ be the simplex whose faces are parallel to the faces of a simplex σ and pass through the vertex of σ that is opposite the parallel face. Then

$$\text{vol}(\overline{\sigma}) = d^d\,\text{vol}(\sigma),$$

where d is the dimension of σ.

Proof. Because the ratio of volumes is an affine invariant, we may assume that the $d+1$ vertices of σ are the origin and a point on each of the positive axes of a Cartesian coordinate system. Thus the faces of σ lie in the hyperplanes $x_k = 0, \quad k = 1, 2, \ldots, d$, and $\sum_{k=1}^{d} \frac{x_k}{a_k} = 1$, where the a_k are positive numbers. It is well known that $\text{vol}(\sigma) = \frac{1}{d!}\Pi a_k$. The faces of $\overline{\sigma}$ lie in the hyperplanes $x_k - a_k = 0, \quad k = 1, 2, \ldots, d$, and $\sum_{k=1}^{d} \frac{x_k}{a_k} = 0$. Set $y_k = x_k - a_k, k = 1, 2, \ldots, d$. Then $\overline{\sigma}$ may be considered to be a simplex in the y-system with faces lying in the hyperplanes $y_k = 0, \quad k = 1, 2, \ldots, d$, and $\sum_{k=1}^{d} \frac{y_k}{b_k} = 1$, where $b_k = -da_k$. It follows that

$$\text{vol}(\overline{\sigma}) = \frac{1}{d!}\Pi_{k=1}^{d}|b_k| = d^d\frac{1}{d!}\Pi_{k=1}^{d}a_k = d^d\,\text{vol}(\sigma).$$

Bounded Solution to a Differential Equation **June 2000**

1577. *Proposed by Philip Korman, University of Cincinnati, Cincinnati, OH.*

Consider the differential equation $x''(t) + a(t)x^3(t) = 0$ on $0 \le t < \infty$, where $a(t)$ is continuously differentiable and $a(t) \ge \kappa > 0$.

(a) If $a'(t)$ has only finitely many changes of sign, prove that any solution $x(t)$ is bounded.

(b) If one does not assume that $a'(t)$ has only finitely many sign changes, is $x(t)$ necessarily bounded?

Solution by the proposer.

(a) Define the "energy" function E by

$$E(t) = \frac{1}{2}\left(x'(t)\right)^2 + a(t)\frac{(x(t))^4}{4}. \tag{1}$$

Using the differential equation we find

$$E'(t) = a'(t)\frac{(x(t))^4}{4}. \tag{2}$$

If $a'(t) \le 0$ on an interval $[t_1, t_2]$, then $E'(t) \le 0$ on this interval. It follows that $E(t) \le E(t_1)$ for $t_1 \le t \le t_2$. If $a'(t) \ge 0$ on $[t_1, t_2]$, then from (1) and (2) we conclude that $E'(t) \le a'(t)\frac{E(t)}{a(t)}$ on this interval. Integrating this expression we find that

$$E(t) \le \frac{E(t_1)}{a(t_1)}a(t) \le \frac{a(t_2)}{a(t_1)}E(t_1), \qquad t_1 \le t \le t_2. \tag{3}$$

In particular, $E(t)$ can increase by at most a factor of $a(t_2)/a(t_1)$ on $[t_1, t_2]$. From (1) and (3) we also conclude that

$$\frac{(x(t))^4}{4} \le \frac{E(t_1)}{a(t_1)}, \qquad t_1 \le t \le t_2. \tag{4}$$

Now assume that $a'(t)$ changes sign at points $c_1, c_2, \ldots, c_n$. Because $E(t)$ is non-negative and non-increasing on any interval on which $a'(t) \le 0$, and increases by at most a factor $a(c_{k+1})/a(c_k)$ on any bounded interval $[c_k, c_{c+1}]$ on which $a'(t) \ge 0$, it follows that $E(t)$ remains bounded on $[0, c_n]$. If $a'(t) \le 0$ on $[c_n, \infty)$, then by (2), $E(t)$ is non-increasing on this interval, so remains bounded. This implies that $x(t)$ is bounded on $[0, \infty)$. If $a'(t) \ge 0$ on (c_n, ∞), then by (4), $x^4(t) \le 4E(c_n)/a(c_n)$ for $t \ge c_n$, again showing that $x(t)$ is bounded.

(b) The answer is no. One can construct $a(t)$ which will "pump up" the energy function $E(t)$, and consequently $x(t)$ will become unbounded. We outline the construction. We construct $a(t)$ depending on the behavior of the solution $x(t)$. Start with the initial conditions $x(0) = 1$ and $x'(0) = 1$, and set $a(t) = t$. By (2), the energy is increasing. Slightly before the time $t = 2$ we smoothly change $a(t)$ to a constant function $a(t) = 2$, and keep it constant for a while. This keeps the energy unchanged. It is well known that for constant $a(t)$ solutions of our equation move on closed curves around the origin in (x, x')-plane. Hence at some time $t_1 > 2$ we will have $x(t_1)$ small. Near t_1 we quickly but smoothly decrease $a(t)$ to $a(t) = 1$. This will result in a loss of energy which, by (2), is very small. We now keep $a(t) = 1$, until a time $t_2 > t_1$, at which $x(t)$ is as large possible for this energy level. This happens when $x'(t_2) = 0$. At this time we quickly and smoothly increase $a(t)$ to $a(t) = 2$. This will increase the energy considerably. We continue this process, which increases the energy without bound. Since $a(t) \le 2$, this implies that $x(t)$ becomes unbounded, in particular, at times t when $x'(t) = 0$.

Answers

Solutions to the Quickies from page 240.

A911. Let Q' be the reflection across the plane of Q, so $QR = Q'R$. By the triangle inequality, $|PR - Q'R| \leq PQ'$. The maximal value PQ' is achieved when R is the intersection of line PQ' with the plane. In the event that PQ' is parallel to the plane, the value PQ' is approached as R approaches the point at infinity in the plane that is in the direction of line PQ'.

A912. Let $\mathcal{C}$ denote the circle through A, B, and C. As B moves along $\mathcal{C}$, $\angle ABC$ is unchanged. Because the gradient of a function at a point is normal to the level curve through that point, it follows that the direction of most rapid increase of $\angle ABC$ is achieved when B moves towards the center of $\mathcal{C}$.

In the February 2001 issue, the following readers were inadvertently omitted from the lists of those who had submitted correct solutions:

Problem 1589: *Jean Bogaert (Belgium), Daniele Donini (Italy), Victor Kutsenok, Rick Mabry, José H. Nieto (Venezuela), Sang-il Oum (South Korea), Guo Zi Long (China), Jayavel Sounderpandian, Michael Vowe (Switzerland), David Zhu, and Paul J. Zwier*

Problem 1590: *Jean Bogaert (Belgium), John Christopher, Knut Dale (Norway), Daniele Donini (Italy), Victor Kutsenok, Akalu Tefera and Omer Yayenie, Michael Vowe (Switzerland), Westmont College Problem Solving Group, and David Zhu*

Problem 1591: *Brian D. Beasley, Joel D. Haywood, and David Zhu*

Problem 1592: *Jean Bogaert (Belgium), Daniele Donini (Italy), Victor Kutsenok, and Michael Vowe (Switzerland)*

Problem 1593: *Knut Dale (Norway)*

The editors apologize for the omission.

50 Years Ago in the MAGAZINE

Volume 24, No. 1 (January–February, 1951) included one in a continuing series of articles on "What Mathematics Means to Me." Here is an excerpt from E. T. Bell's offering, which included an elaborate apology for writing in the first person:

> Another thing I got from mathematics has meant more to me than I can say. No man [*sic*] who has not a decently skeptical mind can claim to be civilized. Euclid taught me that without assumptions there is no proof. Therefore, in any argument, examine the assumptions. Then, in the alleged proof, be alert for inexplicit assumptions. Euclid's notorious oversights drove this lesson home. Thanks to him, I am (I hope!) immune to all propaganda, including that of mathematics itself. . . .

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of mathematics literature. Readers are invited to suggest items for review to the editors.

Robinson, Sara, Why mathematicians now care about their hat color, *New York Times* (10 April 2001) D5 (Nat'l Ed.), F5 (City Ed.), `http://www.nytimes.com/2001/04/10/science/10MATH.html`. Ebert, Todd, The Colored Hats Puzzle, `http://www.ics.uci.edu/~ebert/teaching/spring2001/ics151/puzzles.html` and Solution to Colored Hats Puzzle, `http://www.ics.uci.edu/~ebert/coloredHatsSolution.html` . Rudich, Steven, Rudich, Steven, et al. The expressive power of voting polynomials (with J. Aspnes, R. Biegel, and M. Furst), `www.cs.cum.edu/~rudich/papers/voting.ps` .

Todd Ebert (Univ. of California–Irvine) has introduced three new twists on an old logic puzzle, connecting it to unsolved problems in coding theory. The traditional puzzle variously involves smudged foreheads, hats of two different colors, or unfaithful husbands; its roots go back 150 years or so (see 9.C and 9.D of *Sources in Recreational Mathematics: An Annotated Bibliography*, 6th prelim. ed., by David Singmaster, `david.singmaster@sbu.ac.uk`). The participants reason (perfectly) and sequentially determine their own situations from learning that others have figured out their own. Ebert's new ingredients are that the hat colors are determined by coin tosses, that the participants work as a team (including strategizing together in advance), and that without further communication all must simultaneously either guess their own situations or abstain. The team wins if at least one member guesses correctly and no member guesses incorrectly. The problem is to find a strategy that maximizes the team's chances of winning. Surprise! Mathematics turns out to be applicable, in the form of Hamming (and other) error-correcting codes. The players should order themselves in advance, so their hat colors become a binary sequence; their strategy is to hope that the sequence is not a codeword in the smallest Hamming code in which it embeds. In such codes, error words predominate over codewords and an error word can be made into a codeword in only one way, by changing a unique bit. Each player applies the relevant parity-check matrix to the two words that could represent the situation (one with a 1 for that player and the other with a 0); if the result of one is a codeword, the player guesses the other. If the true situation is an error word, only one player will guess, and will guess correctly; in the less likely event that it is a codeword, all will guess wrong. Rudich advises in a posting to the Usenet newsgroup `sci.math` that in 1987 he gave (in an unpublished paper) an optimal solution to the hat puzzle (couched as a voting puzzle) in terms of perfect codes. Meanwhile, other investigators are trying to devise codes that allow a large number of players to win even more frequently.

Houston, Ken (ed.), *Creators of Mathematics: The Irish Connection*, University College Dublin Press/Dufour Editions (Chester Springs, PA 19425–0007), 2001; x + 150 pp, $19.95 (P). ISBN 1–900621–49–5.

How many Irish mathematicians can you name? Hamilton, for sure. But don't forget Boole, Stokes, Thomas Harriot, William Thomson (Lord Kelvin), Edgeworth, Reynolds, Whittaker, and Gosset. The short sketches in this book commissioned by the Irish Royal Academy set out biographical elements of their lives and others'.

Suri, Manil, Adventures of a mathematician on a book tour, *Chronicle of Higher Education* (23 March 2001) B7–B10. Suri, Manil, *The Death of Vishnu*, Norton, 2001.

What do you do in your spare time? If you have ever wondered what life would be like if you had secreted yourself away for years—not to prove Fermat's Last Theorem, like Andrew Wiles, but to write a novel, and it became a big hit in 13 languages—this article (and perhaps mathematician Suri's novel about life in a Bombay apartment house) are for you.

Kirtland, Joseph, *Identification Numbers and Check Digit Schemes*, Mathematical Association of America, 2001; xi + 174 pp, $32.95 (P). ISBN 0–88385–720–0.

"This text is ideal for a liberal arts mathematics course." So states the author, who is right to commend the book to that use, since it will help students to see mathematics at work in the world around them. The book introduces modulo arithmetic and applies it to check-digit schemes for airline tickets, money orders, UPCs, and ISBNs; the concepts are then applied to the RSA cryptography system. To set the stage for hashing and more advanced schemes, a chapter is devoted to sets, functions, and permutations. Two others treat symmetries and group theory, to prepare for the Verhoeff scheme, which employs a dihedral group of permutations.

Stein, Sherman, *How the Other Half Thinks: Adventures in Mathematical Reasoning*, McGraw-Hill, 2001; xi + 165 pp, $18.95. ISBN 0–07–137339–X.

"The other half," refers to *us*(!), the mathematicians and scientists. Stein, whose classic *Mathematics: The Man-Made Universe* remains one of the best books to introduce liberal arts students to mathematics, here again tries to bridge the gap between the "Two Cultures" (humanities vs. sciences). "The alleged gap can be narrowed or completely overcome by anyone... the non-mathematical reader can go far in understanding mathematical reasoning." What is novel and surprising is that each of the eight chapters begins with a simple question about strings made up of two letters. The strings may be produced by chance (throws of a needle, games played until a team wins by two points, streaks and slumps, counting ballots), by labeling of dots, by trying not to repeat a triplet, by avoiding adjacent repetitions, or by just going on forever. Except for the last topic (which serves to introduce infinite sets and questions about them), the topics all have applications, although only some are elaborated. (The author's philosophy: "[W]hile my primary goal is to illustrate the mathematical way of thinking, if a particular result has applications, so much the better.") This is a delightful book that serves very well its intention to illustrate that "Mathematics can tempt all those who possess the spirit of adventure."

Chown, Marcus, The Omega man, *New Scientist* (10 March 2001) 28–31; `http://www.newscientist.com/features/features.jsp?id=ns22811` . Chaitin, G.J., *Exploring Randomness*, Springer-Verlag, 2001; 176 p, $34.95. ISBN 1–85233417–7.

Gregory Chaitin (IBM T.J. Watson Research Center) "has found that the core of mathematics is riddled with holes.... [T]here are an infinite number of mathematical facts but, for the most part, they are unrelated to each other and impossible to tie together with unifying theorems. If mathematicians find any connections, between these facts, they do so by luck." Moreover, according to Chaitin, "Most of mathematics is true for no particular reason. Maths is true by accident." Author Chown's immediate follow-up is that "this is partuclarly bad news for physicists," thereby embracing physicists among his readers but leaving mathematicians unaided in trying to get up off the floor. Chaitin's theory rests on the number Omega, the probability that a randomly selected program on a Turing machine will halt. Omega is uncomputable—even its first bit. Moreover, each bit of Omega also tells whether a particular member of a family of diophantine equations has finitely or infinitely many solutions; hence, the answer for each equation is unknowable and independent of the answer for the others. Says Chaitin, "Randomness is the true foundation of mathematics." Chown stikes a parting blow to mathematicians with "The discovery of Omega has exposed gaping holes in mathematics, making research in the field look like playing a lottery." Though generalizations of Omega have links to questions about real computers, computer scientists may be better at dissociating themselves from their objects of study (they don't feel responsible or offended if a machine fails to produce any output).

41st International Mathematical Olympiad

Taejon, Republic of Korea

July 19 and 20, 2000

edited by Titu Andreescu and Zuming Feng

Problems

1. Two circles ω_1 and ω_2 intersect at M and N. Line ℓ is tangent to the circles at A and B, respectively, so that M lies closer to ℓ than N. Line CD, with C on ω_1 and D on ω_2, is parallel to ℓ and passes through M. Let lines AC and BD meet at E; let lines AN and CD meet at P; and let lines BN and CD meet at Q. Prove that $EP = EQ$.
2. Let a, b, c be positive real numbers such that $abc = 1$. Prove that

$$(a - 1 + \frac{1}{b})(b - 1 + \frac{1}{c})(c - 1 + \frac{1}{a}) \leq 1.$$

3. Let $n \geq 2$ be a positive integer. Initially, there are n fleas on a horizontal line, not all at the same point. For a positive real number λ, define a *move* as follows:

   choose any two fleas, at points A and B, with A to the left of B; let the flea at A jump to the point C on the line to the right of B with $BC/AB = \lambda$.

   Determine all values of λ such that, for any point M on the line and any initial positions of the n fleas, there is a finite sequence of moves that will take all the fleas to positions to the right of M.
4. A magician has one hundred cards numbered 1 to 100. He puts them into three boxes, a red one, a white one and a blue one, so that each box contains at least one card. A member of the audience selects two of the three boxes, chooses one card from each and announces the sum of the numbers on the chosen cards. Given this sum, the magician identifies the box from which no card has been chosen. How many ways are there to put all the cards into the boxes so that this trick always works? (Two ways are considered different if at least one card is put into a different box.)
5. Determine if there exists a positive integer n such that n has exactly 2000 prime divisors and $2^n + 1$ is divisible by n.
6. Let $\overline{AH_1}$, $\overline{BH_2}$, and $\overline{CH_3}$ be the altitudes of an acute triangle ABC. The incircle ω of triangle ABC touches the sides BC, CA, and AB at T_1, T_2, and T_3, respectively. Consider the symmetric images of the lines H_1H_2, H_2H_3, and H_3H_1 with respect to the lines T_1T_2, T_2T_3, and T_3T_1. Prove that these images form a triangle whose vertices lie on ω.

Solutions

1. Let lines AB and MN meet at K. By the **Power of a point theorem**, $AK^2 = KN \cdot KM = BK^2$. Since $\overline{AB} \parallel \overline{PQ}$, $PM = QM$. Hence it suffices to prove that $\overline{EM} \perp \overline{PQ}$.

   Since $\overline{CD} \parallel \overline{AB}$, $\angle EAB = \angle ECM$. Since $\overline{AB}$ is tangent to the circle at A, $\angle BAM = \widehat{AM}/2 = \angle ACM$. Therefore $\angle EAB = \angle BAM$. Similarly $\angle EBA = \angle ABM$ and $\overline{AB}$ bisects both $\angle EAM$ and $\angle EBM$. Hence $AEBM$ is a kite and $\overline{EM} \perp AB$. Since $\overline{PQ} \parallel \overline{AB}$, we obtain $\overline{EM} \perp \overline{PQ}$, as desired.

2. Since $abc = 1$, this non-homogeneous inequality can be transformed into a homogeneous one by a suitable change of variables. In fact, there exist positive real numbers p, q, r such that $a = p/q, b = q/r, c = r/p$. Rewriting the inequality in terms of p, q, r, we obtain

$$(p - q + r)(q - r + p)(r - p + q) \leq pqr, \tag{1}$$

   where $p, q, r > 0$.

   At most one of the numbers $u = p - q + r, v = q - r + p, w = r - p + q$ is negative, because any two of them have a positive sum. If exactly one of the numbers u, v, w is negative, then $uvw \leq 0 < pqr$. If they are all nonnegative, then by the AM-GM inequality, $\sqrt{uv} \leq (u + v)/2 = p$. Likewise, $\sqrt{vw} \leq q$ and $\sqrt{wu} \leq r$. Hence $uvw \leq pqr$, as desired.

3. The answer is $\lambda \geq 1/(n - 1)$.

   (a) *If $\lambda \geq 1/(n - 1)$, then the fleas can all move to the right of M.*

   Assume without loss of generality that the fleas are all at distinct points; otherwise we can attain such an arrangement by repeatedly jumping the leftmost flea at any given time over the rightmost flea at that time. Let k be the original minimum distance between any two adjacent fleas and let D be the original distance from the leftmost position to M, the point we wish to pass.

   Originally, the leftmost flea **L** is at least $k(n - 1)$ distance away from the rightmost flea **R**; have **L** jump over **R**. Then **L** will land at least $k(n - 1) \cdot \lambda \geq k$ to the right of **R**. Thus we have moved the left side of our flea circus at least k distance to the right, while keeping the minimum distance between any two adjacent fleas at least k. Then after at most $\lceil \frac{D}{k} \rceil$ moves of this sort, all the fleas will be to the right of M, as desired.

   (b) *If $\lambda < 1/(n - 1)$, then the fleas cannot always all move to the right of M.*

   As the fleas jump, let $O = (F_1, F_2, \ldots, F_n)$, where the fleas (from left to right) are at points $F_1, F_2, \ldots, F_n$. Then let

$$P(O) = F_1F_n + F_2F_n + \cdots + F_{n-1}F_n.$$

   We claim that if any flea **F** jumps a distance d from its position, then P decreases by at least γd, where

$$\gamma = \frac{1 - (n - 1)\lambda}{1 + \lambda}.$$

   (Notice that γ is positive since $\lambda < 1/(n - 1)$, and that $\gamma < 1$ since $1 - (n - 1)\lambda < 1 < 1 + \lambda$.)

   If **F** lands on or to the left of F_n, then clearly P decreases by exactly $d > \gamma d$.

On the other hand, suppose **F** lands to the right of F_n. Let A be its starting position, B the point it jumps over, and C its landing position, so that A, B, F_n, C are in that order from left to right. Since the distance between each flea (besides **F**) and the rightmost flea increases by F_nC for an total increase of $(n-1)F_nC$; as for **F**, its distance decreases by AF_n since it *becomes* the rightmost flea. Therefore P changes by

$$\begin{aligned}(n-1)F_nC - AF_n &\le (n-1)BC - AB\\ &\le \lambda(n-1)AB - AB = AB[(n-1)\lambda - 1]\\ &= \frac{AC}{1+\lambda}[(n-1)\lambda - 1] = d\left(\frac{(n-1)\lambda - 1}{1+\lambda}\right) = -\gamma d.\end{aligned}$$

Thus indeed P decreases by at least γd in this case as well.

Now suppose we have a configuration O_0 of fleas where the leftmost flea $\mathbf{L}_0$ is at point F_1. Set $P_0 = P(O_0)$, and choose M to the right of F_1 such that $F_1M > P_0/\gamma$. Each time $\mathbf{L}_0$ jumps a distance d he decreases P by at least γd, so if it moves a total distance of D he decreases P by at least γD. Because P must always be nonnegative (since it is the sum of nonnegative distances), flea $\mathbf{L}_0$ can decrease P by at most P_0. Thus

$$P_0 \ge \gamma D \text{ and } D \le \frac{P_0}{\gamma} < F_1M.$$

It follows that $\mathbf{L}_0$ can never jump to the right of M. Therefore, when $\lambda < 1/(n-1)$, it is *not* always possible to make all the fleas move past M.

4. We first claim that 1 and 2 are in different colors. If not, say $1, 2, \ldots, i-1$ are in red, i is white, and j be smallest blue number. We have $i \ge 3$ and $j - 1 \ge i$. But in view of $i + (j-1) = (i-1) + j$, $j-1$ is white, which leads to the fact that the sum $2 + (j-1) = 1 + j$ does not allow the magician to decide on the unpicked box.

   Now let 1 be red, 2 be white, and j be the smallest blue number. We consider the following cases.

   (a) $j = 3$. Since $1 + 4 = 2 + 3$, 4 is red. Similarly, 5 is white, 6 is blue, and so on.

   (b) $j = 100$. Since $2 + 99 = 1 + 100$, 99 is white. If $t > 1$ is red, since $t + 99 = (t-1) + 100$, $t - 1$ is blue, but 100 is the smallest blue number, a contradiction. So $2, 3, \ldots, 99$ are all white.

   (c) $3 < j < 100$. Since $2 + j = 1 + (j+1)$, $j + 1$ is red. Since $3 + j = 2 + (j+1)$, 3 is blue, but j is the smallest blue number, a contradiction.

   Therefore there are three choices of colors for 1, two choices for 2, and two choices for 3. Once these choices are made, the colors for the remaining numbers are determined. Thus the answer to this problem is 12.

5. We start with the following lemma.

   Lemma For any integer $a > 2$ there exists a prime p such that $p \mid (a^3 + 1)$ but $p \nmid (a + 1)$.

   Proof: For the sake of contradiction assume the statement is false for some integer $a > 2$. Since $a^3 + 1 = (a+1)(a^2 - a + 1)$, each prime divisor of $a^2 -$

$a+1$ divides $a+1$. The identity

$$a^2 - a + 1 = (a+1)(a-2) + 3 \tag{1}$$

then shows that 3 is the only prime dividing $a^2 - a + 1$, that is, $a^2 - a + 1$ is a power of 3. Since $3 \mid (a+1)$, $3 \mid (a-2)$. Hence the right-hand side of (1) is divisible by 3 but not by 9. Being a power of 3, $a^2 - a + 1 = 3$ and $a = 2$, a contradiction. Therefore our assumption is false and the original statement is true.■

By the lemma, there exist distinct primes $p_1, p_2, p_3, \ldots, p_{2000}$ such that $p_1 = 3$, $p_2 \neq 3$, $p_2 \mid (2^{3^2} + 1)$, $p_{i+1} \mid (2^{3^{i+1}} + 1)$ and $p_{i+1} \nmid (2^{3^i} + 1)$, for $i = 2, \ldots, 1999$. It is easy to see that $n = p_1^{2000} \cdot p_2 \cdots p_{2000}$ satisfies that condictions of the problem.

6. **First Solution:** Let A_1, B_1, C_1 be the reflections of T_1, T_2, T_3 across the bisectors of $\angle A, \angle B, \angle C$, respectively. Then A_1, B_1, C_1 lie on ω. We prove that they are the vertices of the triangle formed by the images in question, which settles the claim.

By symmetry, it suffices to show that the reflection ℓ_1 of the line H_2H_3 across the line T_2T_3 passes through B_1. Let I be the center of ω. Note that T_2 and H_2 are always on the same side of the line BI, with T_2 closer to the line BI than that of H_2. In the sequel, we consider only the case when C is on the same side of the line BI, as in the figure, i.e., $\angle C \geq \angle A$ (minor modifications are needed if C is on the other side).

Let $\angle A = 2\alpha$, $\angle B = 2\beta$, $\angle C = 2\gamma$. Then $\alpha + \beta + \gamma = 90°$.

Lemma The mirror image of H_2 with respect the line T_2T_3 lies on the line BI.

Proof: Let ℓ be the line passing through H_2 and perpendicular to the line T_2T_3. Denote by S and T the points of intersections of the line BI with $\overline{T_2T_3}$ and ℓ. Note that S also lies on $\overline{BT}$. It is sufficient to prove that $\angle TSH_2 = 2\angle TST_2$.

We have $\angle TST_2 = \angle BST_3 = \angle AT_3S - \angle T_3BS = (90° - \alpha) - \beta = \gamma$. By symmetry across the line BI, $\angle BST_1 = \angle BST_3 = \gamma$. Note that $\angle BT_1S = 180° - \angle BST_1 - \angle T_1BS = 90° + \alpha > 90°$. Therefore C and S are on the same side of of the line IT_1. Then, in the view of the equalities $\angle IST_1 = \angle BST_1 = \gamma = \angle ICT_1$, the quadrilateral SIT_1C is cyclic, so $\angle ISC = \angle IT_1C = 90°$. But then BCH_2S is also cyclic by $\angle BSC = \angle ISC = 90° = \angle BH_2C$. It follows that $\angle TSH_2 = \angle BCH_2 = 2\gamma = 2\angle TST_2$, as desired. ■

Note that the proof of the lemma also gives $\angle BTT_2 = \angle SH_2T_2 = \beta$, by symmetry across the line T_2T_3 and because BCH_2S is cyclic. Then, since B_1 is the reflection of T_2 across the line BI, $\angle BTB_1 = \angle BTT_2 = \beta = \angle CBT$ and $TB_1 \parallel BC$. To prove that B_1 lies on ℓ_1, it now suffices to show that $\ell_1 \parallel BC$.

Suppose that $\beta \neq \gamma$ (otherwise it is trivial that $\ell_1 \parallel BC$); let the line CB meet the lines H_2H_3 and T_2T_3 at D and E, respectively. Note that D and E lie on the line BC on the same side of $\overline{BC}$. We have $\angle BDH_3 = 2|\beta - \gamma|$ and $\angle BET_3 = |\beta - \gamma|$. Therefore $\ell_1 \parallel BC$. The proof is complete.

Second Solution: (by Kiran Kedlaya) Let H and I be the orthocenter and incenter, respectively, of triangle ABC. Since $\angle BH_2C = \angle CH_3B = 90°$, BH_3H_2C is cyclic, so $\angle AH_2H_3 = \angle ABC$. Therefore triangles AH_2H_3 and ABC are oppositely similar. In particular, reflecting the line H_2H_3 across the line T_2T_3, which is perpendicular to the angle bisector AI of A, gives a line parallel to BC.

Therefore the triangle formed by the reflections has sides parallel to the sides of ABC. By looking at the desired result, we realize that it suffices to show that these reflections form the triangle obtained from ABC by the **homothety** with negative

ratio taking the circumcircle of ABC to its incircle. (We take the ratio to be negative because that gives the correct assertion in case ABC is equilateral.) In particular, it suffices to show that the reflection of the line H_2H_3 across the line T_2T_3 intersects ω obtaining a chord C_1B_1 parallel to the line BC, between A and the incenter I, which intercepts an arc of measure $2\angle A$.

The coefficient of similitude between AH_2H_3 and ABC is $AH_3/AC = \cos A$. That means we can obtain AH_2H_3 from ABC by **dilating** towards A with ratio $\cos A$, then reflecting across AI. In particular, the line H_2H_3 is tangent to the circle ω_1, the incircle of triangle AH_2H_3, obtained from the incircle of triangle ABC by **dilating** towards A with ratio of $\cos A$. Let P be the center of ω_1, Q the intersection of the line AI with the line T_2T_3, and R the reflection of P across the line T_2T_3. Then

$$AP = AI\cos A,$$

$$AQ = AT_3\cos A/2 = AI\cos^2 A/2,$$

$$AR = 2AQ - AI = AI(2\cos^2 A/2 - \cos A) = AI.$$

Let $\mathbf{T}$ denote the reflection across the line T_2T_3. Under $\mathbf{T}$, the respective images of the line H_2H_3, ω_1 (radius $r\cos A$), and its center P are the line C_1B_1, circle ω_2 (radius $r\cos A$), and its center I. Since the line H_2H_3 is tangent to ω_1, the line C_1B_1 is tangent to ω_2, i.e., the distance from chord C_1B_1 to I is $r\cos A$. Therefore $C_1B_1 = 2r\sin A$, and so intercepts an arc of measure $2\angle A$ by the **Extended Law of Sines**. Moreover, the line H_2H_3 and A lie on opposite sides of P, so BC and B_1C_1 lie on opposite sides of O.

Thus as noted above, the reflection of the line H_2H_3 contains the image of $\overline{BC}$ under the homothety of negative ratio taking the circumcircle of ABC to its incircle, which suffices to prove the desired result.

2000 Olympiad Results

The top twelve students on the 2000 USAMO were (in alphabetical order):

David G. Arthur	Toronto, ON
Reid W. Barton	Arlington, MA
Gabriel D. Carroll	Oakland, CA
Kamaldeep S. Gandhi	New York, NY
Ian Le	Princeton Junction, NJ
George Lee, Jr.	San Mateo, CA
Ricky I. Liu	Newton, MA
Po-Ru Loh	Madison, WI
Po-Shen Loh	Madison, WI
Oaz Nir	Saratoga, CA
Paul A. Valiant	Belmont, MA
Yian Zhang	Madison, WI

Reid Barton and Ricky Liu were the winners of the Samuel Greitzer-Murray Klamkin award, given to the top scorer(s) on the USAMO. The Clay Mathematics Institute (CMI) award, to be presented for a solution of outstanding elegance, and carrying a $1000 cash prize, was presented to Ricky Liu for his solution to USAMO Problem 3.

The USA team members were chosen based on their combined performance on the 29th annual USAMO and the Team Selection Test that took place at this year's MOSP

held at the University of Nebraska-Lincoln, June 6–July 4, 2000. Members of the USA team at the 2000 IMO (Taejon, Republic of Korea) were Reid Barton, George Lee, Ricky Liu, Po-Ru Loh, Oaz Nir, and Paul Valiant. Titu Andreescu (Director of the American Mathematics Competitions) and Zuming Feng (Phillips Exeter Academy) served as team leader and deputy leader, respectively. The team was also accompanied by Dick Gibbs (Chair, Committee on the American Mathematics Competitions, Fort Lewis College), as the official observer of the team leader.

At the 2000 IMO, gold medals were awarded to students scoring between 30 and 42 points (there were 4 perfect papers on this very difficult exam), silver medals to students scoring between 20 and 29 points, and bronze medals to students scoring between 11 and 19 points. Barton's 39 tied for 5th. The team's individual performances were as follows:

Barton	Homeschooled	GOLD Medalist
Lee	Aragon HS	GOLD Medalist
Liu	Newton South HS	SILVER Medalist
P.-R. Loh	James Madison Memorial HS	SILVER Medalist
Nir	Monta Vista HS	GOLD Medalist
Valiant	Milton Academy	SILVER Medalist

In terms of total score (out of a maximum of 252), the highest ranking of the 82 participating teams were as follows:

China	218	Belarus	165
Russia	215	Taiwan	164
USA	184	Hungary	156
Korea	172	Iran	155
Bulgaria	169	Israel	139
Vietnam	169	Romania	139

The 2001 IMO is scheduled to be held in Washington, DC and Fairfax, VA, USA. For more information about the 2001 IMO, contact Walter Mientka at `walter@amc.unl.edu` or Kiran Kedlaya at `kedlaya@math.berkeley.edu`.

The 2000 USAMO was prepared by Titu Andreescu (Chair), Zuming Feng, Kiran Kedlaya, Alexander Soifer, Richard Stong and Zvezdelina Stankova-Frenkel. The Team Selection Test was prepared by Titu Andreescu and Kiran Kedlaya. The MOSP was held at the University of Nebraska, Lincoln. Titu Andreescu (Director), Zuming Feng, Răzvan Gelca, Kiran Kedlaya, and Zvezdelina Stankova-Frenkel served as instructors, assisted by Melanie Wood and Daniel Stronger.

CONTENTS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, DC 20036