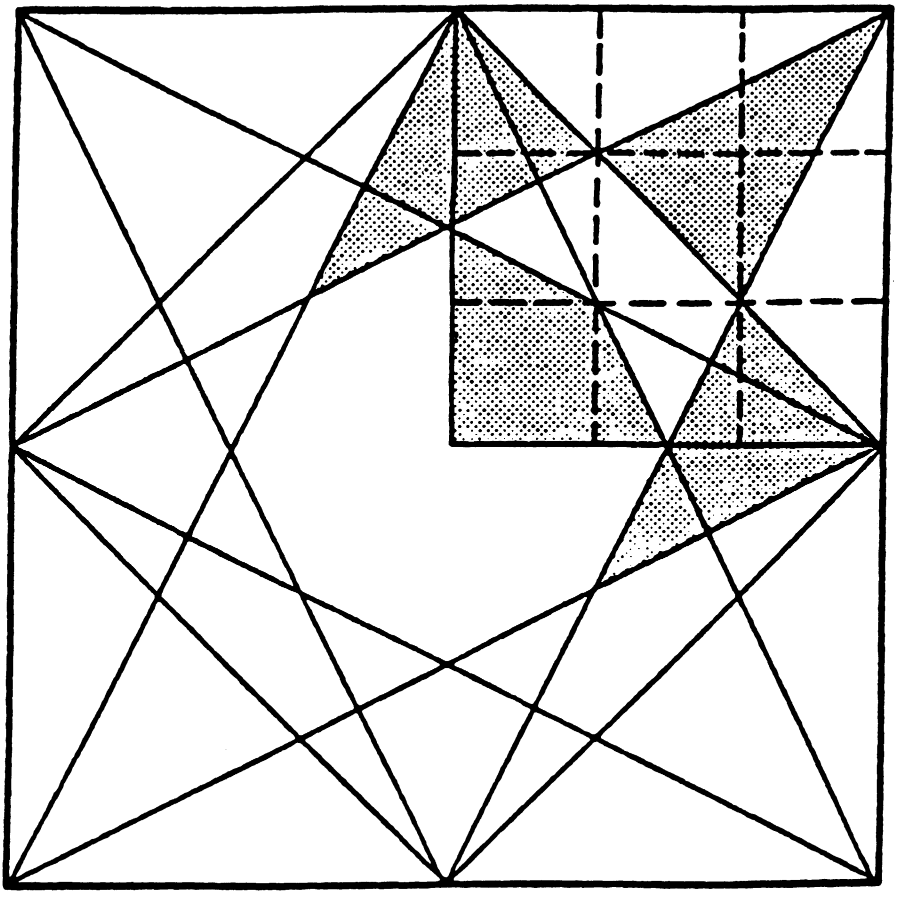
Vol. 67 No. 2 April 1994

# MATHEMATICS MAGAZINE



- Lazzarini's Lucky Approximation of $\pi$
- The Geometry of Harmonic Functions
- Dörrie Tiles and Related Miniatures

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

## EDITORIAL POLICY

The aim of *Mathematics Magazine* is to provide lively and appealing mathematical exposition. This is not a research journal and, in general, the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for an article 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. Articles on pedagogy alone, unaccompanied by interesting mathematics, are not suitable. Neither are articles consisting mainly of computer programs unless these are essential to the presentation of some good mathematics. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

The full statement of editorial policy appears in this *Magazine*, Vol. 64, pp. 71–72, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, nor published by another journal or publisher.

Send new manuscripts to: Martha Siegel, Editor, Mathematics Magazine, Towson State University, Towson, MD 21204. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and two copies and keep one copy. In addition, authors should supply the full five-symbol Mathematics Subject Classification number, as described in *Mathematical Reviews*, 1980 and later. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added. Do not use staples.

## AUTHORS

**Tristan Needham** studied black holes under Roger Penrose, receiving his D.Phil. from Oxford University in 1987. In 1989 he left England to join the happy band of mathematicians at the University of San Francisco. After studying Newton's *Principia*, he gradually realized how Newton's infinitesimal geometry might profitably be applied to the complex plane. The success of this approach in reducing the beautiful truths of elementary complex analysis to visually evident facts led to the writing of his forthcoming book [*Visual Complex Analysis*, Oxford University Press, Oxford]. The present article is based on a section of that work.

**Lee Badger** started life on a farm in the green hills of northern Missouri. After studying logic and foundations at the University of Colorado, and after much mountain climbing, he focused his energy on teaching undergraduate mathematics. This work on Lazzarini started as a simple Math Club talk on the Buffon Needle Problem and exemplifies his belief that the most interesting studies combine several different areas of mathematics and are spiced with history. Today, Dr. Badger's interests include the environment, farming, philosophy, and poetry, and he studies natural resource modeling.

Vol. 67 No. 2 April 1994

# MATHEMATICS MAGAZINE

The *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 the 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 Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.



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

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

PRINTED IN THE UNITED STATES OF AMERICA

# ARTICLES

## Lazzarini's Lucky Approximation of $\pi$

LEE BADGER
Weber State University
Ogden, UT 84408-1702

### 1. Introduction

In 1812 Laplace [14] remarked that one could approximate $\pi$ by performing a Buffon needle experiment. Since then several needle casters claim to have done just that. Lazzarini's 1901 Buffon approximation of $\pi$ [15] was accurate to six decimal places. This work was commended in several publications for illustrating the connectedness of mathematics [13] and validating the laws of probability [3], [7]. However, the 1960 study of Gridgeman [9] suggested that Lazzarini's experiment was not carried out in an entirely legitimate fashion and perhaps didn't warrant the praise it later received. But Gridgeman stopped short of establishing that the experiment was contrived to achieve the desired numerical result.

I will begin by reviewing the history of Lazzarini's experiment and the work of Gridgeman that debunked it. I will then extend Gridgeman's work to virtually rule out any possibility that Lazzarini performed a valid experiment. Some of this work was anticipated by that of O'Beirne [16], however, it also goes beyond that of O'Beirne. In this study elementary applications of probability, recurrence relations, and various numerical techniques are used to look deeper into a small piece of the history of mathematics. For brief expositions of the work of Gridgeman and O'Beirne see also Pilton [17] and Zaydel [18].

### 2. Buffon's Needle

In 1777 Georges-Louis Leclerc, Comte de Buffon, published the results of an earlier study that has come to be known as the Buffon Needle Problem [4]. In its simplest form it assumes that a needle of length $l$ is cast at random on an infinite plane, ruled with parallel lines of uniform separation $d$ where $d > l$. It asks for the probability of the event that the needle intersects one of the lines. Buffon found this probability to be $2l/\pi d$. At this point let us review the technique of solution.

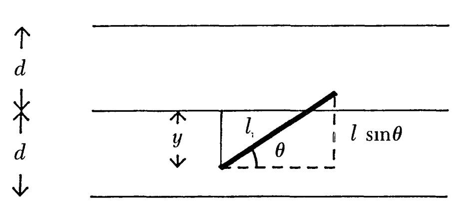



FIGURE 1

Assume the grid is oriented in FIGURE 1, $y$ measures the perpendicular distance from the lower end of the needle to the nearest grid line above it, and $\theta$ measures the smallest counterclockwise angle from the grid direction to the needle. There is clearly a one-to-one correspondence between possible tosses of the needle and ordered pairs $(\theta, y)$, where $0 \le \theta < \pi$ and $0 \le y < d$. The needle hits one of the grid lines if, and only if, $y < l \sin\theta$. A random toss of the needle means that the needle's vertical displacement $(y)$ and orientation $(\theta)$ with respect to the grid are each random and uniformly distributed and so the probability of a hit is the ratio of the area under the curve $y = l\sin\theta$ to the area of the rectangular sample space of FIGURE 2, that is

$$P(\text{Hit}) = \frac{\int_0^{\pi} l \sin\theta \, d\theta}{\pi d} = \frac{2l}{\pi d}.$$

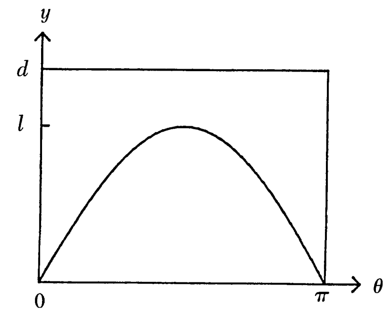



FIGURE 2

Elementary properties alone imply that the probability of a hit is proportional to the length of the needle. This observation opens an interesting "back door" method to obtain the Buffon result (see Gnedenko [8]). Suppose a convex polygon with $n$ sides of length $l_1, l_2, \ldots, l_n$ is tossed at random onto the grid. Also suppose the polygon is of diameter less than $d$ and that a hit occurs if, and only if, exactly two sides hit. Then

$$\begin{aligned} P(\text{polygon hits}) &= \frac{1}{2}\sum_{i=1}^{n} k l_i \\ &= ks/2 \quad \text{where } s \text{ is the perimeter.} \end{aligned}$$

A limiting argument yields the same result for any closed convex curve.

Now let's apply the result when the curve is a circle of radius $r$ with $2r < d$. Looking at where the center of the circle falls at random between grid lines we see that

$$\begin{aligned} P(\text{circle hits}) &= 2r/d, \quad \text{so} \\ k(2\pi r)/2 &= 2r/d \end{aligned}$$

and hence $k = 2/(\pi d)$. So

$$\begin{aligned} P(i\text{th side hits}) &= \sum_{j \ne i} p_{ij} = k l_i \\ &= \frac{2 l_i}{\pi d}, \end{aligned}$$

which is the original Buffon result.

It is intuitively clear that the probability should be an increasing function of $l$ and a decreasing function of $d$, but that the probability depends on $\pi$ is perhaps unexpected and has been a source of many other studies—including this one.

## 3. Lazzarini's "Approximation" of $\pi$

Laplace recognized that Buffon's result could be used to obtain an experimental approximation of $\pi$. If one casts $N$ needles and if $H$ of them hit, then since the theoretical probability of a hit is approximated by the relative frequency, $2l/\pi d \approx H/N$. It follows that

$$\hat{\pi} := \frac{2l}{d}\frac{N}{H} \approx \pi \quad (A := B \text{ means } A \text{ is defined by } B).$$

Several needle casters actually performed this experiment and published their experimental values of $\pi$. These results are summarized in Gridgeman. The focus of this paper is the experiment reported by Mario Lazzarini in 1901. One of Lazzarini's results has been widely quoted; in it $l = 2.5$ cm, $d = 3$ cm, $N = 3408$, and $H = 1808$, so that $\hat{\pi} = (2(2.5)/3)(3408/1808) = 3.1415929\ldots$. Since $\pi = 3.1415926\ldots$, Lazzarini's result gives six-decimal place accuracy.

But something seems a little suspect about those numbers 3408 and 1808. Why cast 3408 needles? Why not a nice round number like 1000 or 3500? Our skepticism increases if we look at what happens when the number of hits is increased or decreased by 1. If $H = 1807$, $\hat{\pi} = 3.1433\ldots$; if $H = 1809$, $\hat{\pi} = 3.1398\ldots$. Lazzarini appears to have been extraordinarily lucky!

Repeated trials are capable of a simple statistical analysis. If, using Lazzarini's grid and needle, one wants to be 95% confident that $|\pi - \hat{\pi}| < 0.5 \times 10^{-6}$ (six-decimal place accuracy), then one needs to cast around 134 trillion needles! To obtain this number, we seek $N$ such that

$$P\left(\left|\pi - \frac{5N}{3H}\right| < 0.5 \times 10^{-6}\right) \geq 0.95.$$

Now $\left|\pi - \frac{5N}{3H}\right| < \epsilon$ if, and only if, $\left|\frac{3\pi}{5} - \frac{N}{H}\right| < \frac{3\epsilon}{5}$ if, and only if, (assuming $1/(x)$ is locally linear near $3\pi/5$) $\left|\frac{5}{3\pi} - \frac{H}{N}\right| < \frac{1}{(3\pi/5)^2}\frac{3\epsilon}{5} = \frac{5\epsilon}{3\pi^2}$ if, and only if, $\left|\frac{5N}{3\pi} - H\right| < \frac{5N\epsilon}{3\pi^2}$. Since $H$ is binomially distributed with parameters $N$ and $p = \frac{5}{3\pi}$, its expectation is $Np$ and its variance is $Np(1-p)$. Using the normal approximation, in order to have $P\left(\left|\frac{5N}{3\pi} - H\right| < \frac{5N}{3\pi^2}(0.5) \times 10^{-6}\right) = .95$ we need

$$\frac{\frac{5N}{3\pi^2}(0.5) \times 10^{-6}}{\sqrt{Np(1-p)}} \approx 1.96$$

or $N \approx 134 \times 10^{12}$.

In addition to questioning the number of needles cast, one may also question the accuracy of the measurement of $l$ and $d$. Lazzarini reported that they were measured to be 2.5 cm and 3 cm, but gave no tolerances on his measuring instruments. At the turn of the century, state of the art micrometers had errors of about $\pm 0.0005$ cm **[12]**. Incorporating these best error bounds, one calculates that $3.1404 < \hat{\pi} < 3.1427$. So the last four figures of agreement of Lazzarini's $\hat{\pi}$ with $\pi$ are meaningless. Zaydel **[18]** gives additional analyses of measurement error in needle experiments.

Another source of skepticism emerges when we look more closely at Lazzarini's

$$\hat{\pi} = \frac{2(2.5)}{3}\frac{3408}{1808} = \frac{355}{113}.$$

This fraction is known to number theorists as a convergent in the continued fraction for $\pi$ ([1] and [11]) and historians recognize it as a rational approximation to $\pi$ discovered by the fifth century A.D. Chinese mathematician Tsu Ch'ung-chih [5]. So to many mathematicians—and presumably to Lazzarini—it was a well-known rational approximation of $\pi$. Also, in terms of the magnitude of its denominator, it is an exceedingly accurate approximation of $\pi$. The next smallest denominator that yields a strictly better approximation occurs in the fraction 52,163/16,604.

## 4. A Lesson in "Experimental Design"

Let us consider how we would go about rigging up a good Buffon experimental approximation to $\pi$. To get Lazzarini's approximation, we need to choose $l$, $d$, $N$, and $H$ such that

$$\frac{2lN}{dH}=\frac{355}{113}=\frac{355k}{113k}=\frac{5\cdot 71k}{113k}.$$

A reasonable choice might be $2l=5$ so $l=2.5$ and $d>l$, say $d=3$, resulting in $(N/H)=(213k/113k)$. The net effect is that if, at any multiple of 213 casts, we have the same multiple of 113 hits, we achieve the desired approximation. Lazzarini achieved this at the sixteenth multiple.

In his conclusion Gridgeman suggested that the mysterious 3408 was selected as a stopping point only because it was a potential generator of 355/113 and that by the "very happiest of coincidences" the optimum $H$ was observed. He went on to question whether Lazzarini performed any experiment at all, or if the results were purely mental concoctions.

I will give a more definitive answer to this question. Is there any chance that Lazzarini actually performed an experiment? Assuming that $\hat{\pi}$ is computed after each cast, one could stop at any point at which $\hat{\pi}=355/113$. What is the likelihood of a sequence of casts yielding $\hat{\pi}=355/113$ at *some* stopping point, for his choice of $d=3$ and $l=2.5$?

Let $A_k$ denote the event that $113k$ hits occur during $213k$ casts, and let $a_k := P(A_k)$. If $N$ needles are cast then there are $[N/213] =: m$ such events. We are interested in the value of $P(A_1\cup A_2\cup\cdots\cup A_m) =: P(U_m)$, the probability that the ratio 113/213 is achieved at least once in the first $213m$ casts. We will find out in Section 6 that Lazzarini claimed to have dropped $N=4000$ needles, which corresponds to $m=18$. We will show shortly that $P(U_{18})$ is about 0.30. So there is a good possibility that Lazzarini could actually have performed his experiment and achieved his reported result at some point.

To get this value for $P(U_{18})$ we first use Stirling's approximation [6] on the binomial probabilities $a_k$ and simplify:

$$a_k=\frac{(213k)!}{(113k)!(100k)!}p^{113k}(1-p)^{100k}\sim c\alpha^k/\sqrt{k}$$

where

$$\alpha := \frac{213^{213}}{113^{113}100^{100}}p^{113}(1-p)^{100}=0.99999\,99999\,99132+,$$

$$c:=\sqrt{\frac{213}{113\cdot 2\cdot 100\pi}}=0.05477+\quad\text{and } p=\frac{5}{3\pi},$$

and where we use " $\sim$ " in the standard sense that the ratio of the two sides approaches 1 as $k \to \infty$. The error bounds of Stirling's formula [18] can be used to show that $a_k$ is actually within 1% of this estimate for all $k$. An explanation for $\alpha$'s closeness to 1 is that, considered as a function of $p$, $\alpha$ takes its maximum value of 1 when $p = 113/213$ and $5/3\pi$ is extremely close to 113/213.

Whenever one of the events $A_i$ occurs, we may think of the entire experiment and the accounting of hits as beginning anew. In other words, for $j > i$, $P(A_j|A_i) = P(A_{j-i})$ and so $P(A_j \cap A_i) = a_i a_{j-i}$. Properties of this type will play an important role in our computations. For instance

$$\begin{aligned}\sum_{1 \le i < j < k \le m} P(A_i \cap A_j \cap A_k) &= \sum_{i=1}^{m-2} \sum_{1 \le j < k \le m-i} P(A_i \cap A_{i+j} \cap A_{i+k}) \\ &= \sum_{i=1}^{m-2} \sum_{1 \le j < k \le m-i} P(A_i) P(A_{i+j} \cap A_{i+k} | A_i) \\ &= \sum_{i=1}^{m-2} \sum_{1 \le j < k \le m-i} P(A_i) P(A_j \cap A_k) \\ &= \sum_{i=1}^{m-2} a_i S(2, m-i)\end{aligned}$$

where $S(u, v)$ is the sum of the probabilities of all intersections of $u$ members of $A_1, \dots, A_v$. With $S(0, v) := 1$ for $0 \le v \le m$, we can show by induction on $u$ that $S(u, v) = \sum_{i=1}^{v-u+1} a_i S(u-1, v-i)$ for $v = u, u+1, \dots, m$. By the principle of inclusion-exclusion, $P(U_m) = S(1, m) - S(2, m) + \cdots + (-1)^{m+1} S(m, m)$. This provides a reasonably efficient, $O(m^3/3)$, way to compute $P(U_m)$. Using this algorithm, we obtain $P(U_{18}) \approx 0.3041$.

The sequence $u_m := P(U_m)$ is slow to converge; $u_1 \approx 0.05$, $u_{10} \approx 0.23$, $u_{30} \approx 0.45$, $u_{100} \approx 0.55$, and $u_{500} \approx 0.76$; it is unclear what its limit is.

## 5. Is Ultimate Success a Certainty?

The next question is whether, in an *infinite* sequence of casts, the ratio 355/113 is certain to occur. That is, is $\lim_{m \to \infty} P(U_m) = 1$? The answer is no, and we can argue as follows. The expected number of occurrences of this ratio, $\sum_{k=1}^{\infty} a_k$ is finite because $a_k \sim c\alpha^k/\sqrt{k}$ and $\alpha < 1$. If the desired ratio occurs with probability 1, then it must occur with probability 1 at some finite $A_k$ and then accounting of successes and failures could begin anew and we may argue that with probability 1 it must occur at some subsequent finite $A_k$ and then at yet another and so on, ad infinitum. But since the expected number of such occurrences is finite, this is impossible.

More precisely, let $B_n$ be the event "$n$ or more $A$'s". Then $B_1 = \bigcup_{k=1}^{\infty} A_k$. The sequence $B_n$ is nested decreasing so $B_{n+1} = B_{n+1} \cap B_n$ and $P(B_{n+1}) = P(B_{n+1} \cap B_n) = P(B_n) \cdot P(B_{n+1}|B_n) = P(B_n) \cdot P(B_1)$, and by induction we have $P(B_n) = (P(B_1))^n$. If $P(B_1)$ were equal to 1, then we would have $P(B_n) = 1$ for all $n$ and hence $P(\bigcap_{n=1}^{\infty} B_n) = \lim_{n \to \infty} P(B_n) = 1$. But $\bigcap_{n=1}^{\infty} B_n =$ "infinitely many $A$'s" $= \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k$. So $1 \le P(\bigcup_{k=j}^{\infty} A_k) \le \sum_{k=j}^{\infty} P(A_k)$ for all $j$, contradicting the convergence of $\sum_{k=1}^{\infty} a_k$. Thus $P(B_1) < 1$.

A more detailed analysis will yield a numerical approximation of $f := P(\bigcup_{k=1}^{\infty} A_k) = P(B_1)$, the probability of ultimate success. The value of $f$ is related to that of $a := \sum_{k=1}^{\infty} a_k$. The relation $f = a/(1+a)$ is standard renewal theory and is derived by Feller [6] using generating functions. It can also be derived from the above analysis and the identity, $E(X) = \sum_{n=1}^{\infty} P(X \geq n)$ where $X :=$ "the number of $A_k$ that occur." The following is a more direct, elementary proof.

Let $F_n$ be the event "355/113 occurs first at trial $213n$", i.e. $F_n = A_n \cap A_{n-1}^c \cap \cdots \cap A_1^c$. The $F$'s are mutually exclusive and $A_n = F_1 \cup \cdots \cup F_n$. So $P(\bigcup_{n=1}^{\infty} A_n) = f = P(\bigcup_{n=1}^{\infty} F_n) = \sum_{n=1}^{\infty} f_n$ where $f_n := P(F_n)$.

Also for $k \geq 1$,
$$\begin{aligned} a_k &= P(A_k) \\ &= P(A_k \cap F_1) + P(A_k \cap F_2) + \cdots + P(A_k \cap F_k) \\ (*) \qquad &= P(F_1)P(A_k|F_1) + P(F_2)P(A_k|F_2) + \cdots + P(F_k)P(A_k|F_k) \\ &= f_1 a_{k-1} + f_2 a_{k-2} + \cdots + f_k a_0, \quad \text{where } a_0 = 1. \end{aligned}$$

Expanding, collecting terms with like sums of indices, and using (*) one obtains

$$\begin{aligned} &(f_1 + f_2 + \cdots + f_n)(a_0 + a_1 + a_2 + \cdots + a_n) \\ &\quad = a_1 + a_2 + \cdots + a_n + b_{n+1} + \cdots + b_{2n}, \quad \text{where } 0 < b_k < a_k. \end{aligned}$$

So

$$\sum_{k=1}^{n} a_k < \sum_{k=1}^{n} f_k \cdot \sum_{k=0}^{n} a_k < \sum_{k=1}^{2n} a_k \quad \text{and}$$

$$\sum_{k=1}^{n} a_k \Big/ \left(1 + \sum_{k=1}^{n} a_k\right) < \sum_{k=1}^{n} f_k < \sum_{k=1}^{2n} a_k \Big/ \left(1 + \sum_{k=1}^{n} a_k\right) \quad \text{for } n = 1, 2, \ldots.$$

Letting $n \to \infty$ we obtain that $a < \infty$ implies $f < 1$, as obtained earlier, but furthermore, that $f = a/(1+a)$ for all extended real $a$.

So to estimate $f$, we first need to estimate $a$ and its Stirling approximation, $\sum_{n=1}^{\infty} c\alpha^n/\sqrt{n}$.

Since $\alpha^x/\sqrt{x}$ is decreasing for $x > 0$,

$$\sum_{n=1}^{\infty} \alpha^n/\sqrt{n} > \int_1^{\infty} \alpha^x/\sqrt{x}\, dx$$

and

$$\sum_{n=1}^{\infty} \alpha^n/\sqrt{n} = \alpha + \sum_{n=2}^{\infty} \alpha^n/\sqrt{n} < \alpha + \int_1^{\infty} \alpha^x/\sqrt{x}\, dx.$$

By the change of variable $t := \sqrt{-\ln \alpha}\sqrt{x}$

$$\int_1^{\infty} \alpha^x/\sqrt{x}\, dx = 2/\sqrt{-\ln \alpha} \int_{\sqrt{-\ln \alpha}}^{\infty} e^{-t^2}\, dt.$$

Also,

$$\int_{\sqrt{-\ln \alpha}}^{\infty} e^{-t^2}\, dt < \int_0^{\infty} e^{-t^2}\, dt = \sqrt{\pi}/2$$

and

$$\int_{\sqrt{-\ln \alpha}}^{\infty} e^{-t^2}\, dt = \int_0^{\infty} e^{-t^2}\, dt - \int_0^{\sqrt{-\ln \alpha}} e^{-t^2}\, dt$$

$$> \int_0^\infty e^{-t^2}\,dt - \int_0^{\sqrt{-\ln\alpha}} 1\,dt$$

$$> \sqrt{\pi}/2 - \sqrt{-\ln\alpha}\,.$$

Combining these inequalities we obtain

$$\sqrt{\pi/(-\ln\alpha)} - 2 < \sum_{n=1}^{\infty} \alpha^n/\sqrt{n} < \alpha + \sqrt{\pi/(-\ln\alpha)}\,.$$

With the values of $\alpha$ and $c$ from section 4, we get

$$1{,}902{,}750 < \sum_{n=1}^{\infty} \alpha^n/\sqrt{n} < 1{,}902{,}760$$

and

$$104{,}218 < \sum_{n=1}^{\infty} c\alpha^n/\sqrt{n} < 104{,}219.$$

The bound of Stirling's approximation of factorials [6] reveals that for each $n$,

$$0.9988c\alpha^n/\sqrt{n} < P(A_n) < c\alpha^n/\sqrt{n}$$

and so

$$104{,}092 < \sum_{n=1}^{\infty} P(A_n) = a < 104{,}219.$$

Finally, since $P(\bigcup_{n=1}^{\infty} A_n) = f = a/(1+a)$,

$$0.99999039 < P\left(\bigcup_{n=1}^{\infty} A_n\right) < 0.99999041.$$

We conclude that it is not *certain* that Lazzarini would eventually have obtained the optimum $\hat{\pi}$, but the odds favoring it are overwhelming (assuming no measurement error).

*Note*: The estimation of the above numerical values and that of $\alpha$ originally caused many headaches. While this work was in progress my school got a multiple precision arithmetic package. The headaches went away! Students' reliance on calculators may be shaken by asking them to calculate $\alpha$. Some methods yield results less than 1, some greater than 1 and some yield overflow.

## 6. The Case Against Lazzarini

The result of section 5 suggests that it is at least plausible, ignoring measurement error, that Lazzarini actually performed the experiment. However, Lazzarini did not report just a single experiment of 3408 casts and 1808 hits. He reported a series of casts:

| N | 100 | 200 | 1000 | 2000 | 3000 | 3408 | 4000 |
|---|---|---|---|---|---|---|---|
| H | 53 | 107 | 524 | 1060 | 1591 | 1808 | 2122 |

It is highly suspicious that all the values of $H$ in this series are very close to their expected values $Np$, which are respectively 53.05, 106.10, 530.52, 1061.03, 1591.55, 1808.00, and 2122.07. Ordinarily we expect much greater fluctuations than this in random data. In fact, even if we only look at those hits when $N$ is a multiple of 1000, the probability of being this close to the expected values is exceedingly small. Define $G_k$ to be the event that the number of hits, $H_k$, in $1000k$ casts is at least as close to the expected number of hits as Lazzarini reported. Then

$$G_1 = \{524 \le H_1 \le 537\}$$
$$G_2 = \{1060 \le H_2 \le 1062\}$$
$$G_3 = \{1591 \le H_3 \le 1592\}$$
$$G_4 = \{H_4 = 2122\} \quad \text{and}$$

$$\begin{aligned} &P(G_1 \cap G_2 \cap G_3 \cap G_4) \\ &\quad = P(G_1) \cdot P(G_2|G_1) \cdot P(G_3|G_1 \cap G_2) \cdot P(G_4|G_1 \cap G_2 \cap G_3) \\ &\quad \le P(G_1) \cdot P(G_2|H_1 = 531) \cdot P(G_3|H_2 = 1061) \cdot P(H_4|H_3 = 1592) \\ &\quad \le P(524 \le H_1 \le 537) \cdot P(1060 - 531 \le H_1 \le 1062 - 531) \\ &\qquad \cdot P(1591 - 1061 \le H_1 \le 1592 - 1061) \cdot P(H_1 = 2122 - 1592). \end{aligned}$$

Using the normal approximation to the binomial probabilities we obtain a probability of less than 0.00003. Thus it seems exceedingly unlikely that Lazzarini carried out a random series of tosses with results as nearly optimal as he reported. So it seems likely the experiment was not done—at least not in a random fashion.

## 7. Speaking of Hoaxes

But setting aside measurement error and granting that the experiment was a hoax, one may ponder the quality of the hoax. Here are three hoaxes to compare with that of Lazzarini.

In hoax one, let $d = 10$ cm and $l = 7.1$ cm. It seems not unlikely that a garden variety needle might measure 7.1 cm and a round figure for $d > l$ is $d = 10$ cm. In this hoax any multiple of 250 casts is a potential generator of 355/113—one needs the same multiple of 113 hits. This hoax has the advantage that the optimal stopping points are plausible; for instance, every multiple of 1000 is such a point.

In hoax two, let $l = 7.1$ cm and $d = 11.3$ cm. These seem like less plausible "objective" values of $l$ and $d$, but in the experiment, the occurrence of 355/113 is more assured because every multiple of five casts is a potential generator.

In hoax three, we go for an even more accurate $\hat{\pi}$. As mentioned earlier, the next improvement on 355/113 is 52,163/16,604. This was found by computer and may not have been known to nineteenth-century mathematicians. But continued fraction convergents were available and the next convergent after 355/113 is 103,993/33,102. But 103,993 is a prime and so this convergent is not a good candidate, but the one after that, 104,348/33,215, does lead to a plausible hoax. Since $104{,}348 = 2 \cdot 2 \cdot 19 \cdot 1373$ and $33{,}215 = 5 \cdot 7 \cdot 13 \cdot 73$, we can take $l = 2 \cdot 2 \cdot 1373 \times 10^{-3} = 5.492$ cm, and $d = 7 \cdot 13 \cdot 73 \times 10^{-3} = 6.643$ cm. These aren't particularly plausible but they are within the accuracy that was then measurable and the payoff is that every multiple of 19 casts is a potential generator of $\hat{\pi}$ that misses $\pi$ by less than $3 \times 10^{-10}$ and so is accurate in the ninth decimal place.

The advantage of hindsight (which Lazzarini lacked) allows us to design bogus experiments that foil today's statistical tests designed to expose them. It seems to me that hoax one does just that. This is especially true if only a final value of $\hat{\pi}$ is reported or, if reported, a series of values of hits has sufficiently random dispersion.

Today, one occasionally hears of bogus experiments and/or rigged data [**2**], [**10**]. Presumably modern hoaxers are aware of the type of statistics that exposed the Lazzarini hoax and put enough dispersion in their "data" to avoid the same fate. It will be interesting to see what, if any, (future) statistical or scientific test will be brought to bear on their work, thereby labeling them as only poor hoaxers and properly removing them from the ranks of objective scientists.

**Acknowledgements.** I want to thank the referees for their helpful suggestions and support. I especially want to thank my friend and colleague, Jim Foster. Jim's input helped convert my qualitative findings into quantitative probabilities and accounts for a good deal of the precise character of the results.

## REFERENCES

1. P. Beckmann, *A History of* $\pi$, The Golem Press, New York, 1971.
2. W. Broad and N. Wade, *Betrayers of the Truth*, Simon and Schuster, New York, 1982.
3. J. Bronowski, *The Common Sense of Science*, Harvard University Press, Cambridge, 1953.
4. G. L., C. Buffon, Essai d'arithmétique morale, in Vol. 4 of the *Supplément à l'Histoire Naturelle*, 1777.
5. H. Eves, *An Introduction to the History of Mathematics*, CBS College Publishing, New York, 1983.
6. W. Feller, *An Introduction to Probability Theory and Its Applications*, *Vol. 1*, John Wiley & Sons, Inc., New York, 1957.
7. G. Gamow, *One Two Three* $\cdots$ *Infinity*, The Viking Press, New York, 1961.
8. B. V. Gnedenko, *The Theory of Probability*, *4th ed*, (*English translation*), Chelsea Publishing Co., New York, 1967.
9. N. T. Gridgeman, Geometric probability and the number $\pi$, *Scripta Math*., 25 (1960), 183–95.
10. D. P. Hamilton, Verdict in sight in the "Baltimore Case", *Science*, Vol. 251 (8 March 1991), 1168–72.
11. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, London, 1979.
12. IBM, *Precision Measurement in the Metal Working Industry*, Syracuse U. Press, Syracuse, NY, 1952.
13. E. Kasner and J. Newman, *Mathematics and the Imagination*, Simon and Schuster, New York, 1940.
14. P-S. Laplace, *Théorie Analytique des Probabilités*, 1812.
15. M. Lazzarini, Un' applicazione del calcolo della probabilità alla ricerca sperimentale di un valore approssimato di $\pi$, *Periodico di Matematica*, 4 (1901), 140–3.
16. T. H. O'Beirne, *Puzzles and Paradoxes*, Oxford University Press, London, 1965.
17. B. Pilton, Genuine fakes, in *Seven Years of Manifold*, ed. by Ian Stewart and John Jaworski, Shiva Publications, Nantwich, 1981.
18. A. N. Zaydel, Delusion or fraud, *Quantum*, September/October 1990, 6–9.

---

## WEATHER REPORT ON T.V. NEWS, 11:15 p.m. November 21, 1990 (reported by Ralph Boas)

"It's warmer right now than today's high."

# The Geometry of Harmonic Functions

TRISTAN NEEDHAM
University of San Francisco
San Francisco, CA 94117-1080

## 1. Introduction

Imagine a society in which the citizens are encouraged, indeed compelled up to a certain age, to read (and sometimes write) musical scores. All quite admirable. However, this society also has a very curious (few remember how it all started) and disturbing law: *Music must never be listened to or performed!*

Though its importance is universally acknowledged, for some reason music is not widely appreciated in this society. To be sure, professors still excitedly pore over the great works of Bach, Wagner, and the rest, and they do their utmost to communicate to their students the beautiful meaning of what they find there, but they still become tongue-tied when brashly asked the question, "What's the point of all this?!"

In this parable, it was patently unfair and irrational to have a law forbidding would-be music students from experiencing and understanding the subject directly through "sonic intuition." But in our society of mathematicians we *have* such a law. It is not a written law, and those who flout it may yet prosper, but it says, *Mathematics must not be visualized!*

More likely than not, when a mathematics student today opens a random text on a random subject, he is confronted by abstract symbolic reasoning that is divorced from his sensory experience of the world, *despite* the fact that the very phenomena he is studying were often discovered by appealing to geometric (and perhaps physical) intuition. This reflects the fact that steadily over the last hundred years the honor of visual reasoning in mathematics has been besmirched. Only recently have many mathematicians picked up the gauntlet on its behalf, openly challenging the current dominance of purely symbolic logical reasoning.

Rather than indulge in further pulpit-thumping, we refer the sympathetic reader to a cheering MAA book [1]. The present author has joined the fray by attempting to render palpable the beautiful truths of elementary complex analysis by means of new geometric insights. Both this paper and a previous one [2] arose in connection with that work [3].

Our concern here will be with the various formulae for expressing a harmonic function in the interior of a planar region in terms of its values on the boundary. In place of the usual symbolic arguments, we shall supply simple geometric explanations/derivations of these formulae. But before we begin to visualize these formulae (starting in the next section) we must clarify the nature of the questions to which they are the answers. We begin with Poisson's formula for the disk.

Think of the complex plane as a thermally insulated sheet of metal; heat flows freely within it, but does not leak away into the surrounding space. Now supply heat at a constant rate to various points (*sources*) of the plane, and likewise remove heat at other places (*sinks*). Initially, the temperature of the metal at any given point will vary with time. A small element of the metal plate gains or gives up energy as heat attempts to flow across it from the sources to the sinks. But eventually (quickly, if the thermal conductivity is high) the heat flow will settle down into a steady pattern and the temperature at a point $z$ will likewise settle to a definite value $T(z)$. In this steady state, the global statement of the conservation of energy is that total heat supplied at the sources equals the total heat removed at the sinks (possibly including infinity).

It is shown in elementary physics that $-\nabla T$ is the heat flow vector, and it then follows that the *local* statement of the indestructibility of energy is that, in the steady state, $T(z)$ is *harmonic*: Away from sources and sinks, it satisfies Laplace's equation,

$$\begin{aligned}\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2} &= \nabla\cdot(\nabla T)\\ &= -(\textit{local rate of energy production per unit area})\\ &= 0.\end{aligned}$$

Suppose that we now measure the temperature around the circumference $C$ of a circle of radius $R$, the interior of which is free of sources and sinks, and the center of which we conveniently choose to be the origin. We hope that it may seem physically plausible that these values actually determine the temperature at any interior point $a$. This was confirmed by Poisson in 1820 when he derived an explicit formula for $T(a)$ in terms of $T(C)$.

As $z = Re^{i\theta}$ moves around $C$, we may express the measured temperature as a function of the angle: $T = T(\theta)$. *Poisson's formula* is then

$$T(a)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{R^2-|a|^2}{|z-a|^2}\right]T(\theta)\,d\theta. \tag{1}$$

The quantity in square brackets is called the *Poisson kernel*, and we shall write it $\mathscr{P}_a(z)$. Thus (1) may be roughly paraphrased as saying that the heat of an element of $C$ at $z$ propagates to $a$ with a "facility" $\mathscr{P}_a(z)$ that dies away as the square of the distance between $z$ and $a$.

This formula is connected with an important and difficult issue that engaged an illustrious cast of characters: Riemann, Weierstrass, Schwarz, Klein, Poincaré, and Hilbert. Instead of dealing with a pre-existing harmonic function, *Dirichlet's problem* demands that we arbitrarily (but piecewise continuously) assign values to the boundary of a simply connected region $R$ and then inquire if there always exists a harmonic function in $R$ that takes on these values as the boundary is approached. (This problem is in fact closely related to the equally famous *Plateau's problem*: Given a simple closed curve in space, does there exist a minimal surface that spans it?)

In the case of the disk, H. A. Schwarz demonstrated that not only does the solution to Dirichlet's problem exist, but it is explicitly given by (1). If we are handed the piecewise-continuous values $T(\theta)$ on $C$ then we may construct a function $T(a)$ in the interior according to Poisson's recipe. Schwarz's solution then amounted to showing that $T(a)$ is automatically harmonic, and that as $a$ approaches a boundary point at which $T(\theta)$ is continuous, $T(a)$ approaches the given value $T(\theta)$. The truth of all this will be explained in the next two sections.

If we assume (as was implicit in the previous discussion) that mathematical harmonic functions are *identical* with physical temperature distributions, then both the existence and uniqueness of a mathematical solution to Dirichlet's problem for general regions is assured by Nature's solution to the equivalent physical problem: Heat the boundary points of $R$ to their assigned temperatures; let things settle down; the temperature in the interior is then the desired harmonic function. In like manner, if we assume the identity between soap films and minimal surfaces, we need only dip a bent loop of wire into soapy water in order to solve Plateau's problem.

A sliver of history: Riemann *did* make the above identification (actually, he thought in terms of electricity rather than heat) and thereby reaped a rich harvest of mathematical discoveries based on physical intuition; in particular, later we shall see how it led him to his mapping theorem. Perhaps aware of the audacity of his style of reasoning, Riemann sought to bolster his physical intuition with a more mathematical idea. In ignorance of its earlier use by Gauss and Lord Kelvin, he christened this idea *Dirichlet's Principle*. Roughly, it asserts that if we consider the functions $T(a)$, continuous in the interior of $R$ and taking on prescribed boundary values, then, due to the nonnegative integrand, there *must be* one (don't you think?) that minimizes

$$\int\int_R \left[\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2\right] dx\, dy.$$

But it can then be shown that this minimizing function is also the solution to Dirichlet's problem, the existence of which was sought by Riemann.

However, in 1869 the brilliant but dryly logical Weierstrass threw a wrench into the works when he produced a *counterexample* to the general idea underlying Dirichlet's Principle. As Felix Klein later described the situation, "With this a large part of Riemann's developments came to nought." These clouds of doubt would continue to hang in the air for three decades. Undeterred, mathematicians such as Klein—who did not shun Weierstrassian rigor but was himself driven by geometric/physical intuition—continued to expound and extend Riemann's ideas. Only in 1900 did Hilbert finally "resurrect Dirichlet's Principle" (to use his own words) by showing that although Weierstrass had discredited the general idea behind it, this particular instance actually *is* correct.

In fairness, it should be explained that while Weierstrass doubted Riemann's proofs, he believed the results. Indeed, it was at his urging that Schwarz (a former pupil) found the above solution for the disk that did not rely on the suspect principle. Schwarz was then able to use this to show that a solution will also exist if $R$ is any *union* of disks.

It is clear to which camp the author owes his allegiance. Although Weierstrass would not approve, our mode of explanation will remain steadfastly geared to geometric intuition rather than logical rigor.

## 2. Schwarz's Interpretation

There is an exceedingly beautiful geometric interpretation of formula (1), due to Schwarz, which deserves to be far better known than it is. Of the myriad complex analysis texts, we have only found it described in the book by Ahlfors ([4], p.170). Schwarz obtained it, and likewise Ahlfors explains it, as a consequence of Poisson's formula (itself derived by computation). In this section we shall instead demonstrate Schwarz's result directly and geometrically, only then producing the Poisson formula as a consequence of *it*. First we remind the reader of some preliminary facts.

Suppose we measure the temperature at the center 0 of the circle $C$. If half of $C$ were at one temperature while the other half were at another, then symmetry and physical intuition would suggest that at 0 we would find the average of these temperatures. Dividing up $C$ further into arcs of constant temperature, then passing to the limit of infinitesimal arcs $R\, d\theta$ at temperatures $T(\theta)$, we are led to suspect *Gauss' mean value theorem* for harmonic functions:

$$T(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} T(\theta)\, d\theta. \tag{2}$$

This is often proved, as opposed to merely motivated, by appealing to Cauchy's Integral Formula; its consonance with the desired formula (1) is evident upon setting $a = 0$.

FIGURE 1(a) is intended to make the meaning of (2) more vivid. Imagine that there are thermometers placed all along the circle and that you are standing at 0. Turning your head successively through the same small angle marked • you would see the thermometers at the white dots on the boundary. The average of the temperatures you see (as • ⇝ 0) is then the temperature where you stand.

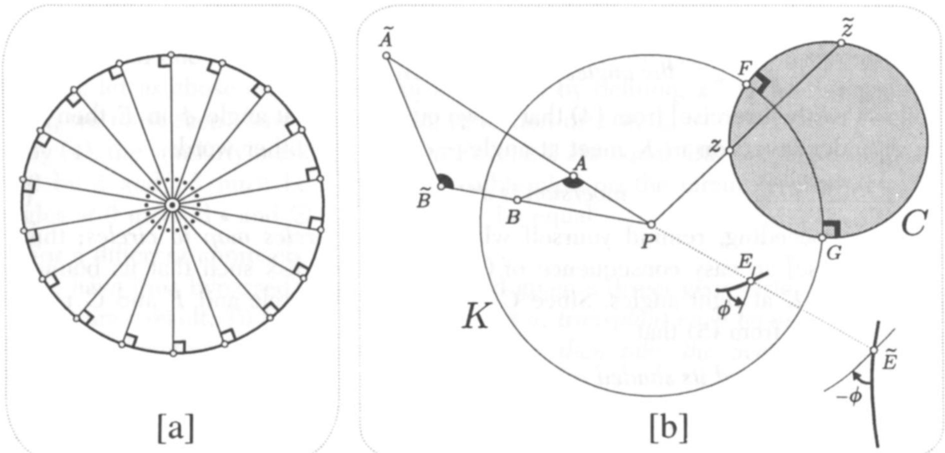



FIGURE 1

Next, recall that analytic mappings are essentially those that are *conformal*, that is, those for which the angle of intersection of two image curves is identical, in magnitude *and* sense, to that of the preimages. This makes it easy to understand that the composition $f[h(z)]$ of two analytic functions is itself analytic, for if two mappings preserve angles then so will their composition.

Since the real and imaginary parts of an analytic function $f(z) = T(z) + iS(z)$ are automatically harmonic by virtue of the Cauchy-Riemann equations, it follows from the above that $T[h(z)]$ is also harmonic. But if we are *given* a harmonic function $T(z)$ in a simply connected region, then it is always possible to find a *harmonic conjugate* $S(z)$ such that $f(z) = T(z) + i\ S(z)$ is analytic; in fact the level curves of $S$ are just the paths of the heat as it flows orthogonally across the isotherms $T = constant$. Hence, if $T(z)$ is any harmonic function and $h(z)$ any conformal mapping, then $T(z^*)$ is automatically harmonic, with $z^* = h(z)$.

Suppose now that $h(z)$ maps the disk to *itself*. If $z = Re^{i\theta}$ lies on $C$ then so does $z^* = Re^{i\theta^*}$, and, since we suppose that we have measured the temperature all around $C$, we therefore know the temperature $T(\theta^*)$ at $z^*$. Having the values of $T(\theta^*)$, we may now compute the integral in (2) for the harmonic function $T[h(z)]$ to obtain

$$T(0^*) = \frac{1}{2\pi}\int_{-\pi}^{\pi} T(\theta^*)\, d\theta, \tag{3}$$

in which it should be stressed that the averaging is still taking place with respect to the angle of $z$, not its image $z^*$.

We may interpret (3) as follows: To obtain the temperature at $0^*$ we need only take the average of the new temperature distribution on $C$ obtained by transplanting the temperature measured at each $z$ to the new location $z^*$. We are now halfway to Schwarz's result. To find the temperature at $a$ we must find a conformal mapping of the disk to itself such that 0 is sent to $a$, then take the average of the new temperature distribution. The chief surprise will be how simply the new distribution is related to the old one.

In order to find the mapping $h(z)$, we require the basic properties of *inversion* that are illustrated in FIGURE 1(b). Readers familiar with inversive geometry may readily skip to FIGURE 2.

Given a circle $K$ of radius $r$ and center $P$, recall that the point $\tilde{A}$ is the *inverse* of $A$ if it lies in the same direction from $P$ as $A$, and $PA \cdot P\tilde{A} = r^2$. If we consider a second point $B$ and its inverse $\tilde{B}$ then $(PA/PB) = (P\tilde{B}/P\tilde{A})$, and therefore the triangles $PAB$ and $P\tilde{B}\tilde{A}$ are similar. Thus,

$$\textit{the angles } PAB \textit{ and } P\tilde{B}\tilde{A} \textit{ are equal.} \tag{4}$$

It follows easily [exercise] from (4) that if two curves meet at angle $\phi$ in $E$ then their images under inversion in $K$ meet at angle $-\phi$ in $\tilde{E}$. In other words,

$$\textit{inversion is anticonformal.} \tag{5}$$

Before proceeding, remind yourself why it is that *circles map to circles*; this is again [exercise] an easy consequence of (4). Consider a disk such that its boundary circle $C$ cuts $K$ at right angles. Since $C$ is mapped to a circle and $F$ and $G$ remain fixed, it follows from (5) that

$$C \textit{ and its shaded interior are mapped onto themselves.} \tag{6}$$

In particular, the figure shows $z$ being mapped to $\tilde{z}$.

Returning to our original problem, the desired conformal mapping $h(z)$ is now within easy reach. See FIGURE 2. Through 0 and $a$ (at which the temperature is sought) draw the line $L$. Through $a$, draw the line perpendicular to $L$, meeting $C$ in $F$ and $G$ (not shown). Letting $P$ be the intersection of the tangents (not shown) at $F$ and $G$, draw the circle $K$ with center $P$ and radius $PF$. Since $K$ is orthogonal to $C$, (6) says the white disk is mapped to itself under inversion in $K$; furthermore, it is easy to see that the points 0 and $a$ are interchanged by the mapping. The only snag is that, by (5), the mapping is anticonformal rather than conformal. However, if we now *reflect in L* then the angle between two curves will be reversed a second time, thereby returning it to its original state. A viable conformal mapping is therefore

$$h(z) = \textit{inversion in } K\textit{, followed by reflection in } L. \tag{7}$$

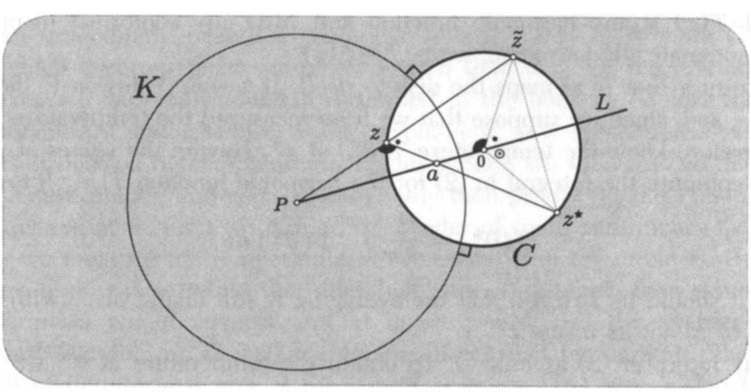



FIGURE 2

For readers with a smattering of hyperbolic geometry (which we will be using shortly) there is a simpler way of looking at $h$. [Excellent introductions to this geometry are [5] and [6].] The intersection point $m \equiv K \cap L$ of the two orthogonal hyperbolic lines $K$ and $L$ is the *midpoint* of the hyperbolic line segment $0a$. Inversion in $K$ corresponds to hyperbolic reflection in $K$, and, just as in Euclidean geometry, successively reflecting across two intersecting lines yields a rotation about their intersection point through double the angle contained by the lines. *Thus $h$ is a rotation of the hyperbolic plane through angle $\pi$ about the midpoint $m$*, making it easy to understand why the ends of the line segment are interchanged.

The geometric key to Schwarz's still unstated result lies in the following splendid fact. Instead of first sending $z$ to $\tilde{z}$ and then reflecting it to $z^*$, we may achieve the same thing in one fell swoop, and without needing $K$, by *projecting $z$ through $a$*. To see this, let us abuse our notation for a moment by defining $z^*$ to be this projected point; we must then show that it is the reflection of $\tilde{z}$ in $L$.

By (4), the similarly marked angles in FIGURE 2 are equal. But the angle subtended at 0 by $\tilde{z}$ and $z^*$ must be double that subtended on the circumference at $z$. The angles at 0 marked • and ⊙ must therefore be equal.

For a different approach, see [3].

We have thus bypassed all calculation and given a direct geometric demonstration of Schwarz's result: *To find the temperature at $a$, transplant each temperature on $C$ to the point directly opposite it as seen from $a$, then take the average of the new temperature distribution on $C$.*

The example in FIGURE 3 illustrates the beauty of this. In FIGURE 3(a), half of $C$ is kept at 100 degrees with steam, while the other half is kept at 0 degrees with ice. Being close to the cold side, we would expect $a$ to be cool. FIGURE 3(b) shows the new temperature distribution obtained by projection through $a$. It is now vividly clear how the distant hot semicircle is 'focused' through $a$ onto a much smaller arc, yielding a low average temperature on $C$ and hence a low temperature at $a$ itself.

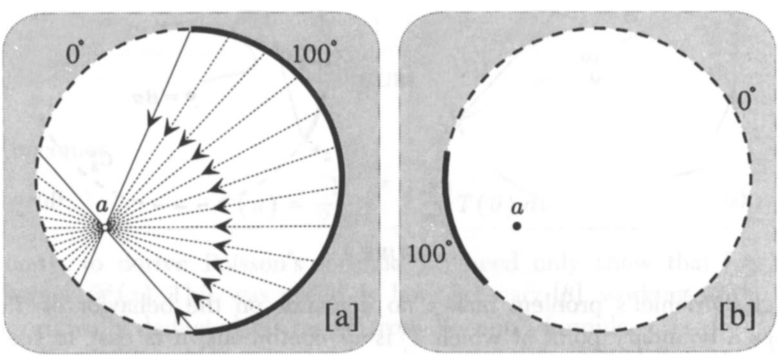



**FIGURE 3**

While we have not found this approach to Schwarz's result elsewhere, it would be surprising if it had been missed. However, let us end this section by pointing out something that initially obscured the issue for the author, and that may have also hindered other writers. Although we have had no need of it, the *formula* for our geometrically natural mapping is

$$h(z) = R^2\left(\frac{z-a}{\bar{a}z - R^2}\right).$$

For some reason, though, the formula that is conventionally used (e.g., [4], p. 167; [7], p. 197) in the calculational proof of (1) is, instead,

$$k(z) = R^2\left(\frac{z+a}{\bar{a}z + R^2}\right).$$

Of course this too has the required property of mapping 0 to $a$, but it is not self-inverse (as $h$ is) and does not send $a$ to 0.

Since the two mappings are related by $k(z)=h(-z)$, we see that $k(z)$ first projects the points of $C$ through 0, then through $a$. See FIGURE 4(a). Short of this figure itself, there now appears to be no simple relationship among the points $z$, $a$, and $k(z)$.

## 3. Dirichlet's Problem for the Disk

Our example in FIGURE 3 was a trifle hasty. For the moment, Schwarz's result merely says how the interior values of a given harmonic function in the disk may be found from the values of $C$. But in FIGURE 3 we blithely assumed that we could also use it to *construct* such a function in the disk, given arbitrary piecewise-continuous boundary values. In other words, we assumed Schwarz's solution of Dirichlet's problem for the disk (outlined in the introduction). We now justify this.

FIGURE 4(b) shows $a$ approaching a boundary point $z$; also shown are the images ($C_1^*$ and $C_2^*$) under projection through $a$ of the two small arcs ($C_1$ and $C_2$) adjacent to $z$. If the given boundary values are continuous at $z$ then $T$ is essentially constant on $C_1 \cup C_2$, and so the new temperature distribution is likewise almost constant on $C_1^* \cup C_2^*$. As required, the constructed function $T(a)$ therefore *does* approach $T(z)$ as $a$ approaches $z$.

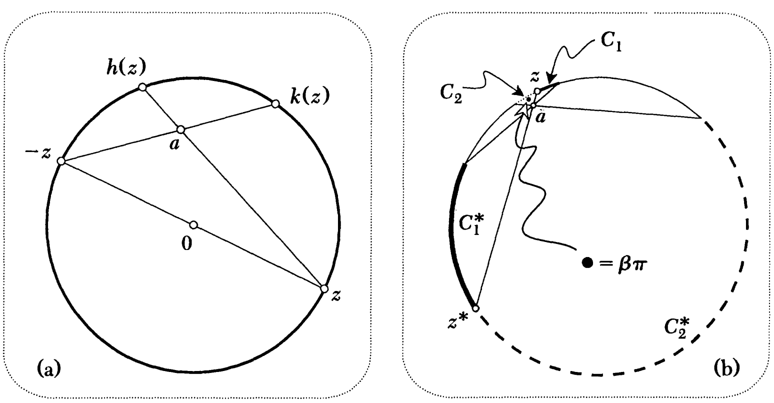



FIGURE 4

Although Dirichlet's problem makes no demands on the behavior of $T(a)$ as $a$ approaches a boundary point at which $T$ is *dis*continuous, it is easy to see (though not to calculate!) what actually happens. Suppose that the boundary temperature jumps from $T_1$ to $T_2$ as we pass from $C_1$ to $C_2$. If $a$ arrives at $z$ while traveling in a direction making an angle $\beta\pi$ with $C_2$, then [exercise] $T(a)$ approaches $[\beta T_1 + (1-\beta)T_2]$. This result is relevant to the representation of discontinuous functions by Fourier series.

It now only remains to show that the constructed function is indeed harmonic. First we shall pause to recover Poisson's formula in its classical form. We begin by noting that (3) may be re-expressed as

$$T(a) = \frac{1}{2\pi}\int_{-\pi}^{\pi} T(\theta)\, d\theta^*. \tag{8}$$

In order to put this into the same form as (1), we now require $d\theta^*$ in terms of $d\theta$. Consider FIGURE 5(a), which shows the movement $R\Delta\theta^*$ of $z^*$ resulting from a movement $R\Delta\theta$ of $z$.

For brevity in this and later arguments, let us employ the following shorthand. If the ratio of $X$ and $Y$ tends to unity as small quantities in a geometric construction tend to zero, we say that $X$ and $Y$ are "ultimately equal", or that $X = Y$ when the small quantities are "infinitesimal." For example, the chord $s$ is ultimately equal to the arc $R\Delta\theta$. It follows from the basic theorems on limits that ultimate equality inherits many of the properties of ordinary equality.

Returning to FIGURE 5(a) the arcs $R\Delta\theta^*$ and $R\Delta\theta$ are ultimately equal to the chords $t$ and $s$, respectively, so $(\Delta\theta^*/\Delta\theta)$ is ultimately equal to $(t/s)$. But $t$ and $s$ are corresponding sides of two similar triangles [shaded], so $(t/s) = (\sigma'/\rho)$. Finally, since $(\sigma'/\rho)$ is ultimately equal to $(\sigma/\rho)$, we obtain

$$\frac{d\theta^*}{d\theta} = \left[\frac{\sigma}{\rho}\right].$$

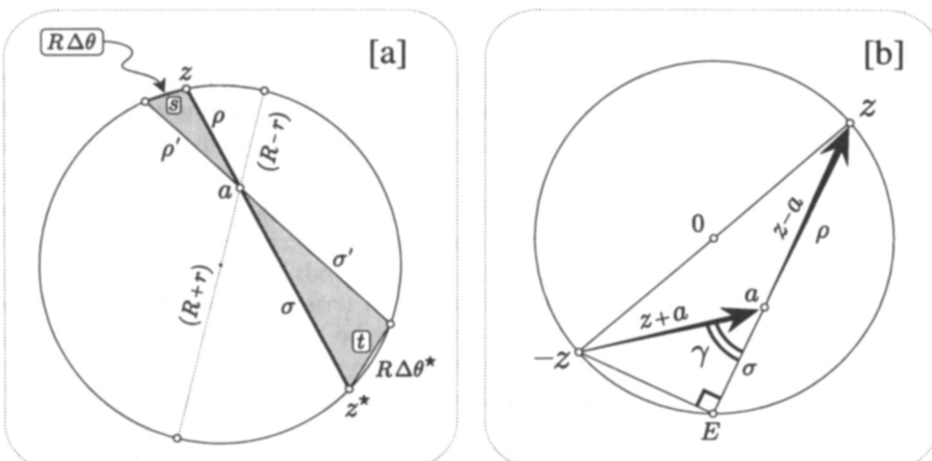



FIGURE 5

Thus (8) becomes

$$T(a) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{\sigma}{\rho}\right]T(\theta)\,d\theta. \tag{9}$$

Consequently, to derive Poisson's formula we need only show that $[\sigma/\rho]$ is the Poisson kernel $\mathscr{P}_a(z)$. This was precisely how Schwarz [8], working in the opposite direction, originally deducted his result *from* Poisson's formula.

Since $\rho\sigma = \rho'\sigma'$ is constant, we may evaluate it for the dotted diameter through $a$ to obtain $\rho\,\sigma = (R^2 - r^2)$, where $r = |a|$. Thus we do indeed find that

$$\left[\frac{\sigma}{\rho}\right] = \left[\frac{R^2 - r^2}{\rho^2}\right] = \mathscr{P}_a(z).$$

As an interesting consequence of the geometric interpretation of the Poisson kernel, we see that (with $z$ fixed) the level curves of $\mathscr{P}_a$ are the circles that are tangent to $C$ at $z$, with $\mathscr{P}_a = 0$ being $C$ itself.

Returning to the issue of harmonicity, we see that if we permit ourselves differentiation under the integral sign of (9), then it is sufficient to show that $[\sigma/\rho]$ is a harmonic function of $a$. To see that it is, consider FIGURE 5(b). Since the angle at $E$ is a right angle, we have

$$\left[\frac{\sigma}{\rho}\right] = \frac{|z+a|\cos\gamma}{|z-a|} = \mathrm{Re}\left(\frac{z+a}{z-a}\right).$$

Because it is the real part of an analytic function of $a$, $[\sigma/\rho]$ is automatically harmonic, and we are done.

This line of reasoning yields a bonus result. Let $S$ be a harmonic conjugate of $T$, so that $f = T + iS$ is an analytic function. This function $f$ is uniquely defined (up to an additive imaginary constant) and so it must be given by

$$f(a) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{z+a}{z-a}\right)T(\theta)\,d\theta,$$

for this is analytic and has $T(a)$ as its real part. This result is called *Schwarz's formula*, and it enables us to resurrect the complete analytic function $f$ from the ashes of its real part on $C$.

## 4. Hyperbolic Geometry

If we specify arbitrary piecewise-continuous temperatures $T(z)$ along the edge (the real axis) of the upper half-plane, then there is another formula due to Poisson that yields the temperature at any point $a = X + iY$ $(Y > 0)$:

$$T(a) = \frac{1}{\pi}\int_{-\infty}^{\infty}\left[\frac{Y}{(X-x)^2+Y^2}\right]T(x)\,dx. \tag{10}$$

We shall explain this result by reinterpreting (8) in terms of elementary hyperbolic geometry. The transition from (1) to (10) will then be seen as nothing more than a transition between the Poincaré and upper half-plane model of the hyperbolic plane. First, however, let us obtain still another geometric interpretation of Poisson's formula.

For simplicity, let us employ the *unit* circle. Consider FIGURE 6. Let the arc $K$ be heated to unit temperature while the rest of $C$ is kept at zero degrees. By Schwarz's result, the temperature at $a$ is $T(a) = (K^*/2\pi)$, while the temperature at the center of the circle is $T(0) = (K/2\pi)$.

Next, imagine yourself standing at $a$, looking out at a vast number of thermometers placed along the circle. As you turn your head through a full revolution (remembering to turn your feet!) let $\langle T\rangle_a$ denote the average (over all directions) of the temperatures you see. For example, Gauss' mean value theorem may be restated as $T(0) = \langle T\rangle_0$.

In FIGURE 6, $\langle T\rangle_a = (\lambda/2\pi)$, where $\lambda$ is the angle subtended by $K$ at $a$. But we see from the figure that

$$\lambda = \tfrac{1}{2}(K^* + K),$$

so $\langle T\rangle_a = \frac{1}{2}[T(a) + T(0)]$: *The average of the boundary temperatures as they appear to you is equal to the average of the temperature where you are and the temperature at the center*. It is then easy to see that this is still true if we instead have many arcs at different temperatures, and ultimately a general piecewise-continuous temperature distribution. Thus Poisson's formula may be re-expressed as

$$T(a) = 2\langle T\rangle_a - T(0).$$

This result is due to Neumann [9]; we merely rediscovered it, as did Duffin [10] from another point of view. For an interesting generalization, see [11].

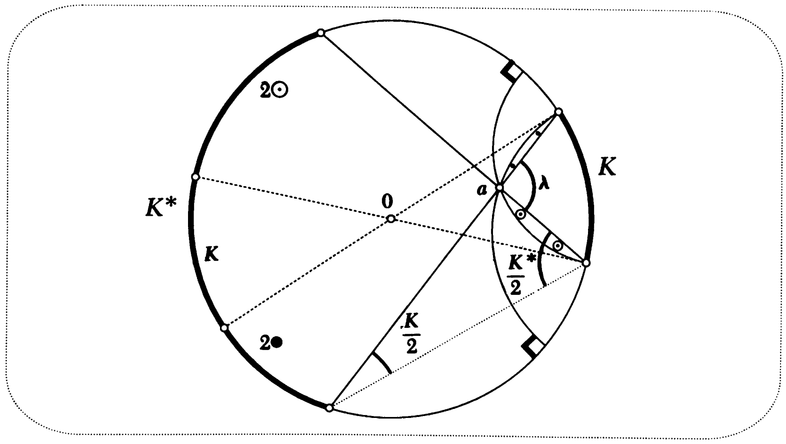



FIGURE 6

FIGURE 7(a) generalizes FIGURE 1(a) and is intended to make this result vivid. Turning your head successively through the same small angle marked • you see the thermometers located at the white dots on the boundary. The average of their temperatures is then a good approximation [exact as • ⇝ 0] to $\langle T \rangle_a$, and hence to the average of the temperature where you stand and the temperature at 0. Note how the white dots became crowded together on the part of the boundary nearest you. As you would expect, this part of the boundary therefore has the greatest influence on the temperature where you stand.

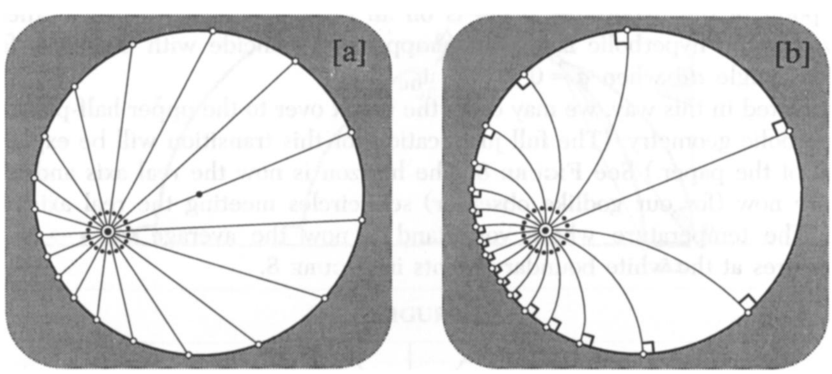



FIGURE 7

To obtain our third and final interpretation of Poisson's formula, let us return to FIGURE 6. Imagine that the disk is the Poincaré model of the hyperbolic plane, and that you are once again standing at the point $a$ looking out to $K$, which is now infinitely far away on the horizon. How big does $K$ appear to you in this distorted geometry? To a godlike observer looking down on this model of the hyperbolic plane, the straight lines along which light travels to you now appear to be arcs of circles orthogonal to $C$, and so *you* see the angular size of $K$ as being

$$\textit{hyperbolic angle} = \lambda + (\bullet + \odot).$$

But we see in the figure that

$$(\bullet + \odot) = \tfrac{1}{2}(K^* - K),$$

and hence we obtain the following remarkable fact:

$$\begin{aligned} \textit{hyperbolic angle} &= \tfrac{1}{2}(K^* + K) + \tfrac{1}{2}(K^* - K) \\ &= K^* \\ &= 2\pi T(a). \end{aligned}$$

The temperature where you are is simply proportional to how big $K$ looks!

Reinterpreting (8), we now see that $d\theta^*$ is simply the hyperbolic angle subtended at $a$ by the element of $C$: *The temperature of each element of $C$ contributes to the temperature at an interior point in proportion to its hyperbolic size as seen from that point.* Much as we did in the Euclidean case, let $\prec T \succ_a$ denote the average of the temperatures you see on the horizon of the hyperbolic plane as you turn your head through a full revolution while standing at $a$. We have found that

$$T(a) = \prec T \succ_a. \tag{11}$$

[Again, we merely rediscovered this: The result (exceeding even the beauty of Schwarz's) is due to Bôcher [12], [13]; the only explicit mention of hyperbolic geometry we have found is Carathéodory's [14].] We have chosen to present (11) as a consequence of Schwarz's result, but at the end of the paper we shall see that it can be understood in a much simpler way.

The analogue of FIGURE 7(a) is now FIGURE 7(b). Standing at the same point as before, and again turning your head successively through the angle •, the figure shows the new locations of the thermometers you see on the boundary. The average of their temperatures is then a good approximation [exact as • ⇝ 0] to $\prec T \succ_a$, and hence to the temperature where you stand. Note how the white dots again become crowded together on the part of the boundary nearest you, so that this part of the boundary has the greatest influence on the temperature where you stand.

From the vantage point of (11), the distinction between (2) and (8) evaporates. Every point of the hyperbolic plane is on an equal footing with every other, it is merely that the hyperbolic angle $d\theta^*$ happens to coincide with the more familiar Euclidean angle $d\theta$ when $a = 0$.

Formulated in this way, we may carry the result over to the upper half-plane model for hyperbolic geometry. (The full justification for this transition will be explained at the end of the paper.) See FIGURE 8. The horizon is now the real axis and 'straight lines' are now (for our godlike observer) semicircles meeting the real axis at right angles. The temperature where you stand is now the average (as • ⇝ 0) of the temperatures at the white boundary points in FIGURE 8.

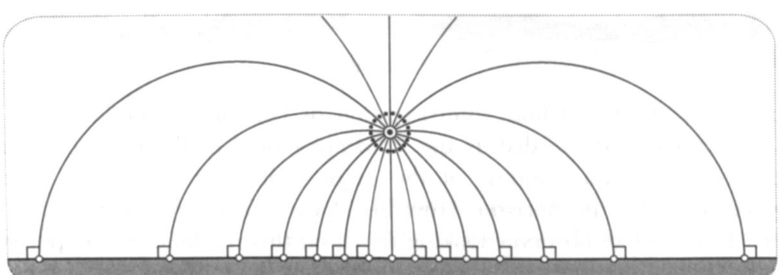


FIGURE 8

FIGURE 9 analyzes this in greater detail. It shows both the hyperbolic angle $\Delta\theta^*$ and the Euclidean angle $\Delta\theta$ subtended at $a$ by the element $\Delta x$ of the horizon. Thinking of $\Delta x$ as sufficiently small that $T(x)$ is essentially constant on it, the contribution to the temperature at $a$ is $(1/2\pi)T(x)\Delta\theta^*$. Integrating along the entire horizon we obtain

$$T(a) = \frac{1}{2\pi}\int_{x=-\infty}^{x=\infty} T(x)\,d\theta^*. \tag{12}$$

In order to put this into precisely the same form as (10), we need to find $(d\theta^*/dx)$. We shall do this via an attractive and rather surprising fact: *The non-Euclidean angle $\Delta\theta^*$ is exactly double the Euclidean angle $\Delta\theta$, even if $\Delta x$ is not small*. To see this, concentrate on the semicircle meeting the axis at $p$. The angle between the dotted tangent at $a$ and the vertical is clearly double that between the chord $ap$ and the vertical. The result then follows immediately.

Now consider FIGURE 10. The small shaded triangle is constructed to be right angled, and it is thus ultimately similar to the large shaded triangle as $\Delta\theta$ shrinks to nothing. Thus $(\xi/\Delta x)$ is ultimately equal to $(Y/\Omega)$. Also, since $\xi$ is like a tiny arc of circle of radius $\Omega$, it is ultimately equal to $\Omega\Delta\theta$. Thus if $\Delta\theta$ is infinitesimal,

$$\frac{\Omega\Delta\theta}{\Delta x} = \frac{\xi}{\Delta x} = \frac{Y}{\Omega}.$$

We can now combine this with the previous result to obtain

$$\frac{d\theta^*}{dx} = 2\left[\frac{d\theta}{dx}\right] = 2\left[\frac{Y}{\Omega^2}\right] = 2\left[\frac{Y}{(X-x)^2+Y^2}\right].$$

Putting this into (12), we obtain (10).

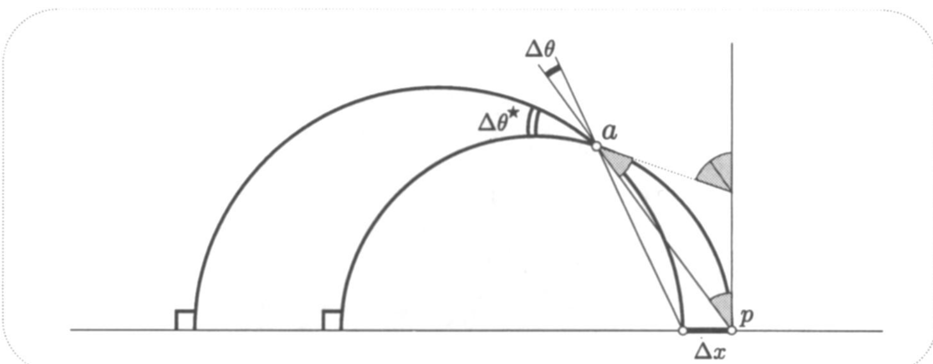



FIGURE 9

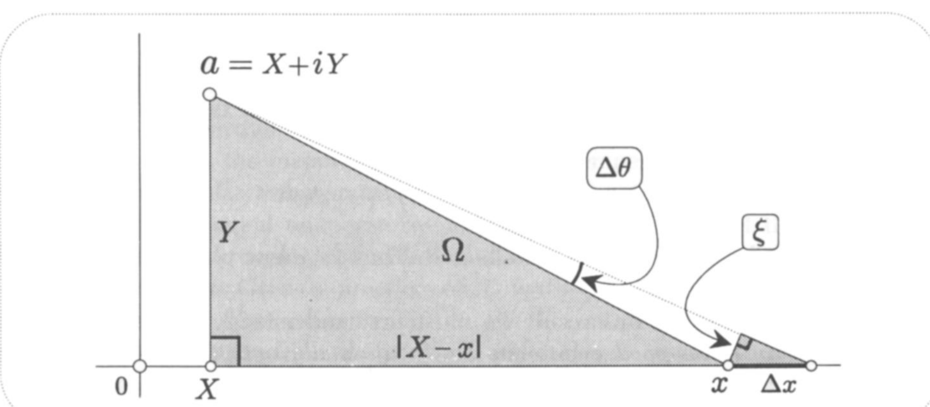



FIGURE 10

While the precise form of the above argument may be new, the basic idea of transferring Bôcher's result from the disk to the half-plane was given by Osgood [15]. For a different but related approach to (10), see [16]. For more on all three of the interpretations thus far obtained, see [11].

## 5. Green's General Formula

Here is the recipe for finding the temperature at the point $a$ inside a simply connected region $R$ in terms of the values $T(z)$ on the boundary $B$.

First, supply heat at the constant rate $2\pi$ to the point $a$ while holding the temperature all round $B$ at the constant value 0. After the heat flow has settled down, the temperature in $R$ will be a well-defined (except at $a$) harmonic function $\mathscr{G}_a(z)$ called the *Green's function* of $R$ with *pole* at $a$. Since $B$ is an isotherm, the heat flow vector $\boldsymbol{H} = -\nabla\mathscr{G}_a$ will be orthogonal to it, and so its magnitude $\mathscr{Q}_a$ (the local heat flux) may be expressed as

$$\mathscr{Q}_a = -\frac{\partial\mathscr{G}_a}{\partial n},$$

where $n$ measures distance in the direction of N, the outward unit normal vector to $B$ (see FIGURE 11).

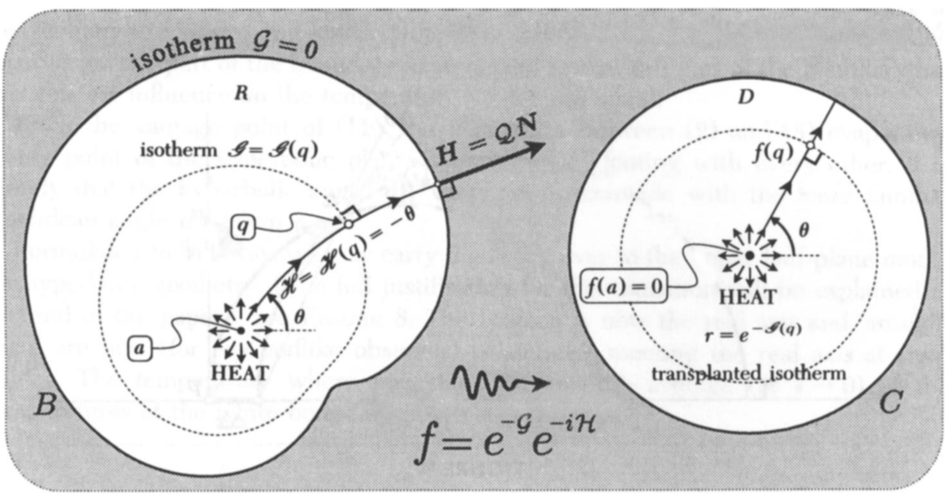



FIGURE 11

Given a harmonic function $T$ on $R$, we may now use $\mathscr{Q}_a$ as a tool with which to find its values inside $R$ in terms of its values $T(z)$ on the boundary $B$. Here is Green's remarkable general formula:

$$T(a) = \frac{1}{2\pi}\oint_B \mathscr{Q}_a(z)T(z)\,ds, \tag{13}$$

where $ds$ is an element of arc length along $B$. Thus $\mathscr{Q}_a$ now plays the same role as the Poisson kernel did in (1).

In order to understand this result we must first understand Riemann's Mapping Theorem: *$R$ may be mapped one-to-one and conformally onto the unit disk $D$*. We shall see in a moment that the existence of such a mapping is equivalent to the existence of the Green's function. In modern texts the mapping theorem is proved independently of physical considerations, the existence of the Green's function then following as a corollary. It was otherwise for Riemann. He seems to have taken the existence of the Green's function to have been guaranteed by Nature, and he was thereby led to his mathematical mapping theorem. The following is one possibility as to how this may have happened. For a better-motivated approach, see [3].

Let us first consider the behavior of the function $\mathscr{G}_a$ in the immediate vicinity of $a$. Physical intuition leads us to expect, irrespective of the temperatures assigned to $B$, that heat will flow out of the source *symmetrically*: $\mathscr{G}_a$ and $\mathscr{Q}_a$ will be nearly constant on a tiny circle $K$ centered there. If the flow were *perfectly* symmetrical, and if the radius of $K$ were $\rho$, then

$$2\pi = \textit{heat supplied to } a = \textit{flux across } K = 2\pi\rho\mathscr{Q}_a = -2\pi\rho\frac{\partial\mathscr{G}_a}{\partial\rho}.$$

Thus, very close to $a$, the temperature behaves like $-\ln\rho$. The precise statement is that

$$\mathscr{G}_a = -\ln\rho + g_a, \tag{14}$$

where $g_a$ is harmonic throughout $R$.

Let $\mathscr{H}_a$ be a harmonic conjugate of $\mathscr{G}_a$, so that (except at $a$) $\mathscr{F} \equiv \mathscr{G}_a + i\mathscr{H}_a$ is conformal. Since $\ln\rho$ is the real part of $\log(z-a)$ it follows from (14) that $\mathscr{F}(z) = -\log(z-a) + F$, where $F$ is conformal throughout $R$. Since the imaginary part of $F$ is only determined up to a constant, we may choose $\operatorname{Im} F(a) = 0$. Thus, very close to $a$,

$$\mathscr{H}_a = -\arg(z-a).$$

The advantage of this particular choice is that we may then interpret the value of $\mathscr{H}_a$ at a typical point $q$ in the simple manner illustrated in FIGURE 11. Follow the flow of heat back from $q$ to $a$; the angle at which it enters $a$ is then $-\mathscr{H}_a(q)$. We thus have clear physical interpretations for both the real and imaginary parts of the mapping $\mathscr{F}$.

With $\mathscr{F}$ in hand, we are now close to Riemann's theorem. We require a mapping that is conformal throughout $R$, but while the mapping $\mathscr{F}$ is otherwise conformal, it has a *logarithmic singularity* at $a$. In order to undo this singularity, we are therefore *forced* to compose $\mathscr{F}$ with the exponential mapping. The mapping $f$ so obtained is then conformal everywhere in $D$:

$$z \rightsquigarrow w = f(z) = e^{-\mathscr{F}(z)} = e^{-\mathscr{G}_a}e^{-i\mathscr{H}_a}.$$

As illustrated in FIGURE 11, and as was desired, $f$ maps $R$ conformally to $D$: The pole at $a$ maps to 0; the dashed isotherm at temperature $\mathscr{G}(q)$ maps to the dashed circle of radius $e^{-\mathscr{G}_a(q)}$; the streamline $\mathscr{H}_a = \mathscr{H}_a(q) = -\theta$ entering $a$ at angle $\theta$ maps to the ray entering 0 at angle $\theta$.

We make a few further observations before returning to the explanation of (13). Now that we possess the mapping $f$, any harmonic temperature distribution $T(z)$ on $R$ may be *conformally transplanted* to a harmonic function $\tilde{T}(w)$ on $D$ (and vice versa) by assigning equal temperatures to corresponding points of the two regions $\tilde{T}[f(z)] \equiv T(z)$. In particular, the values of $T$ on $B$ are transplanted to $C$.

Next, consider the Green's function of $D$ with pole at 0. See FIGURE 11. On grounds of symmetry, the isotherms must be circles centered at 0, and the streamlines must be rays emanating from there. More precisely, it should be clear that if $w$ denotes a point of $D$ then the Green's function is $-\ln|w|$. But this means [make sure you can see this] that each point of $D$ is at the same temperature as its preimage in $R$; in other words, $f$ conformally transplants the Green's function $\mathscr{G}_a(z)$ of $R$ with pole at $a$ to the Green's function of $D$ with pole at $f(a) = 0$.

This is no accident. More generally, let $J(z)$ be a one-to-one conformal mapping of $R$ to some other simply connected region $S$ with boundary $Y$. Then $J$ conformally transplants the Green's function $\mathcal{G}_a(z)$ of $R$ with pole at $a$ to the Green's function of $S$ with pole at $J(a)$. In particular, the streamlines of the flow in $R$ map to the streamlines of the flow in $S$. In this sense, *the concept of the Green's function is conformally invariant.*

The explanation is not difficult. Since $\mathcal{G}_a$ is harmonic except at $a$, its transplant is harmonic except at $J(a)$. Also since $\mathcal{G}_a$ vanishes on $B$, its transplant vanishes on $Y$. To complete the proof, we must show that the source at $a$ transplants to a source of equal strength at $J(a)$. To see this, recall that the local effect of an analytic function $J$ is an expansion by $|J'|$ and a rotation of $arg(J')$. Geometrically, the source of strength $2\pi$ at $a$ is characterized by the fact that the radii of the infinitesimal circular isotherms round $a$ are proportional to $e^{-temperature}$. The mapping $J$ merely expands these by $|J'(a)|$ to produce infinitesimal circular isotherms round $J(a)$ having (by definition) the same temperatures as the originals. Thus the radii of the transplanted isotherms round $J(a)$ are again proportional to $e^{-temperature}$, so we have a source of strength $2\pi$ at $J(a)$. Done.

We are now in a position to explain the general formula in a beautifully simple way. To begin with, imagine that $T(z)$ is a given harmonic function in $R$ whose value $T(a)$ at an interior $a$ we wish to determine from the boundary values. See FIGURE 12. Just as in FIGURE 7(a), imagine standing inside $R$ at $a$ and turning your head successively through the small angle •. But now *suppose that light travels along the illustrated streamlines of the heat flow* $\boldsymbol{H} = -\nabla\mathcal{G}_a$ *associated with the Green's function*. You would then see the thermometers at the illustrated points on the boundary.

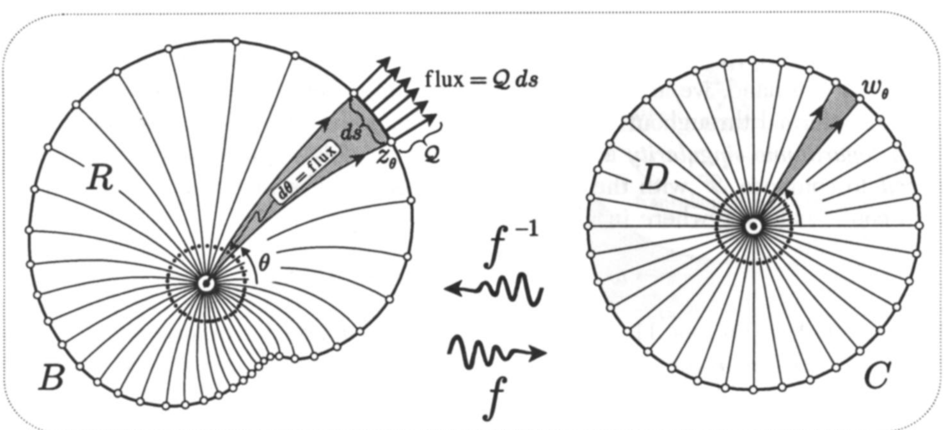



FIGURE 12

The key observation is that (even without passing to the limit of vanishing •) *the average of the observed temperature is conformally invariant*. As before, let $J(z)$ be a one-to-one conformal mapping of $R$ to some other simply connected region $S$ with boundary $Y$. Just as we did with $f$, let us choose $J$ so that the directions of curves through $a$ are preserved (i.e., $\arg[J'(a)] = 0$). Let $z_\theta$ denote the point on $B$ that you see when you look in the direction $\theta$, and let $w_\theta \equiv J(z_\theta)$ be its image on $Y$.

By the conformal invariance of the Green's function, the image of the streamline leaving $a$ at angle $\theta$ is the streamline leaving $J(a)$ at the same angle. Thus $w_\theta$ is not only the image of $z_\theta$, it is also the boundary point which an observer at $J(a)$ sees when looking in the direction $\theta$. But, by definition, the temperature at each point $z_\theta$

on $B$ is transplanted to $w_\theta$ on $Y$, so the observer $J(a)$ sees exactly the same temperatures on $Y$ as the original observer at $a$ saw on $B$.

Passing to the limit of vanishing •, the conformal invariance of this average may be expressed as

$$\frac{1}{2\pi}\oint_B T(z_\theta)\,d\theta=\frac{1}{2\pi}\oint_Y \tilde{T}(w_\theta)\,d\theta.$$

FIGURE 12 illustrates the particular case where $J=f$ is the previously constructed function that maps $R$ to $D$ and $a$ to 0. The virtue of this special case is that the conformally invariant average may now be *evaluated*. By Gauss' mean value theorem, the average of $\tilde{T}(w_\theta)\equiv T(z_\theta)$ on $C$ is the temperature $\tilde{T}(0)\equiv T(a)$ at the center:

$$\frac{1}{2\pi}\oint_B T(z_\theta)\,d\theta=\frac{1}{2\pi}\oint_C \tilde{T}(w_\theta)\,d\theta=T(a).$$

Thus, returning to the left-hand side of FIGURE 12, we have found that in the limit of vanishing • *the average of the temperatures you see on the boundary of R is the temperature where you stand!*

It now remains to show that this geometric construction is equivalent to the conventional formula (13). On the left of FIGURE 12, consider the shaded region—let us call it a "tube"—between two of the streamlines leaving $a$ with infinitesimal angular separation $d\theta$, and let the tube intercept $B$ at $z_\theta$ in an element of length $ds$. Since $2\pi$ of heat flux emerges symmetrically from $a$, the flux emitted into the tube is equal to $d\theta$. Furthermore, since $\mathscr{G}_a$ is harmonic, no heat is created or destroyed in the tube, so all the heat that enters the tube at the source must emerge at the other end. Thus $d\theta=\mathscr{Q}_a(z_\theta)\,ds$, and

$$T(a)=\frac{1}{2\pi}\oint_B T(z_\theta)\,d\theta=\frac{1}{2\pi}\oint_B \mathscr{Q}_a(z)T(z)\,ds,$$

as was to be shown. Surprisingly, we have not found this interpretation and explanation elsewhere in the literature. We hope you will agree that it contrasts strikingly with the conventional approach (c.f. [7], p. 209).

This line of reasoning also explains the stronger result that (13) solves Dirichlet's problem for $R$. Using $f$ to conformally transplant the given boundary values from $B$ to $C$, we know that Poisson's formula allows us to construct the solution to Dirichlet's problem in $D$. Transferring this solution back from $D$ to $R$ with $f^{-1}$, we have found the harmonic function $T$ in $R$, and its value at $a$ must then be given by (13). You can now understand why we lavished so much attention on the special case of the disk.

We can also use this conformally invariant average to understand much of what has gone before in a simple, unified way. For example, the average $\prec T \succ_a$ depicted in FIGURE 7(b) is merely the special case in which $R$ is the unit disk. To see this, recall (7). The conformal mapping $h(z)$ of $D$ to itself interchanges 0 and $a$ and maps circles to circles. It follows [why?] that it interchanges FIGURE 7(b) and FIGURE 1(a). Similarly by employing an inversion in a circle centered on the real axis it is possible (see [5]) to construct an analogous (conformal and circle-preserving) mapping from the upper half-plane to the unit disk. It follows that the result in FIGURE 8 is also a special case.

Although we hope you will agree that this is all delightfully intuitive, one could still wish for an explanation of (13) that dealt directly with $R$ rather than requiring the assistance of the disk. It does not appear to be widely known, but such an explanation is indeed possible. Working with electricity rather than heat, James Clerk Maxwell [17] was able to give a direct explanation of (13) by arguing in terms of electrostatic energy. But because the required reasoning is *purely* physical, we shall not enter into this matter here.

**Acknowledgments.** I should like to thank Prof. Roger Penrose for suggesting future lines of investigation in this area, and Dr. George Burnett-Stuart for explaining to me Maxwell's splendid work on this subject. I also thank the Faculty Development Committee of the University of San Francisco for making possible the trip to Oxford during which the conversations with the above gentlemen took place.

## REFERENCES

1. W. Zimmermann, S. Cunningham (Editors), *Visualization in Teaching and Learning Mathematics*, MAA Notes 19, Washington, DC, 1991.
2. T. Needham, Newton and the transmutation of force, *Amer. Math, Monthly*, **100** (1993), 119–137.
3. T. Needham, *Visual Complex Analysis*, Oxford University Press, Oxford, (to appear).
4. L. Ahlfors, *Complex Analysis*, 3rd edition, McGraw-Hill Book Co., New York, 1979.
5. J. Stillwell, *Geometry of Surfaces*, Springer-Verlag, New York Inc., New York 1992.
6. W. Thurston, *Three-Dimensional Geometry and Topology*, draft available from the Geometry Center, University of Minnesota.
7. R. Boas, *Invitation to Complex Analysis*, Random House, Inc., New York, 1987.
8. H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, Vol. 2, Chelsea Publishing Co., Inc., New York, 1972, p. 360.
9. C. Neumann, *Vorlesungen über Riemann's Theorie der Abel'schen Integrale*, 2nd edition, B G. Teubner, 1884, p. 410.
10. R. J. Duffin, A note on Poisson's integral, *Quarterly of Applied Math*., **15** (1957), 109–111.
11. F. W. Perkins, An intrinsic treatment of Poisson's integral, *Amer. Jour. of Math*., **50** (1928), 389–414.
12. M. Bôcher, Note on Poisson's Integral, *Bull. Amer. Math. Soc*., **4** (1898), 424–426.
13. M. Bôcher, On harmonic functions in two dimensions, *Proc. Amer. Acad. Arts & Sci*., **41** (1906), 577–583.
14. C. Carathéodory, *Theory of Functions of a Complex Variable*, Vol. 1, Chelsea Publishing Co., 1954, p. 156.
15. W. F. Osgood, *Lehrbuch der Funktionentheorie*, 5th edition, Vol. 1, Chelsea Publishing Co., 1928, pp. 666–670.
16. R. Lange, R. A. Walsh, A heuristic for the Poisson integral for the half-plane and some caveats, *Amer. Math. Monthly*, **92** (1985), 356–358.
17. J. C. Maxwell, *A Treatise on Electricity and Magnetism*, 3rd edition, Vol. 1, Dover Publications, Mineola, NY, 1954, pp. 103–107.

---

## 50 YEARS AGO...

Grace Chisholm Young died just before she was to receive an honorary degree from the Fellows of Girton College, Cambridge. Grace Chisholm was born near London in 1868, was educated at home, and entered Girton, the first English institution to allow women to receive a university education. She attained a superior score on the Cambridge Tripos and continued her education at Göttingen, earning her Ph.D. in 1895—the first woman to receive a German doctorate in mathematics through the regular procedure!

from Victor J. Katz, *A History of Mathematics*, Harper Collins, 1993.

# NOTES

## Introducing... Cwatsets!

GARY J. SHERMAN*
Rose-Hulman Institute of Technology
Terre Haute, IN 47803

MARTIN WATTENBERG*
Brown University
Providence, RI 02912

**Introduction** They lurk in $\mathbf{Z}_2^n$—more than a subset, at most a subgroup. Here's one:

$$\begin{aligned} T &= \{000, 110, 101\}, \\ T + 110 &= \{110, 000, 011\} = \{000, 110, 011\} = T^{(1,2)}, \\ T + 101 &= \{101, 011, 000\} = \{000, 011, 101\} = T^{(1,3)}. \end{aligned} \tag{1}$$

In this array, $T + 110$ and $T + 101$ are right cosets of $T$ by 110 and 101, respectively, and $T^{(1,2)}$ and $T^{(1,3)}$ are results of transposing the indicated components of each element of $T$. While $T$ is not a subgroup, it is 'structured': Each coset of $T$ by an element of $T$ can be obtained by rearranging the components of the elements of $T$. What is $T$? It's a *cwatset*—a term that we are hereby coining.

To formalize our definition of a cwatset, we need to make precise the notation $T^{(1,2)}$ and $T^{(1,3)}$. Let $\mathbf{b} = b_1 b_2 \dots b_n$ be an element of $\mathbf{Z}_2^n$, let $\sigma$ be an element of $S_n$ (the symmetric group on $n$ symbols) and let $j = i\sigma$. Then $\mathbf{b}^\sigma$ is the element of $\mathbf{Z}_2^n$ whose $j$-th component is the $i$-th component of $\mathbf{b}$.; i.e.,

$$\mathbf{b}^\sigma = b_{1\sigma^{-1}} b_{2\sigma^{-1}} \dots b_{j\sigma^{-1}} \dots b_{n\sigma^{-1}}.$$

Informally speaking, $\mathbf{b}^\sigma$ is the result of rearranging the components of $\mathbf{b}$ according to the definition of $\sigma$. For example, $(11010)^{(1,5,4)(2,3)} = 10101$ because the action of the permutation $(1,5,4)(2,3)$ is read from left to right. Of course, a permutation acts on a subset of $\mathbf{Z}_2^n$ by acting on each element of the subset; i.e.,

$$B^\sigma = \{\mathbf{b}^\sigma | \mathbf{b} \in B\}.$$

*Definition.* A subset, $C$, of $\mathbf{Z}_2^n$ is a *cwatset* if for each element, $\mathbf{c}$, of $C$, there exists a permutation, $\sigma$, of $S_n$ such that $C + \mathbf{c} = C^\sigma$; i.e., $C$ is **c***losed* **w***ith* **a** **t***wist*.

Cwatsets are interesting combinatorial objects—or so we hope to convince you with this paper. Indeed, our purpose is to suggest the 'Theory of Cwatsets' as a research topic suitable for undergraduate students taking combinatorics, algebra or discrete mathematics courses.

**Notation** As in the introduction, $\mathbf{b} = b_1 b_2 \dots b_n$ denotes an element of $\mathbf{Z}_2^n$. In

*Work supported by National Science Foundation grant DMS-8922674

particular, $\mathbf{0} = 00\ldots0$ and $\mathbf{1} = 11\ldots1$. Lower-case Greek letters $(\sigma, \pi, \tau)$ denote elements of $S_n$. Obviously, $\mathbf{0}^\sigma = \mathbf{0}$ and $\mathbf{1}^\sigma = \mathbf{1}$. The 'action' of $\sigma$ on $\mathbf{Z}_2^n$ is clearly one-to-one and onto. Moreover, since addition in $\mathbf{Z}_2^n$ is done component-wise, it is easy to check that

$$(\mathbf{x} + y)^\sigma = \mathbf{x}^\sigma + \mathbf{y}^\sigma.$$

Thus, depending on one's point of view, $\sigma$ determines an automorphism of the group $\mathbf{Z}_2^n$ or an invertible linear transformation of the vector space $\mathbf{Z}_2^n$ of dimension $n$ over the finite field $\mathbf{Z}_2$. We will trade on the group-theoretic version later.

The number of ones appearing in the binary representation of $\mathbf{b}$ is referred to as the *weight* of $\mathbf{b}$ and is denoted by $w(\mathbf{b})$. Notice that $w(\mathbf{x} + \mathbf{y})$ and $w(\mathbf{x}) + w(\mathbf{y})$ have the same parity,

$$w(\mathbf{x} + \mathbf{y}) \equiv w(\mathbf{x}) + w(\mathbf{y}) \pmod 2, \tag{2}$$

and that

$$w(\mathbf{b}^\sigma) = w(\mathbf{b}). \tag{3}$$

Notice also that $w(\mathbf{x} + \mathbf{y})$ is just the number of places in which $\mathbf{x}$ and $\mathbf{y}$ differ. This is commonly referred to as the *distance*, $d(\mathbf{x}, \mathbf{y})$, between $\mathbf{x}$ and $\mathbf{y}$.

**Statistically speaking** Cwatsets are useful in statistics. Indeed, the defining condition for a cwatset is motivated by the fact that it guarantees that a cwatset may be used to construct confidence intervals for the location parameter (mean or median) of a continuous, symmetric random variable [2]. To illustrate, let $T$ be as in (1), let $\mathbf{Y}$ be a continuous random variable, symmetric about $\mu$, and let $\mathbf{Y}_1$, $\mathbf{Y}_2$, and $\mathbf{Y}_3$ be a random sample from $\mathbf{Y}$. A subsample mean may be associated with each nontrivial element of $T$:

$$\mathbf{M}_{110} = (\mathbf{Y}_1 + \mathbf{Y}_2)/2 \quad \text{and} \quad M_{101} = (\mathbf{Y}_1 + \mathbf{Y}_3)/2.$$

The confidence intervals for $\mu$ are constructed from the ordered listing of these subsample means. In this example

$$\begin{gathered}(-\infty, \min\{\mathbf{M}_{110}, \mathbf{M}_{101}\})\\ (\min\{\mathbf{M}_{110}, \mathbf{M}_{101}\}, \max\{\mathbf{M}_{110}, \mathbf{M}_{101}\}).\\ (\max\{\mathbf{M}_{110}, \mathbf{M}_{101}\}, \infty)\end{gathered}$$

are each $100/3\%$ confidence intervals for $\mu$. The construction of confidence intervals using cwatsets generalizes the discussion of Hartigan's group-theoretic approach to resampling [3] that recently appeared in this MAGAZINE [6].

**Cwatsets are codes** We opened with group-theoretic jargon and will eventually suggest that the best way to think of, and generalize, cwatsets is in terms of groups. Yet, a coding theorist's version of our opening sentence might be, "They lurk in $\mathbf{Z}_2^n$—more than a nonlinear code, at most a linear code." Of course, a linear code is just a subgroup of $\mathbf{Z}_2^n$ and a nonlinear code is just a subset of $\mathbf{Z}_2^n$ ([4]).

Well, almost 'just' a subset of $\mathbf{Z}_2^n$. Coding theorists like to assume, as a matter of convention, that $\mathbf{0}$ is an element of any nonlinear code $N$. This is because their main concern is with $d(\mathbf{x}, \mathbf{y})$. In general, for a given number of codewords (elements of $\mathbf{Z}_2^n$), a coding theorist wants the minimum value of $d(\mathbf{x}, \mathbf{y})$, over all pairs of words in the code, to be as large as possible. Since

$$\begin{aligned} d(\mathbf{x}+\mathbf{c}, \mathbf{y}+\mathbf{c}) &= w(\mathbf{x}+\mathbf{c}+\mathbf{y}+\mathbf{c}) \\ &= w(\mathbf{x}+\mathbf{y}) \\ &= d(\mathbf{x}+\mathbf{y}), \end{aligned}$$

this minimum value, indeed the distribution, of $d(\mathbf{x}, \mathbf{y})$ is unchanged if $N$ is replaced by $N+\mathbf{c}$. So, if $\mathbf{0}$ is not an element of $N$, just replace $N$ with $N+\mathbf{c}$ where $\mathbf{c}$ is any element of $N$. No matter, any cwatset contains $\mathbf{0}$—a fact that we will prove a bit later.

The distribution of $d(\mathbf{x}, \mathbf{y})$ is also invariant under the action of $\sigma$:

$$\begin{aligned} d(\mathbf{x}^\sigma, \mathbf{y}^\sigma) &= w(\mathbf{x}^\sigma + \mathbf{y}^\sigma) \\ &= w\big((\mathbf{x}+\mathbf{y})^\sigma\big) \\ &= w(\mathbf{x}+\mathbf{y}) \\ &= d(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Thus, from a coding theorist's point of view, $N$ and $(N+c)^\sigma$ are essentially the same. More precisely, two nonlinear codes $M$ and $N$ are *equivalent* if there exist $\tau \in S_n$ and $\mathbf{c} \in N$ such that $M = (N+\mathbf{c})^\tau$. It is an exercise in notation to verify that we have defined an equivalence relation. For instance,

$$M = (N+\mathbf{c})^\tau \text{ if, and only if, } N = (M+\mathbf{c}^\tau)\tau^{-1}.$$

Notice that our statistically motivated definition of a cwatset is hidden in this definition of equivalence for nonlinear codes. If $C$ (remember, $C$ is a cwatset) and $M$ are equivalent as codes, then $M = (C+\mathbf{c})^\tau$, where $\mathbf{c} \in C$, which implies $M = (C^\sigma)^\tau = C^{\sigma\tau}$ for some $\sigma$; i.e., the equivalence class of $C$ is made up of 'rearrangements' of $C$. Conversely, suppose the only codes equivalent to $N$ are 'rearrangements' of $N$. Then, for each $\mathbf{c} \in N$ and any $\sigma$, there exists $\tau$ such that $N^\tau = (N+\mathbf{c})^\sigma$. This implies $N^{\tau\sigma^{-1}} = N+\mathbf{c}$; i.e., $N$ is a cwatset. Thus, we could have defined a cwatset to be a nonlinear code whose equivalence class consists of 'rearrangements'.

**Examples, constructions, and results** The restriction of the coding definition of equivalence to cwatsets is an equivalence relation on cwatsets and the equivalence class of $C$ is just

$$[C] = \{C^\tau | \tau \in S_n\}.$$

Notice that $\{C + c | c \in C\}$ is a subset of $[C]$ due to the very definition of cwatset. Fortunately,

FACT 1. *Each element of* $[C]$ *is a cwatset.*

*Proof.* For $\mathbf{c} \in C$, $C^\tau + \mathbf{c}^\tau = (C+\mathbf{c})^\tau = (C^\sigma)^\tau = (C^\tau)^{\tau^{-1}\sigma\tau}$.

For example, the equivalence class of the cwatset $T$ (see (1)) is

$$\{\{000, 110, 101\}, \{000, 011, 110\}, \{000, 101, 011\}\}.$$

Thus, given a cwatset it is easy to construct others that are essentially the same; i.e., equivalent. Where do we get that 'given' cwatset?

FACT 2. *Any subgroup of* $\mathbf{Z}_2^n$ *is a cwatset.*

*Proof.* A subgroup is closed under addition so that the trivial 'twist' is sufficient to close it.

But, the set of subgroups of $\mathbf{Z}_2^n$ is not the only source of cwatsets because the cwatset $T$ is not a subgroup of $\mathbf{Z}_2^n$. Here's how to construct $T$. Choose $\mathbf{b} = 110$ and $\sigma = (1,2,3)$ and obtain a sequence of elements from $\mathbf{Z}_2^3$ as follows;

$$\mathbf{b}_1 = \mathbf{b} = 110,$$

$$\mathbf{b}_2 = (\mathbf{b}_1)^\sigma + \mathbf{b}_1 = (110)^{(1,2,3)} + 110 = 011 + 110 = 101,$$

$$\mathbf{b}_3 = \left((\mathbf{b}_1)^\sigma + \mathbf{b}_1\right)^\sigma + \mathbf{b}_1 = (\mathbf{b}_1)^{\sigma^2} + (\mathbf{b}_1)^\sigma + \mathbf{b}_1 = (101)^{(1,2,3)} + 110 = 110 + 110$$
$$= 000,$$

Thus, $T = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

We will refer to $T$ as the *cyclic cwatset* generated by $\sigma$ and $\mathbf{b}$ and we will write

$$\begin{aligned} T &= C(\sigma, \mathbf{b}) \\ &= \left\{\mathbf{b}_k | \mathbf{b}_1 = \mathbf{b} \textit{ and } \mathbf{b}_k = (\mathbf{b}_{k-1})^\sigma + \mathbf{b} \textit{ for } k \geq 2\right\} \\ &= \left\{\mathbf{b}^{\sigma^k} + \cdots + \mathbf{b}^\sigma + \mathbf{b} | k \geq 0\right\}. \end{aligned}$$

Here is another cyclic cwatset in $\mathbf{Z}_2^3$:

$$S = C((1,2,3), 100) = \{100, 110, 111, 011, 001, 000\}.$$

Our reference to $T$ and $S$ as cwatsets is justified:

PROPOSITION 1. *Let $\sigma \in S_n$ and let $\mathbf{b} \in \mathbf{Z}_2^n$. $C = C(\sigma, \mathbf{b}) =$ is a cwatset.*

*Proof.* First, notice that $C(\sigma, \mathbf{b}) = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ because $\mathbf{Z}_2^n$ is finite. Moreover, we may take $m$ to be the least-positive integer such that $(\mathbf{b}_m)^\sigma + \mathbf{b} = \mathbf{b}$; i.e., $\mathbf{b}_m = \mathbf{0}$. This follows because if $(\mathbf{b}_m)^\sigma + \mathbf{b} = \mathbf{b}_p$ for some $p$ such that $1 < p < m$, then

$$\begin{aligned} \left((\mathbf{b}_{m-1})^\sigma + \mathbf{b}\right)^\sigma + \mathbf{b} &= (\mathbf{b}_{p-1})^\sigma + \mathbf{b}, \\ \left((\mathbf{b}_{m-1})^\sigma + \mathbf{b}\right)^\sigma &= (\mathbf{b}_{p-1})^\sigma, \\ (\mathbf{b}_{m-1})^\sigma + \mathbf{b} &= \mathbf{b}_{p-1}, \end{aligned}$$

which contradicts the minimality of $m$. In fact, it is clear that

$$\mathbf{b}_i = \mathbf{b}_j \text{ if, and only if, } i \equiv j \pmod{m}.$$

Thus, for $1 \leqslant k \leqslant m$,

$$\begin{aligned} C + \mathbf{b}_k &= \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\} + \mathbf{b}_k \\ &= \{\mathbf{b}_{k+1}, \mathbf{b}_{k+2}, \dots, \mathbf{b}_{k+m}\} + \mathbf{b}_k \\ &= \left\{(\mathbf{b}_1)^{\sigma^k}, (\mathbf{b}_2)^{\sigma^k}, \dots, (\mathbf{b}_m)^{\sigma^k}\right\} \\ &= C^{\sigma^k}. \end{aligned}$$

PROPOSITION 2. *The order of a cyclic cwatset divides twice the group-theoretic order of the generating permutation.*

*Proof.* Let $|\langle\sigma\rangle| = r$ and let $|C(\sigma, \mathbf{b})| = m$. Then,

$$\mathbf{b}_{2r} = \sum_{i=0}^{2r-1} \mathbf{b}^{\sigma^i}$$

$$= \sum_{i=0}^{r-1} \mathbf{b}^{\sigma^i} + \sum_{i=r}^{2r-1} \mathbf{b}^{\sigma^i}.$$

However, since $\sigma^i = \sigma^{i+r}$ (the order of $\sigma$ is $r$),

$$\sum_{i=0}^{r-1} \mathbf{b}^{\sigma^i} = \sum_{i=r}^{2r-1} \mathbf{b}^{\sigma^i}.$$

Thus, $\mathbf{b}_{2r} = 0$, which implies $2r \equiv 0 \pmod m$; i.e., $m$ divides $2r$.

We note in passing that the constructions of $T$ and $S$ generalize to show that $\mathbf{Z}_2^n$ always contains cyclic cwatsets of orders $n$ and $2n$.

Now for a noncyclic cwatset, say $D$:

| | | |
|---|---|---|
| 000100 | 101100 | 110100 |
| 000110 | 101110 | 110110 |
| 000111 | 101111 | 110111 |
| 000011 | 101011 | 110011 |
| 000001 | 101001 | 110001 |
| 000000 | 101000 | 110000. |

$D$ is a cwatset because it is the 'direct sum' of the cwatsets $T$ and $S$: $D = T \oplus S$. A formal definition of 'direct sum' is more notational clutter than content. It suffices to say that the direct sum of cwatsets uses their cartesian product as the underlying set and that addition of binary $n$-tuples and the action of permutations on these $n$-tuples is done in the 'appropriate' component-wise manner. Let's illustrate.

$$\begin{aligned} T \oplus S + 101001 &= (T + 101) \oplus (S + 011) \\ &= T^{(1,2,3)^2} \oplus S^{(1,2,3)^4} \\ &= T^{(1,3,2)} \oplus S^{(1,2,3)} \\ &= (T \oplus S)^{(1,3,2)(1,2,3)'} \\ &= (T \oplus S)^{(1,3,2)(4,5,6)}. \end{aligned}$$

The notation $(1,2,3)'$ means the permutation $(1,2,3)$ translated to the appropriate components of the direct sum. In this case, $(1,2,3)' = (4,5,6)$.

$D$ is not a cyclic cwatset. Suppose to the contrary that $D = C(\sigma, \mathbf{b})$ for some $\sigma \in S_6$ and some $\mathbf{b} \in Z_2^6$. Then, Proposition 2 implies that 18 divides $2|\langle\sigma\rangle|$, which contradicts the fact that the maximum order of an element of $S_6$ is six.

QUESTION. Can any cwatset, $C$, in $\mathbf{Z}_2^n$ be expressed as

$$C = (H \oplus C_1 \oplus C_2 \oplus \cdots \oplus C_m)^\sigma$$

where $H$ is a subgroup and the $C_i$'s are cyclic cwatsets? We have yet to find a cwatset that is not of this form.

What about special elements in cwatsets? The presence of inverses in any subset of $\mathbf{Z}_2^n$ is guaranteed by the underlying structure ($\mathbf{b} + \mathbf{b} = \mathbf{0}$ implies $\mathbf{b} = -\mathbf{b}$) and has nothing to do with the definition of cwatset. But, as promised earlier and as illustrated by $S$, $T$, and $D$:

FACT 3. *A cwatset contains* $\mathbf{0}$.

*Proof.* For $\mathbf{c} \in C$, $\mathbf{0} = \mathbf{c} + \mathbf{c} \in C + \mathbf{c} = C^\sigma$. Thus $\mathbf{0} = \mathbf{0}^{\sigma^{-1}} \in C^{\sigma\sigma^{-1}} = C$.

PROPOSITION 3. *If a cwatset contains* $\mathbf{1}$ *or an element of odd weight, then the order of the cwatset is even.*

*Proof.* For $\mathbf{c} \in C$, choose $\sigma$ such that $C + \mathbf{c} = C^\sigma$. If $\mathbf{1} \in C$, then $\mathbf{1} \in C + \mathbf{c}$ because $\mathbf{1} = \mathbf{1}^\sigma$. But $\mathbf{1} \in C + \mathbf{c}$ implies that $\mathbf{1} + \mathbf{c} \in C$. Now $\mathbf{1} + \mathbf{c} \neq \mathbf{c}$, so $C$ must have even order because it is the disjoint union of sets of the form $\{\mathbf{c}, 1 + \mathbf{c}\}$. On the other hand, it is convenient to identify the elements of even weight (say $C_e$) in $C$. For, if $\mathbf{c} \in C - C_e$, then $(C + \mathbf{c})_e = (C - C_e) + \mathbf{c}$ as $w(\mathbf{c}) \equiv 1 \pmod 2$. Thus,

$$|C_e| = |(C^\sigma)_e| = |(C + \mathbf{c})_e| = |(C - C_e) + \mathbf{c}| = |(C - C_e)|,$$

which implies that $2 \cdot |C_e| = |C|$.

In view of Proposition 3, all the elements of a cwatset or exactly one-half the elements of a cwatset are of even weight. Moreover, the elements of even weight behave just as they do in a subgroup of $\mathbf{Z}_2^n$:

FACT 4. *The elements of even weight in a cwatset form a (sub)cwatset.*

*Proof.* Let $\mathbf{e} \in C_e$ (remember, $C_e$ is not empty because $\mathbf{0} \in C_e$). Since $C$ is a cwatset there exists $\sigma \in S_n$ such that $C + \mathbf{e} = C^\sigma$. Thus, $(C + \mathbf{e})_e = (C^\sigma)_e$, which, in view of (2) and (3), implies that $C_e + \mathbf{e} = (C_e)^\sigma$, i.e.; $C_e$ is a cwatset.

PROPOSITION 4. *Let $C$ be a cwatset in $\mathbf{Z}_2^n$ such that $|C| > 2^{n-1}$. Then, $C$ contains an element of weight $k$, for $k = 0, 1, \ldots, n$.*

*Proof.* We prove the contrapositive: If, for some $k$, $C$ has no elements of weight $k$, then $|C| \leqslant 2^{n-1}$. So, fix $k$, assume $C$ has no elements of weight $k$ and let $\mathbf{b}_k$ be any element of $\mathbf{Z}_2^n$ of weight $k$. It follows that $C \cap (C + \mathbf{b}_k) = \varnothing$. Otherwise, there exists $\mathbf{c} \in C$ such that $\mathbf{c} + \mathbf{b}_k \in C$. In this case, the fact that $C$ is a cwatset implies that for some $\sigma$,

$$\mathbf{c} + (\mathbf{c} + \mathbf{b}_k) \in C^\sigma,$$
$$\mathbf{b}_k \in C^\sigma,$$
$$(\mathbf{b}_k)^{\sigma^{-1}} \in C,$$

which contradicts our assumption that $C$ has no elements of weight $k$. Thus, since $|C + \mathbf{b}_k| = |C|$, there are at least as many elements of $\mathbf{Z}_2^n$ not in $C$ as there are in $C$; i.e., $|C| \leqslant 2^{n-1}$.

Proposition 2, Proposition 3, Proposition 4, and some immediate corollaries,

$$\textit{if } |C| \neq |C_e|, \textit{ then } |C| = 2|C_e|, \text{ and}$$
$$\textit{if } |C| > 2^{n-1}, \textit{ then } |C| \textit{ is even},$$

suggest that cwatsets aren't completely free of arithmetical constraints. But, the cwatsets $T$ and $S$ guarantee that the most important arithmetical constraint of elementary group theory,

LAGRANGE'S THEOREM. *The order of a subgroup divides the order of the group,*

doesn't have an exact analogue in cwatset theory. Here's the happy medium.

PROPOSITION 5. *The order of a cwatset of $\mathbf{Z}_2^n$ divides $n! \cdot 2^n$.*

The proof of this 'Lagrange-like' theorem is best made, a bit later, as an application of Lagrange's Theorem. Of course, this is possible only if we have a group-theoretic explanation of cwatsets.

**Group-theoretically speaking** Let's collect our permutations and $n$-tuples in the set $W = S_n \times \mathbf{Z}_2^n$. In the previous sentence the emphasis is on the word "set": $W$ is not an algebraic structure, in particular it is not a 'product' of two groups—yet. Our definition of a cwatset requires not just two actions (the application of permutations to $n$-tuples and the addition of $n$-tuples), but also the 'interaction' of these two actions. This is because $C + \mathbf{c} = C^\sigma$ is equivalent to $C^\sigma + \mathbf{c} = C$; i.e., cwatsets involve elements of $\mathbf{Z}_2^n$ that look like $\mathbf{b}^\sigma + \mathbf{c}$. Thus, we define an *action* of $W$ on $\mathbf{Z}_2^n$ by

$$\mathbf{b}^{(\sigma,\mathbf{x})} = \mathbf{b}^\sigma + \mathbf{x}$$

in an attempt to capture the essence of this interaction. This new 'iterated' action is intimately related to our definition of cwatset because

$$C^{(\sigma,c)} = C \text{ is equivalent to } C + \mathbf{c} = C^\sigma$$

where, of course, $C^{(\sigma,c)} = \{\mathbf{b}^{(\sigma,\mathbf{c})} | \mathbf{b} \in C\}$.

Here's what happens if we follow the action of $(\sigma, \mathbf{x})$ with the action of $(\tau, \mathbf{y})$.

$$\begin{aligned} \mathbf{b}^{(\sigma,\mathbf{x})(\tau,\mathbf{y})} &= \left(\mathbf{b}^{(\sigma,\mathbf{x})}\right)^{(\tau,\mathbf{y})} \\ &= \left(\mathbf{b}^\sigma + \mathbf{x}\right)^{(\tau,\mathbf{y})} \\ &= \left(\mathbf{b}^\sigma + \mathbf{x}\right)^\tau + \mathbf{y} \qquad (4) \\ &= \mathbf{b}^{\sigma\tau} + \mathbf{x}^\tau + \mathbf{y} \\ &= \mathbf{b}^{(\sigma\tau,\mathbf{x}^\tau + \mathbf{y})}. \end{aligned}$$

It's almost the same as if we had performed the actions componentwise—except for the 'twist' in the binary component. Do you see what we are up to? We are going to define an operation on $S_n \times \mathbf{Z}_2^n$ that reflects the calculations in (4).

*Definition.* The *product* of two elements, $(\sigma, \mathbf{x})$ and $(\tau, \mathbf{y})$, of $S_n \times \mathbf{Z}_2^n$ is defined by

$$(\sigma,\mathbf{x})(\tau,\mathbf{y}) = (\sigma\tau, \mathbf{x}^\tau + \mathbf{y}). \qquad (5)$$

Now $S_n \times \mathbf{Z}_2^n$ has algebraic structure: It is straightforward to verify that $S_n \times \mathbf{Z}_2^n$, together with the binary operation defined in (5), is a group. This group is denoted by $S_n \wr \mathbf{Z}_2$ and is referred to as the *wreath product* of $\mathbf{Z}_2$ by $S_n$ (see [5], pages 219–228). Notice in particular that if $\iota\delta$ denotes the identity of $S_n$, then the identity of $S_n \wr \mathbf{Z}_2$ is $(\iota\delta, \mathbf{0})$ and $(\sigma, \mathbf{x})^{-1} = (\sigma^{-1}, \mathbf{x}^{\sigma^{-1}})$. The subgroups of $S_n \wr \mathbf{Z}_2$ and their projections in $\mathbf{Z}_2^n$ are of particular interest to us. By the *projection* in $\mathbf{Z}_2^n$ of a subgroup $P$ of $S_n \wr \mathbf{Z}_2$ we mean

$$\{\mathbf{b} | (\sigma, \mathbf{b}) \in P \text{ for some } \sigma \in S_n\}.$$

PROPOSITION 6. *A subset of $\mathbf{Z}_2^n$ is a cwatset if, and only if, it is the projection of a subgroup of $S_n \wr \mathbf{Z}_2$.*

*Proof.* Let $C$ be a cwatset and define a subset, $P(C)$, of $S_n \wr \mathbf{Z}_2$ by

$$P(C) = \{(\sigma, \mathbf{c}) | C^{(\sigma,\mathbf{c})} = C \text{ for some } \sigma \in S_n \text{ and some } \mathbf{c} \in C\}.$$

The following two claims establish that $P(C)$ is a subgroup of $S_n \wr \mathbf{Z}_2$. Clearly, $C$ is the projection of $P(C)$ in $\mathbf{Z}_2^n$.

*Claim.* $(\sigma, \mathbf{c}) \in P(C)$ implies $(\sigma^{-1}, \mathbf{c}^{\sigma^{-1}}) \in P(C)$. This follows from two reformulations of $C^{(\sigma, \mathbf{c})} = C$:

$$\text{(i) } (C + \mathbf{c})^{\sigma^{-1}} = C; \text{ i.e., } \mathbf{c}^{\sigma^{-1}} \in C \text{ because } \mathbf{0} \in C.$$

$$\text{(ii) } C + \mathbf{c}^{\sigma^{-1}} = C^{\sigma^{-1}}; \text{ i.e., } C^{(\tau, \mathbf{b})} = C \text{ where } \tau = \sigma^{-1}, \mathbf{b} = \mathbf{c}^{\sigma^{-1}}.$$

*Claim.* $(\sigma, \mathbf{x})$, $(\tau, \mathbf{y}) \in P(C)$ implies $(\sigma\tau, \mathbf{x}^\tau + \mathbf{y}) \in P(C)$. It follows from $C^{(\tau, \mathbf{y})} = C$ that $\mathbf{x}^\tau + \mathbf{y} \in C$. Thus,

$$\begin{aligned} C + \mathbf{x}^\tau + \mathbf{y} &= (C + \mathbf{y}) + \mathbf{x}^\tau \\ &= C^\tau + \mathbf{x}^\tau \\ &= (C + \mathbf{x})^\tau \\ &= C^{\sigma\tau}, \end{aligned}$$

and so, $C^{(\sigma\tau, \mathbf{x}^\tau + \mathbf{y})} = C$.

For the converse, let $P$ be a subgroup of $S_n \wr \mathbf{Z}_2$ and let $B$ be the projection of $P$ in $\mathbf{Z}_2^n$. Now, for each $\mathbf{b} \in B$ choose some $\sigma_\mathbf{b} \in S_n$ such that $(\sigma_\mathbf{b}, \mathbf{b}) \in P$ and fix $\mathbf{c} \in B$. Since $P$ is a subgroup we have

$$\begin{aligned} P &\supseteq \{(\sigma_\mathbf{b}, \mathbf{b})(\sigma_\mathbf{c}, \mathbf{c}) | \mathbf{b} \in B\} \\ &= \{(\sigma_\mathbf{b}\sigma_\mathbf{c}, \mathbf{b}^{\sigma_\mathbf{c}} + \mathbf{c}) | \mathbf{b} \in B\}, \end{aligned}$$

which implies that $\{\mathbf{b}^{\sigma_\mathbf{c}} + \mathbf{c} | \mathbf{b} \in B\} \subseteq B$. Thus, $\{\mathbf{b}^{\sigma_\mathbf{c}} + \mathbf{c} | \mathbf{b} \in B\} = B$, because $|\{\mathbf{b}^{\sigma_\mathbf{c}} + \mathbf{c} | \mathbf{b} \in B\}| = |B|$, and so $B$ is a cwatset.

As an application of Proposition 6, note that the cyclic cwatset $C(\sigma, \mathbf{b})$ is just the projection of the cyclic subgroup generated by the element $(\sigma, \mathbf{b})$ of $S_n \wr \mathbf{Z}_2$. Moreover, we may use Proposition 6 to establish a generalization of Proposition 2 and Proposition 5.

PROPOSITION 7. *If $H$ is a subgroup of $S_n \wr \mathbf{Z}_2$ and $C$ is the projection of $H$ in $\mathbf{Z}_2^n$, then the order of $C$ divides the order of $H$.*

*Proof.* By Lagrange's theorem it suffices to show that $|C| = [H:K]$ for some subgroup $K$ of $H$. To this end, notice that $H$ is the disjoint union of precisely $|C|$ sets of the form

$$H(\mathbf{c}) = \{(\sigma, \mathbf{c}) | C^{(\sigma, \mathbf{c})} = C \text{ for some } \sigma \in S_n\}; \text{ i.e., } H = \bigcup_{\mathbf{c} \in C} H(\mathbf{c}).$$

It is easy to see that $H(\mathbf{0})$ is a subgroup of $H$—it plays the role of $K$. And, each $H(\mathbf{c})$ is just a right coset of $H(\mathbf{0})$ in $H$. To see this, fix $(\tau, \mathbf{c}) \in H(\mathbf{c})$ and let $(\sigma, \mathbf{0}) \in H(\mathbf{0})$. Then

$$C^{(\sigma, \mathbf{0})(\tau, \mathbf{c})} = C^{(\sigma\tau, \mathbf{c})} = C^{\sigma\tau} + \mathbf{c} = C^\tau + \mathbf{c} = C;$$

i.e., the right coset of $H(\mathbf{0})$ by an element of $H(\mathbf{c})$ is contained in $H(\mathbf{c})$. This means that $H(\mathbf{c})$ is a union of such cosets. But any two elements of $H(\mathbf{c})$ determine the same right coset of $H(\mathbf{0})$: $(\tau, \mathbf{c})$, $(\pi, \mathbf{c}) \in H(\mathbf{c})$ implies that

$$(\tau, \mathbf{c})(\pi, \mathbf{c})^{-1} = (\tau, \mathbf{c})(\pi^{-1}, \mathbf{c}^{\pi^{-1}}) = (\tau\pi^{-1}, \mathbf{c}^{\pi^{-1}} + \mathbf{c}^{\pi^{-1}}) = (\tau\pi^{-1}, \mathbf{0}).$$

Therefore, $|H| = [H\colon H(\mathbf{0})] \cdot |H(\mathbf{0})| = |C| \cdot |H(\mathbf{0})|$.

Proposition 5 is an immediate corollary of Proposition 7 because Lagrange's theorem implies that $|H|$ divides $|S_n \wr \mathbf{Z}_2|$.

QUESTION. As a group, $\mathbf{Z}_2^n$ is 'converse-Lagrange'; i.e., if $k$ divides $2^n$, then $\mathbf{Z}_2^n$ has subgroups of order $k$. Is $\mathbf{Z}_2^n$ 'converse-Proposition 5'; i.e., if $k$ divides $n! \cdot 2^n$ and $k \leqslant 2^n$, does $\mathbf{Z}_2^n$ contain cwatsets of order $k$?

**A generalization** What made our construction of $S_n \wr \mathbf{Z}_2$ work is that each element of $S_n$ determines an automorphism of the group $\mathbf{Z}_2^n$ and that the automorphism determined by the composition of two permutations is the composition of the automorphisms determined by each permutation. This is just a wordy way to say that there is a homomorphism from $S_n$ into the automorphism group of $\mathbf{Z}_2^n$.

So, let $H$ and $K$ be finite groups and let $\phi$ be a homomorphism from $H$ into the automorphism group of $K$. As with the wreath product construction above, we collect the elements of $H$ and $K$ in $H \times K$ and then define a binary operation of $H \times K$ by

$$(w, y)(x, z) = (wx, y^x z),$$

where $y^x$ means the image of $y$ under the automorphism of $K$ determined by $x$. The set $H \times K$ together with this binary operation is a group, denoted by $H \times_\phi K$ and referred to as the semi-direct product of $K$ by $H$ (see [1], pages 133–134, and [5], pages 208–215), whose structure depends on $\phi$. For example, $\mathbf{Z}_{10}$ and $D_5$ (the dihedral group on five symbols), are semi-direct products of $\mathbf{Z}_5$ by $\mathbf{Z}_2$.

QUESTION. What can one say about 'cwatsets' in semidirect products?

**Acknowledgement.** The authors are grateful to the referees for their suggestions.

## REFERENCES

1. M. A. Armstrong, *Groups and Symmetry*, Springer-Verlag, New York, Inc., New York, 1988.
2. J. E. Atkins and G. J. Sherman, Sets of typical subsamples, *Statistics and Probability Letters* 14 (1992), 115–117.
3. J. A. Hartigan, Using subsample values as typical values, *Amer. Stat. Assoc. J.* 64 (1969), 1303–1317.
4. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, New York, 1978.
5. J. S. Rose, *A Course on Group Theory*, Cambridge University Press, New York, 1978.
6. G. J. Sherman, Confidence intervals from groups, this MAGAZINE 65 (1992), 118–122.

## 150 YEARS AGO...

Jules Henri Poincaré was born in Nancy, France. After graduating from the École Polytechnique in 1875, Poincaré took a degree in mining engineering at École des Mines (1879) and a doctorate in science from the University of Paris in the same year. He was truly a person who knew all of mathematics. At the Sorbonne, he lectured each year on a different topic in pure or applied mathematics! In addition, he wrote many popular works, explaining mathematics and science to people of varied interests. For his writing he was elected a member of the literary section of the French Institut, one of France's highest honors.

from Howard Eves, *Introduction to the History of Mathematics*, 5th Edition, Saunders College Publishing, 1983.

# An Introduction to Combinatorial Existence Theorems

MARTIN ERICKSON
Northeast Missouri State University
Kirksville, MO 63501

If $f$ is a function from the set $\{1,\dots,10\}$ into the set $\{1,\dots,9\}$, then $f(a)=f(b)$ for some $a$ and $b$ with $1 \leq a < b \leq 10$. Many readers will instantly recognize this statement as an example of the famous pigeonhole principle of combinatorial theory, a principle cited and used in various forms in [1], [4], and [6]. Applied to this example, the principle says that if 10 objects are placed in nine pigeonholes, then one pigeonhole will contain at least two objects.

The first purpose of this article is to introduce the reader to a precise and fairly general version of the pigeonhole principle and to some combinatorial existence theorems that follow from it; namely, we will prove Erdős and Szekeres' theorem on monotonic subsequences, Dilworth's lemma on partial orders, and Ramsey's theorem on monochromatic subgraphs. The statements of these theorems and proofs appear among [1], [2], [3], [4], and [8]. Even mathematicians who have employed these theorems may be unaware that they are logically related. Therefore, the second purpose of this paper is to show (as indicated in [4]) that Dilworth's Lemma is a generalization of Erdős and Szekeres' Theorem and Ramsey's Theorem is a generalization of Dilworth's Lemma. We will also examine the connection of these results to other existence results, including Sperner's Theorem on antichains and Schur's Lemma on partitions of sets of integers.

THE PIGEONHOLE PRINCIPLE. *Suppose $A$ and $B$ are finite nonempty sets and $f$ maps $A$ to $B$. Then for some $j \in B$, the inverse image of $j$ under $f$ contains at least $|A|/|B|$ elements.*

*Proof.* All proofs of the pigeonhole principle proceed by contradiction, so suppose the inverse image $f^{-1}(j)$ has fewer than $|A|/|B|$ elements for every $j$ in $B$. Then $|A| = \Sigma_{j \in B}|f^{-1}(j)| < \Sigma_{j\in B}|A|/|B| = |B| \cdot |A|/|B| = |A|$, which is a contradiction.

The statement and proof of the pigeonhole principle are simple. The trick is knowing when and how to apply the principle. As the first application, we shall give a proof of Erdős and Szekeres' Theorem essentially found in [8]. We say that a sequence $x_1,\dots,x_n$ of real numbers is *monotonically increasing* if $x_1 \leq \cdots \leq x_n$, *monotonically decreasing* if $x_1 \geq \cdots \geq x_n$, and *monotonic* if it is monotonically increasing or monotonically decreasing.

ERDŐS AND SZEKERES' THEOREM. *Suppose $a, b \in \mathbf{N}$, $n = ab+1$, and $x_1,\dots,x_n$ is a sequence of $n$ real numbers. This sequence contains a monotonically increasing (decreasing) subsequence of $a+1$ terms or a monotonically decreasing (increasing) subsequence of $b+1$ terms.*

For example, let $a=3$, $b=4$, and $n=13$. The theorem says that any sequence of 13 real numbers contains a monotonically increasing subsequence of length 4 or a monotonically decreasing subsequence of length 5. Thus the sequence

$$1, 3, \sqrt{2}, \pi, 2e, 0, -\pi, 2, 1, -3, \tfrac{1}{2}, 0, 9$$

contains $1, \sqrt{2}, \pi, 9$ as a monotonically increasing subsequence of length 4.

*Proof of Erdős and Szekeres' Theorem* (See [8].) Suppose the theorem is false for some $a, b$ and some sequence $x_1, \ldots, x_n$, for $n = ab + 1$. By assumption, all monotonically increasing subsequences of this sequence have at most $a$ terms and all monotonically decreasing subsequences have at most $b$ terms. Therefore, we may define a function $f$: $\{1, \ldots, ab+1\} \to \{1, \ldots, a\} \times \{1, \ldots, b\}$ by $f(j) = (i_j, d_j)$, where $i_j$ is the length of the longest monotonically increasing subsequence beginning with $x_j$ and $d_j$ is the length of the longest monotonically decreasing subsequence beginning with $x_j$. Because $\{1, \ldots, a\} \times \{1, \ldots, b\}$ has $ab$ elements, the pigeonhole principle tells us that $f^{-1}((m, n))$ contains at least $(ab+1)/ab$ elements for some $(m, n) \in \{1, \ldots, a\} \times \{1, \ldots, b\}$. Since $(ab+1)/ab > 1$, this means that $f(j) = f(k)$ for two distinct elements $j, k$ $(j < k)$ of $\{1, \ldots, ab+1\}$. Finally, $f(j) = f(k)$ means $i_j = i_k$ and $d_j = d_k$; but if $x_j \leq x_k$ then $i_j > i_k$, which is a contradiction, while if $x_j \geq x_k$ then $d_j > d_k$, which is also contradiction.

Erdős and Szekeres' Theorem is a "best possible" result in the sense that it is false if the value of $n$ is decreased from $ab + 1$ to $ab$. To see this, note that the following sequence of $ab$ integers contains neither a monotonically increasing subsequence of $a + 1$ terms nor a monotonically decreasing subsequence of $b + 1$ terms:

$$ba + 1, ba + 2, \ldots, ba + a,$$

$$(b-1)a + 1, (b-1)a + 2, \ldots, (b-1)a + a, \ldots,$$

$$2a + 1, 2a + 2, \ldots, 2a + a$$

$$a + 1, a + 2, \ldots, a + a.$$

If $a = b$ then Erdős and Szekeres' Theorem guarantees a monotonic subsequence of length $a + 1$ in any sequence of $n = a^2 + 1$ real numbers. This observation leads to an interesting proposition about total orders on finite sets. Recall that $\leq$ is a *partial order* on a set $S$ if $\leq$ is a relation on $S$ for which the following three properties hold:

1. (reflexivity) $a \leq a$ for all $a \in S$,
2. (antisymmetry) $a \leq b$ and $b \leq a$ imply $a = b$, and
3. (transitivity) $a \leq b$ and $b \leq c$ imply $a \leq c$.

Furthermore, $\leq$ is a *total order* on $S$ if it also satisfies

4. (comparability) For any $a, b \in S$, either $a \leq b$ or $b \leq a$.

Thus the real numbers are totally ordered by the usual less than or equal to relation $\leq$ explicit in Erdős and Szekeres' Theorem. The only property of $\leq$ used in our proof of Erdős and Szekeres' Theorem is that $\leq$ is a total order on the set $\{x_1, \ldots, x_n\}$. Therefore, the theorem remains true if $\leq$ is replaced by $\leq_1$, an arbitrary total order on $\{x_1, \ldots, x_n\}$.

A second total order is implicit in Erdős and Szekeres' Theorem. The elements of $\{x_1, \ldots, x_n\}$ are totally ordered by the order in which they appear in the sequence. In the following version of Erdős and Szekeres' Theorem, we replace this order by $\leq_2$, another arbitrary total order on $\{x_1, \ldots, x_n\}$.

Now Erdős and Szekeres' Theorem (in the case $a = b$) may be restated: *Suppose $n = a^2 + 1$ and $\leq_1$ and $\leq_2$ are two total orders on the set $S = \{x_1, \ldots, x_n\}$. Then for some subset $T$ of $S$ of size $a + 1$, one of the following two statements holds*:

1. *For all $x, y \in T$, $x \leq_1 y$ if, and only if, $x \leq_2 y$.*
2. *For all $x, y \in T$, $x \leq_1 y$ if, and only if, $y \leq_2 x$.*

In other words, the two orders $\leq_1$ and $\leq_2$ are in complete agreement or complete disagreement on a subset of size $a+1$. As an example with $a=3$, consider the set $\{1,\dots,10\}$ totally ordered in the following two ways:

$$10,8,6,1,3,4,5,7,9,2$$
$$7,10,3,4,1,8,9,6,5,2.$$

There must be a subset of size $a+1=4$ on which these two orders completely agree or completely disagree. Indeed, the subset $\{2,5,6,8\}$ appears as a decreasing subsequence in both orders.

Let us now consider the combinatorial gem known as Dilworth's Lemma.

DILWORTH'S LEMMA. *Suppose $a,b\in\mathbf{N}$, $n=ab+1$, and $\le$ is a partial order on the set $S=\{x_1,\dots,x_n\}$. Then $S$ contains a subset $T$ of size $a+1$ on which $\le$ is a total order, or else $S$ contains a subset $T$ of size $b+1$ in which no two elements are related by $\le$.*

A set that is totally ordered by a partial order is called a *chain* and a set in which no two elements are comparable is called an *antichain*. Dilworth's Lemma may be restated using these terms: *If $\le$ is a partial order on a set $S$ of $ab+1$ elements, then $S$ contains a chain of size $a+1$ or an antichain of size $b+1$.*

*Proof* (See [2].) Suppose $S$ contains no chain of size $a+1$. We will conclude that $S$ contains an antichain of size $b+1$. We define $f\colon S\to\{1,\dots,a\}$ by $f(x_j)=L_j$, where $L_j$ is the size of the longest chain in $S$ with $x_j$ as the least element. The pigeonhole principle tells us that $f^{-1}(r)$ has at least $(ab+1)/a$ elements for some $r\in\{1,\dots,a\}$. Since $(ab+1)/a>b$ and $|f^{-1}(r)|$ is an integer, it must be at least $b+1$. No two comparable elements of $S$ are members of $f^{-1}(r)$, since $x_j\le x_k$ implies $L_j>L_k$. Hence $f^{-1}(r)$ contains at least $b+1$ elements, no two of which are comparable.

Dilworth's Lemma is "best possible" in the sense that $ab+1$ may not be replaced by $ab$. For example, FIGURE 1 shows a partial order on six elements $u$, $v$, $w$, $x$, $y$, and $z$ that does not contain a chain of size 3 or an antichain of size 4. The arrow from $u$ to $x$ means $x\le u$.

When $\le$ is a partial order on $S$ we say that the size of the longest chain is the *length* of $S$ and the size of the largest antichain is the *width* of $S$. Suppose $S$ has cardinality $N$, length $L$, and width $W$. By the following argument, Dilworth's Lemma is easily seen to be equivalent to the inequality $N\le LW$.

If $N\ge LW+1$ then Dilworth's Lemma implies that $S$ contains a chain of size $L+1$ or an antichain of size $W+1$, contradicting the definition of $L$ or $W$. On the other hand, if $N\ge ab+1$ and $S$ does not contain a chain of size $a+1$ or an antichain of size $b+1$, then $N\ge ab+1\ge LW+1>LW$, which contradicts the inequality $N\le LW$.

Let $S$ be the power set of a $t$-element set $T$, and let the elements of $S$ be partially ordered by set inclusion. Clearly, $N=2^t$ (the number of subsets of $T$) and $L=t+1$ (the length of a chain starting with the empty set and adding one element at a time until $T$ is exhausted), but what is $W$? This question is answered by the following theorem. ($[x]$ is the greatest integer less than or equal to $x$.)

SPERNER'S THEOREM. *The maximum cardinality of a collection of subsets of a $t$-element set $T$, none of which contains another, is the binomial coefficient $C(t,[t/2])$.*

*Proof* (Also see [1] and [3].) First note that the family of $[t/2]$-element subsets provides a family of $C(t,[t/2])$ subsets of $T$, none of which contains another. Now suppose that $A_1,\dots,A_m$ are subsets (with cardinalities $a_1,\dots,a_m$), none of which contains another. Each $A_i$ is contained in $a_i!(t-a_i)!$ chains of length $t+1$ (which start with the empty set and add one element at a time until $T$ is exhausted).

The chains coming from different $A_i$ are different (no $A_i$ contains an $A_j$), so $\sum_{i=1}^{m} a_i!(t-a_i)! \le t!$, since $t!$ is the total number of chains of length $t+1$. This inequality may be rewritten as $\sum_{i=1}^{m} c(t,a_i)^{-1} \le 1$. Therefore,

$$m = C(t,[t/2]) \sum_{i=1}^{m} C(t,[t/2])^{-1} \le C(t,[t/2]) \sum_{i=1}^{m} C(t,a_i)^{-1} \le C(t,[t/2]),$$

the first inequality holding because the largest binomial coefficient(s) in a row of Pascal's triangle is (are) the one(s) in the middle. Hence $W = C(t,[t/2])$.

To discuss Ramsey's Theorem, we abandon partial orders for the moment and turn our attention to graphs. Following [5], we take a *graph* to mean a finite undirected graph without loops or multiple edges. The complete graph with $n$ vertices is denoted $K_n$. A *two-coloring* of a graph $G$ is an assignment of each edge of $G$ to one of two *color classes*. A two-coloring may be viewed as a function $f$: $E(G) \to \{\text{green}, \text{red}\}$, where $E(G)$ is the edge set of $G$. A subgraph $H$ of $G$ is *monochromatic* if $f$ is constant on $E(H)$. If $H$ is monochromatic then we call it *green* or *red*.

We now present one of the many versions of Ramsey's Theorem. For other versions, as well as an introduction to Ramsey theory, the reader should consult [4].

RAMSEY'S THEOREM. *For each $m, n \in \mathbf{N}$ with $m, n \ge 2$ there exists a least integer $R(m,n)$ such that no matter how $K_{R(m,n)}$ is two-colored, it will contain a green subgraph $K_m$ or a red subgraph $K_n$.*

*Furthermore, $R(m,n) \le R(m-1,n) + R(m,n-1)$ if $m,n \ge 3$.*

*Proof* (See [8] for a similar proof.) We employ induction on $m+n$. The theorem is trivially true for $m=2$ or $n=2$, as $R(2,n)=n$ and $R(m,2)=m$ (for example, if not all the edges of $K_n$ are red then at least one is green). Now suppose the existence of $R(m-1,n)$ and $R(m,n-1)$; we will prove the existence of $R(m,n)$. Consider a two-coloring of the complete graph on $R(m,n-1)+R(m-1,n)$ vertices. Designate one vertex of this graph by $v$. Since $R(m,n-1)+R(m-1,n)-1$ edges emanate from $v$, the pigeonhole principle guarantees that either the number of green edges is at least $R(m-1,n)$ or the number of red edges is at least $R(m,n-1)$. Without loss of generality, let's suppose at least $R(m-1,n)$ edges are green, and let $T$ denote the set of vertices to which $v$ is joined by green edges. Since $T$ has at least $R(m-1,n)$ vertices, it contains a red $K_n$ and we are done, or else it contains a green $K_{m-1}$. In the latter case, the green $K_{m-1}$ together with $v$ and all the edges between the two constitute a green $K_m$. Therefore, $R(m,n)$ exists and it satisfies

$$R(m,n) \le R(m-1,n) + R(m,n-1).$$

The values of $R(m,n)$ are known as *Ramsey numbers*. Clearly $R(m,n) = R(n,m)$. Only eight Ramsey numbers are known with $3 \le m \le n$ (see [4] and [7]). The value $R(3,3)=6$ is easily deduced. Ramsey's Theorem says that $R(3,3) \le R(3,2)+R(2,3) = 3+3$ (trivial values) $= 6$, and FIGURE 2, with solid lines representing green and dashed lines representing red, shows that $R(3,3) > 5$.

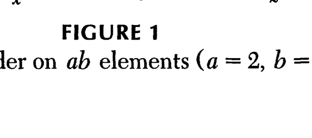



**FIGURE 1**

A partial order on $ab$ elements ($a=2$, $b=3$).

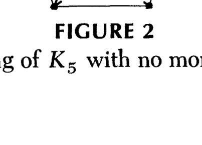


**FIGURE 2**

A two-coloring of $K_5$ with no monochromatic $K_3$.

A *digraph* is a graph in which each edge has been replaced by a directed edge (see [5]). A *transitive digraph* is a digraph where the adjacency relation is transitive. Ramsey's Theorem is easily seen (e.g., [4]) to be equivalent to the following statement: *For every* $m \in \mathbf{N}$ *there exists an* $n \in \mathbf{N}$ *such that any complete digraph on* $n$ *vertices contains a complete transitive subdigraph on* $m$ *vertices.*

Let $R_2(m) = R(m, m)$ and recursively define $R_k(m) = R(m, R_{k-1}(m))$ for $k > 2$. A $k$*-coloring* of $G$ is an assignment of each edge of $G$ to one of $k$ *color classes*. If this assignment is a constant function on $E(H)$, for a subgraph $H$ of $G$, then $H$ is *monochromatic*. Ramsey's Theorem may be generalized (see [8]) as follows: *For each* $m \in \mathbf{N}$, *no matter how a complete graph on* $R_k(m)$ *vertices is* $k$*-colored, there exists a monochromatic complete subgraph* $K_m$. This generalization may be invoked to prove Schur's Lemma concerning partitions of sets of integers.

Schur's Lemma. *For each* $k \in \mathbf{N}$ *there exists a least integer* $s(k)$ *such that no matter how the set of integers* $\{1, \ldots, s(k)\}$ *is partitioned into* $k$ *classes, one class must contain integers* $x, y, z$ *such that* $x + y = z$. *(*$x$ *and* $y$ *may be equal.)*

*Proof* (See [4], [6], and [8] for similar proofs.) Suppose the set $\{1, \ldots, R_k(3) - 1\}$ is partitioned into $k$ classes. The partition yields a $k$-coloring of the complete graph on vertices $1, \ldots, R_k(3)$ as follows: Edge $\{i, j\}$ is assigned to the same class as the integer $|i - j|$. As indicated in the generalization of Ramsey's Theorem, there exists a monochromatic triangle on, say, vertices $a, b, c$ $(a > b > c)$. Since the triangle is monochromatic, $a - b$, $b - c$, and $a - c$ belong to the same class. Since $(a - b) + (b - c) = a - c$, $s(k)$ exists and $s(k) \leq R_k(3) - 1$.

The only known exact values of $s(k)$ are $s(1) = 2$, $s(2) = 5$, $s(3) = 14$, and $s(4) = 45$ (see [4]).

Erdős and Szekeres' Theorem is a theorem about a sequence of $ab + 1$ real numbers and Dilworth's Lemma is a statement about a partial order on a set of $ab + 1$ elements. The common value $ab + 1$ is not a superficial similarity between the two theorems. In fact, Dilworth's Lemma is a generalization of Erdős and Szekeres' Theorem.

Theorem. *Dilworth's Lemma implies Erdős and Szekeres' Theorem.*

*Proof* (See [4].) Assume $x_1, \ldots, x_n$ is a sequence of $n$ real numbers, where $n = ab + 1$. We define a partial order $\leq'$ on the set $S = \{x_1, \ldots, x_n\}$ by $x_i \leq' x_j$ if, and only if, $i \leq j$ and $x_i \leq x_j$. In other words, $x_i \leq' x_j$ means $x_i$ is less than or equal to $x_j$ and $x_i$ appears to the left of $x_j$ in the sequence. Dilworth's Lemma implies that $S$ contains a chain of size $a + 1$ or an antichain of size $b + 1$. A chain of size $a + 1$ corresponds to a monotonically increasing subsequence of length $a + 1$ and an antichain of size $b + 1$ corresponds to a monotonically decreasing subsequence of length $b + 1$.

Ramsey's Theorem cannot be used to prove Dilworth's Lemma even though Ramsey's Theorem is a generalization of Dilworth's Lemma. The reason is that Dilworth's Lemma depends upon the transitivity of a partial order, while transitivity is absent from the relation in Ramsey's Theorem. However, Ramsey's Theorem guarantees that some value of $n$ satisfies the conclusion of Dilworth's Lemma.

Theorem. *Ramsey's Theorem implies that for* $a, b \in \mathbf{N}$ *there exists an integer* $n$ *such that if* $\leq$ *is a partial order on the set* $S = \{x_1, \ldots, x_n\}$, *then* $S$ *contains a chain of size* $a + 1$ *or an antichain of size* $b + 1$.

*Proof* (See [4].) Let $a, b \in \mathbf{N}$ and $n = R(a + 1, b + 1)$. Given a partial order $\leq$ on $S = \{x_1, \ldots, x_n\}$, with the listing of the $x_i$ consistent with $\leq$, we define a coloring on

$K_n$ as follows: Suppose $i < j$. Edge $\{i, j\}$ is colored green if $x_i \leq x_j$ and red otherwise. Ramsey's Theorem guarantees that $K_n$ contains a green $K_{a+1}$ or a red $K_{b+1}$. A green $K_{a+1}$ corresponds to a subset of size $a + 1$ of $S$ in which every pair of elements is related, in other words, a chain of size $a + 1$. A red $K_{b+1}$ corresponds to a subset of size $b + 1$ of $S$ in which no two elements are related, in other words, an antichain of size $b + 1$.

Dilworth's Lemma supplies the best possible value $n = ab + 1$, whereas the previous theorem merely provides the value $n = R(a + 1, b + 1)$. As mentioned above, this shortcoming is due to the lack of transitivity in the hypothesis of Ramsey's Theorem.

A reader who is comfortable with Erdős and Szekeres' Theorem, Dilworth's Lemma, and Ramsey's Theorem can study and appreciate a host of other combinatorial existence theorems. These include van der Waerden's Theorem, Rado's Theorem, the Hales-Jewett Theorem, and Folkman's Theorem, which all appear in [4].

## REFERENCES


1. M. Aigner, *Combinatorial Theory*, Springer-Verlag New York, Inc., New York, 1979.
2. B. Bollobás, *Graph Theory, an Introductory Course*, Springer-Verlag New York, Inc., New York, 1979.
3. B. Bollobás, *Random Graphs*, Academic Press, London, 1985.
4. R. L. Graham, B. L. Rothschild, J. H. Spencer, *Ramsey Theory*, John Wiley & Sons, Inc., New York, 1990.
5. F. Harary, *Graph Theory*, Addison-Wesley Publishing Co., Reading, MA, 1972.
6. R. Honsberger, *Mathematical Gems* III, MAA, Washington, DC, 1985.
7. B. D. McKay and K. M. Zhang, "The Value of the Ramsey Number $R(3,8)$", *J. of Graph Theory*, to appear.
8. A. P. Street and W. D. Wallis, *Combinatorial Theory, an Introduction*, Charles Babbage Research Center, Winnipeg, Canada, 1977.


## Puzzle

Starting with a square sheet of paper, fold it to produce a square having three-fourths its area. Only five folds are allowed.

from *Mathematical Brain Benders*, by Stephen Barr, Dover Publications, NY, 1982.

# Perturbation of a Tridiagonal Stability Matrix

M. ELIZABETH MAYFIELD
Hood College
Frederick, MD 21701

It is well known that the eigenvalues of a real, symmetric matrix are real, and in some cases it is fairly easy to compute these eigenvalues. But what happens if the matrix is not quite real and symmetric? What do the eigenvalues look like then? We investigate here what happens when a well-known tridiagonal matrix is perturbed slightly, representing, for example, an unusual boundary condition for a partial differential equation.

The matrix

$$T = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & \bigcirc & \\ & & & & & \\ & \bigcirc & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$$

is called *tridiagonal* because its only nonzero entries occur on the main diagonal and on the super- and sub-diagonals immediately above and below it. In addition, this kind of matrix is called *Toeplitz* because each of its diagonals is constant [**1**, p. 27]. These two properties make $T$ a particularly easy matrix to work with; $T$ turns out to be especially useful in finding numerical solutions to partial differential equations.

Partial differential equations are often solved numerically by methods known as *finite-difference methods*. If we wish to solve an equation involving a function $u$ of the space variable $x$ and the time variable $t$, for example, we discretize the $tx$-plane into rectangles of width $\Delta t$ and height $\Delta x$. Partial derivatives of $u$ can then be approximated by "finite differences," such as

$$\frac{\partial u}{\partial x} \approx \frac{1}{\Delta x}\left[u(t, x + \Delta x) - u(t, x)\right].$$

We often think of solving the partial differential equation by finding approximate values for $u$ as we "march out" in time, so that $u_{j+1}$, the solution at time $t = j + 1$, depends in some way on the solution $u_j$ at time $j$. If the dependence is explicit, say

$$u_{j+1} = Au_j,$$

then the finite-difference method is said to be *explicit*; if the relationship instead looks like

$$Au_{j+1} = Bu_j,$$

then the method is *implicit*. (Here $u$ is an $n$-dimensional column vector, and $A$ and $B$ are $n \times n$ matrices.)

As a first example, consider the parabolic heat equation

$$u_t = u_{xx}, \quad 0 < x < 1, t > 0.$$

An explicit method for solving this equation can be expressed as

$$u_{j+1} = (I - rT)u_j,$$

where $r$ is the ratio $\Delta t/(\Delta x)^2$ and $T$ is the particular Toeplitz matrix defined earlier. An implicit (Crank-Nicolson) method for solving the same equation is:

$$(2I + rT)u_{j+1} = (2I - rT)u_j,$$

where $r$ and $t$ are defined as above.

As another example, consider the hyperbolic wave equation

$$u_{tt} = u_{xx}, \quad 0 < x < 1, t > 0.$$

An implicit method for solving this equation is

$$\left(I + \frac{1}{2}rT\right)u_{j+1} = \left(I - \frac{1}{2}rT\right)u_j - b_j$$

where $T$ is the same Toeplitz matrix, but $r$ is now $(\Delta t)^2/(\Delta x)^2$ and $b_j$ is a known vector [**2**].

In each of these examples, the *stability* of the finite-difference scheme is determined by a matrix; in particular, it is determined by the eigenvalues of the matrix $A$ that appear when the scheme is written in the form $u_{j+1} = Au_j + v_j$. The three schemes may be written

1) $u_{j+1} = (I - rT)u_j$,
2) $u_{j+1} = (2I + rT)^{-1}(2I - rT)u_j$,
3) $u_{j+1} = (I + \frac{1}{2}rT)^{-1}(I - \frac{1}{2}rT)u_j + (I + \frac{1}{2}rT)^{-1}b_j$.

If the eigenvalues of $A$ have moduli less than or equal to 1, the scheme is *stable* (error will not propagate exponentially in the scheme); otherwise, the scheme is *unstable*. In each of these three examples, we can determine the range of $r$ that will guarantee stability of the numerical scheme.

Since $T$ is a real, symmetric matrix, its eigenvalues are known to be real. Since it is in fact an $n \times n$ Toeplitz tridiagonal matrix, its eigenvalues are known exactly [**2**, p. 86]:

$$\lambda_j = 4\cos^2\left[\frac{j\pi}{2(n+1)}\right]$$

for $j = 1, 2, \dots n$. And since each of the error-propagating matrices,

1) $A_1 = (I - rT)$,
2) $A_2 = (2I + rT)^{-1}(2I - rT)$,
3) $A_3 = (I + \frac{1}{2}rT)^{-1}(I - \frac{1}{2}rT)$,

in the three given examples, is an algebraic expression in $T$, its eigenvalues $\mu_j$ are also known [**2**, p. 85]:

1) $\mu_j = 1 - r\lambda_j$,
2) $\mu_j = (2 - r\lambda_j)/(2 + r\lambda_j)$,
3) $\mu_j = (1 - \frac{1}{2}r\lambda_j)/(1 + \frac{1}{2}r\lambda_j)$.

In each case, then, it is a fairly easy task to determine which values of $r$ will yield a stable scheme for solving the corresponding partial differential equation. Implicit finite-difference schemes are almost always stable for all values of the ratio $r$, while explicit schemes are usually stable only for values of $r$ in some range (like $r < 1$).

We now turn our attention to what happens if we perturb the matrix $T$ slightly. Suppose we add the complex quantity $\varepsilon$ to the last entry of the matrix to obtain

$$T' = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & \bigcirc & \\ & & & & & \\ & \bigcirc & & -1 & 2 & -1 \\ & & & & -1 & 2+\varepsilon \end{bmatrix}.$$

We wish to see how this change affects the eigenvalues of $T$ and hence the error-propagating matrix $A$. The following theorem shows that perturbing $T$ in this manner pushes all of its eigenvalues off the real axis and into the complex plane—either all above the real axis, or all below. Although the result can be proved easily with quadratic forms, this approach shows how elementary theory from several branches of mathematics can be combined to solve a problem in partial differential equations.

THEOREM. *If the last entry of the special Toeplitz tridiagonal matrix $T$ is perturbed by the complex number $\varepsilon$ to obtain $T'$, then all eigenvalues of $T$ become complex, and*

*i) if $Re(\varepsilon) > 0$, then all eigenvalues of $T'$ lie above the real axis;*
*ii) if $Re(\varepsilon) < 0$, then all eigenvalues of $T'$ lie below the real axis.*

*Proof.* The matrix $T'$ can be written in the form $T' = T + \varepsilon U$ where $U$ is the matrix

$$U = \begin{bmatrix} 0 & 0 & . & . & . & 0 & 0 \\ . & & & & & & \\ . & & & & & & \\ . & & & & & & \\ 0 & 0 & . & . & . & 0 & 1 \end{bmatrix}.$$

We write the eigenvalues and eigenvectors of $T'$ as power series in $\varepsilon$:

$$\mu_j = \lambda_j + \varepsilon\lambda_{1j} + \varepsilon^2\lambda_{2j} + \dots,$$

$$y^j = x^j + \varepsilon x^{j1} + \varepsilon^2 x^{j2} + \dots,$$

where $\mu_j$ is the eigenvalue of $T'$ corresponding to $\lambda_j$ of $T$, $y^j$ is the eigenvector of $T'$ corresponding to the orthonormal eigenvector $x^j$ of $T$, the $\lambda_{ij}$'s are unknown complex numbers, and the $x^{ji}$'s are unknown vectors. (This is a standard approach to such a perturbation problem; see, for example [3, pp. 61–63].) It is known that, for sufficiently small values of $\varepsilon$, these power series will converge [4, p. 67]. In fact, we shall see that, for this example, the "power series" for $\mu_j$ are really just binomials. Following the example of [3], we now write the characteristic equation for $T'$:

$$T'y = \mu y$$

and substitute the expressions above for $T'$, $\mu$, and $y$:

$$(T + \varepsilon U)(x^j + \varepsilon x^{j1} + \dots) = (\lambda_j + \varepsilon\lambda_{1j} + \dots)(x^j + \varepsilon x^{j1} + \dots).$$

Multiplying out, collecting like terms, and equating powers of $\varepsilon$ yield the following equations for each $j$,

$$Tx^j = \lambda_j x^j, \text{ the characteristic equations for } T; \tag{1}$$

and

$$(T - \lambda_j I)x^{j1} = -(U - \lambda_{1j} I)x^j. \tag{2}$$

Since $T$ is symmetric and the vectors $x^j$ are orthonormal, the inner product of each side of equation (2) with the vector $x^j$ gives

$$\lambda_{1j} = \left(x_n^j\right)^2,$$

where $x_n^j$ is the last component of the eigenvector $x^j$. (This real number is known to be $\sin(nj\pi/(n+1))$; note in particular that it is never zero [**2**, p. 115].) Hence

$$\mu_j = \lambda_j + \left(x_n^j\right)^2 \varepsilon + \alpha, \tag{3}$$

where $\alpha$ is a complex number that includes the terms of the power series in $\varepsilon^2$, $\varepsilon^3$, and so forth. We will now show that $\alpha = 0$.

Recall that the *trace* of a matrix is the sum of its diagonal entries; it is also the sum of its eigenvalues [**1**, p. 40]. Then

$$\text{Trace}\,(T) = \Sigma\lambda_j = 2n, \tag{4}$$

$$\text{Trace}\,(T') = \Sigma\mu_j = 2n + \varepsilon. \tag{5}$$

First we show that the imaginary part of $\alpha$ is zero. From equation (5) we see that

$$\Sigma\text{Imag}(\mu_j) = \text{Imag}(\varepsilon).$$

But from (3) we know that

$$\Sigma\text{Imag}(\mu_j) = \left[\Sigma\left(x_n^j\right)^2\right] \cdot \text{Imag}(\varepsilon) + n \cdot \text{Imag}(\alpha).$$

But $\Sigma(x_n^j)^2 = 1$. (Since the eigenvectors $x^j$ are orthonormal, the matrix formed by using these vectors as columns is orthogonal, as is its transpose [**1**, p. 67].) And so we see that $\text{Imag}(\alpha) = 0$. Similarly, it can be shown that the real part of $\alpha$ is zero. Therefore, we have

$$\mu_j = \lambda_j + \left(x_n^j\right)^2 \varepsilon,$$

where $\lambda_j$ is a real number, and $(x_n^j)^2$ is a positive real number. The sign of the imaginary part of $\varepsilon$, then, completely determines the sign of the imaginary part of each of the eigenvalues $\mu_j$ of the perturbed matrix $T'$.

**Acknowledgement.** The author is grateful to Professors John Papadakis and Vassilios Dougalis of the University of Crete, whose insights into this problem made a journey to Greece even more enjoyable.

## REFERENCES

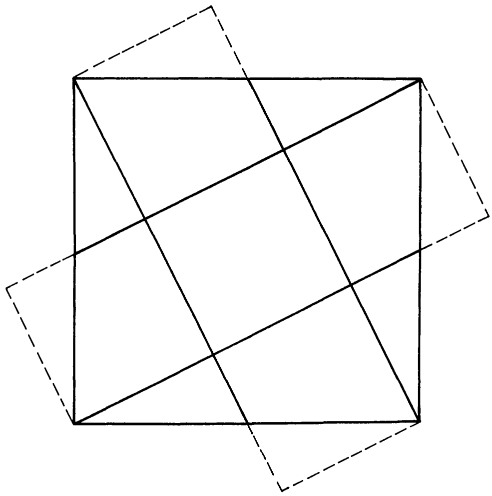
1. Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1987.
2. G. D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 2nd edition, Clarendon Press, Oxford, 1978.
3. Richard Bellman, *Introduction to Matrix Analysis*, 2nd edition, McGraw-Hill Book Co., New York, 1970.
4. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1988.

# Dörrie Tiles and Related Miniatures

EDWARD KITCHEN
Santa Monica, CA 90402

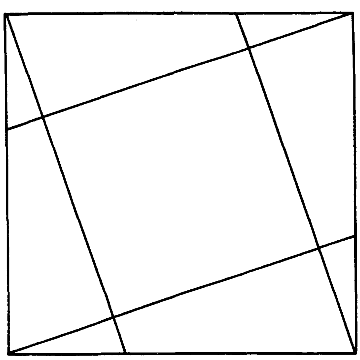
In recent years this MAGAZINE has encouraged readers to submit *Proofs without Words*. This note consists of just such material in the form of tile problems together with look-see solutions.

We begin with a well-known configuration [1, 2, 3] that seems to have originated with H. Dörrie: The vertices of the unit square are joined to the midpoints of the sides, as shown in Tile 1. Then the area of the smaller square is $1/5$.

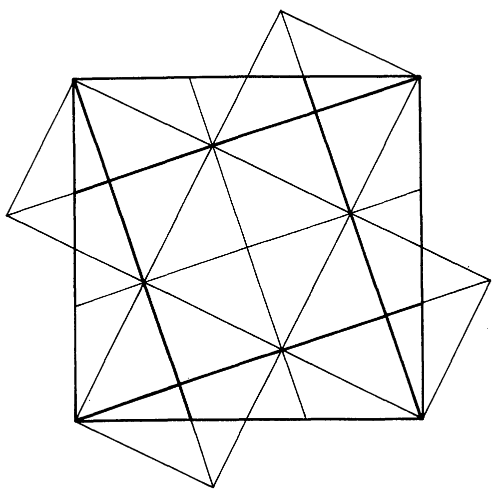


**TILE 1**
Dörrie's tile.

We propose the following related puzzle. Suppose that vertices are joined to points of trisection (Tile 2). By superimposing Tile 1 on Tile 2, we see that the area of the smaller square in Tile 2 is $2/5$ (Tile 3).

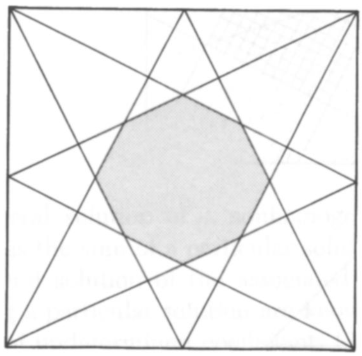


**TILE 2**

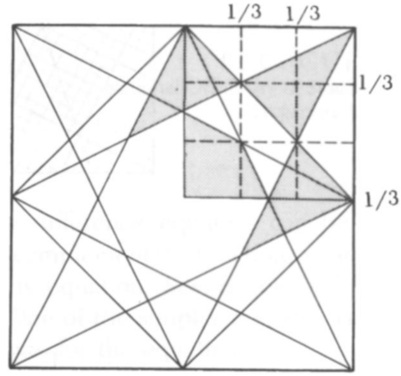


**TILE 3**

Another tile due to Dörrie [1, 3] is the following. Suppose that each vertex of the unit square is joined to the midpoints of the corresponding two opposite sides (Tile 4): Then the area of the centrally symmetric octagon (shaded) is $1/6$. Our solution in Tile 5 confirms this. The area of the shaded octagon in Tile 4 is $1/9 + 1/18 = 1/6$.

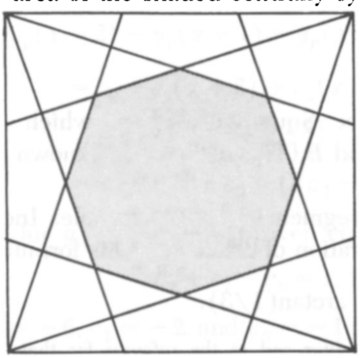


**TILE 4**
Dörrie's second tile.

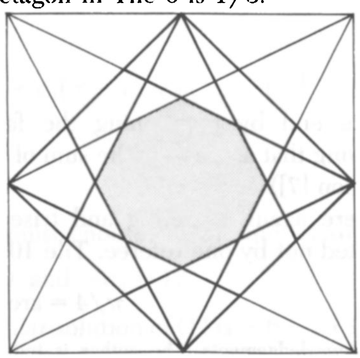



**TILE 5**

As with Tile 2, we now consider the case when each vertex is joined to trisection points of the corresponding two opposite sides (Tile 6). Then Tile 5 and Tile 7 show that the area of the shaded centrally symmetric octagon in Tile 6 is $1/3$.

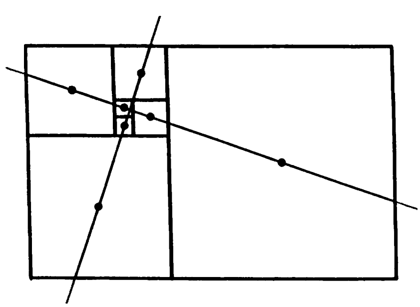


**TILE 6**

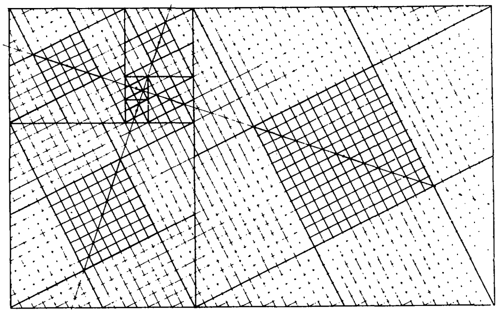


**TILE 7**

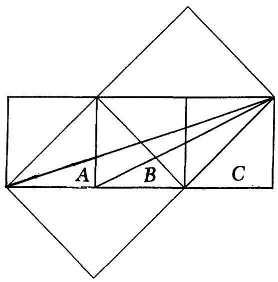
Recall that the Fibonacci sequence is defined by $F_1 = 1$, $F_2 = 1$, $F_n + F_{n+1} = F_{n+2}$. There is a well-known result [4, 5, 6] that when Fibonacci squares $1 \times 1$, $1 \times 1$, $2 \times 2$, $3 \times 3$, $5 \times 5$, $8 \times 8, \ldots$ are fitted together, their centers lie on two perpendicular lines whose slopes are 3 and $-1/3$ when the spiral of squares is oriented counterclockwise (Tile 8). The close relation between Tile 1 and Tile 2 as seen in Tile 3 underscores our proof without words of this result (Tile 9).



**TILE 8**



TILE 9




TILE 10

We end by mentioning the familiar Three Squares Problem, which involves showing that angle $C$ is the sum of angles $A$ and $B$ (Tile 10 with well-known look-see solution [7]).

Here again, bisecting and trisecting line segments play a key role. Indeed, as pointed out by one referee, Tile 10 is an illustration of Euler's famous formula for $\pi$:

$$\pi/4 = \arctan(1/2) + \arctan(1/3).$$

**Acknowledgements.** The author is indebted to the editor and to the referees for their insightful recommendations and guidance that transformed the original rambling material into a more coherent and cohesive piece. Many thanks also to Mark Billy for his careful diagrams.

## REFERENCES

1. H. S. M. Coxeter, *Introduction to Geometry*, John Wiley & Sons, Inc., New York, 1969, p. 212.
2. Ross Honsberger, *Mathematical Morsels*, Dolciani No. 3, MAA, Washington, DC, 1978, p. 204.
3. H. Dörrie, *Mathematische Miniaturen*, Breslau, 1943.
4. Herbert L. Holden, Fibonacci tiles, *The Fibonacci Quarterly* 13, (1975), 45–49.
5. Tony Gardiner, The Fibonacci spiral, *Mathematical Gazette*, (1984), 120.
6. Problem 1295, this MAGAZINE 62 (1989), 140–141.
7. Martin Gardner, *Mathematical Circus*, Spectrum Series, MAA, Washington, DC, 1992, p. 125.

# On Linear Difference Equations with Constant Coefficients: An Alternative to the Method of Undetermined Coefficients

R. C. GUPTA
National University of Singapore
Singapore 0511

The general solution of a nonhomogeneous linear difference equation can be expressed as the sum of a particular solution and the complementary function, which is the general solution of the associated homogeneous equation. Several methods for obtaining a particular solution are known [**1**, **2**, **3**]. One of the simpler methods is the method of undetermined coefficients [**3**]. For example, for the equation

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

an appropriate trial solution $y_p(x)$ is $c_0+c_1x+c_2x^2$ (see [**3**]); and the constant coefficients $c_0$, $c_1$, and $c_2$ need to be determined so that $y_p(x)$ satisfies the given equation for all $x$. Now

$$\begin{aligned} y_p(x+2) &- y_p(x+1) - y_p(x) \\ &= c_0+c_1(x+2)+c_2(x+2)^2-\left\{c_0+c_1(x+1)+c_2(x+1)^2\right\} \\ &\quad -\left\{c_0+c_1x+c_2x^2\right\} \\ &= 3c_2+c_1-c_0+(2c_2-c_1)x-c_2x^2. \end{aligned}$$

In order for $y_p$ to satisfy the given equation, we must choose $c_0$, $c_1$, and $c_2$ so that

$$3c_2+c_1-c_0=1,\ 2c_2-c_1=0, \text{ and } -c_2=1.$$

Thus $c_0=-6$, $c_1=-2$, and $c_2=-1$ and a particular solution of (1) is $-6-2x-x^2$.

Therefore, the method of undetermined coefficients is effected by appropriately choosing a trial solution containing unknown constants. Determination of these constants so that the trial solution satisfies the difference equation for every nonnegative integer $x$ leads to the required particular solution. The procedure is usually involved and one often ends up in tedious algebra. However, let us rewrite (1) in terms of the linear forward difference operator defined by $\Delta y(x)=y(x+1)-y(x)$ as

$$(\Delta^2+\Delta-1)y(x)=x^2+1, \tag{1$'$}$$

and operate on both sides of this equation by $\Delta$ repeatedly until the right side becomes a constant. Thus

$$(\Delta^3+\Delta^2-\Delta)y(x)=(x+1)^2+1-(x^2+1)=2x+1, \tag{2}$$

$$(\Delta^4+\Delta^3-\Delta^2)y(x)=2(x+1)+1-(2x+1)=2. \tag{3}$$

Clearly, $\Delta^2y(x)=-2$ is a solution of (3), then $\Delta^4y(x)=0=\Delta^3y(x)$. Using these in (2) we obtain $\Delta y(x)=-3-2x$. Finally, from (1$'$) one obtains

$$y(x)=-2+(-3-2x)-(x^2+1)=-6-2x-x^2,$$

which, as expected, is the same as the solution obtained earlier.

Motivated by this illustration, we present a very simple method for finding a particular solution of a nonhomogeneous linear difference equation to which the method of undetermined coefficients is applicable. Unlike the method of undetermined coefficients [3], this does not involve any chain of rules for a trial solution or the solution of simultaneous equations, but uses elementary algebra and the process of differencing.

**The procedure** The method consists in reducing the problem by the principle of superposition to finding a particular solution of

$$b_0y(x+n)+b_1y(x+n-1)+\cdots+b_ny(x)=a^xf(x),$$

which is the same as

$$\sum_{i=0}^{n} b_iE^{n-i}y(x)=a^xf(x). \tag{4}$$

Here $x$ belongs to the discrete set $N=\{0,1,2,\dots\}$, $E$ is the shift operator $Ey(x)=y(x+1)$, $f(x)$ is a polynomial of finite degree $n$ in $x$, and $a, b_i$, $0\leqslant i\leqslant n$, $b_0\neq 0$, $b_n\neq 0$ are (real or complex) constants.

To simplify the problem further, set $y(x)=a^xu(x)$ in (4). Since $E(a^xu(x))=a^{x+1}u(x+1)=a^x(aE)u(x)$, simple induction implies $E^k(a^xu(x))=a^x(aE)^ku(x)$ for any integer $k\geq 0$. Thus (4) becomes $\sum_{i=0}^{n}b_i(aE)^{n-i}u(x)=f(x)$ (see [2, p. 375]). Next we express the difference equation in terms of $\Delta$, by replacing $E$ by $\Delta+1$, and operate on both sides of this equation repeatedly by $\Delta$ until the right side becomes a constant. This equation can always be solved by inspection. Finally, by backward substitution a particular solution is obtained.

To illustrate, consider the equation

$$(E^2+2E-5)y(x)=3^x. \tag{5}$$

Set $y(x)=3^xu(x)$ to transform this equation to

$$(9E^2+6E-5)u(x)=1.$$

In terms of $\Delta$, this can be expressed as

$$(9\Delta^2+24\Delta+10)u(x)=1.$$

Clearly this equation is satisfied by $u(x)=\frac{1}{10}$, $\Delta u(x)=0=\Delta^2u(x)$. A $u(x)=\frac{1}{10}$ leads to $y(x)=\frac{1}{10}3^x$ as a particular solution of (5).

However, to make the operation on $f(x)$ by $\Delta$ easier it will be convenient if $f(x)$ is expressed as a polynomial in factorial powers defined by

$$x^{(n)}=x(x-1)(x-2)\cdots(x-n+1);\ n=1,2,3,\dots;\ x^{(0)}=1. \tag{6}$$

This can be accomplished easily by noting that $x=x^{(1)}$, $x^2=x^{(2)}+x^{(1)}$, $x^3=x^{(3)}+3x^{(2)}+x^{(1)}$, etc. In general, such conversions can be readily effected using Stirling numbers of the second kind $S_i^k$ (see [2]), which express the powers $x^k$ in terms of factorial powers as

$$x^k=\sum_{i=1}^{k}S_i^kx^{(i)}. \tag{7}$$

The coefficients $S_i^k$ can be found through a table-look-up, or from the recursive relation

$$S_i^{k+1} = S_{i-1}^k + iS_i^k,\ k > 0$$

with

$$S_k^k = 1,\ S_i^k = 0,\ i \leqslant 0,\ i \geqslant k+1. \tag{8}$$

Using relation (7) for $k = 0, 1, \dots, n$ one can easily express $f(x)$ as $\sum_{i=0}^n c_i x^{(i)}$, with constants $c_i, 0 \leqslant i \leqslant n$ depending on Stirling numbers of the second kind. Since

$$\begin{aligned}\Delta x^{(n)} &= (x+1)^{(n)} - x^{(n)} = (x+1)x(x-1)\dots(x+1-n+1) - x(x-1)\dots(x-n+1)\\ &= x(x-1)\dots(x-n+2)(x+1-(x-n+1)) = x(x-1)\dots(x-(n-1)+1)n,\end{aligned}$$

the factorial powers (6) satisfy

$$\Delta x^{(n)} = nx^{(n-1)}. \tag{9}$$

This procedure completely replaces the method of undetermined coefficients and the following examples manifest its versatility.

*Example* 1. The equation

$$y(x+2) - y(x) = 12x^2$$

is the same as

$$(E^2 - 1)y(x) = 12x^2, \tag{10}$$

which, in turn, is

$$(\Delta^2 + 2\Delta)y(x) = 12(x^{(2)} + x^{(1)}). \tag{10.1}$$

Therefore,

$$(\Delta^3 + 2\Delta^2)y(x) = 12(2x^{(1)} + 1), \text{ and} \tag{10.2}$$

$$(\Delta^4 + 2\Delta^3)y(x) = 24. \tag{10.3}$$

Obviously, a solution of (10.3) is $\Delta^3 y(x) = 12$, $\Delta^4 y(x) = 0$. Thus (10.2) gives $\Delta^2 y(x) = 12x$; consequently (10.1) gives $\Delta y(x) = 6x^{(2)}$, which in view of (9) leads to $y(x) = 2x^{(3)} = 2x(x-1)(x-2)$ as a particular solution of the given equation.

*Example* 2. The equation

$$y(x+2) - 5y(x+1) + 6y(x) = 4^x(x^2 - x + 5)$$

can be rewritten as

$$(E^2 - 5E + 6)y(x) = 4^x(x^2 - x + 5). \tag{11}$$

Set $y(x) = 4^x u(x)$ to transform this equation to

$$(16E^2 - 20E + 6)u(x) = x^{(2)} + 5, \tag{11.1}$$

which is the same as

$$(16\Delta^2 + 12\Delta + 2)u(x) = x^{(2)} + 5. \tag{11.2}$$

This gives

$$(16\Delta^3 + 12\Delta^2 + 2\Delta)u(x) = 2x^{(1)}, \text{ and} \tag{11.3}$$

$$(16\Delta^4 + 12\Delta^3 + 2\Delta^2)u(x) = 2.$$

From this we have $\Delta^2 u(x) = 1$, and $\Delta^3 u(x) = 0 = \Delta^4 u(x)$. Then (11.3) gives $\Delta u(x) = x - 6$. Thus from (11.2) we obtain

$$2u(x) = x^2 - x + 5 - 16 - 12(x-6) = x^2 - 13x + 61,$$

which gives $y(x) = \frac{1}{2}(x^2 - 13x + 61)4^x$ as a particular solution of (11).

*Example* 3. The equation

$$y(x+2) + y(x) = \sin x$$

in terms of the shift operator is

$$(E^2 + 1)y(x) = \sin x. \tag{12}$$

Since $\sin x$ is the imaginary part of $e^{ix}$, we solve

$$(E^2 + 1)z(x) = e^{ix}, \tag{12.1}$$

and then $y(x) = \operatorname{Im} z(x)$. To solve (12.1), set $z(x) = e^{ix}v(x)$ in (12.1) to obtain

$$\left((e^i E)^2 + 1\right)v(x) = 1,$$

or, equivalently,

$$\left(e^{2i}(\Delta^2 + 2\Delta) + (e^{2i} + 1)\right)v(x) = 1.$$

An obvious solution is

$$v(x) = \frac{1}{e^{2i} + 1}, \quad \Delta v(x) = 0 = \Delta^2 v(x),$$

and thus

$$y(x) = \operatorname{Im} \frac{\cos x + i \sin x}{(1 + \cos 2) + i \sin 2} = \frac{\sin(x-2) + \sin x}{2(1 + \cos 2)}$$

solves (12).

*Example* 4. To solve $y(x+2) - 2y(x+1)\cos\frac{1}{2} + y(x) = \cos\frac{x}{2}$, rewrite this as

$$\left(E^2 - 2(\cos\tfrac{1}{2})E + 1\right)y(x) = \cos\frac{x}{2}. \tag{13}$$

Here we solve

$$\left(E^2 - 2(\cos\tfrac{1}{2})E + 1\right)z(x) = e^{ix/2} \tag{13.1}$$

and $y(x) = \operatorname{Re} z(x)$. Set $z(x) = e^{ix/2}v(x)$ in (13.1) to get

$$\left(e^i\Delta^2 + 2\sin\tfrac{1}{2}(-\sin\tfrac{1}{2} + i\cos\tfrac{1}{2})\Delta\right)v(x) = 1.$$

This leads to $\Delta v(x) = -(\sin\frac{1}{2} + i\cos\frac{1}{2})/2\sin\frac{1}{2}$, $\Delta^2 v(x) = 0$. Therefore

$$v(x) = -x(\sin\tfrac{1}{2} + i\cos\tfrac{1}{2})/2\sin\tfrac{1}{2}$$

and finally, as a particular solution to (13), we get

$$y(x) = \operatorname{Re}\left(-\frac{x}{2\sin\frac{1}{2}}\left(\sin\frac{1}{2} + i\cos\frac{1}{2}\right)\left(\cos\frac{x}{2} + i\sin\frac{x}{2}\right)\right)$$

$$= x\sin\left(\frac{1}{2}(x-1)\right)/2\sin\frac{1}{2}.$$

*Example* 5. In order to solve

$$y(x+2)-2y(x+1)+4y(x)=-2^x\left(6\cos\frac{x\pi}{3}+2\sqrt{3}\sin\frac{x\pi}{3}\right),$$

express this as

$$(E^2-2E+4)y(x)=-2^x\left(6\cos\frac{x\pi}{3}+2\sqrt{3}\sin\frac{x\pi}{3}\right). \tag{14}$$

Setting $y(x)=2^x u(x)$ we obtain

$$(4E^2-4E+4)u(x)=-\left(6\cos\frac{x\pi}{3}+2\sqrt{3}\sin\frac{x\pi}{3}\right).$$

Since right side is $-(6\,\mathrm{Re}\,e^{ix\pi/3}+2\sqrt{3}\,\mathrm{Im}\,e^{ix\pi/3})$, we solve

$$4(E^2-E+1)z(x)=-e^{ix\pi/3} \tag{14.1}$$

and $u(x)=6\,\mathrm{Re}\,z(x)+2\sqrt{3}\,\mathrm{Im}\,z(x)$. As before, to solve (14.1) set $z(x)=e^{i\pi x/3}v(x)$ to obtain

$$4(e^{2i\pi/3}E^2-e^{i\pi/3}E+1)v(x)=-1.$$

This is equivalent to

$$4\left(-e^{2i\pi/3}\Delta^2+\left(3\cos\frac{\pi}{3}-i\sin\frac{\pi}{3}\right)\Delta\right)v(x)=1.$$

Its solution is $\Delta v(x)=\dfrac{1}{6-i2\sqrt{3}}$ and $\Delta^2 v(x)=0$. We, therefore, have

$$v(x)=x/(6-i2\sqrt{3})\ \textit{and}\ z(x)=xe^{i\pi x/3}/(6-i2\sqrt{3}).$$

Finally

$$y(x)=2^x\left(6\,\mathrm{Re}\,z(x)+2\sqrt{3}\,\mathrm{Im}\,z(x)\right)=2^x x\cos\frac{\pi x}{3}$$

is a particular solution of (14).

Comments from the referees are gratefully acknowledged.

## REFERENCES


1. B. Friedman, Difference equations, Chapter 4 in *Lectures on Applications–Oriented Mathematics*, Holden-Day, Inc., Oakland, CA, 1969.
2. Brand, L., *Differential and Difference Equations*, John Wiley & Sons, Inc., New York, 1966.
3. Goldberg, S., *Introduction to Difference Equations*, John Wiley & Sons, Inc., New York, 1967.

# A Note on the History of the Cantor Set and Cantor Function

JULIAN F. FLERON
SUNY at Albany
Albany, New York 12222

A search through the primary and secondary literature on Cantor yields little about the history of the Cantor set and Cantor function. In this note, we would like to give some of that history, a sketch of the ideas under consideration at the time of their discovery, and a hypothesis regarding how Cantor came upon them. In particular, Cantor was not the first to discover "Cantor sets." Moreover, although the original discovery of Cantor sets had a decidedly geometric flavor, Cantor's discovery of the Cantor set and Cantor function was neither motivated by geometry nor did it involve geometry, even though this is how these objects are often introduced (see e.g. [1]). In fact, Cantor may have come upon them through a purely arithmetic program.

The systematic study of point set topology on the real line arose during the period 1870–1885 as mathematicians investigated two problems:

1) conditions under which a function could be integrated, and
2) uniqueness of trigonometric series.

It was within the framework of these investigations that the two apparently independent discoveries of the Cantor set were made; each discovery linked to one of these problems.

Bernhard Riemann (1826–1866) spent considerable time on the first question, and suggested conditions he thought might provide an answer. Although we will not discuss the two forms his conditions took (see [2, pp. 17–18]), we note that one of these conditions is important as it eventually led to the development of measure theoretic integration [2, p. 28]. An important step in this direction was the work of Hermann Hankel (1839–1873) during the early 1870s. Hankel showed, within the framework of Riemann, that the integrability of a function depends on the nature of certain sets of points related to the function. In particular, "a function, he [Hankel] thought, would be Riemann-integrable if, and only if, it were *pointwise discontinuous* [2, p. 30]," meaning, in modern terminology, that for every $\sigma > 0$ the set of points $x$ at which the function oscillated by more than $\sigma$ in every neighborhood of $x$ was nowhere dense. Basic to Hankel's reasoning was his belief that sets of the form $\{1/2^n\}$ were prototypes for all nowhere dense subsets of the real line. Working under this assumption Hankel claimed that all nowhere dense subsets of the real line could be enclosed in intervals of arbitrarily small total length (i.e. had zero outer content) [2, p. 30]. As we shall see, this is not the case. (See also [3].)

Although Hankel's investigation into the nature of certain point sets would become extremely important, "as was the case with Dirichlet and Lipschitz, it was the inadequacy of his understanding of the possibilities of infinite sets—in particular, nowhere dense sets—that led him astray. It was not until it was discovered that nowhere dense sets can have positive outer content that the importance of negligible sets in the measure-theoretic sense was recognized [2, p. 32]." The discovery of such sets, nowhere dense sets with positive outer content, was made by H. J. S. Smith (1826–1883), Savilian Professor of Geometry at Oxford, in a paper [4] of 1875. After an exposition of the integration of discontinuous functions, Smith presented a method for constructing nowhere dense sets that were much more "substantial" than the set

$\{1/2^n\}$. Specifically, he observed the following:

> Let $m$ be any given integral number greater than 2. Divide the interval from 0 to 1 into $m$ equal parts; and exempt the last segment from any subsequent division. Divide each of the remaining $m-1$ segments into $m$ equal parts; and exempt the last segments from any subsequent subdivision. If this operation be continued *ad infinitum*, we shall obtain an infinite number of points of division $P$ upon the line from 0 to 1. These points lie in loose order... [4, p. 147].

In modern terminology Smith's 'loose order' is what we refer to as nowhere dense. Implicit in Smith's further discussion is the assumption that the exempted intervals are open, so the resulting set is closed. Today this set would be known as a general Cantor set, and this seems to be the first published record of such a set.

Later in the same paper, Smith shows that by dividing the intervals remaining before the $n$th step into $m^n$ equal parts and exempting the last segment from each subdivision we obtain a nowhere dense set of positive outer content. Smith was well aware of the importance of this discovery, as he states, "the result obtained in the last example deserves attention, because it is opposed to a theory of discontinuous functions, which has received the sanction of an eminent geometer, Dr. Hermann Hankel [4, p. 149]." He continues by explaining the difficulties in the contemporary theories of integration that his examples illuminate.

It is interesting to note that an editor's remark at the conclusion of Smith's paper states "this paper, *though it was not read*, was offered to the society and accepted in the usual manner." (Emphasis added.)[1] In fact, this paper went largely unnoticed among mathematicians on the European continent and unfortunately Smith's crucial discoveries lay unknown. It took the rediscovery, almost a decade later, of similar ideas by Cantor to illuminate the difficulties of contemporary theories of integration and to begin the evolution of measure-theoretic integration.

Georg Cantor (1845–1918) came to the study of point set topology after completing a thesis on number theory in Berlin in 1867. He began working with Eduard Heine (1821–1881) at the University of Halle on the question of the uniqueness of trigonometric series. This question can be posed as follows:

> If for all $x$ except those in some set $P$ we have
>
> $$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) = 0$$
>
> must all the coefficients $a_n$ and $b_n$ be zero?

Heine answered the question in the affirmative "when the convergence was *uniform in general* with respect to the set $P$, which is thus taken to be finite [2, p. 23]," meaning, by definition, that the convergence was uniform on any subinterval that did not contain any points of the finite set $P$.

Cantor proceeded much further with this problem. In papers [5, 6] of 1870 and 1871, he removed the assumption that the convergence was "uniform in general" and began to consider the case when $P$ was an infinite set. In doing so he began to look at what we now consider the fundamental point set topology of the real line. In a paper [8] of 1872, Cantor introduced the notion of a *limit point* of a set that he defined as we do today, calling the limit points of a set $P$ the *derived set*, which he denoted by $P'$. Then $P''$ was the derived set of $P'$, and so on. Cantor showed that if the set $P$ was

[1] It is possible that "not read" simply meant that the paper was not presented at a meeting of the London Mathematical Society. However, in weighing the significance of this note, one must consider that in vols. 3–10 of the *Proceedings of the London Mathematical Society* (1871–1879), and perhaps even further, no other paper was similarly noted.

such that $P^{(n)}=\varnothing$ for some integer $n$ and the trigonometric series $\frac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos(nx)+b_n\sin(nx))=0$, except possibly on $P$, then all of the coefficients had to be zero. Cantor's work on this problem was "decisive" [**9**, p. 49], and doubly important as his derived sets would play an important role in much of his upcoming work.

In the years 1879–1884 Cantor wrote a series of papers entitled "Über unendliche, lineare Punktmannichfaltigkeiten [**10–15**]," that contained the first systematic treatment of the point set topology of the real line.[2] It is the introduction of three terms in this series that concerns us most here. In the first installment of this series Cantor defines what it meant for a set to be *everywhere dense* (literally "überall dicht"), a term whose usage is still current. He gives a few examples, including the set of numbers of the form $2^{2^n+1}/2^m$ where $n$ and $m$ are integers, and continues by noting the relationship between everywhere dense sets and their derived sets. Namely, $P\subseteq(\alpha,\beta)$ is everywhere dense in $(\alpha,\beta)$ if [and only if] $P'=(\alpha,\beta)$ [**10**, pp. 2–3]. In the fifth installment of this series Cantor discusses the partition of a set into two components that he terms *reducible* and *perfect* [**14**, p. 575]. His definition of a perfect set is also still current: A set $P$ is perfect provided that $P=P'$.

After introducing the term *perfect* in the fifth installment, Cantor states that perfect sets need not be everywhere dense [**14**, p. 575]. In the footnote to this statement Cantor introduces the set that has become known as the *Cantor (ternary) set*: The set of real numbers of the form

$$x=\frac{c_1}{3}+\cdots+\frac{c_\nu}{3^\nu}+\cdots$$

where $c_\nu$ is 0 or 2 for each integer $\nu$. Cantor notes that this is an infinite, perfect set with the property that it is not everywhere dense in any interval, regardless of how small the interval is taken to be. We are given no indication of how Cantor came upon this set.

During the time Cantor was working on the 'Punktmannichfaltigkeiten' papers, others were working on extensions of the Fundamental Theorem of Calculus to discontinuous functions. Cantor addressed this issue in a letter [**18**] dated November 1883, in which he defines the Cantor set, just as it was defined in the paper [**14**] of 1883 (which had actually been written in October of 1882). However, in the letter he goes on to define the Cantor function, the first known appearance of this function. It is first defined on the complement of the Cantor set to be the function whose values are

$$\frac{1}{2}\left(\frac{c_1}{2}+\cdots+\frac{c_{\mu-1}}{2^{\mu-1}}+\frac{2}{2^\mu}\right)$$

for any number between

$$a=\frac{c_1}{3}+\cdots+\frac{c_{\mu-1}}{3^{\mu-1}}+\frac{1}{3^\mu}\quad\text{and}\quad b=\frac{c_1}{3}+\cdots+\frac{c_{\mu-1}}{3^{\mu-1}}+\frac{2}{3^\mu},$$

where each $c_\nu$ is 0 or 2. Cantor then concludes this section of the letter by noting that this function can be extended naturally to a continuous increasing function on $[0,1]$. That serves as a counterexample to Harnack's extension of the Fundamental Theorem of Calculus to discontinuous functions, which was in vogue at the time (see e.g. [**2**, p. 60]). We are given no indication of how Cantor came upon this function.

There are two other topics that interested Cantor that we would like to mention because they are indicative of Cantor's facility with arithmetic constructions and it is

[2] In addition, these papers contained many other topics that had far reaching implications (see [**16, 17**]), including Cantor's investigation of higher order derived sets that marked the "beginnings of Cantor's theory of transfinite numbers [**2**, p. 72]."

possibly within this setting that Cantor came upon the Cantor set and Cantor function. First, Cantor spent some time in the mid 1870s considering the possible existence of a bijective correspondence between a line and a plane, a question most of his contemporaries had dismissed as absurd. In 1877, in a letter to Richard Dedekind (1831–1916), Cantor explained that he had found such a correspondence. This "correspondence" can be expressed as follows:

> Let $(x_1, x_2)$ be a point in the unit square, and let $0.x_{1,1}x_{1,2}x_{1,3}\cdots$ and $0.x_{2,1}x_{2,2}x_{2,3}\cdots$ be decimal expansions of $x_1$ and $x_2$ respectively. Map the point $(x_1, x_2)$ to the point on the real line whose decimal expansion is $0.x_{1,1}x_{2,1}x_{1,2}x_{2,2}\cdots$ (See e.g. [**19**, p. 187].)

Dedekind pointed out that there was a problem with this approach. The decimal expansions of rationals are not unique, so to avoid duplication we must not allow expansions of some type, say expansions that contain infinite strings of zeros. However, by disallowing expansions with infinite strings of zeros, the irrational number $0.110102010102010102\ldots$ could never be obtained under Cantor's correspondence.

This reasoning does however give us an injection of $[0,1]\times[0,1]$ into $[0,1]$. It is trivial to find an injection of $[0,1]$ into $[0,1]\times[0,1]$. These two facts, together with the Schroeder-Bernstein Theorem (if there are injections of the set $A$ into the set $B$ and $B$ into $A$ respectively, then there is a bijective correspondence between $A$ and $B$; see e.g. [**20**]), allow us to conclude that there is a bijective correspondence between $[0,1]$ and $[0,1]\times[0,1]$. However, set theory was in its infancy during the period in question and it would be 20 years before E. Schroeder and Felix Bernstein independently proved the theorem that bears their names [**16**, p. 172–173] and occasionally Cantor's name as well (e.g. [**21**, **22**]). So this was not an option for Cantor.

Instead, Cantor needed to explicitly exhibit a bijection. To do this he modified his previous approach to use continued fractions [**23**]. Denote the continued fraction

$$\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\cdots}}}\quad \text{by } [a_1, a_2, a_3, \ldots] \quad \text{where} \quad a_1, a_2, a_3, \ldots > 0 \text{ are integers.}$$

Since a continued fraction is infinite if, and only if, it represents an irrational number, in which case the representation is unique [see e.g. **24**], Cantor could set up the correspondence

$$([a_{1,1}, a_{1,2}, \ldots], [a_{2,1}, a_{2,2}, \ldots], \ldots, [a_{n,1}, a_{n,2}, \ldots])$$
$$\leftrightarrow [a_{1,1}, a_{2,1}, \ldots a_{n,1}, a_{1,2}, a_{2,2}, \ldots, a_{n,2}, \ldots]$$

between $n$-tuples of irrationals in $(0,1)^n = (0,1)\times(0,1)\times\cdots\times(0,1)$ and irrationals in $(0,1)$. This avoids the difficulties of the previous approach and gives a bijective correspondence between $([0,1]-\mathbf{Q})^n$ and $[0,1]-\mathbf{Q}$. Cantor then took great lengths to prove there was a bijective correspondence between $[0,1]$ and $[0,1]-\mathbf{Q}$. Repeated application of this fact combined with the previous correspondence gives a bijective correspondence between $[0,1]^n$ and $[0,1]$.

Secondly, it is known that Cantor studied binary expansions. In fact:

> Cantor recognised that the power of the linear continuum, denoted by **o**, could be represented as well by [the power of] the set of all representations:
>
> $$x = \frac{f(1)}{2} + \cdots + \frac{f(\nu)}{2^\nu} + \cdots,$$
>
> where $f(\nu) = 0$ or $1$ [for each integer $\nu$] [**19**, p. 209].

There is, so it seems, no substantive evidence about how Cantor came upon the Cantor set and Cantor function. However, given Cantor's route into point set topology, his arithmetic introduction of the Cantor set and Cantor function, and his facility with arithmetic methods, as we have just illustrated, it is feasible that it is within the arithmetic framework of binary and ternary expansions that Cantor came upon the Cantor set and Cantor function.

## REFERENCES


1. R. L. Wheeden and A. Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Marcel Decker, Inc., New York, 1977, p. 35.
2. T. Hawkins, *Lebesgue's Theory of Integration: Its Origins and Development*, Chelsea Publishing Co., Madison, WI, 1975.
3. H. L. Royden, *Real Analysis*, Macmillan Publishing Co., New York, 1988, p. 64.
4. H. J. S. Smith, On the integration of discontinuous functions, *Proc. London Math. Soc.* (*1*) 6 (1875), 140–153.
5. G. Cantor, Beweis, daß eine für jeden reellen Werth von $x$ durch eine trigonometrische Reihe gegebene Function $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen läßt, Part 1, *Crelle Jl. Math.* 72 (1870), 139–142. Reprinted [7, pp. 80–83].
6. G. Cantor, Beweis, daß eine für jeden reellen Werth von $x$ durch eine trigonometrische Reihe gegebene Function $f(x)$ sich nur auf eine einzige Weise in dieser Form darstellen läßt, Part 2, *Crelle Jl. Math.* 73 (1871), 294–296. Reprinted [7, pp. 84–86].
7. G. Cantor, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, E. Zermelo (ed.), Springer-Verlag, New York, 1980.
8. G. Cantor, Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen, *Math. Ann.* 5 (1872), 123–132. Reprinted [7, pp. 92–102].
9. W. Purkert, Cantor's Philosophical Views, in *The History of Modern Mathematics*, Vol. 1: *Ideas and Their Reception*, D. E. Rowe and J. McCleary (eds.), Academic Press, Boston, 1989.
10. G. Cantor, Über unendliche, lineare Punktmannichfaltigkeiten, Part 1, *Math. Ann.* 15 (1879), 1–7. Reprinted [7, pp. 139–145].
11. G. Cantor, Über unendliche, lineare Punktmannichfaltigkeiten, Part 2, *Math. Ann.* 17 (1880), 355–358. Reprinted [7, pp. 145–148].
12. G. Cantor, Über unendliche, lineare Punktmannichfaltigkeiten, Part 3, *Math. Ann.* 20 (1882), 113–121. Reprinted [7, pp. 149–157].
13. G. Cantor, Über unendliche, lineare Punktmannichfaltigkeiten, Part 4, *Math. Ann.* 21 (1883), 51–58. Reprinted [7, pp. 157–164].
14. G. Cantor, Über unendliche, lineare Punktmannichfaltigkeiten, Part 5, *Math. Ann.* 21 (1883), 545–591. Reprinted [7, pp. 165–209].
15. G. Cantor, Über unendliche, lineare Punktmannichfaltigkeiten, Part 6, *Math. Ann.* 23 (1884), 453–488. Reprinted [7, pp. 210–246].
16. J. W. Dauben, *Georg Cantor: His Mathematics and Philosophy of the Infinite*, Princeton University Press, Princeton, NJ, 1990.
17. W. Purkert and H. J. Ilgauds, *Georg Cantor: 1845–1918*, Birkhäuser Verlag, Basel, 1987.
18. G. Cantor, De la puissance des ensembles parfaits de points, *Acta Math.* 4 (1884), 381–392. Reprinted [7, pp. 252–260].
19. J. W. Dauben, The development of Cantorian set theory, in *From the Calculus to Set Theory: 1630–1910*, I. Grattan-Guinness (ed.), Gerald Duckworth & Co., London, 1980.
20. G. F. Simmons, *Introduction to Topology and Modern Analysis*, R. E. Krieger Publishing Co., Malabar, FL, 1983, pp. 28–30.
21. K. Hrbacek and T. Jech, *Introduction to Set Theory*, Marcel Decker, Inc., New York, 1984, p. 72.
22. G. Takeuti and W. M. Zaring, *Introduction to Axiomatic Set Theory*, Springer-Verlag, New York, 1982, p. 86.
23. G. Cantor, Ein Beitrag zur Mannigfaltigkeitslehre, *Crelle Jl. Math.* 84, (1878), 242–258. Reprinted [7, pp. 119–133].
24. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon Press, Oxford, 1989, pp. 129–140.

# Covers of a Finite Set

ANTHONY J. MACULA
SUNY College at Geneseo
Geneseo, NY 14454

**1. Introduction** Let $X$ be a finite set with cardinality $N$. In a recent note in this MAGAZINE [1], Nelsen and Schmidt enumerate the chains in the power set of $X$ (ordered by set inclusion). It came as somewhat of a surprise to me that the analogous question, about the number of anti-chains in $\mathscr{P}(X)$, is open [2]. This note grows out of an inquiry into this latter question. An *anti-chain* in $\mathscr{P}(X)$ is a family of subsets of $X$ such that for any two members, one is not a subset of the other. If for a given $X$, one could count all the anti-chains $\beta$ for which $\bigcup\{S\colon S \in \beta\} = X$, then one could count *all* the anti-chains by summing over all nonempty subsets of $X$. We have not succeeded in counting this class of anti-chains, but the above observation leads one to consider a broader class of families of subsets of $X$ called the *covers of* $X$, denoted by $C(X)$. (See [3].) A family $\gamma$ of subsets of $X$ is called a *cover* of $X$, if $\bigcup\{S\colon S \in \gamma\} = X$. In this note, an application of the inclusion-exclusion principle is used to compute the cardinality of $C(X)$, a large number even for relatively small $X$. In fact, there are 32,297 covers of a four element set. For example $\{\{1\}, \{1,2\}, \{2,4\}, \{1,2,3\}, \{1,3,4\}, \{2,3,4\}\}$ is a proper nonminimal cover (see below) of $X = \{1,2,3,4\}$. However, more than half are trivial since any $\gamma$ with $X \in \gamma$ is a cover of $X$ (see Table 1). A computation of the cardinality of $C(X)$, in and of itself, is not so fundamental, and in all likelihood, may already be known, however we haven't been able to locate one (see [3], [4], [5], or [6]), and indeed, the one presented here may be the first. Our method has applications to counting some more interesting subclasses of $C(X)$. The most well-known subclass is the class of partitions $P(X)$ of $X$, the cardinality of which is given by the $N$th *Bell number* [4]. Some other subclasses are listed below; in all instances, $\gamma \in C(X)$:

(1) *proper covers*: $C'(X) = \{\gamma \in C(X)\colon X \notin \gamma\}$;

(2) *anti-chain covers*: $AC(X) = \{\gamma \in C(X)\colon \gamma$ is an anti-chains$\}$;

(3) *minimal covers*: $MC(X) = \{\gamma \in C(X)\colon$ For each $U \in \gamma$, there is an $x \in U$ such that $x \notin V$ for all $V \in \gamma$ with $V \neq U\}$, i.e., $\gamma$ is a *minimal cover* if the removal of one member destroys the covering property;

(4) *k-covers*: $k\text{–}C(X) = \{\gamma \in C(X)\colon$ For all $x \in X$, $x$ is in *at least* $k$ members of $\gamma\}$;

(5) *$k^*$-covers*: $k^*\text{–}C(X) = \{\gamma \in C(X)\colon$ For all $x \in X$, $x$ is in *exactly* $k$ members of $\gamma\}$.

The cardinality of $C'(X)$ is computed here. We have counted $|MC(X)|$ in [7], and used the results in this paper to count $2\text{–}C(X)$. It's worth noting that $MC(X) \subset AC(X) \subset C'(X)$, hence $|C'(X)|$ and $|MC(X)|$, respectively, give upper and lower bounds for $|AC(X)|$. Both inclusions are strict, as for $X = \{1,2,3\}$, the family $\{\{1,2\},\{1,3\},\{2,3\}\}$ is a nonminimal anti-chain cover, while $\{\{1\},\{1,3\},\{2,3\}\}$ is a proper cover that is not an anti-chain. Also note that a 1-cover is just a cover, and a $1^*$-cover is just a partition; so $|1^*\text{–}C(X)|$ is the $N$th *Bell number* $\sigma(N)$. In this light, the numbers $|k^*\text{–}C(X)|$ are of particular interest as generalizations of the Bell numbers.

**2. Families of subsets** Throughout this paper, $X$ denotes the finite set $\{1,2,\ldots,N\}$, $|S|$ denotes the number of elements in the finite $S$, $\mathscr{P}(X)$ denotes the power set of $X$, and we let $\bigcup\gamma$ denote the subset of $X$ covered by $\gamma$ (i.e., $\bigcup\gamma=\bigcup_{S\in\gamma}S$). From basic set theory we have

$$|\mathscr{P}(X)|=2^{|X|}=2^N. \tag{1}$$

We let $\mathscr{P}_2(X)$ denote the collection of all nonempty families of nonempty subsets of $X$. $\mathscr{P}_2(X)=\mathscr{P}(\mathscr{P}(X)-\{\varnothing\})-\{\varnothing\}$, and from (1), we have

$$|\mathscr{P}_2(X)|=2^{2^N-1}-1=\frac{2^{2^N}}{2}-1. \tag{2}$$

**3. The number of covers of X** If $\gamma$ is not a cover of $X$, there is an $x\in X$ that is not covered by $\gamma$, so $\bigcup\gamma\neq X$. Therefore,

$$\begin{aligned}|C(X)|&=|\mathscr{P}_2(X)-\{\gamma\in\mathscr{P}_2(X)\colon \bigcup\gamma\neq X\}|\\&=|\mathscr{P}_2(X)-\{\gamma\colon \bigcup\gamma\subseteq X-\{x\},\text{ for some }x\in X\}|\\&=|\mathscr{P}_2(X)|-|(\textstyle\bigcup_{x\in X}\mathscr{P}_2(X-\{x\}))|.\end{aligned}$$

Let $\eta(X)=|(\bigcup_{x\in X}\mathscr{P}_2(X-\{x\}))|$. Then $|C(X)|=\left(\frac{2^{2^N}}{2}-1\right)-\eta(X)$. To find $\eta(X)$ we use the inclusion-exclusion principle:

*Inclusion-Exclusion Principle.* If $A_1,A_2,\ldots,A_k$ are finite sets, then

$$|\bigcup A_i|=\sum_{i=1}^{k}(-1)^{i+1}\cdot\xi_i, \tag{3}$$

where, $\xi_i$ is the sum of the cardinalities of the intersections of the sets taken $i$ at a time. See [8].

Since, for $K\subseteq X$, with $|K|=k$, $\bigcap_{x\in K}\mathscr{P}_2(X-\{x\})=\mathscr{P}_2(X-K)$ and, by (2), $|\mathscr{P}_2(X-K)|=\frac{2^{2^{N-k}}}{2}-1$; an application of (3) gives:

$$\eta(X)=\sum_{k=1}^{N}(-1)^{k+1}\cdot\binom{N}{k}\cdot\left(\frac{2^{2^{N-k}}}{2}-1\right).$$

And since

$$\sum_{k=1}^{N}(-1)^{k+1}\binom{N}{k}=1,$$

$$\eta(X)=\left(\frac{1}{2}\sum_{k=1}^{N}(-1)^{k+1}\cdot\binom{N}{k}\cdot 2^{2^{N-k}}\right)-1.$$

Therefore,

$$|C(X)|=\left(\frac{2^{2^N}}{2}-1\right)-\left(\frac{1}{2}\sum_{k=1}^{N}(-1)^{k+1}\cdot\binom{N}{k}\cdot 2^{2^{N-k}}\right)+1,$$

and placing $\frac{2^{2^N}}{2}$ inside the sum we get:

$$|C(X)|=\frac{1}{2}\sum_{k=0}^{N}(-1)^{k}\cdot\binom{N}{k}\cdot 2^{2^{N-k}}.$$

A family $\gamma$ is called *proper* if $X \notin \gamma$. We let $\mathscr{P}_2'(X)$ denote the collection of all proper families in $\mathscr{P}_2(X)$, i.e., $\mathscr{P}_2'(X) = \mathscr{P}(\mathscr{P}(X) - \{\varnothing, X\}) - \{\varnothing\}$, and

$$|\mathscr{P}_2'(X)| = 2^{2^N - 2} - 1 = \frac{2^{2^N}}{4} - 1.$$

The set of covers of $X$ that have $X$ as a member is clearly $\mathscr{P}_2(X) - \mathscr{P}_2'(X)$, which has cardinality

$$\left(\frac{2^{2^N}}{2} - 1\right) - \left(\frac{2^{2^N}}{4} - 1\right) = \frac{2^{2^N}}{4}.$$

Therefore, $C'(X)$, the set of *proper covers*, has cardinality

$$|C'(X)| = \left(\frac{1}{2} \sum_{k=0}^{N} (-1)^k \cdot \binom{N}{k} \cdot 2^{2^{N-k}}\right) - \frac{2^{2^N}}{4}.$$

Some values of $|C(X)|$ and $|C'(X)|$ are displayed in Table 1. Now a cover may have any size from 1 ($\gamma = \{X\}$) to $2^N - 1$ ($\gamma = \mathscr{P}(X) - \varnothing$). For technical reasons in counting other subclasses of $C(X)$, we need to enumerate the set of *proper* covers of length $n$ (denoted by $C_n'(X)$), where $1 \le n \le 2^N - 1$. This method is essentially the same as the one above and is outlined below. We first count the set of *all* covers of length $n$ (denoted by $C_n(X)$).

TABLE 1. Cardinalities of some subclasses of $C(X)$ where $|X| = N$.

| | partitions | covers | proper covers |
|---|---|---|---|
| $N=1$ | 1 | 1 | 0 |
| $N=2$ | 2 | 5 | 1 |
| $N=3$ | 5 | 109 | 45 |
| $N=4$ | 15 | 32,297 | 15,913 |
| $N=5$ | 52 | 2,147,321,017 | 1,073,579,193 |
| $N=6$ | 203 | $\approx 9.223 \times 10^{18}$ | $\approx 4.612 \times 10^{18}$ |
| $N=7$ | 877 | $\approx 1.701 \times 10^{39}$ | $\approx 8.501 \times 10^{38}$ |

Let $\mathscr{P}_{2,n}(X) = \{\gamma \in \mathscr{P}_2(X) : |\gamma| = n\}$. Then

$$|\mathscr{P}_{2,n}(X)| = \binom{2^N - 1}{n}.$$

Analogously

$$C_n(X) = \mathscr{P}_{2,n}(X) - \{\gamma \in \mathscr{P}_{2,n}(X) : \cup\gamma \ne X\}.$$
$$= \mathscr{P}_{2,n}(X) - \left(\bigcup_{x \in X} \mathscr{P}_{2,n}(X - \{x\})\right).$$

Let $\eta_n(X) = |(\bigcup_{x \in X} \mathscr{P}_{2,n}(X - \{x\}))|$. Then

$$|C_n(X)| = \binom{2^N - 1}{n} - \eta_n(X).$$

To find $\eta_n(X)$, we use the inclusion-exclusion principle as before. Since, for $K \subseteq X$ ($|K| = k$), $\bigcap_{x \in K} \mathscr{P}_{2,n}(X - \{x\}) = \mathscr{P}_{2,n}(X - K)$ and, $|\mathscr{P}_{2,n}(X - K)| = \binom{2^{N-k} - 1}{n}$. An application of (3) gives

$$\eta_n(X) = \sum_{k=1}^{N} (-1)^{k+1} \cdot \binom{N}{k} \cdot \binom{2^{N-k} - 1}{n}.$$

Therefore,

$$|C_n(X)| = \binom{2^N - 1}{n} - \sum_{k=1}^{N} (-1)^{k+1} \cdot \binom{N}{k} \cdot \binom{2^{N-k} - 1}{n}$$
$$= \sum_{k=0}^{N} (-1)^k \cdot \binom{N}{k} \cdot \binom{2^{N-k} - 1}{n}.$$

Let $\mathscr{P}'_{2,n}(X) = \{\gamma \in \mathscr{P}'_2(X) \colon |\gamma| = n\}$. Then $|\mathscr{P}'_{2,n}(X)| = \binom{2^N - 2}{n}$. The set of covers of size $n$ that contain $X$ is

$$\mathscr{P}_{2,n}(X) - \mathscr{P}'_{2,n}(X),$$

which has cardinality $\binom{2^N - 2}{n-1}$. Hence

$$|C'_n(X)| = \left[\sum_{k=0}^{N} (-1)^k \cdot \binom{N}{k} \cdot \binom{2^{N-k} - 1}{n}\right] - \binom{2^N - 2}{n-1}.$$

*Author's note.* Recently, we discovered that this enumeration appears in Louis Comlet, *Advanced Combinatorics*, D. Reidel, Dordrecht, 1974.

**Acknowledgement.** The author is most grateful for the insightful and thoughtful suggestions of the referees on the style and content of this note.

## REFERENCES

1. R. B. Nelsen and H. Schmidt, Jr., Chains in power sets, this MAGAZINE, 64(1991), 23–31.
2. I. Rival, Unsolved problems, *Order*, 1(1984), 103–105.
3. L. L. Dornhoff and F. E. Hohn, *Applied Modern Algebra*, Macmillan Publishing Co., New York, 1978.
4. N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, Orlando, FL, 1973.
5. C. Greene and D. J. Kleitman, Proof techniques in the theory of finite sets, *Studies in Combinatorics*, G. Rota ed. MAA, Washington, DC, 22–79, 1978.
6. P. Erdos and D. J. Kleitman, Extremal problems among subsets of a set, *Discrete Mathematics*, 8(1974), 281–294.
7. A. J. Macula, Venn Diagrams and Minimal Covers, manuscript.
8. N. L. Biggs, *Discrete Mathematics*, Oxford University Press, Oxford, 1985.

# PROBLEMS

LOREN C. LARSON, *editor*
St. Olaf College

GEORGE GILBERT, *associate editor*
Texas Christian University

## Proposals

*To be considered for publication, solutions should be received by September 1, 1994.*

**1443.** *Proposed by Allen J. Schwenk, Western Michigan University, Kalamazoo, Michigan.*

A river flows from Town A to Town B and has the property that any point on either of its banks is no farther than 100 yards from some point on the other bank.

a. A boat sails down the river, trying to stay always within a distance $d$ from both banks. For what values of $d$ can we guarantee that such a trip is possible?

b. How far can a swimmer in the river be from the nearest bank?

**1444.** *Proposed by Cristian Turcu, London, England.*

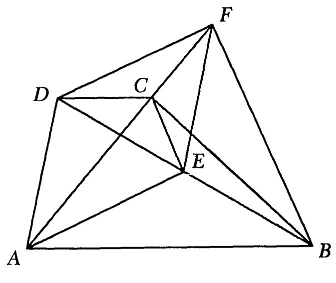
In the following figure, $ABCD$ is a trapezoid, with $AB$ parallel to $CD$, and the length of $AB$ is the sum of the lengths of $AC$ and $CD$. $E$ is the midpoint of $BD$, and $F$ is a point on $AC$ such that $BF$ is parallel to $CE$.

Prove that

a. $AE$ and $DF$ are perpendicular to $BF$;

b. $C$ is the incenter of triangle $DEF$ if, and only if, $AD$ is perpendicular to $AB$;

c. $EF$ is parallel to $AD$ if, and only if, the length of $AB$ is 3 times the length of $CD$.




ASSISTANT EDITORS: CLIFTON CORZAT, BRUCE HANSON, RICHARD KLEBER, KAY SMITH, *and* THEODORE VESSEY, *St. Olaf College and* MARK KRUSEMEYER, *Carleton College. We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) next to a problem number indicates that neither the proposer nor the editors supplied a solution.*

*Solutions should be written in a style appropriate for Mathematics Magazine. Each solution should begin on a separate sheet containing the solver's name and full address.*

*Solutions and new proposals should be mailed in duplicate to Loren Larson, Department of Mathematics, St. Olaf College, 1520 St. Olaf Ave., Northfield, MN 55057-1098 or mailed electronically via fax: (507) 663-3549 or e-mail: larson@stolaf.edu.*

**1445.** *Proposed by Ilya V. Burkov, St. Petersburg Technical University, St. Petersburg, Russia.*

Let $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ and $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ be orthonormal right-oriented triplets of vectors in $\mathbf{R}^3$, and $k_1, k_2, k_3$ be nonzero real numbers with different absolute values such that

$$k_1(\mathbf{r}_1 \times \mathbf{s}_1) + k_2(\mathbf{r}_2 \times \mathbf{s}_2) + k_3(\mathbf{r}_3 \times \mathbf{s}_3) = \mathbf{0}.$$

Prove that $\mathbf{r}_i$ is parallel to $\mathbf{s}_i$, $i = 1, 2, 3$.

**1446.** *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Canada.*

Determine the least number of times the graph of

$$y = \frac{a^2}{x^2-1} + \frac{b^2}{x^2-4} + \frac{c^2}{x^2-9} - 1$$

intersects the $x$-axis ($a, b, c$ are nonzero real constants).

**1447.** *Proposed by Florin S. Pîrvănescu, Slatina, Romania.*

Let $M$ denote an arbitrary point inside or on a tetrahedron $A_1A_2A_3A_4$, and let $B_i$ be a point on the face $F_i$ opposite vertex $A_i$, $i = 1, 2, 3, 4$. For each $i$, let $M_i$ be the point where the line through $M$ parallel to $A_iB_i$ intersects $F_i$. Show that

$$\min_{1 \le i \le 4} A_iB_i \le \sum_{i=1}^{4} MM_i \le \max_{1 \le i \le 4} A_iB_i.$$

# Quickies

*Answers to the Quickies are on page 152.*

**Q817.** *Proposed by Robert B. McNeill, Northern Michigan University, Marquette, Michigan.*

Determine all positive integers that are expressible in the form $a^2 + b^2 + c^2 + c$, where $a$, $b$, and $c$ are integers.

**Q818.** *Proposed by Jiro Fukuta, Kamimakuwa, Shinsei-cho, Gifu-ken, Japan.*

In a quadrangle $ABCD$ inscribed in a circle with center $O$, reflect four segments of the circle that are out of $ABCD$ in $AB$, $BC$, $CD$, and $DA$. Prove that if $O$ is an interior point of $ABCD$, the aggregate of the four reflected segments cannot cover the quadrangle.

**Q819.** *Proposed by L. Van Hamme, Vrije Universiteit Brussel, Brussels, Belgium.*

Prove that the residue of the function $1/(1-e^z)^n$ at zero is equal to $-1$ for all positive integers $n$.

# Solutions

**Euclidean construction** **April 1993**

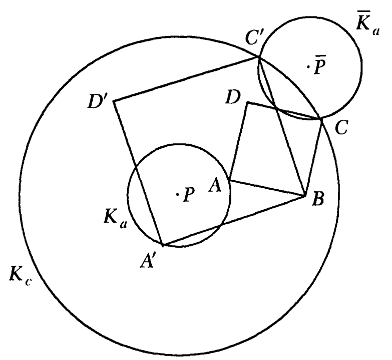
**1418.** *Proposed by Irvin Roy Hentzel and Richard H. Sprague, Iowa State University, Ames, Iowa.*

Given three distances $a, b, c$ construct (using straightedge and compass and without analytic geometry) a square $ABCD$ and a point $P$ such that $|PA| = a$, $|PB| = b$, and $|PC| = c$.

*Solution by László Szücs, Fort Lewis College, Durango, Colorado.*

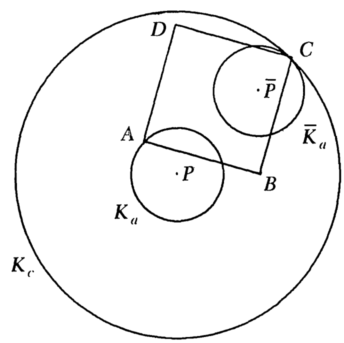
Choose an arbitrary point $P$, and any point $B$ at a distance $b$ from $P$. Draw concentric circles $K_a$ and $K_c$ with center at $P$ and radii $a$ and $c$ respectively. Rotate $K_a$ about $B$ through an angle of $90°$, and denote the rotated circle by $\overline{K}_a$, and its center by $\overline{P}$. The rotation is to be performed in the clockwise or counterclockwise direction according as the vertices $A, B, C, D$ of the square are to be oriented counterclockwise or clockwise, respectively. There are three cases to consider.

*Case* 1. $|c - a| < \sqrt{2}\,b < c + a$. In this case, $\overline{K}_a$ intersects $K_c$ in two points $C$ and $C'$. The inverse image of $C$ under this rotation is a point $A$ on $K_a$ such that $\angle ABC = 90°$ and $|AB| = |BC|$. The completion of the isosceles right triangle $ABC$ to the desired square $ABCD$ is straightforward (see figure). Similarly, using the point $C'$, we can construct a second square $A'BC'D'$ with the required properties.




In the degenerate case in which $a = b = c$, one of the points of intersection is the point $B$, so in this case, only one point of intersection can be used as $C$, and only one square can be constructed for each orientation.

*Case* 2. $\sqrt{2}\,b = |c - a|$ or $\sqrt{2}\,b = c + a$. In this case, $\overline{K}_a$ is tangent to $K_c$ at a point $C$. We proceed as in Case 1 and obtain just one square $ABCD$.



*Case* 3. $\sqrt{2}b < |c-a|$ or $\sqrt{2}b > c+a$. In this case, $\overline{K}_a$ does not intersect $K_c$, and there is no square $ABCD$ having the required properties. For if there were such a square, then rotating $K_a$ about $B$ through an angle of 90° would have to map $A$ to $C$, hence $\overline{K}_a$ would have to intersect $K_c$ at $C$.

*Also solved by R. Akhlaghi and R. Dai, Ching Avery (Hong Kong), Michael Bertrand, Gerald D. Brown, Con Amore Problem Group (Denmark), Milton P. Eisner, Wee Liang Gan (student, Singapore), Robert Geretschläger (Austria), David Hankin, Hans Kappus (Switzerland), Detlef Laugwitz (Germany), Maria Ascensión López (Spain), O. P. Lossers (The Netherlands), Helen M. Marston, Brad Parsons, Waldemar Pompe (Poland), Werner Raffke (Germany), Ruedi Suter (Switzerland), Kao H. Sze and Irene C. Sze, Michael E. Seidler, and the proposers.*

**Primitive roots modulo $p^n$** **April 1993**

**1419.** *Proposed by William P. Wardlaw, U.S. Naval Academy, Annapolis, Maryland.*

Show that for each odd prime $p$, there is an integer $g$ such that $1 < g < p$ and $g$ is a primitive root modulo $p^n$ for every positive integer $n$.

*Solution by Richard Holzsager, The American University, Washington, DC.*

As is well known, the multiplicative group $\{1, 2, \dots, p-1\}$ modulo $p$ is cyclic, so there is a $g$ such that $1 < g < p$ and $g$ is a primitive root modulo $p$. If $g^{p-1}$ is not congruent to 1 modulo $p^2$, say $g^{p-1} = kp + 1$ with $k$ prime to $p$, then, using the binomial expansion, $g^{r(p-1)} \equiv krp + 1 \pmod{p^2}$, so $g$ is a primitive root modulo $p^2$. In fact, again from the binomial expansion, $g^{p(p-1)} \equiv kp^2 + 1 \pmod{p^3}$. This prepares the next step of an induction that uses the same argument to show that for $r = 1, 2, \dots, n-1$ and all $i$, $p^{rp^i(p-1)} \equiv krp^i + 1 \pmod{p^{i+1}}$. This says that $g$ is the primitive root we are looking for.

We still have to handle the case $g^{p-1} \equiv 1 \pmod{p^2}$. Let $h < p$ be the inverse of $g$ modulo $p$. Since $gh < p^2$, $gh = kp + 1$, with $0 < k < p$. Therefore, $g^{p-1}h^{p-1} \equiv k(p-1)p + 1 \pmod{p^2}$, from which it follows that $h^{p-1}$ is not congruent to 1 modulo $p^2$. Using $h$ in place of $g$ and the results of the previous paragraph, we are done.

*Also solved by R. Akhlaghi and R. Dai, David Callan, Hugh Edgar, Stephen I. Gendler, David W. Koster, Kee-Wai Lau (Hong Kong), Peter W. Lindstrom, O. P. Lossers (The Netherlands), David E. Manes, Neville Robbins, Richard F. Ryan, Michael J. Semenoff, Trinity University Problem Group, Vu Ha Van (student, Hungary), and the proposer.*

Several readers noted that this result is standard material in number theory textbooks; for example, see Apostol's *Introduction to Analytic Number Theory*, Springer-Verlag, 1984, Section 10.6, pp. 208–210. *Harris Kwong*, SUNY College at Fredonia, mentioned the following result by D. Kruyswijk ("On the congruence $u^{p-1} \equiv 1$ modulo $p^2$," *Mathematisch Centrum Amsterdam*, 1966; see *Mathematical Reviews*, No. 3995, April 1967, p. 676). Define a primitive root of $p$ to be *strong* if it is also a primitive root of $p^2, p^3, \dots$. For every $p$, at least half of all primitive roots on $[0, p]$ are strong.

**Analytic trigonometry** **April 1993**

**1420.** *Proposed by Cristian Turcu, London, England.*

If $\alpha, \beta, \gamma, \delta$ are real numbers, $n$ is an odd integer, $\cos\alpha + \cos\beta + \cos\gamma + \cos\delta = 0$, and $\sin\alpha + \sin\beta + \sin\gamma + \sin\delta = 0$, prove that $\cos n\alpha + \cos n\beta + \cos n\gamma + \cos n\delta = 0$ and $\sin n\alpha + \sin n\beta + \sin n\gamma + \sin n\delta = 0$.

I. *Solution by Reiner Martin, student, University of California at Los Angeles, Los Angeles, California.*

Let $\omega_1 = \cos\alpha + i\sin\alpha, \dots, \omega_4 = \cos\delta + i\sin\delta$. These are numbers on the complex unit circle, and our assumptions translate into $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0$. We have to show that $\omega_1^n + \omega_2^n + \omega_3^n + \omega_4^n = 0$, for odd $n$.

If $\omega_1 = -\omega_2$ than $\omega_3 = -\omega_4$, and we are done. Otherwise the unit circle centered at $z = -\omega_1 - \omega_2$ intersects the unit circle centered at zero in at most two points, and so it is clear that the condition $\omega_3 + \omega_4 = z$ determines $\{\omega_3, \omega_4\}$ uniquely. Thus, $\{\omega_3, \omega_4\} = \{-\omega_1, -\omega_2\}$, and again we are done.

II. *Solution by Thomas Jager, Calvin College, Grand Rapids, Michigan.*

Using the notation of the preceding solution, the fact that $\omega_i\overline{\omega}_i = 1$, and the assumption that $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0$, we have

$$(\omega_1+\omega_2)(\omega_1+\omega_3)(\omega_1+\omega_4) = \omega_1^2(\omega_1+\omega_2+\omega_3+\omega_4) + \omega_1\omega_2\omega_3\omega_4\left(\overline{\omega}_1+\overline{\omega}_2+\overline{\omega}_3+\overline{\omega}_4\right) = 0.$$

Without loss of generality, we may conclude that $\omega_2 = -\omega_1$. Hence $\omega_4 = -\omega_3$, and thus, for odd $n$, $\omega_1^n + \omega_2^n + \omega_3^n + \omega_4^n = 0$.

III. *Solution by Hans Kappus, Rodersdorf, Switzerland.*

Let $u, v, w$ be complex numbers satisfying

$$u + v + w = -1, \qquad |u| = |v| = |w| = 1. \tag{1}$$

Taking the conjugate of (1) and multiplying both sides by $t = uvw$ we obtain

$$vw + wu + uv = -t.$$

Thus, $u, v, w$ are the roots of the cubic equation

$$z^3 + z^2 - tz - t = 0.$$

But $z^3 + z^2 - tz - t \equiv (z+1)(z^2 - t)$ and consequently, $u = -1$ and $w = -v$, say. Therefore,

$$u^n + v^n + w^n = -1, \quad \text{for odd } n.$$

The statement of the problem now follows upon taking $u = e^{i(\beta-\alpha)}$, $v = e^{i(\gamma-\alpha)}$, and $w = e^{i(\delta-\alpha)}$.

IV. *Solution by Nick Lord, Tonbridge School, Kent, England.*

Let $a = e^{i\alpha}$, $b = e^{i\beta}$, $c = e^{i\gamma}$, $d = e^{i\delta}$. Then, by adding and subtracting the given relations, we deduce that $a + b + c + d = 0$ and $a^{-1} + b^{-1} + c^{-1} + d^{-1} = 0$. Thus,

$$(1+at)(1+bt)(1+ct)(1+dt) = (1-at)(1-bt)(1-ct)(1-dt)$$

since the coefficients of $t$ $(\pm(a+b+c+d))$ and of $t^3$ $(\pm abcd(a^{-1}+b^{-1}+c^{-1}+d^{-1}))$ on each side are zero. Hence,

$$\begin{aligned} 0 &= \ln\left(\frac{1+at}{1-at}\cdot\frac{1+bt}{1-bt}\cdot\frac{1+ct}{1-ct}\cdot\frac{1+dt}{1-dt}\right) \\ &= \ln\left(\frac{1+at}{1-at}\right) + \ln\left(\frac{1+bt}{1-bt}\right) + \ln\left(\frac{1+ct}{1-ct}\right) + \ln\left(\frac{1+dt}{1-dt}\right) \\ &= 2\sum_{n \text{ odd}} (a^n + b^n + c^n + d^n)\frac{t^n}{n}, \qquad \text{if } |t| < 1 \end{aligned}$$

using the expansion of $\ln((1+at)/(1-at))$, etc. Equating coefficients of $t^n$ shows that $a^n+b^n+c^n+d^n=0$, which, by de Moivre's Theorem, give the desired result.

*Also solved by Sinefakopoulos Achilleas (student, Greece), R. Akhlaghi and R. Dai, Ricardo Alfaro, S. J. Becker, Francisco Bellot (Spain), Manjul Bhargava (student), J. C. Binz (Switzerland), David Callan, Con Amore Problem Group (Denmark), Miguel Amengual Covas (Spain), Robert L. Doucette, Diane Dowling and Roy Dowling (Canada), Kevin Ford (student), Dr. Freidkin, Jiro Fukuta (Japan), Wee Liang Gan (student, Singapore), G. A. Heuer, Murray S. Klamkin (Canada), David W. Koster, Kee-Wai Lau (Hong Kong), Peter W. Lindstrom, O. P. Lossers (The Netherlands), Helen M. Marston, MATC Problem Solving Group (two solutions), W. Weston Meyer, Kandasamy Muthuvel, Istvan Nemes (Austria), New Mexico Tech Problem Solving Group, Edward D. Onstott and Wayne Dennis, Allan Pedersen (Denmark), F. C. Rembis, Michael J. Semenoff, John S. Sumner and Kevin L. Dove, Ruedi Suter (Switzerland), Kao H. Sze and Irene C. Sze, University of Wyoming Problem Circle, Vu Ha Van (student, Hungary), Michael Vowe (Switzerland), A. N.'t Woord (The Netherlands), Robert L. Young, Sammy (age 13) and Jimmy (age 11) Yu, and the proposers.*

*Klamkin* found an equivalent form of the problem in H. S. Carslan's *Plane Trigonometry*, MacMillan, London, 1948, Exercise 5, p. 199.

**Concyclic circumcenters** **April 1993**

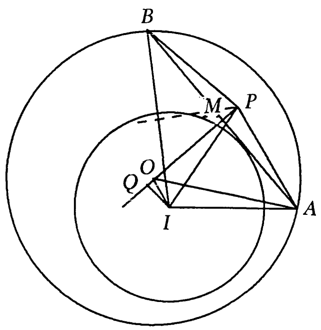
**1421.** *Proposed by Jiro Fukuta, Motosu-gun, Gifu-ken, Japan.*

If a polygon $A_1A_2\ldots A_n$ has an inscribed circle with center $I$ and a circumcircle with center $O$, and $C_i$ is the circumcenter of the triangle $IA_iA_{i+1}$ ($i=1,2,\ldots,n$, where $A_{n+1}=A_1$), prove that the $C_i$'s are concyclic.

I. *Solution by Kiran Kedlaya, student, Harvard University, Cambridge, Massachusetts.*

Let $R$ be the radius of the circumcircle, $r$ the radius of the incircle, and $d$ the distance $OI$. We will show that the centers $C_i$ lie on a circle centered at $O$ with radius $(R^2-d^2)/(2r)$.

Let $AB$ be a side of the polygon with midpoint $M$. Let $P$ be the circumcenter of triangle $IAB$, and $Q$ the projection of $I$ onto $OM$.




Of course, $P$ lies on the perpendicular bisector of $AB$, which is $OM$. Using directed lengths along the line $OM$, we find that

$$\begin{aligned}OA^2-OI^2&=(OA^2-PA^2)+(PI^2-OI^2)\\&=(OM^2+MA^2-PM^2-MA^2)+(PQ^2+QI^2-OQ^2-QI^2)\\&=(OM-PM)(OM+PM)+(PQ-OQ)(PQ+OQ)\\&=(OP)(OM+PM)+(PO)(PQ+OQ)\\&=(OP)(OM+PM+QP+QO)\\&=(OP)(2QM).\end{aligned}$$

In other words, $OP=(R^2-d^2)/(2r)$ as claimed.

II. *Solution by O. P. Lossers, Eindhoven University of Technology, Eindhoven, The Netherlands.*

Without loss of generality we may assume that the incircle is the unit circle $x^2+y^2-1=0$ with center $I=(0,0)$, and that the circumcircle $C$ has the equation

$$C\colon (x-a)^2+y^2-R^2=0.$$

An arbitrary line $L_\varphi$ that is tangent to the incircle can be written as

$$L_\varphi\colon x\cos\varphi+y\sin\varphi-1=0.$$

Any circle passing through the points of intersection of $C$ and $L_\varphi$ has an equation of the form

$$(x-a)^2+y^2-R^2+\lambda(x\cos\varphi+y\sin\varphi-1)=0.$$

This circle passes through (0,0) if, and only if, $\lambda=a^2-R^2$. So any of the triangles $IA_iA_{i+1}$ has a circumcircle of the form

$$(x-a)^2+y^2-R^2+(a^2-R^2)(x\cos\varphi+y\sin\varphi-1)=0,$$

that is,

$$x^2-2x\left(a-\frac{1}{2}(a^2-R^2)\cos\varphi\right)+y^2-2y\left(-\frac{1}{2}(a^2-R^2)\sin\varphi\right)=0.$$

Their centers are evidently on the circle

$$(x-a)^2+y^2=\frac{1}{4}(a^2-R^2)^2.$$

This proves the assertion.

*Also solved by R. Akhlaghi and R. Dai, Richard Holzsager, Wee Liang Gan (student, Singapore), and the proposer.*

**Proportional subdeterminants imply equal row spaces** **April 1993**

**1422.** *Proposed by David Callan, University of Wisconsin, Madison, Wisconsin.*

Let $A$ and $B$ be $r\times n$ matrices with $r\le n$ and rank $A=r$. Suppose that there is a nonzero constant $k$ such that the determinant of every one of the $\binom{n}{r}$ $r$-square submatrices of $B$ is $k$ times the corresponding subdeterminant of $A$. Show that $A$ and $B$ have the same row space.

*Solution by Thomas Jager, Calvin College, Grand Rapids, Michigan.*

Since rank $A=r$, there is an $r\times r$ submatrix $A^*$ of $A$ that is nonsingular. The condition on the $r\times r$ subdeterminants implies that the corresponding submatrix $B^*$ of $B$ is also nonsingular. Assume $A^*$ consists of $\vec{a}_1,\dots,\vec{a}_r$, the first $r$ columns of $A$. If $X=A^*(B^*)^{-1}$, then $\vec{a}_i=X\vec{b}_i$ for $i\le r$. If $i>r$, by Cramer's rule and the condition on subdeterminants,

$$X\vec{b}_i=X\sum_{j=1}^{r}\frac{\det B_{ij}^*}{\det B^*}\vec{b}_j=X\sum_{j=1}^{r}\frac{\det A_{ij}^*}{\det A^*}\vec{b}_j=\sum_{j=1}^{r}\frac{\det A_{ij}^*}{\det A^*}X\vec{b}_j=\sum_{j=1}^{r}\frac{\det A_{ij}^*}{\det A^*}\vec{a}_j=\vec{a}_i$$

where $B_{ij}^*$ and $A_{ij}^*$ are the $r\times r$ matrices resulting from replacing the $j$th columns of $B^*$ and $A^*$ with $\vec{b}_i$ and $\vec{a}_i$. Hence, $A=XB$ where $X$ is invertible. Thus, $A$ and $B$ have the same row space.

*Also solved by Sinefakopoulos Achilleas (student, Greece), R. Akhlaghi and R. Dai, Seung-Jin Bang, Manjul Bhargava (student), Con Amore Problem Group (Denmark), Michael K. Kinyon, John W. Krussel, Peter W. Lindstrom, Ahmad Muchlis (Indonesia), J. M. de Olazábal (Spain), Harvey Schmidt, Jr., Ruedi Suter (Switzerland), University of Wyoming Problem Circle, Vu Ha Van (student, Hungary), and the proposer.*

# Answers

*Solutions to the Quickies on page 146.*

**A817.** We note that $k = a^2 + b^2 + c^2 + c$ if, and only if, $4k + 1 = (2a)^2 + (2b)^2 + (2c+1)^2$. Since only positive integers of the form $4^n(8j+7)$ are not expressible as the sum of three integer squares, $4k+1$ is so expressible. A parity check shows that exactly two of those squares are even. We conclude, therefore, that every positive integer is expressible in the form $a^2 + b^2 + c^2 + c$.

**A818.** There is at least one interior angle of $ABCD$ larger than or equal to a right angle. Let $\angle A$ be such an interior angle. Then there are infinitely many interior points near $A$ not covered by reflected segments through $A$.

Also, reflected segments through $C$ do not contain $A$, because $O$ is an interior point of triangle $BCD$. Therefore we see that there are interior points (near $A$) not contained in the four reflected segments.

Note that it is easy to prove the following: In a simple quadrangle $ABCD$, draw semi-circles with diameters $AB$, $BC$, $CD$, and $DA$ inside the quadrangle, respectively. Then the aggregate of the four semi-circles entirely covers the quadrangle.

**A819.** Let $R_n$ denote the residue of the function $1/(1-e^z)^n$. The residue of the derivative $ne^z/(1-e^z)^{n+1}$ is zero since the derivative of a Laurent series contains no term in $1/z$. Comparing the residues of

$$\frac{ne^z}{(1-e^z)^{n+1}} = -\frac{n}{(1-e^z)^n} + \frac{n}{(1-e^z)^{n+1}}$$

we see that $R_n = R_{n+1}$. Since clearly $R_1 = -1$ the result is proved.

*Remark.* The residue of $1/(\tanh z)^n$ can be calculated in the same way.

# REVIEWS

PAUL J. CAMPBELL
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 the mathematics literature. Readers are invited to suggest items for review to the editors.*

Cipra, Barry, Fermat proof hits a stumbling block, *Science* 262 (24 December 1993) 1967–1968. Peterson, I., Fermat proof flaw: Fixing the details, *Science News* 144 (18 & 25 December 1993) 406.

Andrew Wiles has not yet succeeded in proving the Taniyama-Shimura conjecture for the specific elliptic curves involved in Fermat's Last Theorem. But he is confident that he can do so: "During the review process a number of problems emerged, most of which have been resolved, but one in particular I have not yet settled. The key reduction of (most cases of) the Taniyama-Shimura conjecture to the calculation of the Selmer group is correct. However the final calculation of a precise upper bound for the Selmer group in the semistable case (of the symmetric square representation associated to a modular form) is not yet complete as it stands. I believe that I will be able to finish this in the near future using the ideas explained in my Cambridge lectures. ...In my course in Princeton beginning in February I will give a full account of this work."

Friedman, Avner, James Glimm, and John Lavery, *The Mathematical and Computational Sciences in Emerging Manufacturing Technologies and Management Practices*, SIAM, 1992; vi + 87 pp, free. ISBN 0–89871–307–2

What will be the mathematics used in the manufacturing of the future? This report identifies "areas of the mathematical and computational sciences that are needed in cutting-edge manufacturing" and provides "information to the manufacturing management community on the role of quantitative methodologies in solving the problems they encounter." The report describes specific successes of quantitative methodologies (in advanced materials, manufacturing processes, process control, quality improvement, cost-based performance measures, and benchmarking), opportunities for contributions in emerging manufacturing technologies (intelligent manufacturing, solid modeling, rapid prototyping, molecular manufacturing, biomanufacturing, and environmentally benign manufacturing), and opportunities in emerging management practices (performance measures, information management, flexible manufacturing, budgeting for flexibility, and integrated manufacturing).

Friedman, Avner, and John Lavery, *How to Start an Industrial Mathematics Program in the University*, SIAM, 1993; v + 37 pp, free. ISBN 0–89871–327–7

Does your institution have courses, much less a program, in industrial mathematics? Author Avner Friedman, Director of the Institute for Mathematics and Its Applications (IMA) at the University of Minnesota, relates experience and plans at that institution. Undergraduate courses are in place, a master's program is to start Fall 1994, a doctoral program is being planned, and postdoctoral and research programs are described. The call, though, is for other institutions to follow that lead, in providing a richer spectrum of opportunities for students and consequent greater contributions in all areas of science.

Singmaster, David, *Sources in Recreational Mathematics: An Annotated BIbliography*, 2 vols., 6th preliminary edition, 1993; from the author (School of Computing, Information Systems and Mathematics, South Bank University, London, SE1 0AA, UK); 456 pp, £20 donation to South Bank University, £25 for airmail outside of Europe, or equivalent payment in US $ made out (for convenience in exchange) to David Singmaster.

The author compares this work to Dickson's *History of the Theory of Numbers*, in the effort to be exhaustive of the field and authoritative. This bibliography classifies entries under Biographical Material, General Puzzle Collections, General Material, Mathematical Games, Combinatorial Recreations, Geometric Recreations, Arithmetic & Number-Theoretic Recreations (one-third of the book), Probability Recreations, Logical Recreations, Physical Recreations, and Topological Recreations. After eleven years and six preliminary editions, the author muses (without conviction), "I would like to think that I am about half way through the relevant material."

Davis, Donald M., *The Nature and Power of Mathematics*, Princeton University Press, 1993; xi + 389 pp, $60, $24.95 (P). ISBN 0-691-02562-2 (P)

The major topics of this fresh introduction to mathematics for the general reader (or liberal arts student) are non-Euclidean geometry, number theory and cryptography, and fractals. "The theme that binds them together is the unexpected applications of pure mathematics"; the "application" of fractals is their *beauty*. There are exercises, and the emphasis is on the importance of understanding a concept by struggling with proofs that involve it. This is a book from which students can learn that mathematics is important for the beauty of its ideas.

Nelsen, Roger B. (ed.), *Proofs without Words: Exercises in Visual Thinking*, MAA, 1993; xi + 152 pp, $23 (P). ISBN 0-88385-700-6

A proof without words is a picture or diagram that helps show why a particular result is true, as well as how to go about proving it. "[T]he emphasis is on providing *visual* clues to stimulate mathematical thought." This is the first volume in a new MAA book series, on Classroom Resource Materials; the contents come mostly from MAA journals.

Higham, Nicholas J., *Handbook of Writing for the Mathematical Sciences*, SIAM, 1993; xii + 241 pp, $21.50. ISBN 0-89871-314-5

As an editor and referee, I see lots of manuscripts. Many suffer not from poor mathematics but from poor writing. Their authors could benefit from reading and absorbing the principles enunciated in this book. Higham recommends standard sources, then treats mathematical writing, English usage, difficulties for non-native speakers, the mechanics of writing a paper, how to revise, the editorial/publishing process at a journal, how to write a talk, and computer aids (such as TEX). Appendices recap the Greek alphabet, TEXand LATEXsymbols, and GNU Emacs commands; there are also lists of mathematical organizations and winners of prizes for expository writing.

Kanigel, Robert, Bubble, bubble: Jean Taylor and the mathematics of minimal surfaces, *The Sciences* (May/June 1993) 32–38.

Here is a popular sketch of the work in minimal surfaces—and a little of the life—of Rutgers mathematician Jean Taylor, by the author of *The Man Who Knew Infinity: The Life of the Genius Ramanujan*.

Traub, Joseph F., and Henryk Woźniakowski, Breaking intractability, *Scientific American* (January 1994) 102–107.

Some problems that would otherwise be intractable or unsolvable—such as multivariate definite integration and surface reconstruction—can now be "solved." The tradeoff is that we must give up a guaranteed bound on the error for only the "weaker guarantee that the error is probably no more than $\epsilon$." One approach is randomization, such as the Monte Carlo technique for integration, which picks random points at which to evaluate the function. Another approach attributes an a priori distribution to the problems (e.g., integrands distributed according to a Wiener measure) and then determines optimal deterministic points at which to evaluate the function. This second approach leads to what the authors call *tractable on average*, meaning that we can guarantee a bound on the expected error. Their vein of work has led the authors "to speculate that it might be possible to prove formally that certain scientific questions are unanswerable. The proposed attack is to prove that the computing resources (time, memory, energy) do not exist in the universe to answer such questions. ... [W]e would like to establish a physical Gödel's theorem. ... Such questions include when the universe will stop expanding and what the average global temperature will be in the year 2001." (The accompanying cartoon artwork is stupid and should have been replaced by figures comparing different methods of integration or of reconstruction of medical images.)

McCrone, John, Computers that listen, *New Scientist* (4 December 1993) 30–35.

"The battle to make computers smart enough to understand us when we talk to them seems to have been won at last—by a Russian mathematician who's been dead since 1922." Guess who? dots Markov. Modeling speech as a Markov chain of phonemes in context is proving to be the key behind recent advances in dictation and voice recognition machines. Why wasn't this approach tried earlier? The heavy computation involved is now feasible on desktop machines; and, as one expert claims, linguists didn't have enough math training to appreciate the potential of the approach.

Rombix. Mathematical puzzle and game, devised by Alan Schoen. $10 plus shipping, from Petrick's Sales, Inc., De Pere WI 54115. Inquiries or comments to Rombix U.S.A., Inc., P.O. Box 3553, Carbondale, IL 62902–3553.

This puzzle's 16 plastic pieces consist of four noncongruent rhombuses ("keystones") of different colors, with all sides the same length, plus every possible nonconvex piece formed by joining two keystones along an edge. As with Tangrams, one can rearrange the pieces to depict animals and other objects. They can also be used to tile edge-to-edge a regular 16-gon in a variety of challenging ways, to make "ladders," and to play solitaire and two-person games. Hinted at in the accompanying booklet, by mathematician Schoen, is a connection with Penrose tilings.

Cajori, Florian, *A History of Mathematical Notations* (2 vols. bound as one), Dover, 1993; xvi + 451 pp, xii + 367 pp, $19.95 (P). ISBN 0–486–67766–4

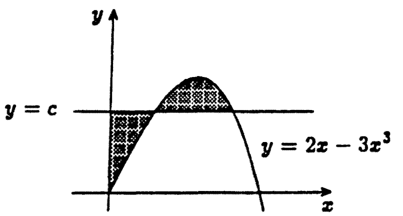
It is a great pleasure to have this classic study from 1928–1929 available again. The first volume is devoted to symbols in elementary arithmetic, algebra, and geometry, while the second treats symbols used in more advanced mathematics. You will certainly learn something here! For example, what do you tell a student who asks how the $\partial$ notation came to distinguish partial derivatives? With this book at hand, you will know that it was Jacobi who established this standard in 1841 (vol. 2, p. 236).

# NEWS AND LETTERS

## 54th ANNUAL WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

***These solutions have been compiled and prepared by Loren Larson, St. Olaf College.***

**A-1.** **The horizontal line $y = c$ intersects the curve $y = 2x - 3x^3$ in the first quadrant as in the figure. Find $c$ so that the areas of the two shaded regions are equal.**




**Solution. The value of $c$ is 4/9.**

**Let $(b, c)$ denote the second intersection point. We wish to find $c$ so that**

$$\int_0^b \left(c - (2x - 3x^3)\right)\, dx = 0.$$

**This leads to $cb - b^2 + (3/4)b^4 = 0$. After substituting $c = 2b - 3b^3$ and solving, we find that $b = 2/3$ and the result follows.**

**A-2.** **Let $(x_n)_{n\geq 0}$ be a sequence of nonzero real numbers such that**

$$x_n^2 - x_{n-1}x_{n+1} = 1, \quad \text{for } n = 1, 2, 3, \ldots .$$

**Prove there exists a real number $a$ such that $x_{n+1} = ax_n - x_{n-1}$ for all $n \geq 1$.**

**Solution.** **It is equivalent to show that**

$$\frac{x_{n+1} + x_{n-1}}{x_n}$$

**is independent of $n$. This follows (by induction) from**

$$\frac{x_{n+2} + x_n}{x_{n+1}} - \frac{x_{n+1} + x_{n-1}}{x_n} =$$

$$= \frac{(x_n x_{n+2} + x_n^2) - (x_{n+1}^2 + x_n x_{n+1})}{x_n x_{n+1}}$$

$$= \frac{-(x_{n+1}^2 - x_n x_{n+2}) + (x_n^2 - x_{n-1}x_{n+1})}{x_n x_{n+1}}$$

$$= \frac{-1 + 1}{x_n x_{n+1}} = 0.$$

**A-3.** Let $\mathcal{P}_n$ be the set of subsets of $\{1, 2, \ldots, n\}$. Let $c(n, m)$ be the number of functions $f : \mathcal{P}_n \to \{1, 2, \ldots, m\}$ such that $f(A \cap B) = \min\{f(A), f(B)\}$. Prove that

$$c(n, m) = \sum_{j=1}^{m} j^n.$$

**Solution.** We induct on $m$. It is clear that $c(n, 1) = 1$ for all $n$. Assume

$$c(n, m-1) = \sum_{j=1}^{m-1} j^n.$$

Suppose $f$, not identically 1, is such a function. Let $S_f = \bigcap_{f(A)\geq 2} A$. The condition implies that $f(A) \geq 2$ if and only if $A \supseteq S_f$.

Given a subset $S$ of $\{1, 2, \ldots, n\}$, to count the allowable $f$'s such that $f(A) \geq 2$ if and only if $A \supseteq S_f$, we note the one-to-one correspondence between the sets $A \supseteq S$ and the subsets of $\{1, 2, \ldots, n\} - S$. These sets map to $\{2, 3, \ldots, m\}$ and the rest to 1. Hence

$$\begin{aligned} c(n, m) &= 1 + \sum_S c(n - |S|, m - 1) \\ &= 1 + \sum_{k=0}^{n} \binom{n}{k} c(n - k, m - 1) \\ &= 1 + \sum_{k=0}^{n} \binom{n}{k} \sum_{j=1}^{m-1} j^{n-k} \\ &= 1 + \sum_{j=1}^{m-1} \sum_{k=0}^{n} \binom{n}{k} j^{n-k} \\ &= 1 + \sum_{j=1}^{m-1} (1 + j)^n = \sum_{j=1}^{m} j^n, \end{aligned}$$

completing the induction.

**A-4.** Let $x_1, x_2, \ldots, x_{19}$ be positive integers each of which is less than or equal to 93 Let $y_1, y_2, \ldots, y_{93}$ be positive integers each of which is less than or equal to 19. Prove that there exists a (nonempty) sum of some $x_i$'s equal to a sum of some $y_j$'s.

**Solution.** We consider numbers $-19, -18, -17, \dots, 0, 1, 2, \dots, 93$ and a pebble is moved among these numbers according to the following algorithm: Initially, the pebble is at number 0. At the first step, we move the pebble to an arbitrary number $x_m$ and we discard $x_m$. At the $i$th step, for $i \geq 1$, suppose the pebble is placed at number $t$ and we move the pebble according to the value of $t$. If $t > 0$, we move the pebble to $t - y$ where $y$ is an unused $y_k$ for some $k$ (and we then discard $y_k$). If $t < 0$, the pebble is moved to $t + x$, where $x$ is an unused $x_m$ for some $m$ (and we then discard $x_m$). We cannot move the pebble to $-19$ without having moved the pebble back to 0 first, hence, there is always an unused $x_m$ when $t < 0$ and no number has been revisited. Similarly, there is always an unused $y_k$ when $t > 0$ and no number has been revisited. Thus, we may continue this process until the pebble is placed at the same number a second time. When this happens, we will have obtained a sum of some $x_i$'s equal to the sum of some $y_j$'s.

**A-5.** Let $Q(x) = \left(\dfrac{x^2 - x}{x^3 - 3x + 1}\right)^2$. Show that

$$\int_{-100}^{-10} Q(x)\,dx + \int_{\frac{1}{101}}^{\frac{1}{11}} Q(x)\,dx + \int_{\frac{101}{100}}^{\frac{11}{10}} Q(x)\,dx$$

is a rational number.

**Solution.** Observe first that the roots of $x^3 - 3x + 1$ can be isolated away from the given intervals; that is, there are sign changes in the intervals $[-2, -1]$, $[1/3, 1/2]$, $[3/2, 2]$. Hence the integrand is defined and continuous throughout.

Set

$$f(t) = \int_{-100}^{t} Q(x)dx + \int_{\frac{1}{101}}^{1/(1-t)} Q(x)dx + \int_{\frac{101}{100}}^{1-1/t} Q(x)\,dx$$

for $-100 \leq t \leq -10$. We wish to compute $f(-10)$. By the Fundamental Theorem of Calculus,

$$f'(t) = Q(t) + Q\left(\frac{1}{1-t}\right)\frac{1}{(1-t)^2} + Q\left(1 - \frac{1}{t}\right)\frac{1}{t^2}.$$

We find that $Q(1/(1-t)) = Q(1 - 1/t) = Q(t)$, so that $f(-10) =$

$$\int_{-100}^{-10} Q(x)\left(1 + \frac{1}{x^2} + \frac{1}{(1-x)^2}\right)dx.$$

But, noting that

$$\frac{1}{Q(x)} = \left(x + 1 - \frac{1}{x} - \frac{1}{x-1}\right)^2,$$

we see that the last integral is of the form $\int du/(u^2)$. Hence, its value is

$$-\frac{x^2 - x}{x^3 - 3x + 1}\Big|_{-100}^{10},$$

which is rational.

**A-6.** The infinite sequence of 2's and 3's

$$2, 3, 3, 2, 3, 3, 3, 2, 3, 3, 3, 2, 3, 3, 2, 3, 3, 3, 2,$$
$$3, 3, 3, 2, 3, 3, 3, 2, 3, 3, 2, 3, 3, 3, 2, \dots$$

has the property that, if one forms a second sequence that records the number of 3's between successive 2's, the result is identical to the given sequence. Show that there exists a real number $r$ such that, for any $n$, the $n$th term of the sequence is 2 if and only if $n = 1 + \lfloor rm \rfloor$ for some nonnegative integer $m$. (Note: $\lfloor x \rfloor$ denotes the largest integer less than or equal to $x$.)

**Solution.** We show that the conclusion holds with $r = 2 + \sqrt{3}$.

Assuming the result, we first derive the value of $r$. Observe that, asymptotically, the proportion of 2's in the first $n$ terms is $1/r$. Thus, assuming there are about $m$ 2's in the first $n \approx rm$ terms, there should be about $(r-1)m$ 3's. These numbers give the approximate number of 3's in the intervals following the first $n$ 2's, namely $2m + 3(r-1)m = (3r-1)m$. Hence, the proportion of 2's in the first $rm + (3r-1)m = (4r-1)m$ terms is $rm/(4r-1)m = r/(4r-1)$. So we want $r$ to satisfy $1/r = r/(4r-1)$, or $r^2 - 4r + 1 = 0$. Since $r$ must exceed 1, $r = 2 + \sqrt{3}$.

Let $a_n$ denote the $n$th term of the sequence. Observe that the first term is 2 and $1 + \lfloor 0r \rfloor = 1$. We prove, by induction on $k$, that the $k$th 2 is $a_{1+\lfloor r(k-1)\rfloor}$. Suppose this holds for $k \leq m$.

Let $n = 1 + \lfloor r(m-1) \rfloor$, so that $a_n$ is the $m$th 2. The $(m+1)$st 2 is either $a_{n+3}$ or

$a_{n+4}$, the former if and only if $a_m = 2$. Also, $1 + \lfloor rm \rfloor$ is $n+3$ or $n+4$ since $3 < r < 4$. We must show then that $a_m = 2$ if and only if $n+3 = 1 + \lfloor rm \rfloor$. By the induction hypothesis, $a_m = 2$ if and only if $m = 1 + \lfloor rs \rfloor$ for some $s$. Hence, it suffices to show that there exists an $s$ such that

$$m - 1 = \lfloor rs \rfloor \iff \lfloor rm \rfloor = \lfloor r(m-1) \rfloor + 3.$$

But

$$m - 1 = \lfloor rs \rfloor \iff m - 1 < rs < m$$

$$\iff$$

$$4m - 4 - r(m-1) < s < 4m - rm,$$

since $r^{-1} = 4 - r$. Therefore, there exists an integer $s$ such that $m - 1 = \lfloor rs \rfloor$ if and only if there is an integer between $rm$ and $r(m-1) + 4$, which happens if and only if $\lfloor rm \rfloor = \lfloor r(m-1) \rfloor + 3$.

**B-1.** Find the smallest positive integer $n$ such that for every integer $m$, with $0 < m < 1993$, there exists an integer $k$ for which

$$\frac{m}{1993} < \frac{k}{n} < \frac{m-1}{1994}.$$

**Solution.** First, it is easily verified that

$$\frac{m}{1993} < \frac{2m+1}{1993+1994} < \frac{m-1}{1994},$$

so $n = 1993 + 1994 = 3987$ suffices. Now consider $m = 1992$ and suppose

$$\frac{1992}{1993} < \frac{k}{n} < \frac{1993}{1994}$$

Since $x/(x+1)$ is strictly increasing for $x > 0$, we must have $k \leq n - 2$ (note: $n > 1994$). However,

$$\frac{1992}{1993} < \frac{n-2}{n}$$

implies $3986 < n$, so $n \geq 3987$, completing the proof.

**B-2.** Consider the following game played with a deck of $2n$ cards numbered from 1 to $2n$. The deck is randomly shuffled and $n$ cards are dealt to each of two players, $A$ and $B$. Beginning with $A$, the players take turns discarding one of their remaining cards and announcing its number. The game ends as soon as the sum of the numbers on the discarded cards is divisible by $2n+1$. The last person to discard wins the game. Assuming optimal strategy by both $A$ and $B$, what is the probability that $A$ wins?

Solution. The probability that $A$ wins is 0.

Clearly, $A$ cannot win on the first turn. Assume $B$ is to play, and that the total of announced numbers is $T$, and that $A$ has cards $x_1, x_2, \ldots, x_k$, and $B$ has cards $y_1, y_2, \ldots, y_{k+1}$. Because the integers $T + y_1, \ldots T + y_{k+1}$ have distinct remainders upon division by $2n+1$, at least one has a remainder other than $2n+1-x_1, \ldots, 2n+1-x_k$. If $B$ discards that $y_i$, it is impossible for $A$'s next discard to make the total divisible by $2n-1$. Therefore, $A$ cannot win under optimal play by $B$.

**B-3.** Two real numbers $x$ and $y$ are chosen at random in the interval $(0,1)$ with respect to the uniform distribution. What is the probability that the closest integer to $x/y$ is even?

**Express the answer in the form $r + s\pi$, where $r$ and $s$ are rational numbers.**

**Solution. The probability is $(5-\pi)/4$ (that is, when $r = 5/4$, $s = -1/4$).**

For $0 \leq a < b < 1$,

$$\Pr(a < x/y < b) = \Pr(ay < x < by) = \int_0^1 \int_{ay}^{by} 1\,dx\,dy = b/2 - a/2.$$

For $1 < a < b$,

$$\Pr(a < x/y < b) = Pr(x/b < y < x/a) = \int_0^1 \int_{x/b}^{x/a} 1\,dy\,dx = 1/(2a) - 1/(2b).$$

Note that the probability that $x/y$ is exactly half an odd integer is 0, so we may safely ignore this possibility. Thus, the probability we desire is

$$\Pr(0 < x/y < 1/2) + \sum_{n=1}^{\infty} \Pr(2n - 1/2 < x/y < 2n + 1/2)$$

$$= \frac{1}{4} + \sum_{n=1}^{\infty} \left(\frac{1}{4n+1} - \frac{1}{4n-1}\right).$$

$$= \frac{1}{4} + (1 - \arctan 1)$$

$$= \frac{5}{4} - \frac{\pi}{4}$$

**B-4.** The function $K(x,y)$ is positive and continuous for $0 \leq x \leq 1$, $0 \leq y \leq 1$, and the functions $f(x)$ and $g(x)$ are positive and continuous for $0 \leq x \leq 1$. Suppose that for all $x$, $0 \leq x \leq 1$,

$$\int_0^1 f(y)\,K(x,y)\,dy = g(x)$$

and

$$\int_0^1 g(y)\,K(x,y)\,dy = f(x).$$

Show that $f(x) = g(x)$ for $0 \le x \le 1$.

**Solution.** Let $m$ and $M$, respectively, be the minimum and maximum values of $f(x)/g(x)$ for $0 \le x \le 1$. We must show $m = M = 1$.

Multiplying the terms in $mg(y) \le f(y) \le Mg(y)$ by $K(x,y)$ and integrating over $0 \le y \le 1$, yields

$$mf(x) \le g(x) \le Mf(x),$$

or equivalently,

$$\frac{1}{M} \le \frac{f(x)}{g(x)} \le \frac{1}{m} \qquad (*)$$

with strict inequalities *for every* $x$, unless $m = M$. By $(*)$, $1/M \le m$ and $1/m \ge M$, and clearly these inequalities cannot be strict. Hence $M = m$ and $Mm = 1$. Therefore, $m = M = 1$.

**B-5.** Show there do not exist four points in the Euclidean plane such that the pairwise distances between the points are all odd integers.

**Solution.** For real numbers $x$ and $y$, and for an integer $n$, define $x \equiv y \pmod n$ if $(x-y)$ is an integer divisible by $n$. Let the points be $(0,0), (a,0), (r,s), (x,y)$. We may assume $a > 0$ (and odd), $s \ge 0$ and $y \le s$. Since squares of odd integers are congruent to $1 \pmod 8$, we have

$$r^2 + s^2 \equiv 1 \pmod 8$$
$$(r-a)^2 + s^2 \equiv 1 \pmod 8$$
$$x^2 + y^2 \equiv 1 \pmod 8$$
$$(x-a)^2 + y^2 \equiv 1 \pmod 8$$
$$(x-r)^2 + (y-s)^2 \equiv 1 \pmod 8$$

Subtracting the first two yields $2ar \equiv a^2 \pmod 8$. In particular, $r$ is a rational number and 2 divides the denominator of $r$, which divides $2a$. Since an identical result holds for $x$, multiplying all coordinates by $a$ yields an example with the denominator of $r$ equal to the denominator of $x$ equal to 2. Thus, we may assume $2r$ and $2x$ are integers congruent to $a \pmod 8$, or, $r \equiv x \equiv (a/2) \pmod 4$. We obtain $s^2 \equiv y^2 \equiv 1 - (a^2/4) \pmod 4$. Thus, $4s^2 \equiv 4y^2 \equiv 4 - a^2 \pmod{16}$, hence $\equiv 3 \pmod 8$. Therefore $s$ and $y$ are of the form $\pm(\sqrt{8k+3})/2$ for some integer $k$. Now the above implies $(x-r)^2$ is an integer divisible by 16, hence $(y-s)^2$ is an integer $\equiv 1 \pmod 8$. All this implies $y = qs$ for some rational $q$ (which we've assumed is $< 1$). Also, $q = u/v$ with $u, v$ odd. The above yields $(u-v)^2(4s^2) \equiv 4v^2 \pmod{32}$. Thus, $(u-v)^2(4-a^2) \equiv 4v^2 \pmod{32}$, $((u-v)/2)^2(4-a^2) \equiv v^2 \pmod 8$. Therefore, $(u-v)/2$ is odd, hence $((u-v)/2)^2 \equiv v^2 \equiv 1 \pmod 8$ while $4 - a^2 \equiv 3 \pmod 8$, a contradiction.

**B-6.** **Let $S$ be a set of three, not necessarily distinct, positive integers. Show that one can transform $S$ into a set containing 0 by a finite number of applications of the following rule: Select two of the three integers, say $x$ and $y$, where $x \le y$, and replace them with $2x$ and $y - x$.**

**Solution.** Say the numbers are $a, b, c$. First, we reduce to the case that exactly one of $a, b, c$ is odd. Namely: (i) if two are odd, apply the rule with those two, and none is odd; (ii) if none is odd, divide all numbers by 2 and apply induction; (iii) if three are odd, apply the rule once, and exactly one is odd. Once exactly one is odd, this will remain so.

Say $a$ is odd and $b$ and $c$ even. We aim to make the power of 2 dividing $b + c$ as large as possible. If $b$ and $c$ have the same number of factors of 2, then applying the rule to those two will yield both divisible by a higher power of 2, or one will have fewer factors of 2 than the other. Since $b + c$ is constant here, after a finite number of applications of the rule, $b$ and $c$ will not have the same number of factors of 2. Also, it is easy to see that, possibly after some additional moves, one has either $bc = 0$ (in which case one stops), or, the one of $b$ and $c$ divisible by the smaller power of 2 is also smaller; say it is $b$, so that $b < c$.

*Case 1:* $a > b$. Now work with $a, b$. Then $a$ remains odd, $b$ is doubled, and $b + c$ is divisible by a higher power of 2.

*Case 2:* $a < b$. Apply the rule first to $a$ and $b$, and then to $b - a$ and $c$. (Note that $c > b > b - a$.) One obtains

$$\begin{matrix} a & b & c \\ 2a & b-a & c \\ 2a & 2b-2a & c-b+a \end{matrix}$$

**Now the odd number is $c - b + a$, and the sum of the even numbers is $2b$, which has more factors of 2 than $b + c$.**

# No matter how you express it, it still means DERIVE® is half price.

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}\qquad \lim_{x\to 0}\frac{x}{\sin(2x)}\qquad \frac{1}{2}$$

$$50\%\qquad 0.5$$

$$\sum_{n=1}^{\infty}\frac{1}{2^{n+1}}\qquad \int_0^1 x\,dx$$

DERIVE®

The *DERIVE A Mathematical Assistant* program lets you express yourself symbolically, numerically and graphically, from algebra through calculus, with vectors and matrices too—all displayed with accepted math notation, or 2D and 3D plotting. *DERIVE* is also easy to use and easy to read, thanks to a friendly, menu-driven interface and split or overlay windows that can display both algebra and plotting simultaneously. Better still, *DERIVE* has been praised for the accuracy and exactness of its solutions. But, best of all the suggested retail price is now only $125. Which means *DERIVE* is now half price, no matter how you express it.

**System requirements**

*DERIVE:* MS-DOS 2.1 or later, 512K RAM, and one 3½" disk drive. Suggested retail price now **$125 (Half off!).**

*DERIVE ROM card:* Hewlett Packard 95LX & 100LX Palmtop, or other PC compatible ROM card computer. Suggested retail price now **$125!**

*DERIVE XM (eXtended Memory):* 386 or 486 PC compatible with at least 2MB of *extended* memory. Suggested list price now $250!



Soft Warehouse INC
HONOLULU • HAWAII

Soft Warehouse, Inc. • 3660 Waialae Ave. Ste. 304 • Honolulu, HI, USA 96816-3236
Ph: (808) 734-5801 • Fax: (808) 735-1105

# CONTENTS

## ARTICLES



## NOTES



## PROBLEMS



## REVIEWS



## NEWS AND LETTERS




THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, D.C. 20036