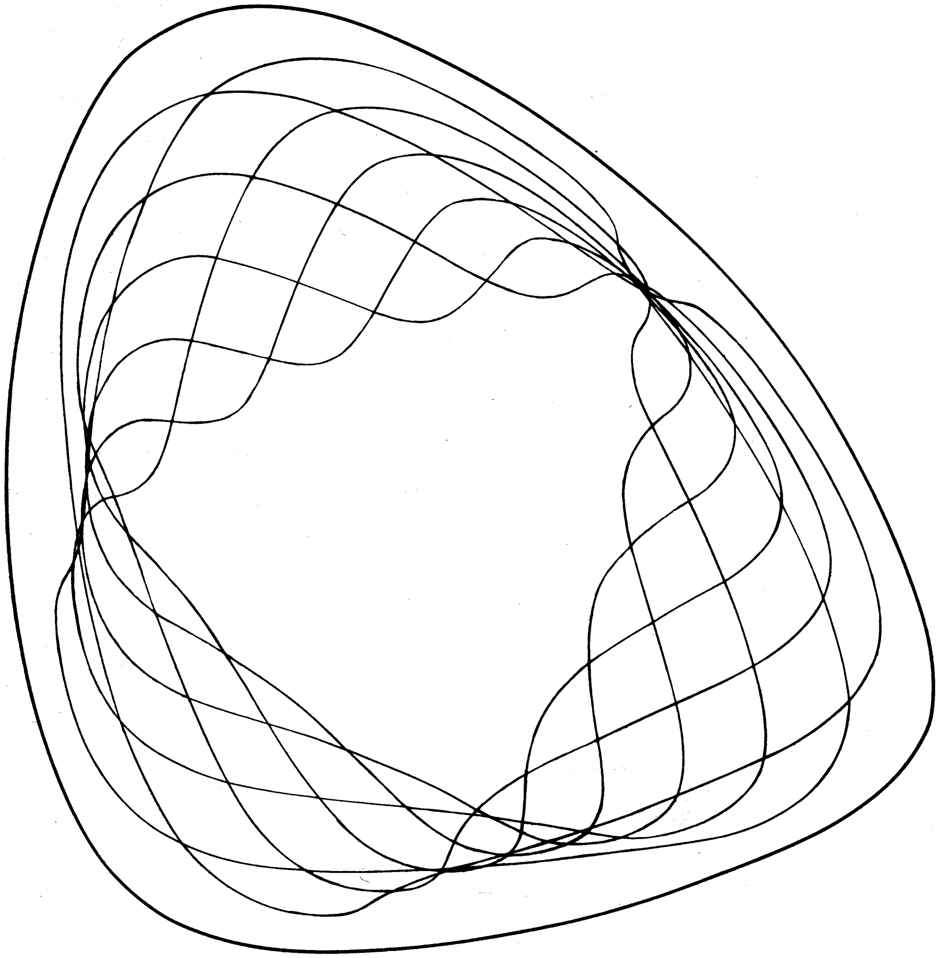
MATHEMATICS MAGAZINE



Vol. 54, No. 3
May, 1981

HOLDITCH'S THEOREM • EXPLORING A RECTANGLE
GÖDEL'S THEOREM • MAGIC CIRCLES • MOUSE TRAP

MATHEMATICS MAGAZINE

Volume 54, Number 3, May 1981
(ISSN—0025-570X)

COVER: A convex curve encloses 7 "Holditch companion curves" where ratios a:b are 2:14, 4:12, 6:10, 8:8, 10:6, 12:4, 14:2. Design by W. Bender, using his 'tracer' (p. 128).

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

EDITORIAL POLICY

Mathematics Magazine is a journal of collegiate mathematics which aims to provide inviting, informal mathematical exposition of interest to undergraduate students. Manuscripts accepted for publication in the *Magazine* should be written in a clear and lively expository style and stocked with appropriate examples and graphics. Our advice to authors is: say something new in an appealing way or say something old in a refreshing way. The *Magazine* is not a research journal and so the style, quality, and level of articles submitted for publication should realistically permit their use to supplement undergraduate courses. The editor invites manuscripts that provide insight into the history and application of mathematics, that point out interrelationships between several branches of mathematics and that illustrate the fun of doing mathematics.

New manuscripts should be sent to: Doris Schattschneider, Editor, Mathematics Magazine, Moravian College, Bethlehem, Pennsylvania 18018. Manuscripts should be prepared in a style consistent with the format of *Mathematics Magazine.* They should be typewritten and double spaced on $8\frac{1}{2}$ by 11 paper. Authors should submit the original and one copy and keep one copy as protection against possible loss. Illustrations should be carefully prepared on separate sheets of paper in black ink, the original without lettering and two copies with lettering added.

Authors planning to submit manuscripts should read the full statement of editorial policy which appears in the News and Letters section of this *Magazine,* Vol. 54, pp. 44–45. Additional copies of the policy are available from the Editor.

AUTHORS

Arne Broman ("Holditch's Theorem") has taught since 1951 at Chalmers University of Technology, at Göteborg, Sweden. He is the author of several publications on a wide range of subjects, and has produced a film on four-dimensional geometry (Göteborg 1976). He is now engaged in a project on smooth alignment of road lines and road surfaces, and in 1980 joined an expedition from Chalmers to Peru to study the roads built during the Inca period. His interest in Holditch's theorem arose during a seminar series at Western Washington University in the spring of 1977, devoted to a book by Luis Santaló: *Integral Geometry and Geometric Probability* (Addison-Wesley, 1976).

Arthur Charlesworth ("A Proof of Gödel's Theorem in Terms of Computer Programs") received his bachelor's degree from Stetson, his doctorate from Duke, and has been on the faculty of the University of Richmond since 1976. He is also the author of "Infinite Loops in Computer Programs" (this *Magazine*, November 1979). These articles resulted from a project to develop relatively simple proofs of major results in the foundations of mathematics and computer science. The approach which he has taken to Gödel's Theorem is to replace technical details by general conditions in an axiomatic system.

ILLUSTRATIONS

Figures which show a curve and a 'companion' curve traced out by a point on a chord of fixed length ("Holditch's Theorem") were constructed by the author using different methods: careful measuring and drawing, computer work, and use of a mechanical device (see p. 128).

Candy Baker, Bethlehem, PA, illustrated "A Proof of Gödel's Theorem in Terms of Computer Programs" pp. 118, 119; she also sketched the talented skater on p. 130.

Ann Litz, Moravian College student, produced accurate curves (p. 130) using computer graphics; calligraphy was added by the Editor.

All other illustrations were provided by the authors.

BUSINESS INFORMATION. Mathematics Magazine is published by the Mathematical Association of America at Washington, D. C., five times a year in January, March, May, September, and November. Ordinary subscriptions are $18 per year. Members of the Mathematical Association of America or of Mu Alpha Theta may subscribe at special reduced rates. College and university mathematics departments may purchase bulk subscriptions for distribution to undergraduate students. A bulk subscription consists of five or more copies of each issue to a single address for one year; the annual bulk subscription rate is $5.00 per copy in the package, with a minimum $25.

Subscription correspondence and notice of change of address should be sent to A. B. Willcox, Executive Director, Mathematical Association of America, 1529 Eighteenth Street, N. W., Washington, D.C. 20036. Back issues may be purchased, when in print, from P. and H. Bliss Co., Middletown, Connecticut 06457.

Advertising correspondence should be addressed to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, New York 14214.

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

ARTICLES

Holditch's Theorem

A fresh look at a long-forgotten theorem.

ARNE BROMAN
Chalmers University of Technology
S-412 96 Göteborg, Sweden

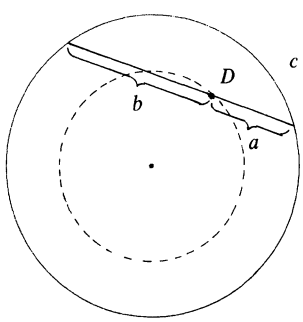
Hamnet Holditch, president of Cajus College in Cambridge during the middle part of the last century, discovered a remarkable property of a curve traced out by a point on a chord of fixed length that slides around with both endpoints on a convex curve. His proof, published in 1858 [**4**], makes a number of unstated assumptions and utilizes some notions unfamiliar to the modern reader. Yet the elementary part of the theorem is accessible to a good high school student, and a much more general result can be proved as a rather elementary application of line integrals in the plane.

In this paper we shall discuss the original presentation of Holditch and point out several limitations of his approach. We shall then describe a much more general theorem which is considerably easier to prove, and conclude by giving some applications to kinematics and to a polygonal version of the theorem in which the curve traced out by the point on the chord of fixed length is apparently of a sort quite different from that originally envisioned by Hamnet Holditch.

The material in this article was the subject of lectures by the author in March 1979 at the University of Washington and Simon Fraser University.

Holditch's Theorem, the Classical Case.

We begin with a problem, elementary enough for a good high school student. In a circle c of radius r, a chord is divided into parts of length a and b by point D. As the chord makes a complete sweep around the circle, an inner circle, the locus of D, is traced out. (See FIGURE 1(a), where the dotted curve is the locus.) *Find the area of the ring-shaped region between the circle c and the locus of D.*



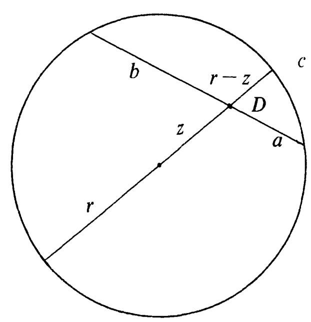
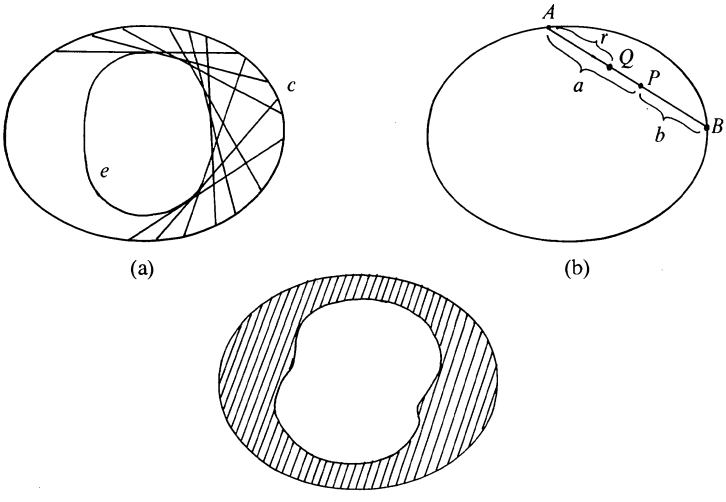


FIGURE 1

For the solution, we refer to the notation in FIGURE 1(b). Since in the circle c, two intersecting chords are divided at D into segments of lengths a, b, and $r+z$, $r-z$, respectively, it follows (see [3]) that $a \cdot b = (r+z) \cdot (r-z) = r^2 - z^2$. Thus $\pi ab = \pi r^2 - \pi z^2$, which shows that the area sought is πab.

It is interesting to note that the answer to this problem is in terms of the segments a and b of the divided chord, and is wholly independent of r, the radius of the circle. Holditch observed that when the outer circle was replaced by a more general curve, the result remained correct.

HOLDITCH'S THEOREM (his own formulation [4]). *If a chord of a closed curve, of constant length $a+b$, be divided into two parts of lengths a, b, respectively, the difference between the areas of the closed curve, and of the locus of the dividing point, will be πab.*



(a) (b) (c)

FIGURE 2

FIGURE 2(a) shows a closed convex curve c and some of its chords of a fixed length. In his proof, Holditch made the assumption that for the given curve c, all of the chords of fixed length $a+b$ would be tangent to another curve e called the envelope of the family of chords. He let Q denote the point where the chord AB touches this envelope, and he let r denote the length of the segment AQ (which can vary with the position of the chord). Denoting P as the dividing point of the chord, so $|AP| = a$, $|PB| = b$, (see FIGURE 2(b)), he then claims that the areas swept out by the segments AQ, BQ, and QP during one revolution of the chord can be written as

$$I_1 = \frac{1}{2}\int_0^{2\pi} r^2 d\theta,\quad I_2 = \frac{1}{2}\int_0^{2\pi} (a+b-r)^2 d\theta,\quad I_3 = \frac{1}{2}\int_0^{2\pi} (a-r)^2 d\theta$$

respectively. The first two integrals are equal since AQ and BQ sweep out the same annular region between e and c. The equation $I_1 = I_2$ and a short computation show that $\int_0^{2\pi} r\,d\theta = \pi(a+b)$. From this, it is readily shown that $I_2 - I_3$ (i.e., the area of the annular region swept out by AP and shaded in FIGURE 2 (c)) is πab. This is the desired result.

Comments on Holditch's Version. A Modern Reformulation.

The careful reader has already noticed several implicit assumptions made by Holditch in his formulation and proof. First, generalizing from the case of a circle, Holditch probably took c to be a convex curve, but he does not mention this assumption.

Next, we note that the integral representations in his proof are known to be valid for the area swept out by a vector based at a fixed point and making one revolution counterclockwise, the length of the vector varying continuously. Although this property can be extended to a vector based at a variable point that goes around a convex curve, the vector always having the direction of the positive half-tangent of the curve, Holditch takes this extension for granted without comment. (Perhaps in Holditch's day these results were well known.)

We have already noted that Holditch assumed the chords of length $a+b$ are all tangent to the envelope e; actually, Holditch does not use the term "envelope." He says instead, in the language of that day: "Let Q be the point in which the chord intersects its consecutive position." This terminology is probably influenced by Newton.

A rather serious omission from the statement of the theorem is his failure to mention that some upper bound for the length of the chord AB is needed. To see this need, let a chord AB of length $\sqrt{2}$ glide along a rectangle $KLMN$ whose sides have the lengths 1 and 2 (see FIGURE 3), and study the locus of the midpoint D of the chord. The chord is too long to slide along the entire rectangle in such a way that each point on the rectangle is hit by both A and B. For example, if the chord AB begins in the position shown in FIGURE 3(a) and moves counterclockwise, both B and A travel along the upper edge of the rectangle until B reaches L, then B travels down the left side of the rectangle to M (FIGURE 3(b)). Now, as endpoint B travels toward N, endpoint A travels back to K (retracing its path) along the upper edge of the rectangle (FIGURE 3(c)). The chord can continue its travel around the rectangle, A sliding down the right edge to N and B retracing its path back to M. Continuing until all possible paths have been traced, the locus of D consists of four line segments (two of them coincide) and four circle arcs. (The locus of D is shown by the dotted curve in FIGURE 3(d).) The point set between the rectangle and the locus has the area $1/2+\pi/2$. This differs from the area given by the conclusion of Holditch's theorem (i.e., from the area $\pi/2$). It is possible that Holditch assumed that the chord AB could travel in a simple way (always moving in the same direction) around the curve c, with both endpoints A and B eventually hitting every point on c.

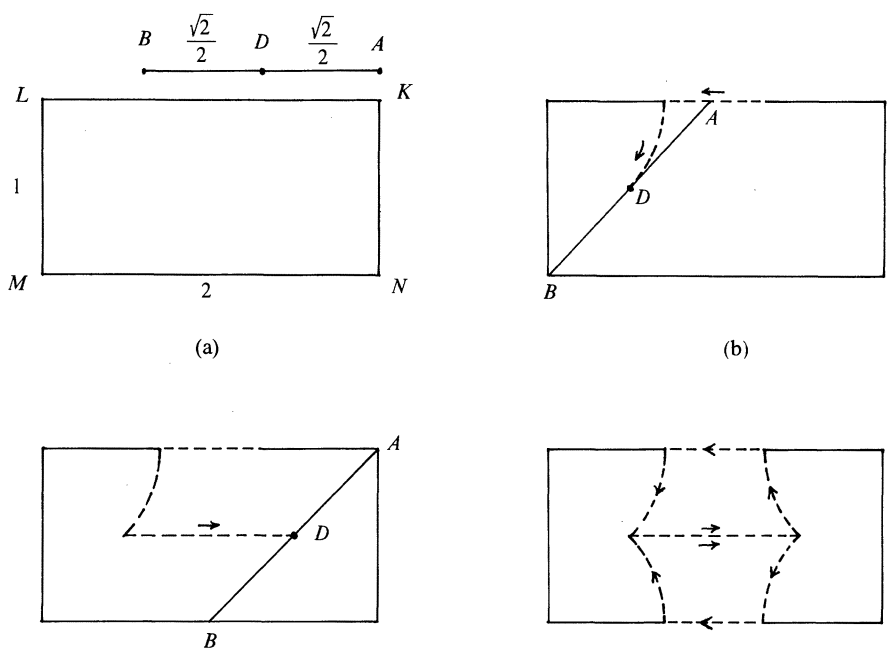


FIGURE 3

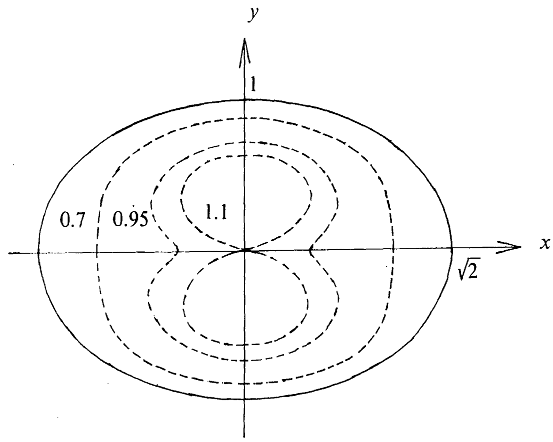


FIGURE 4

The need for an upper bound for the chord length is also seen in FIGURE 4, which depicts the ellipse $(x^2/2)+y^2=1$ and the locus of D in the cases $a=b=0.7$, $a=b=0.95$, $a=b=1.1$. In this situation, the locus of D is a simple closed curve for $a=b<1$, and it is eight-shaped for $a=b\geqslant 1$; the conclusion in Holditch's theorem holds only for $a=b\leqslant 1$. This example shows that Holditch probably assumed that the locus of D was a simple closed curve.

Finally, Holditch overlooks the possibility that the envelope of the chords may not be of the simple type illustrated in FIGURE 1(a). (The assumption that this envelope is a convex curve is implicit in his derivation of the integrals in his proof.) Consider, for example, the chords of length 1 in the unit square (see FIGURE 5). Their envelope consists of several branches; one branch is the arc in the first quadrant of the astroid $x^{2/3}+y^{2/3}=1$; other branches are obtained by successive rotations of this arc 90° around the point (1/2, 1/2). Branches corresponding to chords parallel to a coordinate axis do not exist. Holditch's proof can be construed as covering the situation in FIGURE 5, provided a suitable (but not standard) definition of envelope is considered. The dotted circle arcs in FIGURE 5 are included in the locus of the chord midpoints. They cut off, from the region within the square, a point set of area $\pi/4$; this is in accordance with Holditch's theorem.

In order to make explicit the assumptions Holditch apparently made, and to clarify the conditions under which his theorem holds, we wish to carefully restate the theorem. Some definitions are needed first. A **curve** is a planar point set having a parametric representation of the

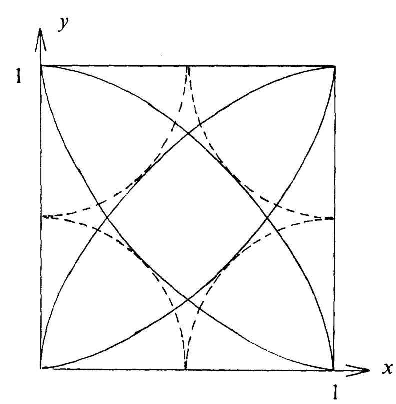


FIGURE 5

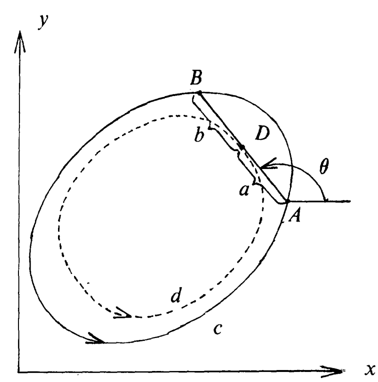


FIGURE 6

form $x=x(t)$, $y=y(t)$, $0 \leqslant t \leqslant 1$, where the functions $x(t)$ and $y(t)$ are continuous. A **simple closed curve** is a curve such that distinct t-values give distinct points of the curve with one exception: $t=0$ and $t=1$ give the same point. A **convex closed curve** c is a simple closed curve such that, if A and B are two points of c, either no interior point of the line segment AB lies on c or every point of the line segment AB lies on c. The expression "the region within a simple closed curve" has a precise meaning, coinciding with the intuitive notion. (This follows from the Jordan curve theorem. There are several proofs of this theorem in the literature. All of them are hard.)

THEOREM 1 (A Modern Version of Holditch's Theorem). *Let c be a convex closed curve, and suppose that $A=A(t)$ and $B=B(t)$, $0 \leqslant t \leqslant 1$, are points that traverse c counterclockwise one revolution as t increases from 0 to 1, so that AB is a chord of c with a constant length $a+b$. Let $D=D(t)$ be the dividing point of AB so that $|AD|=a$. Suppose that the direction angle $\theta=\theta(t)$, $0 \leqslant t \leqslant 1$, of the chord AB (measured as in* FIGURE 6) *is an increasing continuous function of t with $\theta(1)=\theta(0)+2\pi$. Let d be the locus of D, and assume that d is a simple closed curve. Then the area of the region between c and d is equal to πab.*

We do not give a separate proof of this theorem here, since it is a special case of our more general result proved in the next section. It should be pointed out that the existence of the function $\theta(t)$ in the theorem seems essential [2].

Holditch's Theorem. A Generalized Version.

Although the assumptions that c is convex and d is a simple closed curve seem essential to Holditch's result, we can relax these requirements and still obtain a formula which measures the "area" between the two curves. Holditch's theorem is then just a special case. Recall that a function $f(t)$ is of **bounded variation** on an interval $[u,v]$ if there is a positive number M such that for every partition $\{u=t_0<t_1<\cdots<t_n=v\}$, M is an upper bound for the sum

$$\sum_{i=1}^{n} |f(t_i)-f(t_{i-1})|.$$

THEOREM 2 (A Generalization of Holditch's Theorem). *Let α be a closed rectifiable curve with parametric representation $x=x(t)$, $y=y(t)$, $0 \leqslant t \leqslant 1$. Let $\theta=\theta(t)$, $0 \leqslant t \leqslant 1$, be a continuous function of bounded variation with $\theta(1)=\theta(0)+n\cdot 2\pi$, n denoting an integer. Let a, b be positive numbers. Let point $A=A(t)$, $0 \leqslant t \leqslant 1$, traverse α, and for each t, let $B=B(t)$ be the point such that AB is a line segment with length $a+b$ and direction angle $\theta=\theta(t)$. Let $D=D(t)$ be the point of AB at the distance a from A. Denote by β and δ the curves traced out by B and D respectively, as A traverses α. Set*

$$I_\alpha=\int_\alpha x\,dy,\ I_\beta=\int_\beta x\,dy,\ I_\delta=\int_\delta x\,dy.$$

Then

$$I_\delta=\frac{b}{a+b}I_\alpha+\frac{a}{a+b}I_\beta-n\pi\cdot ab.$$

Refer to FIGURE 7 for an illustration of the conditions given in the theorem. The assumptions (that α is rectifiable and that θ is of bounded variation) imply that β and δ are also rectifiable, hence I_α, I_β, I_δ are defined. The condition $\theta(1)=\theta(0)+n\cdot 2\pi$ obviously has the effect that the line segment AB comes back to the starting position when the point A has gone around the curve α. The integer n is known as the **winding number** of AB [**1**]. If α is a simple curve, traversed counterclockwise, then Green's theorem says that I_α is the area of the region within α; similarly for I_β and I_δ. If α is traversed clockwise, then $-I_\alpha$ is the corresponding area.

If α and β are the same curve (as in Holditch's original formulation) then Theorem 2 says

$$I_\delta=I_\alpha-n\pi\cdot ab.$$

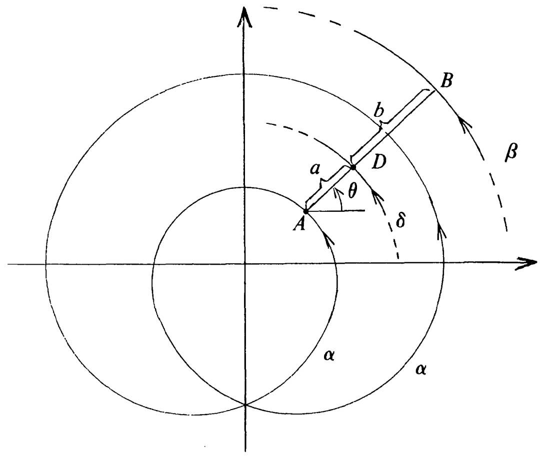


FIGURE 7

If also $n=1$, then

$$I_\delta = I_\alpha - \pi ab. \tag{1}$$

In particular, when α is a simple closed curve, $n=1$, so equation (1) implies Theorem 1. However, this last equation gives more; for example, retrograde motions are allowed when A and B go around α. (Retrograde motion means that as the chord AB traverses α, one endpoint is forced to reverse its direction of travel and retrace portions of its path on α in order to allow the other endpoint to continue to travel around α.) Retrograde motions occur in FIGURE 3; also in FIGURE 4 if $a=b\geqslant 1$.

In the following proof, our operation on line integrals may be unusual. However, the operation is legitimate, for all functions under consideration can be uniformly approximated by continuously differentiable functions. Hence, we use properties of line integrals that follow from analogous properties of the ordinary integral of a continuous function on a closed interval.

Proof of Theorem 2. We have

$$I_\beta = \int_\beta x\,dy = \int_\delta (x + b\cdot\cos\theta)\,d(y + b\cdot\sin\theta).$$

Set

$$I_1 = \int_\delta x\,d(\sin\theta) + \cos\theta\,dy \quad\text{and}\quad I_\theta = \int_\delta \cos\theta\,d(\sin\theta).$$

Then

$$I_\beta = I_\delta + b\cdot I_1 + b^2 I_\theta.$$

Analogously

$$I_\alpha = I_\delta - a\cdot I_1 + a^2 I_\theta.$$

It follows that

$$b\cdot I_\alpha + a\cdot I_\beta = (a+b)I_\delta + ab(a+b)I_\theta.$$

Now observe that

$$I_\theta = \int_\delta \cos^2\theta\,d\theta = n\pi.$$

The conclusion follows.

Some Applications to Problems.

Theorem 2 can be applied to solve several interesting problems.

PROBLEM. Find the area swept out by the point D on the piston rod in FIGURE 8(a).

SOLUTION. The point A travels up and down with the motion of the piston, so the curve α which it traces is rectifiable. Theorem 2 gives, in the notation in FIGURE 8(b),

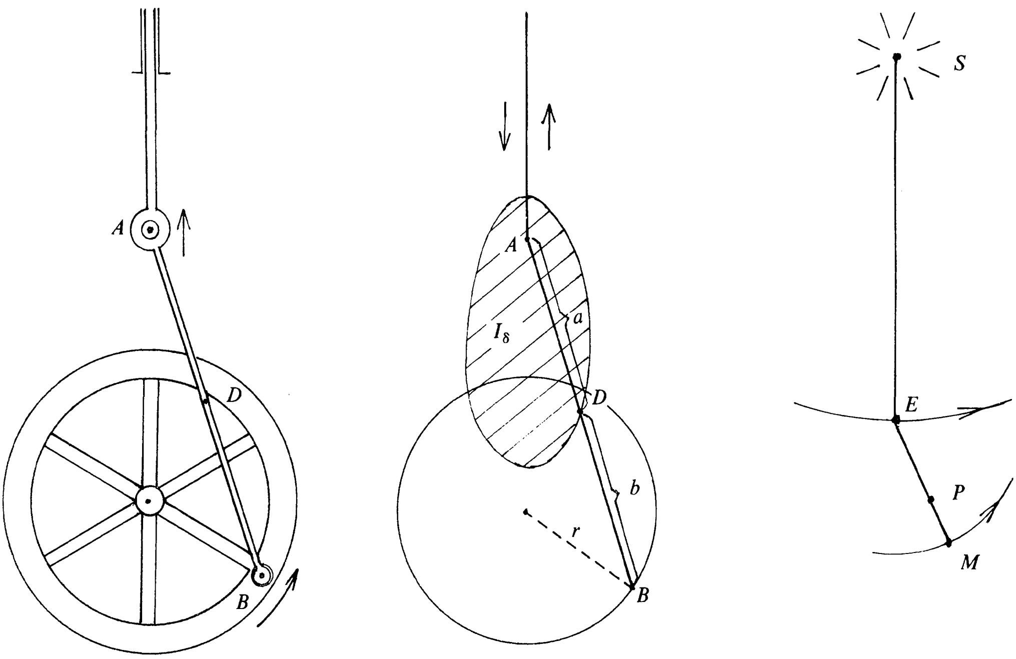
$$I_\delta = \frac{b}{a+b}\cdot 0 + \frac{a}{a+b}\cdot \pi r^2 - 0\cdot ab = \frac{a}{a+b}\cdot \pi r^2,$$

where r is the radius of the circle about which B traces.

EXAMPLE. Assume that the orbit of the earth, E, around the sun, S, is an ellipse, that the orbit of the moon, M, around E is a circle, and that P is a particle, always situated at the point where the attractions from E and M cancel (see FIGURE 8(c)). Assume that S, E, M, P always are in one plane and that the motion has the period 19 years $=254$ sidereal months. Then Theorem 2 is applicable with I_α equal to 19 times the area of the ellipse (the interpretation of I_δ and I_β is a little more complicated) and $n=254$. (In reality,

$$19\times\frac{1 \text{ sidereal year}}{1 \text{ sidereal month}} = 19\times\frac{365.25636 \text{ days}}{27.32165 \text{ days}} = 254.006.$$

Nor are the other assumptions quite realistic.)



(a) (b) (c)

FIGURE 8

PROBLEM (Suggested by John Reay, Western Washington University). Let a triangle be given, along with a line segment shorter than any altitude of the triangle. Let point D divide the line segment into lengths a, b, and let the line segment glide with its end points on the triangle. Assume the line segment is so short that d, the locus of D, is a simple curve (consisting of three curved parts and three line segments). Find the area $\mathbf{A}$ of the point set between the triangle and d (i.e., the sum of the areas of the regions shaded in FIGURE 9(a)).

SOLUTION. Note that when the line segment goes around the triangle, each of its end points makes some retrograde motion whenever a vertex of an acute angle is traversed, hence Theorem 1 is not applicable. Theorem 2, on the other hand, covers this situation. It shows that the area $\mathbf{A}$ is πab.

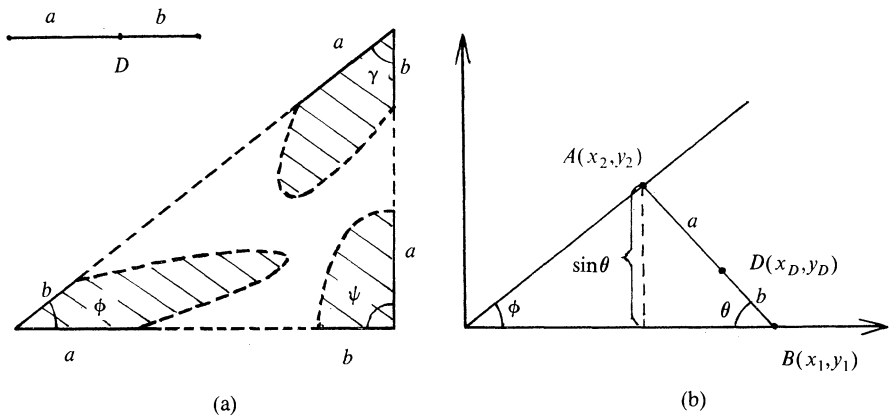


FIGURE 9

We wish to give a separate solution for this problem, since the ideas involved have an interest of their own. Assume $a+b=1$, which is no essential restriction. Let $\mathbf{A}_\phi$ denote the area of the left-hand shaded region in FIGURE 9(a), and let d_ϕ be its boundary. In FIGURE 9(b) we show one position of the line segment AB as it glides around the triangle, near the vertex angle ϕ. Using the notation in FIGURE 9(b), we have the following equations.

$$\begin{aligned}
x_1 &= \cos\theta + \sin\theta\cot\phi, \quad & x_2 &= \sin\theta\cot\phi \\
y_1 &= 0, & y_2 &= \sin\theta
\end{aligned}$$

$$\begin{aligned}
x_D &= a\cdot x_1 + b\cdot x_2 = a\cos\theta + \sin\theta\cot\phi \\
y_D &= a\cdot y_1 + b\cdot y_2 = b\sin\theta
\end{aligned}$$

$$\begin{aligned}
\mathbf{A}_\phi &= \int_{d_\phi} x\,dy \\
&= \int_0^{\pi-\phi} (a\cos\theta + \sin\theta\cot\phi)\,b\cos\theta\,d\theta + \int_{b\sin\phi}^{0} y\cot\phi\,dy.
\end{aligned}$$

A computation of this expression gives

$$\mathbf{A}_\phi = \frac{ab}{2}(\pi-\phi).$$

For a triangle whose angles have the measures ϕ, γ, ψ we then get, defining $\mathbf{A}_\gamma$ and $\mathbf{A}_\psi$ analogously to $\mathbf{A}_\phi$,

$$\mathbf{A} = \mathbf{A}_\phi + \mathbf{A}_\gamma + \mathbf{A}_\psi = \frac{ab}{2}\big((\pi-\phi)+(\pi-\gamma)+(\pi-\psi)\big) = \pi ab.$$

Hence, the claim in Holditch's theorem also holds here.

We note that elimination of θ between the equations for x_D and y_D shows that the curved arcs in FIGURE 9(a) are arcs of ellipses.

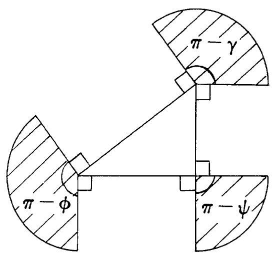


FIGURE 10

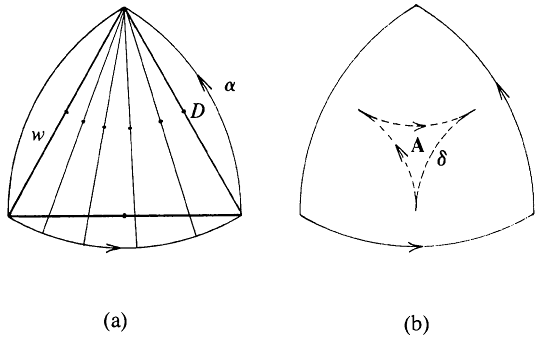


FIGURE 11

We will call the angles $\pi-\phi$, $\pi-\gamma$, $\pi-\psi$ (see FIGURE 10) the **polar angles** of the triangle. The computation of the area A in the above proof can be replaced by the observation that these polar angles have the sum 2π (for the shaded regions in FIGURE 10 can be translated to form a circular disk). It can be shown in an analogous fashion that the formula $\mathbf{A}=\pi ab$ holds when the triangle is replaced by an arbitrary convex polygon (and $a+b$ is sufficiently small).

PROBLEM (Suggested by a referee). Take α to be a Reuleaux triangle of constant width w (the Reuleaux triangle consists of three circular arcs with centers at the vertices of an equilateral triangle whose side has length w). Take δ to be the curve traced out by the midpoint D of a sliding chord of length w (see FIGURE 11). Apply Theorem 2 to determine the area **A** of the region enclosed by δ.

SOLUTION. Define I_α and I_δ as in Theorem 2, assuming α is traversed counterclockwise. Then I_α is the sum of the areas of the three segments and the equilateral triangle in FIGURE 11(a):

$$I_\alpha=3\left(\frac{1}{6}\pi w^2-\frac{1}{4}w^2\sqrt{3}\right)+\frac{1}{4}w^2\sqrt{3}=\frac{1}{2}w^2(\pi-\sqrt{3}).$$

Further, $I_\delta=-2\mathbf{A}$, for the interior "triangle" in FIGURE 11(b) is traversed clockwise two times. Equation (1) following Theorem 2 gives

$$-2\mathbf{A}=I_\delta=I_\alpha-\pi\left(\frac{w}{2}\right)^2=\frac{1}{2}w^2(\pi-\sqrt{3})-\frac{1}{4}\pi w^2,$$

and hence

$$\mathbf{A}=\frac{1}{8}w^2(2\sqrt{3}-\pi).$$

The area **A** can of course also be determined by subtracting the areas of three segments from the area of a triangle. On the other hand, **A** cannot be determined by applying the original Holditch theorem.

Is there a Holditch Theorem in $\mathbf{R}^3$?

Green's theorem (which shows that the area within a simple closed rectifiable curve α, traversed counterclockwise, is $\int_\alpha x\,dy$) is an essential tool in our discussion of Theorem 2. It is known that Stokes' theorem in many contexts is a natural counterpart in R^3 (in xyz-space) of Green's theorem in R^2 (in the xy-plane). We might hope that Stokes' theorem and the technique in the proof of Theorem 2, applied to surfaces α, β, δ in R^3, produce a formula such as

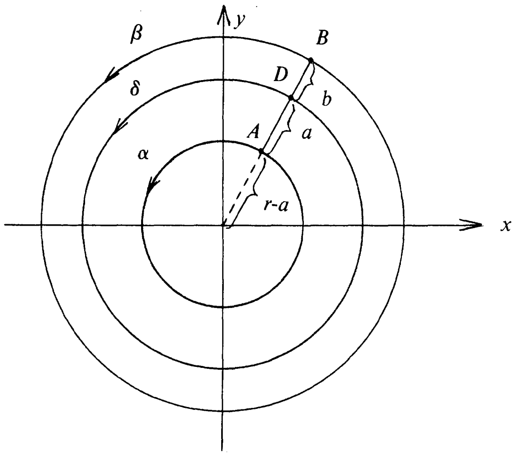


FIGURE 12

$$V_\delta = \frac{b}{a+b} V_\alpha + \frac{a}{a+b} V_\beta - n\frac{4\pi}{3} \cdot ab(a+b),$$

where $V_\delta = \int_\delta x\ dy\ dz$ under certain assumptions is the volume within a closed surface (and similarly for V_α, V_β), n is some "winding number" of a vector-valued function AB in R^3, where $|AB|$ is constant, and $4\pi/3$ is the volume of the unit sphere. Unfortunately, this idea does not lead to the conjectured result. The following example demonstrates this.

Let I_α, I_β, I_δ be the areas within the circles α, β, δ respectively in FIGURE 12. Then (cf. Theorem 2),

$$I_\delta - \frac{b}{a+b} I_\alpha - \frac{a}{a+b} I_\beta = \pi r^2 - \frac{b}{a+b}\pi(r-a)^2 - \frac{a}{a+b}\pi(r+b)^2 = -\pi\frac{a^2b+ab^2}{a+b} = -\pi ab.$$

Rotate the entire figure around the x-axis, and let V_α, V_β, V_δ denote the volumes within the spheres generated by α, β, δ respectively. Then

$$\begin{aligned}\frac{3}{4\pi}\cdot\left(V_\delta - \frac{b}{a+b}V_\alpha - \frac{a}{a+b}V_\beta\right) &= r^3 - \frac{b}{a+b}(r-a)^3 - \frac{a}{a+b}(r+b)^3 \\ &= \frac{b}{a+b}(3r^2a - 3ra^2 + a^3) - \frac{a}{a+b}(3r^2b + 3rb^2 + b^3) \\ &= \frac{-3a^2b - 3ab^2}{a+b} r + \frac{ab(a^2-b^2)}{a+b} = -3abr + ab(a-b).\end{aligned}$$

This last expression varies with r, in contrast to the situation in R^2 where Holditch's theorem yields a constant value. Hence there is probably no Holditch theorem for R^3. An analogous argument indicates that there is probably no Holditch theorem in R^n for $n > 3$.

References

[1] Lars V. Ahlfors, Complex Analysis, 2nd ed., McGraw-Hill, 1966, p. 115.

[2] Arne Broman, Holditch's theorem is somewhat deeper than Holditch thought in 1858, Nordisk Mathematisk Tidskrift, vol. 27, 1979, pp. 89–100. In this article (in Swedish), the author gives a sufficient condition for the existence of such a function $\theta(t)$: Assume that h is the largest real number such that every circle with radius less than h and center on c intersects c at precisely two points; then $a+b<h$ is sufficient. In addition, it follows from the other assumptions that d is a simple closed curve. The proof is rather long.

[3] H. S. M. Coxeter, Introduction to Geometry, 2nd ed., Wiley, New York, 1969, p. 8 (Euclid's Proposition 35, Book III). For Euclid's proof, see Thomas L. Heath, The Thirteen Books of Euclid's Elements, vol. II, Dover Publications, New York, 1956, pp. 71–72.

[4] Rev. Hamnet Holditch, Geometrical theorem, The Quarterly Journal of Pure and Applied Mathematics, vol. 2, London, 1858, p. 38.

A Proof of Gödel's Theorem in Terms of Computer Programs

By considering a few conditions which a universal axiomatic system should satisfy, it is not hard to see that such a system is impossible.

ARTHUR CHARLESWORTH
University of Richmond
Richmond, VA 23173

It would be convenient to have a single axiomatic system within which all mathematical conjectures could be rigorously settled. Kurt Gödel's remarkable Incompleteness Theorem seems to indicate, however, that such a system is impossible. Moreover, the effect of this limitation on Zermelo-Fraenkel set theory (the nearest thing we have to a universal axiomatic system) is being felt increasingly by those who work, not just in the foundations, but in the mainstream of mathematics.

A goal of this article is to provide a proof of Gödel's Theorem which has sufficient depth to permit a healthy understanding and yet which has as few technicalities as possible. In studying Gödel's Theorem, it is natural to search for a loophole in the hypothesis which would permit the quest for a universal axiomatic system to be revived; a second goal of this article is to present the hypothesis in such a general way that no loophole seems to exist.

Let us begin with the following intuitive statement of the theorem.

GÖDEL'S INCOMPLETENESS THEOREM. *Let S be any consistent formal axiomatic system which has a reasonable language, a reasonable concept of theorem, and which is reasonably powerful. Then S is incomplete; that is, there is an assertion A in the language of S such that neither A nor the negation of A is a theorem in S.*

Before we can restate this general theorem more precisely, we must develop an understanding of the need for and the nature of formal axiomatic systems.

Formal axiomatic systems

We are all familiar with the concept of an "axiomatic system": the basic idea is to begin with certain undefined terms and axioms about these terms and proceed to prove theorems from the axioms. In proving theorems we are permitted to use only those properties of the undefined terms which are given by the axioms; we are not permitted to rely on the usual meanings of these terms. This concept of an axiomatic system, of course, has had a long evolution, stimulated largely by Euclid's treatment of geometry. It was David Hilbert, in the late 1800's, who first expressed the concept in the above form, insisting that an axiomatic treatment of geometry should be capable of having the undefined terms "points, lines, and planes" replaced by "tables, chairs, and beer mugs," since the customary meaning of these terms would not be assumed.

Although very useful in many situations, this concept of an axiomatic system suffers from a weakness: in proving theorems from the axioms there are no clear guidelines on just what constitutes a proof. For example, what is there to prevent someone from taking any plausible statement and "proving" it by saying it "obviously follows from the axioms"? In the practice of mathematics this is inhibited by a culturally developed sense of what is acceptable; but mathematical "cultures" vary so that what was acceptable to Euclid was not always acceptable to

Hilbert. Even within the same "culture" it is not uncommon for two mathematicians to disagree on whether a given argument needs to be made more rigorous (either by filling in more details or, if this does not seem possible, by finding a different approach).

What is needed is a clear concept of "proof" which incorporates the essential features of the vague concept, yet which demands so much rigor that virtually all mathematicians can accept it. But what are the essential features of the vague concept of proof? One answer might be that a proof is a finite sequence of statements linked together in a logical way. This suggests that we should make precise the notions of "statement" and "logical way."

A statement in an ordinary proof usually contains, in addition to mathematical terms, certain logical terms such as "implies" and "there exists." Thus, when precisely defined, a "statement" should be capable of containing both mathematical and logical terms. It is customary to use the logical symbols $\neg$, $\vee$, &, $\rightarrow$, $\exists$, and $\forall$ to represent, respectively, "not," "or," "and," "implies," "there exists," and "for every," together with parentheses to clarify their usage. How can these symbols be defined so they have their intended meanings? Except for the last two symbols, this can be done in a concrete way, for example, by using truth tables. But it doesn't seem possible to give the symbols $\exists$ and $\forall$ genuine definitions, especially when they range over infinite sets; it seems that one must first understand the meaning of terms such as "there exists" and "for every." This is similar to the problem of defining "point" in geometry; recall that Hilbert's solution was to consider "point" as undefined and restrict its meaning using axioms. It is natural to use the same approach for $\exists$ and $\forall$; in fact, since all mathematical terms and the logical symbols $\exists$ and $\forall$ will be undefined, it is convenient to add sufficient axioms so that the other logical symbols are undefined as well. (If we chose to define the other logical symbols using truth tables, we would actually be obliged to provide axioms for truth tables or for undefined terms with which we would define truth tables.)

Thus we end up with a kind of system in which the customary meanings of both logical and mathematical notions cannot be used in proofs; only the form of statements is important. For this reason such a system is called a "formal" axiomatic system and the collection of symbols forms the basis for the "formal" language of the system. (Hilbert himself was among those instrumental in the introduction of such formal systems in the early 1900's.)

A word game

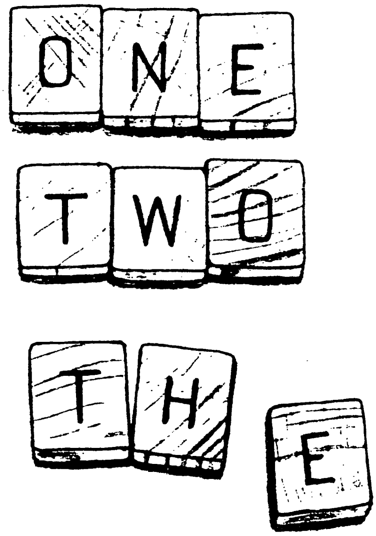
We shall be working with two different levels of languages and deductions: the English language and ordinary mathematical procedures are on one level (which we shall refer to as "metamathematics") and these will be used to examine the second level, which consists of statements and proofs in formal axiomatic systems. Keeping the distinction between these two levels in mind can be tricky without the proper perspective.

To gain a good perspective it will help to consider a variation of the familiar game in which you are asked to transform a word like IRON into a word like GOLD by changing one letter at a time so that each new word created can be found in the dictionary. In particular, let us consider the following variation.

WORD GAME

You are given the words ONE and TWO and are asked to transform them into the word THREE by making a list of dictionary words such that THREE is at the end of the list. To put a word **w** into the list, it must either be one of the given words ONE or TWO or it must satisfy the following rule:

> If **w** were added to the list, there would be two words $\mathbf{w}_1$ and $\mathbf{w}_2$ already on the list such that each letter in **w**, with one possible exception, occurs in the same position in either $\mathbf{w}_1$ or $\mathbf{w}_2$.



One solution to this word game is

(1)	ONE	one of the given words
(2)	TWO	one of the given words
(3)	THE	the rule, using (1) and (2)
(4)	THEE	the rule, using (1) and (3)
(5)	TIE	the rule, using (1) and (3)
(6)	TIRE	the rule, using (4) and (5)
(7)	THREE	the rule, using (4) and (6)

There is a strong analogy between this word game and the kind of formal axiomatic system we have been describing, especially if we extend the game to consider what words, in addition to THREE, can be derived from ONE and TWO. To make the analogy clear, let us say that the above list is a "proof" of THREE from the "axioms" ONE and TWO using the rule. Since there is a proof of THREE, we would say that THREE is a "theorem"; clearly TIRE is also a theorem and there are many other possible theorems. Note how the concept of proof in this game is clearly defined and based entirely on form; we do not rely on the meanings of the words in the game.

A good perspective would be to view this word game as being packed in a cardboard box like a standard table game, with letters of the alphabet on individual wooden squares and with the rule printed on the inside cover of the box. A dictionary should also be packed in the box to make it clear just which words are acceptable. It would then be easy to prove a result such as "there are just finitely many possible theorems in the word game." (For the dictionary contains just finitely many words and every theorem in the game is a dictionary word.) The result in quotes might be called a theorem in metamathematics; it is clearly not a theorem in the word game since it is not a dictionary word. Also notice that our proof of this result was on the metamathematical level; it certainly did not have to satisfy the requirement for a proof in the word game.

An example of a formal axiomatic system

Recall that one way of describing the ordinary concept of proof is to say that a proof is a finite list of statements linked together in a logical way. In the context of formal axiomatic systems we can make the phrase "logical way" precise by requiring that each statement in the proof either be an axiom or that it follow from preceding statements in the proof by one of the rules of the system, which are called "rules of inference." It will help to look at an example of such a formal system.

The type of formal language we shall use in the example is typical of those commonly used when proving Gödel's Theorem. In this article, however, we use this type of formal language merely for illustrative purposes; the hypothesis of the general version of Gödel's Theorem which we consider would be satisfied by a variety of formal languages.

EXAMPLE. The axioms of the formal system S_0 are

1) $\forall x(x+0=x) \& \forall x(0+x=x) \rightarrow \exists y\ (\forall x(x+y=y+x))$,
2) $\forall x(x+0=x)$, and
3) $\forall x(0+x=x)$.

The rules of inference are

1) If A is on the list and B is on the list, then $A \& B$ may be put on the list,
2) If A is on the list and $A \rightarrow B$ is on the list, then B may be put on the list, and
3) If A is on the list and B is a formal statement in S, then $A \vee B$ may be put on the list.

In describing the rules of inference, the letters A and B (which are not in the formal language) are used to represent formal statements in the formal language. This is analogous to the use of the letters $\mathbf{w}$, $\mathbf{w}_1$, and $\mathbf{w}_2$ to represent words in the description of the rule for the word game.

An example of a theorem in the system S_0 is $\exists y\ (\forall x(x+y=y+x))$, since it occurs as the last entry in the following proof.

(1)	$\forall x\ (x+0=x)$	axiom 2
(2)	$\forall x\ (0+x=x)$	axiom 3
(3)	$\forall x\ (x+0=x)\ \&\ \forall x(0+x=x)$	rule 1, using (1) and (2)
(4)	$\forall x\ (x+0=x)\ \&\ \forall x(0+x=x) \rightarrow$ $\exists y\ (\forall x(x+y=y+x))$	axiom 1
(5)	$\exists y\ (\forall x(x+y=y+x))$	rule 2, using (3) and (4)

The language of S_0 is defined by specifying the symbols of the language and then indicating which expressions (finite strings of symbols) will be considered to be formal statements. On the metamathematical level each formal statement has the following important property: whenever the symbols of S_0 are given their intended meanings, a formal statement and its negation correspond to assertions in English which are so well formed and unambiguous that one of them is true. For example, each of the following expressions could be regarded as formal statements: the expression $\forall x(x+0=x)$ (which is intended to mean "for every natural number x, x plus zero equals x"), the expression $1+1=1$ ("one plus one equals one"), and the expression $8=w \rightarrow \exists y\ (w=2 \cdot y)$ ("if eight equals w, then there exists a natural number y such that w equals two times y"). On the other hand, the expressions $x+y=y$ ("x plus y equals y") and $=+x\forall$ ("equals plus x for every") could not be regarded as formal statements.

Of course, we have emphasized that intended meanings are not to be relied on when working within a formal system. Thus it would be necessary to specify, in terms of form alone, which expressions are formal statements. It is not necessary to give a "dictionary" of all formal statements in the language; all that would be required is an algorithm for checking whether or not a particular expression is a formal statement. Since we shall use S_0 only for illustrative purposes, it is not necessary to actually give such an algorithm here. (For those who would like a rough idea: the algorithm could first use a standard procedure, like that in Section 6.6 of [**11**], to check that the expression is "well formed." Such a procedure is analogous to procedures used by computers in checking to see that expressions such as $A=(B+C)*D$ are well formed, for BASIC or FORTRAN, but expressions such as $A=(B+C/*D$ are not. Next, the algorithm could check that all variables in the expression, except possibly one, have all occurrences bound by a quantifier. If there was a variable which was not bound by a quantifier, like w in $8=w\rightarrow\exists y\,(w=2\cdot y)$, it would have to be bound by an initial formal hypothesis such as "$8=w\rightarrow$." Notice that each expression approved by this algorithm would satisfy the important property of formal statements mentioned in the preceding paragraph. It is not crucial, on the other hand, for every expression in S which satisfies this property to be approved by the algorithm.)

Before leaving our discussion of S_0 let us observe that S_0 does not have enough axioms and rules of inference to restrict uniquely the meanings of the symbols. On the one hand, we could interpret the symbols of S_0 as having their usual meanings and consider that they describe addition of natural numbers. Such an interpretation is a satisfactory one because each theorem would be true; i.e., for each theorem, the English assertion which corresponds to the theorem under the interpretation would be true. (To see why each theorem would be true, note that each axiom would be true and each rule of inference would preserve truth; i.e., when applied to a list of formal statements which are all true, each rule would add only true formal statements to the list. It then follows that each formal statement in a proof would be true, so each theorem would be true.) On the other hand, we could interpret $+$ to mean multiplication among natural numbers, 0 to mean the number one, & to mean "implies," $\rightarrow$ to mean "and," with the remaining symbols having their usual meaning. This interpretation of the symbols of S_0 would also be satisfactory since again each theorem would be true (because here also the axioms would be true and the rules of inference would preserve truth).

Conditions for a reasonable formal system S

Rather than put severe restrictions on a formal system S and thus make Gödel's Theorem seem like a narrow result, our approach will be to list general conditions which seem natural for writing mathematical proofs. In doing this we should keep the nature of informal proofs in mind. For instance, since we are accustomed to writing our informal proofs in English, which has less than 100 symbols, it is natural to consider that the formal language of S has just finitely many symbols. (In considering that English has less than 100 symbols, we include the numerals and punctuation but ignore specially defined mathematical notation. Such notation can always be viewed as an abbreviation for English.)

CONDITION 1. *There are just finitely many symbols in the formal language of S. An expression in S is a nonempty string of finitely many symbols of S, including repetition. Each formal statement in S is an expression and there is an algorithm α_1 for checking whether or not any particular expression is a formal statement.*

Although the formal system S_0 has finitely many axioms, there are reasonable formal systems which have infinitely many axioms. For example, one might want every formal statement of the general form $A\ \&\ B\rightarrow A$ to be an axiom, where A and B are formal statements. If S is allowed to have infinitely many axioms, we should still be able to tell which expressions are axioms and which are not. Thus we would expect that there be an algorithm which checks whether or not a given expression is an axiom. For additional generality beyond that of S_0, the axioms of S will not be required to be formal statements, although we will require that the theorems are formal statements.

The rules of inference provided with S should also be clear, so that the correctness of adding a new expression to a proof can be checked by an algorithm. We can summarize these requirements for axioms and rules of inference by requiring that there be an algorithm for checking whether or not a given list of expressions is actually a proof.

CONDITION 2. *A theorem in S is a formal statement which occurs as the last line of a proof. Each proof is a finite list of expressions and there is an algorithm α_2 for checking whether or not any particular finite list of expressions is a proof in S.*

The concept of negation is a fundamental one in informal mathematics and thus it is natural to require that the formal system S have a way of denoting negations. It is not essential that the usual symbol $\neg$ for negation be used in S; the negation of a formal statement could be a more complicated rearrangement of the formal statement than merely putting $\neg$ in front of it. All that is really essential is that there be an algorithm for performing the rearrangement.

In the next condition we refer to the possible meanings of the symbols of S. Just as it would be possible for us to restrict our attention to word games whose axioms are names of metals, we may refer to meanings in the conditions we place on S. The situations in which we must not rely on meanings are those where we are actually constructing a proof within a word game or formal system.

CONDITION 3. *There is an algorithm α_3 which converts each formal statement of S into its negation, which is also a formal statement of S. For each interpretation of the symbols of S which makes all of the theorems true and for each formal statement A and its negation B, either A is true or B is true but not both.*

DEFINITION. A formal axiomatic system S is **consistent** if there is an interpretation of the symbols of S which makes all the theorems true. Otherwise S is **inconsistent**.

Recall that the purpose of using a formal system is to recast informal arguments in a rigorous setting. Thus in practice one chooses a formal system in which all theorems are expected to be true when interpreted in terms of the branch of informal mathematics one has in mind. (It may be hard, of course, to give a truly convincing argument that the formal system actually has this property; we shall return to this observation in the section entitled "Making use of the conditions.")

Notice that if S is consistent and satisfies Condition 3, then there is no formal statement A such that both A and the negation of A are theorems of S.

DEFINITION. A formal axiomatic system S is **complete** if for each formal statement A in S, either A or the negation of A is a theorem in S. Otherwise S is **incomplete**.

A formal language which is used to simulate anything beyond a few small portions of informal mathematics needs to be able to denote any particular natural number. Thus the formal system S should include an algorithm for constructing the appropriate notation for a number, whether the notation used within S is decimal, binary, a tally notation, or even English (in which 1776 is written "seventeen hundred and seventy-six").

CONDITION 4. *There is an algorithm α_4 which constructs the notation in S for a natural number, given the natural number. The notation for each natural number is an expression in S and if the notation for a particular natural number in a formal statement F is replaced by the notation for any other natural number, the resulting modification of F is a formal statement of S.*

The fact that distinct natural numbers must have distinct notations in S will follow from our final condition.

A standard technique in informal mathematics is to prove a result such as "If x is an even number, then x^2 is even" and then to use "If 1776 is an even number, then 1776^2 is even" and "1776 is an even number" to deduce "1776^2 is even." This technique suggests that the formal system S should have expressions analogous to the English phrases "x is an even number" and

"x^2 is an even number," which are not formal statements but which become formal statements when we substitute in the notation for a natural number. For example, the expressions $x=w\to\exists y\,(w=2\cdot y)$ and $x=w\to\exists y\,(w\cdot w=2\cdot y)$ in the language of S_0 would become the formal statements $1776=w\to\exists y\,(w=2\cdot y)$ and $1776=w\to\exists y\,(w\cdot w=2\cdot y)$, when 1776 is substituted for x. We shall refer to such expressions as "number predicates." For simplicity and generality our conditions are stated without referring to the concept of a variable and, as a result, our definition of number predicate actually allows for the possibility of a formal system having the string of symbols "it is an even number" as a number predicate. (For those familiar with predicate logic we should point out that $1776=w\to$ is being used as a convenient device to avoid having to talk about such things as "the formula resulting from $\exists y\,(w\cdot w=2\cdot y)$ when all free occurrences of w are replaced by 1776.")

DEFINITION. Let S be a formal axiomatic system and let F be an expression in S which is not a formal statement in S. Then F is a **number predicate** if F has the following property: F begins with an expression E such that if E is replaced by the notation in S for the number 1, the resulting modification of F is a formal statement in S.

NOTATION. Let F be a number predicate, let n be a natural number, and let E' be the shortest expression within F which could play the role of E. Then $F(n)$ will denote the expression obtained by replacing E' by the notation in S for n.

Notice that the expression $F(n)$ is a formal statement in S, assuming S satisfies Condition 4. Keep in mind that F and $F(n)$ are notations at the metamathematical level, not notations in the system S. (Compare this with the use of $\mathbf{w}$, $\mathbf{w}_1$, and $\mathbf{w}_2$ in describing the word game and the use of A and B in describing the rules of inference of S_0.)

Reasonably powerful formal systems

There are several ways in which the phrase "reasonably powerful" can be made precise. We shall define it in terms of verifying within S the outcome of certain simple computer programs, namely programs which accept just one number as input and eventually stop, printing as output either the word "YES" or the word "NO." To be specific, let us say that the programs are written in the BASIC language. (With only very minor changes in our discussion, we could instead consider FORTRAN, Pascal, or any other general purpose computer language.)

Notice that there exists an informal proof of each such outcome: we merely have to check the steps in the execution of the computer program, assuming the given input, and this check is guaranteed to stop after a finite number of steps. Thus all we are requiring is that S provide us with analogues of certain informal proofs.

CONDITION 5. *For each BASIC program P (of the type which accepts a single number as input and eventually stops, printing either "YES" or "NO" as its output) there is a number predicate F_P in the formal language of S such that for each natural number n, the program P with input n prints "YES" if and only if $F_P(n)$ is a theorem.*

To illustrate the relationship between P and F_P, suppose P is the following program (which answers the question "is N even?").

```
10 INPUT N
20 IF N/2 # INT(N/2) THEN 50
30 PRINT "YES"
40 STOP
50 PRINT "NO"
60 STOP
70 END
```

If the language of S is like that of S_0, the number predicate F_P which represents P might be the expression $x=w\to\exists y\,(w=2\cdot y)$. Assuming that S uses the decimal notation for numbers, $F_P(8)$

would be $8=w \to \exists y\,(w=2\cdot y)$. Condition 5 would require that there be sufficient power in S to prove $F_P(8)$ and, in general, to prove $F_P(n)$ exactly when n is even.

As another example, consider a program which would answer the question "is N prime?" Such a program might look like the one above, with line 20 being replaced by the following lines

```
20 I=2
22 IF N=I THEN 30
24 IF N/I=INT(N/I) THEN 50
26 I=I+1
28 GO TO 22.
```

If the language of S is like that of S_0, for this program P a likely candidate for F_P would be

$$x=w \to (\forall y((1<y \,\&\, y<w) \to \neg\exists z(z\cdot y=w)))$$

so that $F_P(8)$ would be

$$8=w \to (\forall y((1<y \,\&\, y<w) \to \neg\exists z(z\cdot y=w))),$$

assuming the system uses the decimal notation. Condition 5 requires that there be no proof in S of $F_P(8)$, since $F_P(n)$ should be a theorem exactly when n is prime.

Making use of the conditions

Conditions 1 through 5 are natural ones which a formal system S should satisfy if it is both to provide rigor and to permit analogues of at least a few of the techniques used in informal mathematics. If we choose to think of a formal system S which satisfies our conditions as being packaged like a word game in a cardboard box, the box should contain finitely many symbols on wooden squares and we should be free to make copies of these symbols as needed. The box should also contain descriptions of the algorithms α_1, α_2, α_3, and α_4. The task of actually verifying that a formal system S is consistent and satisfies Conditions 1 through 5 might be quite difficult. However, if our desire is to understand the limitation revealed by Gödel's Theorem, this difficulty need not concern us. Roughly stated: if S does not satisfy these conditions, then S is of limited value; but if S does satisfy these conditions, then Gödel's Theorem asserts that S is limited by its incompleteness.

Of course, one could take a formal system S which satisfies these conditions and make a trivial modification in the way number predicates are written (for example, by putting parentheses around each of them) to create a formal system S' which is of value, yet fails to satisfy these conditions. (Condition 5 would fail to hold since there would no longer be expressions which satisfy our definition of number predicate.) But S' would be incomplete, since S is incomplete and S and S' have corresponding theorems.

Let us now see how the algorithms given by the conditions can be combined to give two additional algorithms.

LEMMA 1. *Let S be any consistent formal axiomatic system which satisfies the requirements for reasonableness given by Conditions* 1 *through* 5. *Then there is an algorithm ϕ which generates exactly the theorems in S and there is an algorithm ψ which generates exactly the formal statements $F_1(1)$, $F_2(2)$, $F_3(3),\ldots,$ where the list F_1, F_2, $F_3,\ldots$ includes all the number predicates in S.*

Proof. By Condition 1 there are just finitely many symbols in the formal language for S. Thus there is an algorithm β_1 for generating the expressions E_1, E_2, $E_3,\ldots$. (First generate all strings of one symbol each, then all strings of two symbols each, etc.)

Using the algorithm β_1 we can construct an algorithm β_2 which generates exactly the finite lists L_1, L_2, $L_3,\ldots$ of expressions as follows. The algorithm β_2 has a stage for each natural number n; at the nth stage we simply generate all lists of length $\leqslant n$ which are composed of the expressions E_1, E_2, $E_3,\ldots,E_n$. By using the algorithm α_2 of Condition 2 we can eliminate each L_i which is not a proof and thus obtain an algorithm β_3 which generates exactly the proofs P_1, P_2, $P_3,\ldots$ in S. We can easily modify β_3 (using the algorithm α_1 of Condition 1) to obtain an algorithm ϕ which reports the last formal statement in each proof M_i which ends in a formal statement; clearly ϕ generates exactly the theorems in S.

By combining the algorithm β_1 with the algorithms α_1 and α_4, given by Conditions 1 and 4 respectively, we can obtain an algorithm β_4 which generates exactly the number predicates F_1, F_2, $F_3,\ldots$. (The algorithm β_4 applies the following test to each expression E_i as it is generated by β_1: First α_1 is applied to E_i and if E_i is verified as a formal statement, β_4 does not report E_i. If E_i is not a formal statement, β_4 counts up the number m_i of symbols of E_i and, as j ranges from 1 to m_i, causes the notation in S for 1 to be substituted in place of the substring consisting of the first j symbols of E_i. If any of the resulting modifications of E_i is verified by α_1 to be a formal statement, β_4 reports E_i. Otherwise β_4 does not report E_i.)

It is easy to see how β_4 can be modified, using α_4, to obtain an algorithm ψ which generates the formal statements $F_1(1)$, $F_2(2)$, $F_3(3),\ldots$. □

Church's Thesis

Suppose we are given a formal system S which satisfies Conditions 1 through 5, together with the algorithms which these conditions guarantee. Can the algorithm ϕ, which generates exactly the theorems of S, then be implemented as a BASIC program?

A minor difficulty which would arise in writing such a program is that not all the symbols of S will likely be symbols of BASIC; this difficulty can be easily overcome by adopting a code for the symbols of S. (One simple code would be to represent the symbols of S as square matrices consisting of 0's and 1's which look like the symbols; thus $\neg$ could be coded as a matrix whose top row and rightmost column consist entirely of 1's with the remaining entries being 0's. The number of symbols in S and their similarity to one another would determine how large these matrices should be in order for the program to distinguish between symbols of S.) Another difficulty might be that implementing ϕ requires extensive storage of intermediate results; to handle this, we can imagine that during the course of a computation, the memory of the computer can be enlarged whenever necessary (such as by adding new magnetic disk packs).

Once we have selected a code for the symbols of S and a procedure for handling memory, it is possible to write a BASIC program, based on the algorithm ϕ, which generates the theorems of the formal system S_0. In fact, such a program can be written for any presently known formal system which satisfies Conditions 1 through 5. The question we are interested in is whether this can be done for *any* possible formal system satisfying Conditions 1 though 5. In view of the construction of ϕ, this reduces to the question of whether BASIC is a powerful enough language in which to program the algorithms α_1 and α_2. Since both of these algorithms (as well as α_3 and α_4) are guaranteed to terminate eventually, given any input, we are really asking whether there is a mathematical result which assures us that every algorithm which eventually terminates is programmable in BASIC.

To be precise, such a result would have to have the vague notion of algorithm replaced by a precise notion. There have been several different approaches to define precisely the notion of algorithm, all of which have resulted in the same concept, namely, the concept of a recursive function. It is well known that every recursive function is programmable in BASIC. (See FIGURE 1 of [3] for a simple explanation of this fact, together with a definition of recursive.) Thus we would know that every algorithm is programmable in BASIC if we knew that *every algorithm corresponds to a recursive function*; this assertion is referred to as **Church's Thesis**, since Alonzo Church was one of the first to study it.

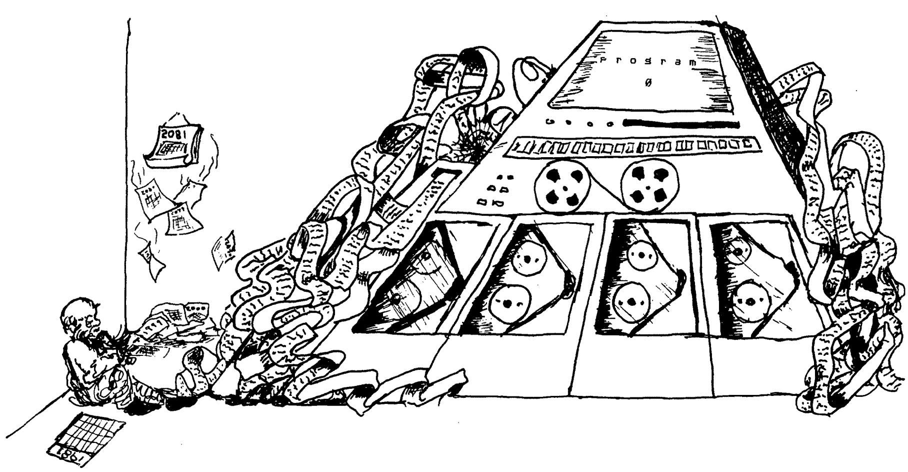
Since it is known that every recursive function corresponds to an algorithm (in fact to a BASIC program), Church's Thesis claims that a particular vague concept (algorithm) actually corresponds to the mathematical concept (recursive function) which was intended to make it precise. One could make a similar assertion about the vague concept of function (a rule which associates with each element of X an element of Y such that no element of X is associated with two different elements of Y) and the mathematical concept of function (a subset of $X \times Y$ such that every element of X is the first coordinate of some member of the subset and no two members of the subset have the same first coordinate). In this case, there is an intuitive argument which shows that anything which satisfies the vague concept corresponds to a set which would satisfy the mathematical concept. (Any such argument can be no more than intuitive, of course, since the vague concept is not itself precise.) In 1936, Alan Turing gave an intuitive argument to show that any algorithm would correspond to a recursive function. (An easily accessible account appears in Section 70 of [**8**].) Since the vague concept of algorithm is not as simple as the vague concept of function, one might worry that there is a type of algorithm which Turing overlooked and which has been overlooked up to the present time. The overall evidence for Church's Thesis, however, is extremely strong (perhaps as strong as evidence can be for such a nontrivial informal assertion) and so Church's Thesis is widely accepted. The case for Church's Thesis is summarized in Sections 62 and 70 of [**8**].

It seems necessary to rely on Church's Thesis when considering a truly general version of Gödel's Theorem. We rely on it for assurance that a BASIC program can be written to implement the algorithms of our conditions; such a program will play a key role in our proof. (Since we know that such a BASIC program can be written for any presently known formal system, it is not necessary to rely on Church's Thesis to prove less general versions of Gödel's Theorem for these formal systems.)

Reliance on Church's Thesis may seem like a "loophole" in the proof of the general Gödel's Theorem, especially to those who are not familiar with Church's Thesis, and this is why we have discussed it at such length. We make one final observation concerning Church's Thesis: if the highly unlikely should happen and an algorithm is discovered which does not correspond to a recursive function, this new kind of algorithm could probably be incorporated into a new, more powerful kind of computer program and our proof of the general Gödel's Theorem would probably survive, with BASIC programs being replaced by the new kind of program.

The proof of Gödel's Theorem

The implications of Lemma 1 should not be misunderstood; it does not give a practical way to settle your favorite unsolved problem. You could indeed take your problem and make it a formal statement A in a formal language, state your axioms in the formal language, select some standard rules of inference, write up the algorithm ϕ as a computer program and wait to see whether A appears in the generated list of theorems. But what is likely to happen is that, after watching the output for a hundred years without having A appear, you will not know whether it is because you haven't waited long enough or because A is not a theorem. In other words, if A fails to be a theorem, you will not learn this fact from ϕ.



However, let us observe that if S is not only consistent but complete, then ϕ will indeed indicate whether or not A is a theorem. For you could use the algorithm α_3 to generate the negation of A and then look at the output of ϕ until either A or the negation of A appears. If A appears, then A is a theorem; if the negation of A appears, then A is not a theorem. (Of course, this algorithm would take too long to be really practical.) We state this observation as another lemma.

LEMMA 2. *Let S be any consistent formal axiomatic system which satisfies the requirements for reasonableness given by Conditions* 1 *through* 5. *If S is complete, then there is an algorithm for determining whether or not any given formal statement in S is a theorem.*

We now restate and prove Gödel's Theorem using a Cantor diagonal argument.

GÖDEL'S INCOMPLETENESS THEOREM. *Let S be any consistent formal axiomatic system which satisfies the requirements for reasonableness given by Conditions* 1 *through* 5. *Then S is incomplete.*

Proof. Suppose that S is complete. Assuming Church's Thesis, a BASIC program P can be constructed as follows. Given the natural number n as input, P uses the algorithm ψ of Lemma 1 to generate the formal statement $F_n(n)$ and then uses the algorithm of Lemma 2 to determine whether or not $F_n(n)$ is a theorem of S. If $F_n(n)$ is a theorem, P prints "NO"; if $F_n(n)$ is not a theorem, P prints "YES." After printing this output, P stops. In summary,

P with input n prints "YES" if and only if $F_n(n)$ is not a theorem of S. (1)

By Condition 5, there is a number predicate F_P such that P with input n prints "YES" if and only if $F_P(n)$ is a theorem of S. Now F_P is one of the number predicates $F_1, F_2, F_3, \ldots$ mentioned in Lemma 1; to be specific, let us say F_P is F_m. Then we have

P with input n prints "YES" if and only if $F_m(n)$ is a theorem of S. (2)

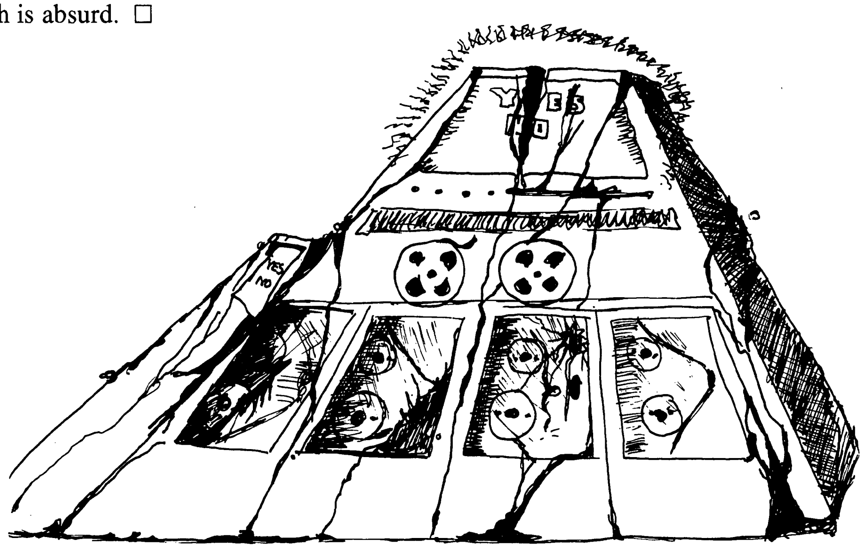
Since (1) and (2) hold for all natural numbers n, they hold for the particular number m. Thus (1) gives

P with input m prints "YES" if and only if $F_m(m)$ is not a theorem of S, (3)

and (2) gives

P with input m prints "YES" if and only if $F_m(m)$ is a theorem of S. (4)

Therefore we may conclude that $F_m(m)$ is not a theorem of S if and only if $F_m(m)$ is a theorem of S, which is absurd. □



Let S satisfy the hypothesis of Gödel's Theorem, let A be a formal statement in S such that neither A nor the negation of A is a theorem in S, and take the intended interpretation of the symbols of S (for which all the theorems are true). Then, by Condition 3, either A or the negation of A is true. Thus there is a formal statement in S which is true, under the intended interpretation of its symbols, but which is not a theorem in S. Sometimes it is this result which is referred to as Gödel's Theorem.

Formal systems which satisfy Conditions 1 through 5

Although our proof of Gödel's Theorem has now been completely presented, you might wonder whether Conditions 1 through 5 (though they seem to be quite reasonable) are actually so strong that no formal axiomatic system can satisfy them and, as a result, the general Gödel's Theorem holds because its hypothesis cannot be satisfied. This is not the case; in fact, there are several formal systems which satisfy these conditions. Recall that in illustrating the relationship that Condition 5 requires between a program P and the number predicate F_P, it was natural for us to consider expressions which had not only the usual logical symbols and symbols for variables but also the symbols for elementary arithmetic among natural numbers, such as 0, $+$, $\cdot$, and $=$. Another useful symbol is the symbol $'$ for the successor of a natural number; it permits the natural numbers to be denoted by 0, $(0)'$, $((0)')'$, etc. There are formal systems, inspired by the informal Peano axioms, which use these symbols and have sufficient axioms and rules of inference to satisfy our conditions. One such system, commonly called "formal arithmetic," is given in Chapter 4 of [**8**]. It is relatively straightforward to verify that formal arithmetic satisfies the first four conditions, but quite tedious and technical to verify that formal arithmetic satisfies Condition 5; this is typical of proofs of Gödel's Theorem for particular formal systems.

Loosely speaking, any formal axiomatic system which is powerful enough to permit analogues of the symbols, axioms, and rules of inference of formal arithmetic will also satisfy our conditions. Zermelo-Fraenkel set theory together with the Axiom of Choice (which we refer to as ZFC) is such a formal system. An example of an axiom of ZFC is the expression

$$\forall z((z \in x \rightarrow z \in y) \& (z \in y \rightarrow z \in x)) \rightarrow x = y,$$

which corresponds, under the usual interpretation of the symbols, to the assertion in English that two sets having exactly the same elements are equal. (The axioms of ZFC, which can be formulated so as to satisfy our conditions, are in [**11**]; see also [**1**], page 330.) The system ZFC is the most widely used formal axiomatic system today, largely because nearly every mathematical conjecture can be posed as a statement about sets and phrased in the language of ZFC and virtually every acceptable mathematical technique has an analogue in ZFC.

There is widespread agreement that the axioms of ZFC are true under the usual interpretation of the symbols and that the rules of inference preserve truth. For this reason, in the remainder of the article we shall assume that ZFC is consistent. Let us say that a formal statement A in ZFC is **undecidable** in ZFC if neither A nor the negation of A is a theorem in ZFC. Since ZFC satisfies Conditions 1 through 5, Gödel's Theorem tells us that there is a formal statement which is undecidable in ZFC. The fact that undecidable formal statements occur in ZFC might not be troublesome if such expressions, when stated in English using the usual interpretation of the symbols of ZFC, became assertions which were outside the mainstream of mathematics or otherwise of little mathematical interest. However, quite a few undecidable assertions are being found which are very natural and simple to state, and which would have seemed amenable to accepted mathematical techniques and sufficient cleverness.

Here are four examples of such undecidable assertions. If you are unfamiliar with the terms used in these examples, you need not be concerned. Just keep in mind the fact that examples (2) and (3) are standard concepts of general topology and example (4) uses standard concepts of homological algebra.

(1) *Any infinite set of real numbers can be placed into one to one correspondence with either the set of natural numbers or the set of all real numbers.* (See [**5**].)
This, of course, is the Continuum Hypothesis of Cantor. There are many conjectures which, like the Continuum Hypothesis, clearly involve more than one level of infinity (perhaps with extra topological, algebraic, or analytical conditions included) and which are known to be undecidable in ZFC. (For some representative examples, see [**2**] and [**4**], [**17**] and [**13**].) Our other three examples are not of this type.

(2) *Every countably compact, perfectly normal, T_1 topological space is compact.* (See [**9**], [**15**], and [**16**].)

(3) *Every compact Hausdorff perfectly normal topological space is separable.* (See [**12**].)

(4) *If A is an abelian group, Z is the group of integers, and* $\operatorname{Ext}(A,Z)=0$, *then A is a free group.* (See [**7**].)

Although Gödel's Theorem has been known for fifty years, these particular assertions were shown to be undecidable by a variety of relatively new techniques. (See the references for further details.)

In the past, many mathematicians may have felt that the limitation revealed by Gödel's Theorem would not affect their own particular field of research. Now that interesting conjectures in several different fields have been shown to correspond to undecidable assertions in ZFC, it is no longer as easy to maintain this point of view. Of course, Gödel's Theorem does not rule out the possibility of a better formal system than ZFC. For example, a statement could eventually be regarded as true, which, when added as an axiom to ZFC, would create a formal system in which some or all of conjectures (1) through (4) can be settled. It is not clear, however, that there will ever be such a formal system. Even if there were, because of Gödel's Theorem, we know that such a formal system would still have "undecidables" of its own.

Work on this paper was partially supported by a 1978 Faculty Summer Research Fellowship from the University of Richmond.

References

[1] J. Barwise, Handbook of Mathematical Logic, North Holland, New York, 1977.
[2] J. E. Baumgartner and K. Prikry, Singular cardinals and the generalized continuum hypothesis, Amer. Math. Monthly, 84 (1977) 108–113.
[3] A. Charlesworth, Infinite loops in computer programs, this MAGAZINE, 52 (1979) 284–291.
[4] A. T. Charlesworth, R. E. Hodel, and F. D. Tall, On a theorem of Jones and Heath concerning separable normal spaces, Colloq. Math., 34 (1975) 33–37.
[5] P. J. Cohen, Set Theory and the Continuum Hypothesis, Benjamin, New York, 1966.
[6] M. Davis, The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems, and Computable Functions, Raven Press, Hewlett, New York, 1965. (Contains a translation of Gödel's original proof, together with related papers of Gödel, Church, Turing, and others.)
[7] P. C. Eklof, Whitehead's problem is undecidable, Amer. Math. Monthly, 83 (1976) 775–788.
[8] S. C. Kleene, Introduction to Metamathematics, American Elsevier, New York, 1971.
[9] A. J. Ostaszewski, On countable compact, perfectly normal spaces, J. London Math. Soc., 14 (1976) 505–516.
[10] C. Reid, Hilbert, Springer-Verlag, New York, 1970.
[11] J. R. Shoenfield, Mathematical Logic, Addison-Wesley, Reading, 1967.
[12] F. D. Tall, Set-theoretic Consistency Results and Topological Theorems Concerning the Normal Moore Space Conjecture and Related Problems, Thesis, University of Wisconsin, Madison, 1969.
[13] F. D. Tall, How separable is a space? That depends on your set theory!, Proc. Amer. Math. Soc., 46, no. 2 (1974) 310–314.
[14] A. M. Turing, On computable numbers, with an application to the Entscheidungsproblem, Proc. London Math. Soc., 2nd series, 42 (1936–7) 230–265; a correction appears in 43 (1937) 544–546.
[15] W. Weiss, Countably compact spaces and Martin's axiom, Canad. J. Math., 30, no. 2 (1978) 243–249.
[16] W. A. R. Weiss, Countably compact, perfectly normal spaces may or may not be compact, Notices Amer. Math. Soc., 22, no. 2 (1975), Abstract 75T-G32, p. A-334.
[17] A. Wilansky, How separable is a space?, Amer. Math. Monthly, 79 (1972) 764–765.

NOTES

Magic Circles

DAVID A. JAMES
University of Michigan, Dearborn
Dearborn, MI 48128

In this paper, we will gather together fragments of information about magic circles which have already appeared, give and justify a streamlined algorithm for producing magic circles, and outline the rules for a game which we call Galois Field Hockey. The use of Galois fields in producing magic circles provides a concrete application of Galois theory, and serves as a pleasant contrast to the customary abstract classroom applications.

On the circle FIGURE 1 we have placed the tag 1 on the segment directed clockwise from node A to node B, 2 from C to D, 3 from B to C. From A to C the circular sum is $1+3=4$, from B to D the circular sum is $3+2=5$, from A to D is 6, from D to E is 7, and so forth. In fact, each positive integer appears once and only once as a (clockwise) sum of one or more consecutive tags; for instance 118 is found by starting at D and circling three and a half times. Such arrangements, reminiscent of graceful graphs, are called **magic circles**. A related form of this idea first appeared in print in 1857 [**3**].

Recent irresponsible rumor has it that the original magic circle dates back to Evariste Galois (1811–1832). In a letter kept secret by the French Army for the past century and a half, the rumor goes, the politically radical young Galois issued a challenge to a platoon of 18 soldiers stationed near his school in Paris to perform a particular field maneuver. This consisted of positioning themselves around the edge of a certain pond, having circumference 307 paces, in such a way that no two soldiers were the same number of paces from each other as were any other two. This amounts to finding a magic circle with 18 nodes (soldiers), where the tags represent the number of paces between consecutive soldiers. After several months the supercilious Galois approached the members of the platoon, now exhausted by their incessant drilling on this original Galois field, and announced that even if they tried a new formation each second nonstop for a trillion years, they still would have essentially no chance of success. It may have been this, and not jealousy over a woman, which provoked a soldier to challenge him to the unfortunate duel.

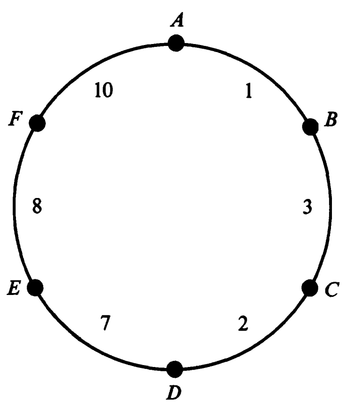


FIGURE 1

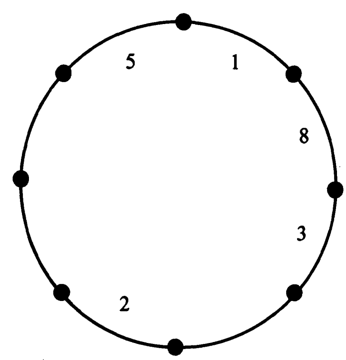


FIGURE 2

This rumor suggests the game of Galois Field Hockey, the object of which is to find, given an integer N, a magic circle with N nodes. This circle would have tags between the N nodes labelled in such a manner that each positive integer appears once and only once as a (clockwise) circular sum of one or more consecutive tags.

A more competitive form of Galois Field Hockey can be created for more than one player. For instance, for two, three, or four players, an integer N between 6 and 20 is selected, and a circle is drawn having N nodes. Players take turns attaching an integer "tag" to any empty space between two adjacent nodes, subject to two rules: first, no number can appear more than once as either a tag or as a sum of any string of adjacent tags; and second, the first tag played must be 1, and thereafter the value of any tag played cannot exceed twice the number of tags already played. When a player believes he cannot move, he drops from the competition, and the next person plays if he is able. The object is to make the last move. Thus in FIGURE 2, the next tag played cannot be $5,1,8,3,2,5+1=6,\ 1+8=9,\ 8+3=11,\ 5+1+8=14,\ 1+8+3=12$, or $5+1+8+3=17$, nor can it induce any circular sum having any of these values. Further, the tag must not exceed 10, since five tags are present. The only play possible is a 10 by the 5, and this play wins the game since no further moves are possible. Of course if the previous player had played, instead of the 8, a 4 to the left of the 5, the game would have ended at that point.

Information relating to magic circles is limited to side comments in a few papers. The most important occur in the 1938 article by James Singer [**6**] in which he proved the existence of **perfect difference sets**: *If $N=p^s+1$ (a power of a prime plus one), then there exist integers $k_1,k_2,\ldots,k_N$ such that the $N(N-1)$ differences k_i-k_j with $i\neq j$, fill all the nonzero congruence classes modulo* $N(N-1)+1$. As one of several applications, Singer distributed these integers, one to a node, clockwise around a circle, and gave the value $k_{i+1}-k_i$ to the tag between the node bearing k_i and the node bearing k_{i+1}. (The last tag, from k_N to k_1, was k_1-k_N.) There are $N(N-1)$ circular sums of tags having less than a full revolution, and they fill, by his theorem, all the nonzero congruence classes modulo $N(N-1)+1$.

Much later Leech [**4**], following methods of Rényi and Rédei [**5**], was searching for minimal difference bases, and as a minor step reduced the k_i modulo $N(N-1)+1$ and then arranged them in ascending order. In recent work by Golay [**1**] on the same problem, it is implicit that if in addition to Singer's and Leech's ideas, the tag from the node bearing k_N to the neighboring node bearing k_1 is taken as $k_1-k_N+N(N-1)+1$ instead of k_1-k_N, then the circular sums in Singer's circles precisely exhaust $1,2,\ldots,N(N-1)$ instead of merely being congruent to this set, and that the full sum of all N tags is $N(N-1)+1$. Golay realized that each positive integer would appear once and only once as a (clockwise) circular sum of consecutive tags, and coined the term "magic circle."

We shall construct a magic circle with 18 nodes, and in the process illustrate Singer's general algorithm as simplified by Halberstam and Laxton [**2**, pp. 80–84], here further streamlined to fit our purpose. If $N=p^s+1$, we use the Galois fields $GF(p^s)$ and $GF(p^{3s})$, that is, the unique fields having respectively p^s and p^{3s} elements. $GF^*(p^s)$ and $GF^*(p^{3s})$ denote the cyclic multiplicative groups which result when zero is removed. In concrete terms, GF(17) is the field of integers modulo 17, and lies in $GF(17^3)$, the field of polynomials $a+bx+cx^2$ where a,b,c range from 0 through 16. Sums and products of such polynomials are carried out as usual, but then powers of x greater than two are reduced by replacing x^3 by 3, and as the last step all coefficients are reduced modulo 17. This produces a field because x^3-3 is irreducible over the integers modulo 17, as is easily verified by checking $x^3\not\equiv 3$ mod 17 for $x=0,1,\ldots,16$. The number of elements in $GF^*(17^3)$ is $17^3-1=2^4(307)$, so any element θ having this order must be a generator. It is an easy computer task to check that with our polynomial multiplication, $(1+x)^k\neq 1$ for $k=2^4 307/2$ and $k=2^4 307/307$, and so $1+x$ generates $GF^*(17^3)$. Next we have the computer calculate a_k,b_k,c_k in the formula $(1+x)^k-(1+x)=a_k+b_kx+c_kx^2$, and print out the integers k for which $b_k\equiv c_k\equiv 0 \bmod 17$, that is, the k such that $\theta^k-\theta$ is in GF(17) where $\theta=1+x$. For $1<k<17^3-1$, there are clearly 17 such k, namely, the k for which θ^k is in the coset $\theta+GF(17)$. These, together with 0, serve as the 18 integers in Singer's theorem.

N	Consecutive tags of a magic circle with $N = p^s + 1$ nodes, p prime, $N \leqslant 48$
3	(S) 1,2,4
4	(S) 1,2,6,4
5	(S) 1,3,10,2,5
6	(S) 1,2,5,4,6,13
8	(S) 1,2,10,19,4,7,9,5
9	(S) 1,2,4,8,16,5,18,9,10
10	(S) 1,2,6,18,22,7,5,16,4,10
12	(S) 1,2,9,8,14,4,43,7,6,10,5,24
14	(S) 1,2,13,7,5,14,34,6,4,33,18,17,21,8
17	(S) 1,2,4,8,16,32,27,26,11,9,45,13,10,29,5,17,18
18	1,37,11,24,26,16,29,23,18,33,6,8,13,4,3,2,10,43
20	1,18,25,16,14,15,36,2,33,7,6,56,10,24,61,26,11,12,5,3
24	1,11,23,3,4,74,27,31,2,13,75,5,19,45,25,28,39,43,14,6,16,32,8,9
26	(G) 1,32,50,21,6,14,39,22,15,3,46,2,7,56,4,25,13,30,44,10,16,8,11,12,5,159
28	(G) 1,4,17,53,35,8,16,28,6,14,13,71,18,19,23,7,32,172,3,12,26,25,29,2,9, 36,10,68
30	(G) 1,40,19,47,15,3,33,10,7,20,12,191,9,26,22,71,21,2,29,25,13,63,11,44, 28,6,8,16,75,4
32	(G) 1,4,30,3,44,51,9,23,48,14,53,20,36,16,54,46,25,49,17,10,12,209,7,8,11, 2,29,55,6,18,40
33	(G) 1,20,3,55,4,80,12,198,22,16,9,44,17,15,13,5,88,27,7,30,26,10,29,2,40, 6,99,11,8,35,14,60,51
38	1,36,26,142,69,15,13,12,2,5,42,52,6,16,29,4,17,64,71,23,8,3,56,20,113,10, 70,57,117,24,30,35,43,18,68,9,44,38
42	1,179,29,2,8,55,30,22,49,27,13,152,7,66,4,11,9,23,25,12,67,35,16,3,42,14, 44,6,28,142,74,38,91,83,33,53,46,36,5,21,105,17
44	1,4,38,104,17,51,69,24,48,19,64,18,2,78,90,15,13,60,6,23,11,22,3,27,47, 127,123,126,32,76,21,14,95,39,85,55,16,10,31,75,61,9,37,7
48	1,30,48,4,3,38,15,14,25,10,8,133,100,16,20,98,22,5,19,42,23,127,21,26,33, 17,87,81,34,112,9,28,40,51,63,12,32,62,73,175,6,90,103,101,58,11,2,129
	S = Singer, G = Golay

TABLE 1

Reducing, rearranging, and tagging, as described in the previous paragraph, transforms this Singer circle into the magic circle listed with 18 nodes in TABLE 1.

To justify this algorithm, we first recall the well-known result that if θ generates any finite cyclic group, say $GF^*(p^{3s})$, then any subgroup, say $GF^*(p^s)$, consists of $1, \theta^i, \theta^{2i}, \theta^{3i}, \ldots, \theta^{(n-1)i}$ where $\theta^{ni} = 1$ and i is the smallest positive integer such that θ^i is in that subgroup. Here ni is the number of elements in the larger group, and n is the number in the smaller. Since $GF^*(p^{3s})$ has $p^{3s} - 1$ elements, and $GF^*(p^s)$ has $p^s - 1$ elements, our i is $(p^{3s} - 1)/(p^s - 1)$. Thus θ^j is in $GF^*(17)$ if and only if $j \equiv 0 \mod (p^{3s} - 1)/(p^s - 1)$. In our case, $i = 307$, so θ^j is in $GF^*(17)$ if and only if $j \equiv 0 \mod 307$.

Suppose the algorithm does not produce a Singer circle. Then there are four distinct k's, denoted k, l, m, n, with $k - l \equiv m - n \mod 307$. Since $k - l - m + n \equiv 0 \mod 307$, it follows $\theta^{k-l-m+n}$ is some element h in $GF^*(17)$, thus $\theta^k \theta^n = h\theta^l \theta^m$. The method of choosing the k's implies that there exist a, b, c, d in $GF(17)$ such that $\theta^k - \theta = a$, $\theta^l - \theta = b$, $\theta^m - \theta = c$, and $\theta^n - \theta = d$, so that $(\theta + a)(\theta + d) = h(\theta + b)(\theta + c)$. Since the decomposition into monic polynomials is unique over a given field, $h = 1$ and $a = b$ or $a = c$. Thus θ^k equals θ^l or θ^m, which is impossible since all powers of θ under consideration are distinct, because θ is a generator.

For any prime p and integer s, a magic circle with $N = p^s + 1$ nodes can be produced by this method. Of course, the generator θ and the values for k will vary.

We close with a series of remarks. If a circle is magic in one direction, it will also be magic in the other, so the use of "clockwise" throughout the paper could be replaced by "counterclockwise."

Singer has conjectured, and Golay claims that if $N = p^s + 1$ with p prime, there are $\phi(N(N-1)+1)/3s$ magic circles having N nodes, where $\phi(J)$, the Euler phi-function, denotes the number of positive integers less than J which are relatively prime to J. Almost nothing is known about the existence of magic circles if N cannot be expressed as $p^s + 1$ with p prime. I have verified that no magic circle exists with 7 nodes. Finally, I must admit that the pond-in-the-middle-of-the-original-Galois-field rumor is, alas, totally without basis.

References

[1] M. Golay, Notes on the representation of $1,2,\ldots,n$ by differences, J. London Math. Soc., (2) 4 (1972) 729–734.

[2] H. Halberstam and K. Roth, Sequences, Oxford Univ. Press, London, 1966.

[3] Kirkman, On the perfect r-partitions of $r^2 - r + 1$, Trans. Historical Society of Lancashire and Cheshire, 9 (1857) 127–142.

[4] J. Leech, On the representation of $1,2,\ldots,n$ by differences, J. London Math. Soc., 31 (1956) 160–169.

[5] L. Rédei and A. Rényi, On the representation of $1,2,\ldots,N$ by differences, Mat. Shornik (Recueil Math.), 66 (NS 24) (1949) 385–389 (Russian).

[6] J. Singer, A theorem in finite projective geometry and some applications to Number Theory, Trans. Amer. Math. Soc., 43 (1938) 377–385.

An Exercise Involving Conditional Probability

MARVIN ORTEL

JOHN ROSSI

Department of Mathematics

University of Hawaii

Honolulu, HI 96822

In this note we present an exercise on conditional probability which is perhaps new and which we find useful in a first course based on Kolmogorov's axioms. Standard exercises on conditional probability lead to models involving directed trees and (single step) transition probabilities (the fractions associated to the edges). In this regard it is important to note that a directed tree is a special type of directed graph in which there is at most one sequence of edges connecting any two vertices. Because of this special feature, the transition probabilities completely determine the probability measure for the model. Moreover, in such models, the probability of any complete path is correctly computed merely by multiplying the transition probabilities along the path. A variety of these examples are found in [2] and, if the student formulates explicit models, these examples are instructive. However, because it is very easy to compute the probability of a single path, students often solve the exercises in a mechanical fashion with no reference to the concepts of probability. To correct this, our exercise leads to a directed graph which is *not* a tree. Because the graph is not a tree, the transition probabilities do not completely determine the measure and the procedure of multiplying transition probabilities to find the probability of a path is *not* generally valid. Examples of this type provide a useful contrast which underlines the concepts involved in the standard examples.

Our example, which we entitle "Mouse in a Maze," is described as follows. A mouse runs a maze between feeding stations to amuse the visitors to a laboratory. The following information is posted for the visitors:

Singer has conjectured, and Golay claims that if $N = p^s + 1$ with p prime, there are $\phi(N(N-1)+1)/3s$ magic circles having N nodes, where $\phi(J)$, the Euler phi-function, denotes the number of positive integers less than J which are relatively prime to J. Almost nothing is known about the existence of magic circles if N cannot be expressed as $p^s + 1$ with p prime. I have verified that no magic circle exists with 7 nodes. Finally, I must admit that the pond-in-the-middle-of-the-original-Galois-field rumor is, alas, totally without basis.

References

[1] M. Golay, Notes on the representation of $1,2,\ldots,n$ by differences, J. London Math. Soc., (2) 4 (1972) 729–734.

[2] H. Halberstam and K. Roth, Sequences, Oxford Univ. Press, London, 1966.

[3] Kirkman, On the perfect r-partitions of $r^2 - r + 1$, Trans. Historical Society of Lancashire and Cheshire, 9 (1857) 127–142.

[4] J. Leech, On the representation of $1,2,\ldots,n$ by differences, J. London Math. Soc., 31 (1956) 160–169.

[5] L. Rédei and A. Rényi, On the representation of $1,2,\ldots,N$ by differences, Mat. Shornik (Recueil Math.), 66 (NS 24) (1949) 385–389 (Russian).

[6] J. Singer, A theorem in finite projective geometry and some applications to Number Theory, Trans. Amer. Math. Soc., 43 (1938) 377–385.

An Exercise Involving Conditional Probability

Marvin Ortel

John Rossi

Department of Mathematics

University of Hawaii

Honolulu, HI 96822

In this note we present an exercise on conditional probability which is perhaps new and which we find useful in a first course based on Kolmogorov's axioms. Standard exercises on conditional probability lead to models involving directed trees and (single step) transition probabilities (the fractions associated to the edges). In this regard it is important to note that a directed tree is a special type of directed graph in which there is at most one sequence of edges connecting any two vertices. Because of this special feature, the transition probabilities completely determine the probability measure for the model. Moreover, in such models, the probability of any complete path is correctly computed merely by multiplying the transition probabilities along the path. A variety of these examples are found in [2] and, if the student formulates explicit models, these examples are instructive. However, because it is very easy to compute the probability of a single path, students often solve the exercises in a mechanical fashion with no reference to the concepts of probability. To correct this, our exercise leads to a directed graph which is *not* a tree. Because the graph is not a tree, the transition probabilities do not completely determine the measure and the procedure of multiplying transition probabilities to find the probability of a path is *not* generally valid. Examples of this type provide a useful contrast which underlines the concepts involved in the standard examples.

Our example, which we entitle "Mouse in a Maze," is described as follows. A mouse runs a maze between feeding stations to amuse the visitors to a laboratory. The following information is posted for the visitors:

Our mouse always feeds at station a first. Then he will feed either at station b or c, with probability $\frac{1}{2}$ each. If he feeds at station b, he will progress to station e_1 with probability $\frac{1}{3}$ or to station d with probability $\frac{2}{3}$. If he feeds at station c, our mouse will go to e_4 with probability $\frac{1}{3}$ or to d with probability $\frac{2}{3}$. After feeding at d, his next stop is either e_2 or e_3 with probability $\frac{1}{2}$ each. The mouse retires after feeding at any of the stations e_1, e_2, e_3, e_4.

In return for a donation of five dollars you will receive a three-year subscription to our journal if the mouse runs path $abde_2$.

Since a journal subscription actually costs the lab forty dollars, explain how the lab makes a profit.

The given data is easily summarized on the directed graph in FIGURE 1. If a student proceeds without formulating a model, the rote procedure is to multiply the transition probabilities along the path $abde_2$ and conclude $P\{abde_2\} = \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{2} = \frac{1}{6}$. This implies expected winnings of $40/6$ dollars, which is more than the donation. At this point the student is quite receptive to a complete analysis.

We specify the probability space $\langle \Omega, P \rangle$. The sample space Ω is the set of all complete (nonextendable) directed paths through the graph in FIGURE 1. Explicitly

$$\Omega = \{abe_1, abde_2, abde_3, acde_2, acde_3, ace_4\}.$$

We use the corresponding capital letter to denote the event comprising all such complete paths through a fixed small letter. For example, $A = \{\text{paths through } a\} = \Omega$, $B = \{\text{paths through } b\} = \{abe_1, abde_2, abde_3\}$, $E_2 = \{abde_2, acde_2\}$. Every statement in the text of the problem now has an explicit interpretation as a property of the probability measure P. Thus, in the model we assume the conditional probabilities shown in TABLE 1.

$P(A) = 1$	
$P(B\mid A) = P(C\mid A) = \frac{1}{2}$	$P(D\mid B) = P(D\mid C) = \frac{2}{3}$
$P(E_2\mid D) = P(E_3\mid D) = \frac{1}{2}$	$P(E_1\mid B) = P(E_4\mid C) = \frac{1}{3}$

TABLE 1

Finally, for the prize, we introduce a random variable W given by $W(abde_2) = 40$, $W(\omega) = 0$ for all other ω in Ω.

Now we attempt to compute $P\{abde_2\}$ by means of the multiplication rule (see **[1]**, **[2]** for details). Thus, we note $\{abde_2\} = A \cap B \cap D \cap E_2$ and hence

$$P\{abde_2\} = P(A)P(B\mid A)P(D\mid A \cap B)P(E_2\mid A \cap B \cap D).$$

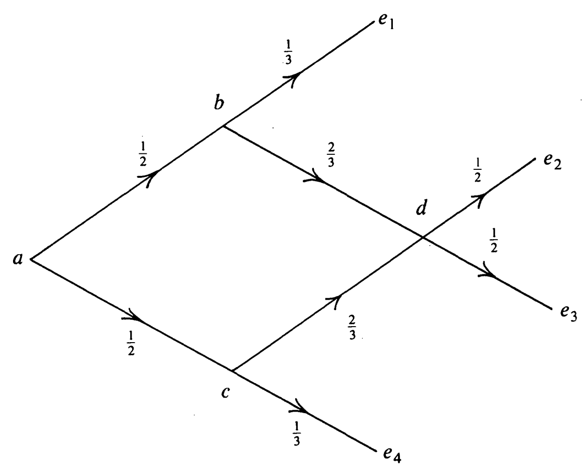


FIGURE 1

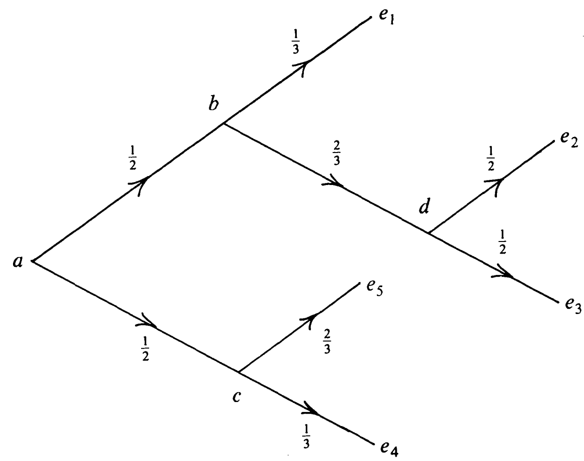


FIGURE 2

We have $P(A)=1$, $P(B|A)=\frac{1}{2}$ given. Also $A \cap B = B$ so $P(D|A \cap B)=P(D|B)=\frac{2}{3}$.

It remains to compute $P(E_3|A \cap B \cap D)$ and it is at this point that our exercise differs from similar exercises involving a tree. To make the comparison, consider the *tree* in FIGURE 2 for the moment. There is only one sequence of edges connecting a to d in FIGURE 2 and consequently, with the same notation, we have $A \cap B \cap D = D = \{abde_2, abde_3\}$. So, in an exercise based on the tree in FIGURE 2, $P(E_2|A \cap B \cap D) = P(E_2|D) = \frac{1}{2}$. But, in our example (FIGURE 1) there are *two* sequences of edges connecting a to d and we see $A \cap B \cap D = \{abde_2, abde_3\}$ while $D = \{abde_2, abde_3, acde_2, acde_3\}$. Therefore, since $A \cap B \cap D \neq D$, the information given by the laboratory does *not* specify $P(E|A \cap B \cap D)$ in any manner.

To see that $P\{abde_2\}$ cannot be computed from the information supplied by the laboratory, note that as long as $0 \leqslant p \leqslant \frac{1}{3}$, the assignment of probability to the six outcomes shown in TABLE 2 is consistent with the information in TABLE 1.

ω	abe_1	$abde_2$	$abde_3$	$acde_2$	$acde_3$	ace_4
$P\{\omega\}$	$\frac{1}{6}$	p	$\frac{1}{3}-p$	$\frac{1}{3}-p$	p	$\frac{1}{6}$

TABLE 2

The consistency is verified directly. For example, TABLE 2 implies

$$P(B \cap D) = P\{abde_2, abde_3\} = \tfrac{1}{3} \text{ and } P(B) = \tfrac{1}{2}. \text{ Hence } P(D|B) = \frac{P(B \cap D)}{P(B)} = \tfrac{2}{3}$$

as in TABLE 1.

This shows that on the basis of the information given by the lab, all that can be concluded is

$$E(W) = 40p \leqslant \frac{40}{3}.$$

But if the lab is to make a profit, it must be that $40p < 5$ or $p < \frac{1}{8}$!

Similar examples may be constructed on any graph in which there are several paths between two vertices. It is often difficult, but always important, to clarify the connection between the abstractions of the theory of probability and the literal expression of a specific problem. The ability to recognize ill-posed problems, or lack of definitive information, is at least as important as the ability to apply standard methods to solve straightforward problems. The paradoxical nature of examples such as the above alerts the students to the necessity of careful analysis of a problem before attempting its solution. Finally, after such an example, the student may be more receptive to a discussion of the connections between probability theory and the statement and solution of problems.

References

[1] W. Feller, An Introduction to Probability Theory and its Applications, vol. 1, 1st ed., chapter 5, Wiley, New York, 1950.

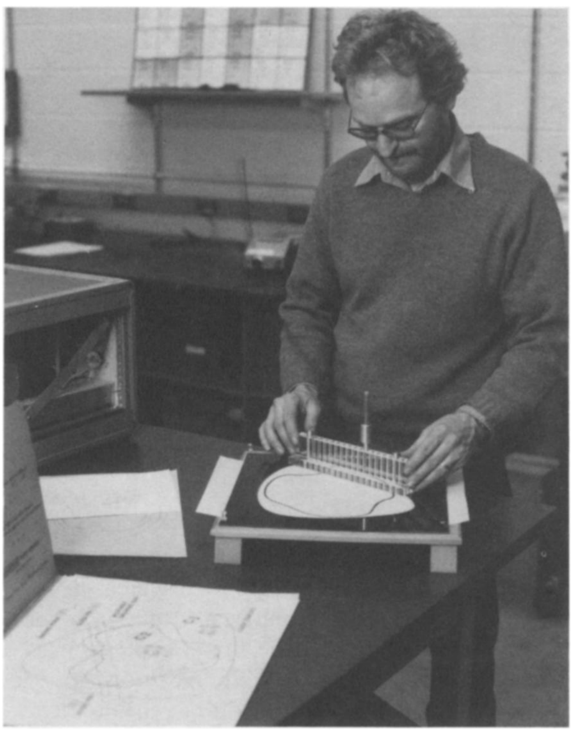
[2] M. Nosal, Basic Probability and Applications, chapters 8–9, Saunders, Philadelphia, 1977.

The Holditch Curve Tracer

WILLIAM BENDER
Western Washington University
Bellingham, WA 98225

In the past, certain geometric theorems have produced ingenious and useful devices to draw lines, to produce specialized curves, or to produce figures in different scale. For example, Peaucellier's cell is an instrument invented in 1864 for the purpose of drawing the inverse of any locus (a different device was invented by H. Hart in 1874 for the same purpose) [**1**]. Holditch's theorem provides another example (perhaps the most recent) of a geometrical theorem which led to the invention of a mechanical device: the Holditch Curve Tracer.

In the summer of 1979, Arne Broman, Professor of Mathematics at Chalmers University of Technology, at Göteborg, Sweden, was a visiting professor at Western Washington University. In



Dan M. Pomeroy operates the Holditch drawing-bar. The guide template is the same as in FIGURE 1; curve c_1 is being drawn.

References

[1] W. Feller, An Introduction to Probability Theory and its Applications, vol. 1, 1st ed., chapter 5, Wiley, New York, 1950.

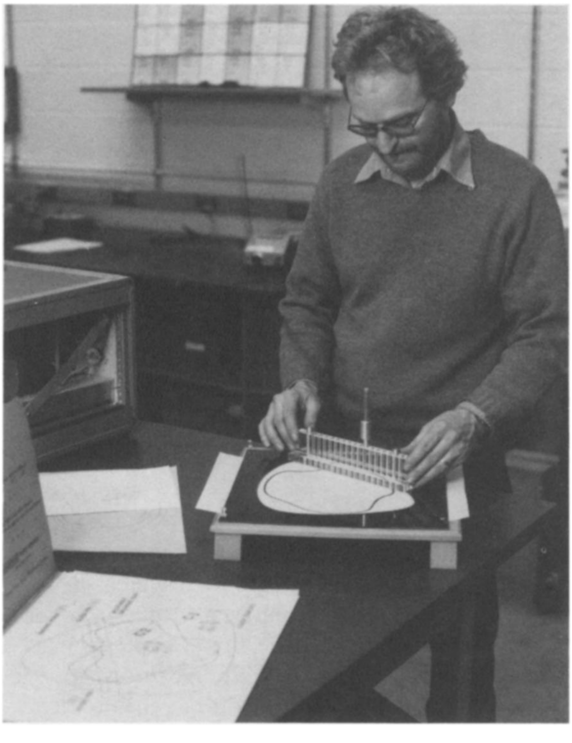
[2] M. Nosal, Basic Probability and Applications, chapters 8–9, Saunders, Philadelphia, 1977.

The Holditch Curve Tracer

WILLIAM BENDER
Western Washington University
Bellingham, WA 98225

In the past, certain geometric theorems have produced ingenious and useful devices to draw lines, to produce specialized curves, or to produce figures in different scale. For example, Peaucellier's cell is an instrument invented in 1864 for the purpose of drawing the inverse of any locus (a different device was invented by H. Hart in 1874 for the same purpose) [1]. Holditch's theorem provides another example (perhaps the most recent) of a geometrical theorem which led to the invention of a mechanical device: the Holditch Curve Tracer.

In the summer of 1979, Arne Broman, Professor of Mathematics at Chalmers University of Technology, at Göteborg, Sweden, was a visiting professor at Western Washington University. In



Dan M. Pomeroy operates the Holditch drawing-bar. The guide template is the same as in FIGURE 1; curve c_1 is being drawn.

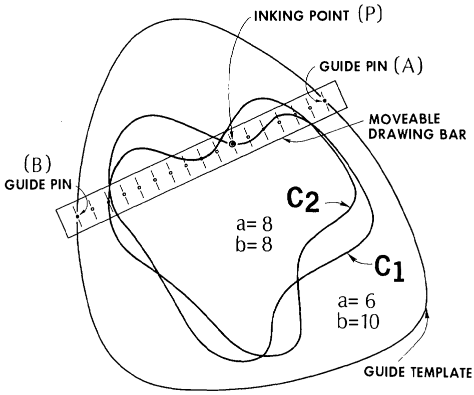


FIGURE 1

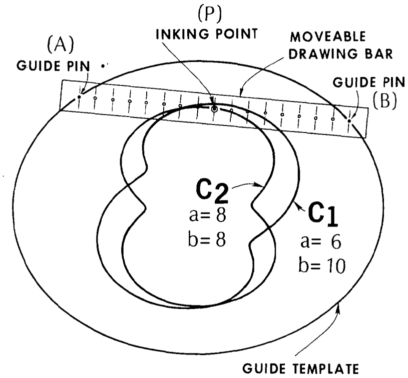


FIGURE 2

a conversation with me, Professor Broman told me of his research which had to do with perfecting Holditch's theorem. Some weeks later it occurred to me that a mechanical device could be made to draw the companion curve to an arbitrary convex curve lying in a plane. I felt challenged to turn a theoretical description into a workable device. Dan Pomeroy, in charge of our machine shop, and John Turner, in charge of our electronics shop (in the Physics and Astronomy Department), have been partners with me in the development of what I have named the *Holditch Curve Tracer*. Mr. Pomeroy's ingenuity in design details has made a smooth working device, purely mechanical in nature. (We are developing an electronic Holditch Curve Tracer.)

The Holditch Curve Tracer works as follows. Begin with an arbitrary closed convex curve c, and cut out a "guide template" (from stiff cardboard or acrylic sheet) with c as its boundary. The movable drawing bar of the tracer has a series of equally spaced holes; two of these are chosen as A, B so that AB is a chord of c. Brass guide pins are dropped through the two holes A and B; these guide pins enable the drawing bar to roll smoothly along c, always keeping the guide pins in contact with c. A pencil or inking point (such as used for computer graphics drawings) is inserted into another hole on the drawing bar designated as point P; it is also fitted snugly and weighted properly in order to draw a smooth curve (the Holditch companion curve). If we denote the distance from A to P as a and the distance from B to P as b, then Holditch's theorem says that the area between c and the companion curve traced by P is πab.

The transparent drawing bar of our instrument has 17 equi-spaced holes (the distance between holes is a half inch; the holes are one-fourth inch in diameter). Any two holes may be chosen for A and B, and any hole between the chosen ones may be designated as the location of P. In FIGURE 1 we show for a single arbitrary convex curve c two different Holditch companion curves produced when P is placed at different locations between A and B. Companion curve c_1 is traced by P when $AP = a = 6$ and $BP = b = 10$; companion curve c_2 is traced by P when $a = 8$, $b = 8$; five additional companion curves join these for the cover design. In FIGURE 2, the "guide curve" c is an ellipse, and companion curves c_1 and c_2 are shown for $a = 6$, $b = 10$, and $a = 8$, $b = 8$ respectively.

The Holditch Curve Tracer appears to have certain engineering uses in preliminary design. Additionally, the Holditch Curve Tracer could produce an interesting parlor game: given a convex closed curve c, pick your points (holes) A, B, P on the drawing bar, and guess the shape of the companion curve which will be traced by the inking point at P, as the drawing bar rolls around c. The participants are in for a surprise!

Reference

[1] H. S. M. Coxeter, Introduction to Geometry, 2nd ed., Wiley, 1968, p. 83.

Curves

A spiral
(curl)
has the sweep
and
swirl
of
a
nebula
out
in
space

A limaçon
(in the lexicon)
is
a
snail
in its carapace

A cardioid
(according to Freud)
is an
amorous
palpitater

A lemniscate
is a
figure
eight
cut
by
a
lemniskater

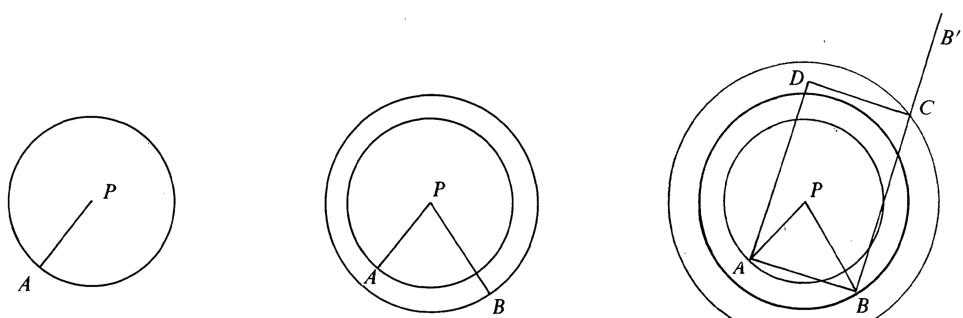
Katharine O'Brien

Exploring a Rectangle Problem

MARION WALTER
University of Oregon
Eugene, OR 97403

If P is in the interior of rectangle ABCD such that $|PA|=3$, $|PB|=4$, $|PC|=5$, *what is* $|PD|$? This problem [2] and variations of it have been considered many times before [5], so finding an answer is not our point. In exploring this problem, we would like to emphasize *how* we look at a problem, the surprises we discover, and how we are led to extend our investigation.

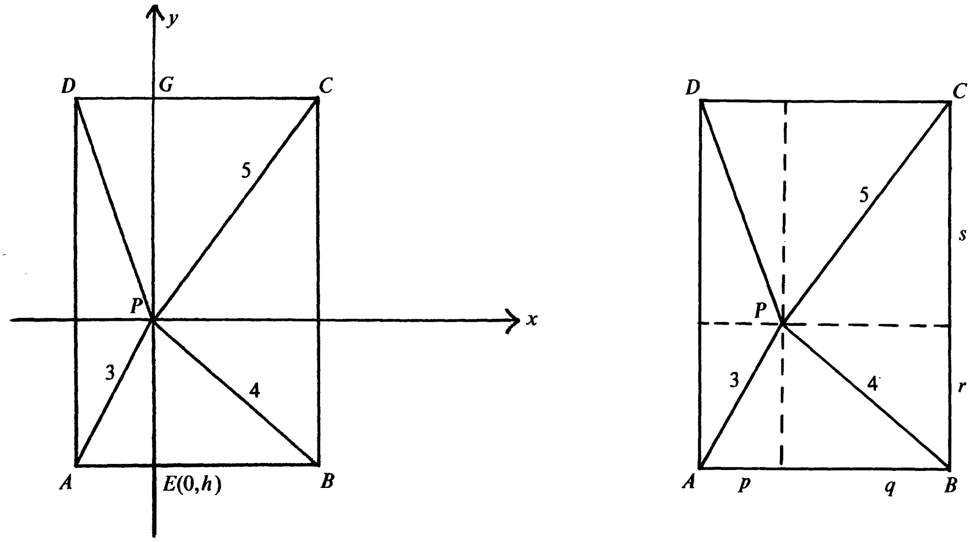
You might first ask: what is the size of $\measuredangle APC$? Is enough information given about the rectangle to solve the problem? Is the rectangle determined by the three segments PA, PB, PC? Suppose we investigate by placing the point A anywhere on a circle with center P and radius 3 (FIGURE 1(a)). Then B must lie on a circle with center P and radius 4 (FIGURE 1(b)). There are an infinite number of possible positions for B though not every position on the circle is allowable if P is to be inside the rectangle. To locate point C draw BB' perpendicular to AB. Since $PC=5$, draw the circle with center P and radius 5, and let it cut BB' at C (FIGURE 1(c)). Since 3 vertices determine a rectangle, $ABCD$ is now fixed. If P is not inside the rectangle, change the position of B until P falls inside. Since B can have an infinite number of positions, there are an infinite number of rectangles satisfying the condition $|PA|=3$, $|PB|=4$, $|PC|=5$. This is, for most people, the first of several surprises. It implies additionally that $\measuredangle APB$ and $\measuredangle APC$ are also not unique.



(a) (b) (c)

FIGURE 1

Let us now look at two solutions to the problem; no doubt you can find others. If you are inclined to use analytic geometry, choose coordinates so that $P=(0,0)$, the line through A, B is horizontal, and $E=(0,h)$ is the point where this line cuts the y-axis. The coordinates for A, B, C are $A=(\mp\sqrt{9-h^2},h)$, $B=(\pm\sqrt{16-h^2},h)$, $C=(\pm\sqrt{16-h^2},y)$. Using the notation in FIGURE 2(a) we calculate, using the Pythagorean Theorem, $|GP|^2=5^2-(16-h^2)$. But also, $|GP|^2=|PD|^2-(9-h^2)$. Combining the two equations, we see that $|PD|^2=18$, so $|PD|=3\sqrt{2}$. Thus $|PD|$ can be calculated easily, even though no information was provided about the sides of the rectangle; this, for many, is another surprise. A second solution (suggested by the editor) which avoids coordinates altogether is as follows. Draw lines through P perpendicular to the sides of the rectangle $ABCD$; one line cuts AB into segments of length p and q, the other line cuts BC into segments of length r and s (FIGURE 2(b)). Using the Pythagorean Theorem (again!) we see that $3^2=p^2+r^2$, $5^2=s^2+q^2$, $4^2=q^2+r^2$, and $|PD|^2=p^2+s^2$. Since the sum of the first two equations equals the sum of the last two equations, we have $3^2+5^2=4^2+|PD|^2$, so as before, $|PD|=3\sqrt{2}$.



(a) (b)

FIGURE 2

If we think about linkages and rigidity in connection with this problem, then we can think of the three segments PA, PB, PC as rods joined at one point P and linked by elastic. The condition that must be met by A,B,C is that they be three vertices of a rectangle. One limiting position for P is when P falls on AD (see FIGURE 3). In this limiting case, it is very easy to calculate $|PD|$. $|AB|^2=4^2-3^2=7=|DC|^2$, $|PD|^2=5^2-7=18$, so $|PD|=3\sqrt{2}$. Here, even though P is not in the interior of the rectangle, the length $|PD|$ remains the same.

What happens if P is outside the rectangle? Rather than answer this question immediately, we raise another question about the assumptions in the original problem. The lengths $3,4,5$ for the segments PA, PB, PC made computation of the solution to the problem easy—but does the solution depend on these lengths? Suppose we let $|PA|=a$, $|PB|=b$, $|PC|=c$, and also remove the assumption that P is in the interior of the rectangle $ABCD$. Following the first method of solution, we again choose $P=(0,0)$, and let $E=(0,h)$ be the point where the horizontal line through A,B cuts the y-axis (FIGURE 4(a)). Now the coordinates for A,B,C are $A=(\mp\sqrt{a^2-h^2},h)$, $B=(\pm\sqrt{b^2-h^2},h)$, $C=(\pm\sqrt{b^2-h^2},y)$. As before, we calculate $|GP|^2=c^2-(b^2-h^2)=|PD|^2-(a^2-h^2)$, so $|PD|^2+b^2=a^2+c^2$. This can be restated as

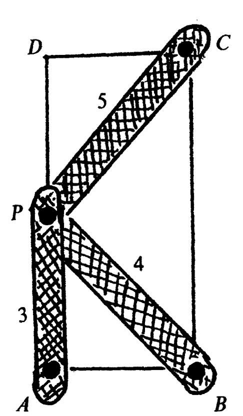
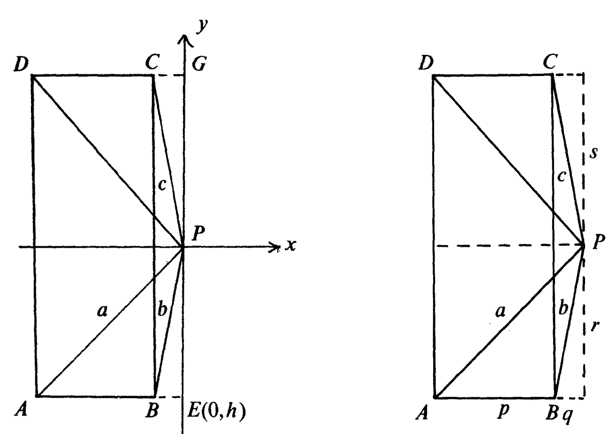


FIGURE 3



(a) (b)

FIGURE 4

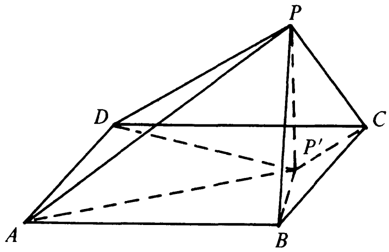


FIGURE 5

$$|PD|^2 + |PB|^2 = |PA|^2 + |PC|^2. \tag{1}$$

So the problem generalizes! There is an added bonus to this calculation—P need not be inside the rectangle at all. The same calculation holds if P is outside, inside, or on the boundary of the rectangle (provided that signs are taken care of). The alternate proof also remains valid, replacing the lengths 3,4,5 with a,b,c, respectively, in FIGURE 2(b). In this case, the sum of the two equations $a^2 = p^2 + r^2$, $c^2 = s^2 + q^2$ equals the sum of the two equations $b^2 = q^2 + r^2$, $|PD|^2 = p^2 + s^2$. This is exactly our result (1) above. A similar argument works even if P is on the boundary or outside of rectangle $ABCD$, (FIGURE 4(b)); the details are left to you.

The illustrations of rectangle $ABCD$ in FIGURES 2 and 4 with point P joined to the corners of the rectangle look like an aerial view of a pyramid with rectangular base. This suggests that we investigate to see if equation (1) remains true if P is not in the plane of the rectangle. FIGURE 5 makes the answer "yes" almost obvious. If P' is the projection of P onto the plane of the rectangle, then PP' is perpendicular to $P'A$, $P'B$, $P'C$, and $P'D$. Once again, the Pythagorean Theorem can be applied: $|P'P|^2 + |P'A|^2 = |PA|^2$, $|P'P|^2 + |P'C|^2 = |PC|^2$, $|PP'|^2 + |P'B|^2 = |PB|^2$, $|P'P|^2 + |P'D|^2 = |PD|^2$. The sum of the first two of these equations equals the sum of the last two (since we know that $|P'A|^2 + |P'C|^2 = |P'B|^2 + |P'D|^2$), thus equation (1) does remain true for this generalization of the problem.

When Ivan Niven looked at the original problem, he observed that one can study it by considering 4 points on the unit circle (FIGURE 6). Choose coordinates so that $P = (0,k)$ and let the 4 vertices A, A', B, B' of the rectangle be on the unit circle so that A and A', B and B' are diametrically placed. Thus $A = (x,y)$, $A' = (-x,-y)$ and $|PA|^2 + |PA'|^2 = x^2 + (y-k)^2 + x^2 + (k+y)^2 = 2(x^2 + y^2) + 2k^2 = 2 + 2k^2$. Similarly $|PB|^2 + |PB'|^2 = 2 + 2k^2$, so $|PA|^2 + |PA'|^2 = |PB|^2 + |PB'|^2$. When P lies on the circle ($k = 1$), $|PA|^2 + |PA'|^2 = |AA'|^2$. When P lies inside the circle ($k < 1$), $|PA|^2 + |PA'|^2 < |AA'|^2$. When P lies outside the circle ($k > 1$) and $|PA|^2 + |PA'|^2 > |AA'|^2$.

This way of considering the problem led Niven to suggest that the problem might generalize to a sphere. Consider the sphere $x^2 + y^2 + z^2 = 1$. Pick any point $P(0,0,k)$ inside the sphere. Choose any point A on the sphere together with its diametrically placed point A' (FIGURE 7). If

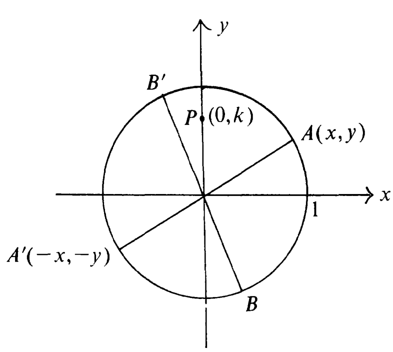


FIGURE 6

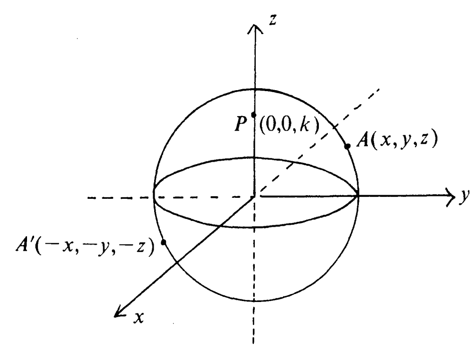


FIGURE 7

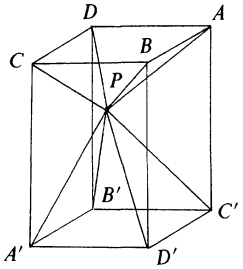


FIGURE 8

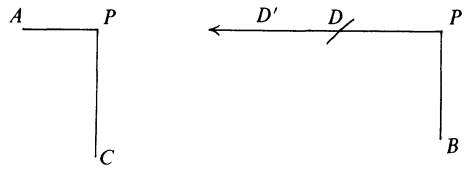


FIGURE 9

$A=(x,y,z)$ and $A'=(-x,-y,-z)$, then $|PA|^2+|PA'|^2=x^2+y^2+(z-k)^2+(-x)^2+(-y)^2+(-z-k)^2=2+2k^2$. Similarly, for any other diametrically opposite points B,B', it is true that $|PB|^2+|PB'|^2=2+2k^2$, so $|PA|^2+|PA'|^2=|PB|^2+|PB'|^2$. We can reinterpret this as a generalization of the original problem: we have a rectangular parallelopiped $ABCDA'B'C'D'$ (which can be inscribed in a sphere) with point P inside and $|PA|^2+|PA'|^2=|PB|^2+|PB'|^2=|PC|^2+|PC'|^2=|PD|^2+|PD'|^2$ (FIGURE 8). Again the calculation does not depend on whether P is inside the box or not.

Equation (1) naturally suggests other problems to investigate. Both equation (1) and the method of solution of the problems considered so far suggest the use of the Pythagorean Theorem to construct segment PD, rather than the technique outlined in FIGURE 1. (Compare this to one solution of the problem "Construct an equilateral triangle equal in area to the sum of two given equilateral triangles" discussed in [3].) If the three segments PA,PB,PC are given, PA and PC can be constructed as legs of a right triangle with hypotenuse AC. Draw (separately) PB with ray PD' perpendicular at P. Now construct PD on the ray PD' so that $|BD|=|AC|$. (See FIGURE 9.) The construction insures that equation (1) is satisfied. Another problem: what happens to equation (1) if we change the metric in $\mathbb{R}^2$? Consider a rectangle in "Taxi-Cab" Geometry for example. Here the distance between two points P,Q is the sum of the lengths of the legs of the right triangle which has PQ as hypotenuse (FIGURE 10). Now equation (1) is replaced by

$$|PD|+|PB|=|PA|+|PC|. \tag{2}$$

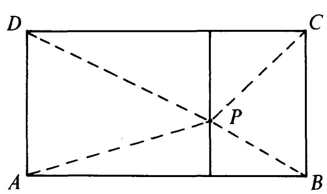


FIGURE 10

It is tempting to ask if there is a geometric analogue to the problem (or another metric on $\mathbb{R}^2$) where the result is $|PD|^{1/2}+|PB|^{1/2}=|PA|^{1/2}+|PC|^{1/2}$ or $|PD|^3+|PB|^3=|PA|^3+|PC|^3$.

There are many other alternatives and new questions that one can ask which can be generated by using the "What-If Not Techniques" described in [1], [3], and [4]. It is clear that there is more to a rectangle than meets the eye!

References

[1] S. Brown and M. Walter, What if not? An elaboration and second illustration, Math. Teaching, 51 (1970) 9–17.

[2] Alan Hoffer, Geometry, Addison-Wesley, 1979, p. 576, problem 45.

[3] M. Walter and S. Brown, Problem posing and problem solving: an illustration of their interdependence, Math. Teacher, vol. 70, no. 1, 4–14.

[4] M. Walter and S. Brown, What if not?, Math. Teaching, 46 (1969) 38–44.

[5] Solutions to Proposal 865, this MAGAZINE, 47(1974) 56–59.

Dice Handicap

B. R. Johnson
University of Victoria
Victoria, B.C., Canada

Consider the following two-player game. Player A rolls $n+1$ fair r-sided dice and scores the sum of the highest n. Player B rolls n fair r-sided dice and scores the sum. The higher score wins, with ties being awarded to Player B. Who is the favorite? Does the answer depend on n and/or r? As a function of n and r, what is the probability that Player A will win?

Motivation for these questions was provided by Proposal 1071 in the Problems Section of this MAGAZINE, submitted by Joseph Browne [**1**]. There, using standard 6-sided dice, the reader was asked first to show that when $n=2$, Player A is the favorite, and then he was asked to find the smallest value of n for which Player B is the favorite. The second part, which was starred, has an interesting answer. Player B is *never* the favorite! (*Editor's note*: Readers' solutions to Proposal 1071 can be found in [**3**], [**4**].)

In the general setting, things are not quite so bleak for Player B. In fact, for 2-sided dice Player B is *always* the favorite! He also has the upper hand when $n=1$ and $r=3$, but that is where his good fortune ends. For all other values of n and r, Player B is the underdog.

When n is large and r is at least 3, it is very laborious to compute the probability that Player A will win. However, we will derive an approximation formula based on Central Limit Theory which gives accurate approximations even when n is small (see TABLE 1 for comparison).

Consider fair r-faced dice with $r \geqslant 2$. Let

$$X_i = \text{number rolled by Player A on } i\text{th die, } 1 \leqslant i \leqslant n+1,$$

$$Y_i = \text{number rolled by Player B on } i\text{th die, } 1 \leqslant i \leqslant n.$$

Then $X_1, X_2, \ldots, X_{n+1}, Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed (i.i.d.) random variables, each having the discrete uniform distribution on $\{1,2,\ldots,r\}$. Set $S_j = X_1 + X_2 + \cdots + X_j$, $T_j = Y_1 + Y_2 + \cdots + Y_j$, $M_j = \min\{X_1, X_2, \ldots, X_j\}$. Then Player A wins if and only if $S_{n+1} - M_{n+1} > T_n$.

The following derivation will be useful for determining who is the favorite as a function of n and r and for obtaining an approximation formula for $P(S_{n+1} - M_{n+1} > T_n)$.

$$\begin{aligned} P(S_{n+1} - M_{n+1} > T_n) &= P(S_n > T_n) + \sum_{j=0}^{r-2} P(S_n = T_n - j, X_{n+1} - M_n > j) \\ &= P(S_n > T_n) + \sum_{j=0}^{r-2} \sum_{l=2}^{r-j} P(S_n = T_n - j, M_n < l, X_{n+1} = l+j) \\ &= P(S_n > T_n) + \sum_{j=0}^{r-2} \sum_{l=2}^{r-j} (1/r)(P(S_n = T_n - j) - P(S_n = T_n - j, M_n \geqslant l)), \end{aligned}$$

because $(S_n = T_n - j, M_n < l)$ and $(X_{n+1} = l+j)$ are independent events and $P(X_{n+1} = l+j) = 1/r$. But since $P(M_n \geqslant l) = P(X_1 \geqslant l, X_2 \geqslant l, \ldots, X_n \geqslant l) = ((r-l+1)/r)^n$, the derivation of a formula for $P(S_{n+1} - M_{n+1} > T_n)$ can be continued as follows.

$$\begin{aligned} P(S_{n+1} - M_{n+1} > T_n) = P(S_n > T_n) &+ \sum_{j=0}^{r-2} \frac{r-j-1}{r} P(S_n = T_n - j) \\ &- \sum_{j=0}^{r-2} \sum_{l=2}^{r-j} \frac{(r-l+1)^n}{r^{n+1}} P(S_n = T_n - j \mid M_n \geqslant l) \end{aligned}$$

$$= P(S_n > T_n) + \sum_{j=0}^{r-2} \frac{r-j-1}{r} P(S_n = T_n - j)$$

$$- \sum_{l=2}^{r} \frac{(r-l+1)^n}{r^{n+1}} \left(\sum_{j=0}^{r-l} P(S_n = T_n - j \mid M_n \geqslant l) \right)$$

$$= P(S_n > T_n) + \sum_{j=0}^{r-2} \frac{r-j-1}{r} P(S_n = T_n - j)$$

$$- \sum_{l=2}^{r} \frac{(r-l+1)^n}{r^{n+1}} P(-(r-l) \leqslant S_n - T_n \leqslant 0 \mid M_n \geqslant l). \tag{1}$$

Since $P(S_n > T_n) + \frac{1}{2} P(S_n = T_n) = 1/2$ and $P(S_n - T_n = 0 \mid M_n \geqslant r) = r^{-n}$, this formula can be rewritten as

$$\begin{aligned} P(S_{n+1} - M_{n+1} > T_n) = \frac{1}{2} + \frac{r-2}{2r} P(S_n - T_n = 0) + \sum_{j=1}^{r-2} \frac{r-j-1}{r} P(S_n - T_n = -j) \\ - r^{-(2n+1)} - \sum_{l=2}^{r-1} \frac{(r-l+1)^n}{r^{n+1}} P(-(r-l) \leqslant S_n - T_n \leqslant 0 \mid M_n \geqslant l), \end{aligned} \tag{2}$$

where it is understood that the two summations do not appear when $r = 2$.

From equation (2) it is clear that Player B is favored when $r = 2$ (coin tossing case) because then $P(\text{Player A wins}) = \frac{1}{2} - 2^{-(2n+1)}$. When $n = 1$, we obtain by direct calculation

$$\begin{aligned} P(S_2 - M_2 > T_1) &= \sum_{j=1}^{r-1} P(T_1 = j, \max\{X_1, X_2\} > j) \\ &= \sum_{j=1}^{r-1} P(T_1 = j)\left(1 - P(\max\{X_1, X_2\} \leqslant j)\right) \\ &= \sum_{j=1}^{r-1} (1/r)(1 - j^2/r^2) \\ &= \frac{2}{3} - \frac{1}{2r} - \frac{1}{6r^2}. \end{aligned} \tag{3}$$

Hence, $P(\text{Player A wins}) = 13/27 < 1/2$ in the case $n = 1$ and $r = 3$. As stated in the following theorem, Player A is favored in every other case.

THEOREM 1. *$P(S_{n+1} - M_{n+1} > T_n) > \frac{1}{2}$ if and only if $r \geqslant 4$, or $r = 3$ and $n \geqslant 2$.*

Proof: Since the "only if" part of the theorem has been established in the previous paragraph, it remains to show the "if" direction. For $n = 1$ the desired result follows immediately from (3).

Suppose $n \geqslant 2$ and $r \geqslant 3$. Since $P(S_n - T_n = -j \mid M_n \geqslant l) \leqslant P(S_n - T_n = -j)$ for nonnegative j, formula (1) yields

$$\begin{aligned} P(S_{n+1} - M_{n+1} > T_n) &\geqslant P(S_n > T_n) + \sum_{j=0}^{r-2} \left(\frac{r-j-1}{r} - \sum_{l=2}^{r-j} \frac{(r-l+1)^n}{r^{n+1}} \right) P(S_n - T_n = -j) \\ &= \frac{1}{2} + \left(\frac{r-2}{2r} - \sum_{l=2}^{r} \frac{(r-l+1)^n}{r^{n+1}} \right) P(S_n - T_n = 0) \\ &\quad + \sum_{j=1}^{r-2} \left(\frac{r-j-1}{r} - \sum_{l=2}^{r-j} \frac{(r-l+1)^n}{r^{n+1}} \right) P(S_n - T_n = -j). \end{aligned} \tag{4}$$

Also,

$$\begin{aligned}\frac{r-2}{2r}-\sum_{l=2}^{r}\frac{(r-l+1)^n}{r^{n+1}}&=\frac{r-2}{2r}-\frac{1}{r}\sum_{l=2}^{r}\left(\frac{r-l+1}{r}\right)^n\\&\geqslant\frac{r-2}{2r}-\frac{1}{r}\sum_{l=2}^{r}\left(\frac{r-l+1}{r}\right)^2\\&=\frac{r^2-3r-1}{6r^2},\end{aligned}$$

which implies

$$\frac{r-2}{2r}-\sum_{l=2}^{r}\frac{(r-l+1)^n}{r^{n+1}}>0 \text{ whenever } r\geqslant 4. \tag{5}$$

If $r=3$, then

$$\frac{r-2}{2r}-\sum_{l=2}^{r}\frac{(r-l+1)^n}{r^{n+1}}=\frac{1}{6}-\frac{1}{3}\left(\frac{2}{3}\right)^n-\left(\frac{1}{3}\right)^{n+1}>0, \text{ provided } n\geqslant 3. \tag{6}$$

Further, if $1\leqslant j\leqslant r-2$,

$$\begin{aligned}\frac{r-j-1}{r}-\sum_{l=2}^{r-j}\frac{(r-l+1)^n}{r^{n+1}}&=\frac{r-j-1}{r}-\frac{1}{r}\sum_{l=2}^{r-j}\left(\frac{r-l+1}{r}\right)^n\\&>\frac{r-j-1}{r}-\frac{1}{r}\sum_{l=2}^{r-j}\frac{r-l+1}{r}\\&=\frac{(r-j)(r-j-1)}{2r^2}>0.\end{aligned}$$

Combine this result with (4), (5), and (6) to obtain $P(S_{n+1}-M_{n+1}>T_n)>\frac{1}{2}$ when $r\geqslant 4$, or $r=3$ and $n\geqslant 3$. For the case $r=3$ and $n=2$, we see by direct calculation that $P(S_3-M_3>T_2)=134/243>1/2$.

All probabilities on the right-hand side of equation (2) converge to zero as $n\to\infty$. This is an immediate consequence of the Central Limit Theorem (also see formula (7) below). Therefore, the game is fair asymptotically, without restriction on r.

THEOREM 2. $\lim_{n\to\infty}P(S_{n+1}-M_{n+1}>T_n)=\frac{1}{2}$ *for every* $r\geqslant 2$.

The convergence to $\frac{1}{2}$ is rapid for $r=2$ (the rate is $2^{-(2n+1)}$), but slow for $r\geqslant 3$. See TABLE 1 for numerical results of the cases $r=3$ and $r=6$.

Since direct computation of $P(\text{Player A wins})=P(S_{n+1}-M_{n+1}>T_n)$ is unwieldy when n is large and $r\geqslant 3$, an accurate approximation formula is needed. The following result is based on a Local Central Limit Theorem for lattice distributions.

THEOREM 3. *As* $n\to\infty$, *the following expression has limit zero*:

$$\sqrt{n}\Bigg(P(S_{n+1}-M_{n+1}>T_n)-\frac{1}{2}+r^{-(2n+1)}$$
$$-\frac{r-2}{2r}\left(\frac{3}{\pi(r^2-1)}\right)^{1/2}-\sum_{j=1}^{r-2}\frac{r-j-1}{r}\left(\frac{3}{\pi(r^2-1)}\right)^{1/2}\exp(-3j^2/n(r^2-1))$$
$$+\sum_{l=2}^{r-1}\frac{1}{r}\left(\frac{r-l+1}{r}\right)^n\left(\frac{6}{\pi((r-l+1)^2+r^2-2)}\right)^{1/2}\sum_{k=l-r}^{0}\exp\left(-\frac{6(k-n(l-1)/2)^2}{n((r-l+1)^2+r^2-2)}\right)\Bigg).$$

Proof: For $1 \leqslant l \leqslant r-1$, the conditional distribution of $(X_1, X_2, \ldots, X_n)$, given $M_n \geqslant l$, has i.i.d. components which are uniformly distributed on $\{l, l+1, \ldots, r\}$. (Note that it is actually an unconditional distribution when $l = 1$.) Therefore, we have the following equations for expected value and variance

$$E(S_n - T_n | M_n \geqslant l) = nE(X_1 - Y_1 | M_n \geqslant l) = n(l-1)/2,$$

$$\operatorname{var}(S_n - T_n | M_n \geqslant l) = n \operatorname{var}(X_1 - Y_1 | M_n \geqslant l) = \frac{n}{12}\left((r-l+1)^2 + r^2 - 2\right).$$

By a Local Central Limit Theorem for lattice distributions (see **[2]**) formula (7) below has limit 0 as $n \to \infty$, uniformly in k:

$$\sqrt{n}\, P(S_n - T_n = k | M_n \geqslant l) - \left(\frac{6}{\pi\left((r-l+1)^2 + r^2 - 2\right)}\right)^{1/2} \exp\left(-\frac{6(k - n(l-1)/2)^2}{n\left((r-l+1)^2 + r^2 - 2\right)}\right). \quad (7)$$

Now apply this result to all probabilities on the right-hand side of equation (2) to obtain the desired result.

The normal approximation for the conditional distribution of $S_n - T_n = \Sigma_{i=1}^{n}(X_i - Y_i)$, given $M_n \geqslant l$, is appropriate even when n is small because the summands are i.i.d. random variables each having a mound-shaped distribution symmetric about $(l-1)/2$. Hence, the convergence to zero of (7) is rapid, and we obtain a very good approximation formula by setting to zero the expression in Theorem 3 and solving for $P(S_{n+1} - M_{n+1} > T_n)$. A comparison with the actual probabilities is given in TABLE 1 for the cases $n = 1,2,3,4,5$ and $r = 3,6$ (standard dice).

$r = 3$			Standard Dice ($r = 6$)	
P (Player A wins)	Approximation	n	P (Player A wins)	Approximation
.4815	.4918	1	.5787	.5985
.5514	.5562	2	.6188	.6250
.5652	.5681	3	.6222	.6258
.5673	.5692	4	.6195	.6220
.5659	.5672	5	.6154	.6172
	.5643	6		.6124
	.5612	7		.6078
	.5582	8		.6036
	.5555	9		.5997
	.5530	10		.5961
	.5508	11		.5928
	.5488	12		.5897
	.5470	13		.5869
	.5453	14		.5843
	.5439	15		.5819
	.5425	16		.5797
	.5413	17		.5776
	.5402	18		.5757
	.5391	19		.5739
	.5381	20		.5722
	.5243	50		.5464
	.5172	100		.5329
	.5077	500		.5148
	.5055	1,000		.5104
	.5017	10,000		.5033
	.5005	100,000		.5010
	.5002	1,000,000		.5003

TABLE 1

APPROXIMATION FORMULA. $P(\text{Player A wins}) \doteq \frac{1}{2} + \frac{r-2}{2r}\left(\frac{3}{n\pi(r^2-1)}\right)^{1/2} - r^{-(2n+1)}$

$$+\sum_{j=1}^{r-2} \frac{r-j-1}{r}\left(\frac{3}{n\pi(r^2-1)}\right)^{1/2} \exp\left(-3j^2/n(r^2-1)\right)$$

$$-\sum_{l=2}^{r-1} \frac{1}{r}\left(\frac{r-l+1}{r}\right)^n \left(\frac{6}{n\pi\left((r-l+1)^2+r^2-2\right)}\right)^{1/2} \sum_{k=l-r}^{0} \exp\left(-\frac{6(k-n(l-1)/2)^2}{n\left((r-l+1)^2+r^2-2\right)}\right).$$

References

[1] J. Browne, Proposal number 1071, this MAGAZINE, 52 (1979) 114.
[2] W. Feller, An Introduction to Probability Theory and its Applications, vol. II, Wiley, New York, 1966, p. 490.
[3] G. A. Heuer, Solution to Proposal 1071, this MAGAZINE, 53 (1980) 247.
[4] J. L. Selfridge, Solution to Proposal 1071, this MAGAZINE, 54 (1981) 141.

E and M

M. R. SPIEGEL
East Hartford, CT 06118

An engineer called E *has students who complain*
That mathematics taught to them is nothing but a pain.

"*Why can't we take a course*," *they ask*, "*a course designed for us?*
A course which helps us in our fields without this kind of fuss.

Why ask us for a proof and make us say, '*I can't, sir*,'
When what we'd really like to know is how to get the answer."

With empathy E *rushed to* M *who gave the ill-famed course*
And told him of the problem and also of the source.

"*The students say the course you give drives them up the wall.*
Epsilons and deltas seem to have no point at all.

They need the kind of math that helps with circuitry and cables.
They need to know how they can use the mathematics tables."

Professor M *was not surprised—he'd heard complaints before.*
"*You realize*," *he said to* E, "*there are things we can't ignore.*

Math requires subtleties—you can't just make a list.
You cannot take derivatives of things that don't exist."

On and on he lectured E *with little hesitation*
On topics such as limits and improper integration,

On necessary and sufficient conditions for existence.
E *tried*, *but could not shake*, *Professor* M'*s insistence.*

There is no end to this debate; E *and* M *cannot agree.*
There are two sides and each will teach as only he can see.

APPROXIMATION FORMULA. $P(\text{Player A wins}) \doteq \frac{1}{2} + \frac{r-2}{2r}\left(\frac{3}{n\pi(r^2-1)}\right)^{1/2} - r^{-(2n+1)}$

$$+\sum_{j=1}^{r-2}\frac{r-j-1}{r}\left(\frac{3}{n\pi(r^2-1)}\right)^{1/2}\exp\left(-3j^2/n(r^2-1)\right)$$

$$-\sum_{l=2}^{r-1}\frac{1}{r}\left(\frac{r-l+1}{r}\right)^{n}\left(\frac{6}{n\pi\left((r-l+1)^2+r^2-2\right)}\right)^{1/2}\sum_{k=l-r}^{0}\exp\left(-\frac{6(k-n(l-1)/2)^2}{n\left((r-l+1)^2+r^2-2\right)}\right).$$

References

[1] J. Browne, Proposal number 1071, this MAGAZINE, 52 (1979) 114.
[2] W. Feller, An Introduction to Probability Theory and its Applications, vol. II, Wiley, New York, 1966, p. 490.
[3] G. A. Heuer, Solution to Proposal 1071, this MAGAZINE, 53 (1980) 247.
[4] J. L. Selfridge, Solution to Proposal 1071, this MAGAZINE, 54 (1981) 141.

E and M

M. R. SPIEGEL
East Hartford, CT 06118

An engineer called E *has students who complain*
That mathematics taught to them is nothing but a pain.

"Why can't we take a course," they ask, "a course designed for us?
A course which helps us in our fields without this kind of fuss.

Why ask us for a proof and make us say, 'I can't, sir,'
When what we'd really like to know is how to get the answer."

With empathy E *rushed to* M *who gave the ill-famed course*
And told him of the problem and also of the source.

"The students say the course you give drives them up the wall.
Epsilons and deltas seem to have no point at all.

They need the kind of math that helps with circuitry and cables.
They need to know how they can use the mathematics tables."

Professor M *was not surprised—he'd heard complaints before.*
"You realize," he said to E, *"there are things we can't ignore.*

Math requires subtleties—you can't just make a list.
You cannot take derivatives of things that don't exist."

On and on he lectured E *with little hesitation*
On topics such as limits and improper integration,

On necessary and sufficient conditions for existence.
E *tried, but could not shake, Professor* M*'s insistence.*

There is no end to this debate; E *and* M *cannot agree.*
There are two sides and each will teach as only he can see.

PROBLEMS

DAN EUSTICE, Editor
LEROY F. MEYERS, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before November 1, 1981.

1122. Let $x_1, x_2, \ldots, x_n$ be n positive numbers, where n is even. Define

$$f(x_1, x_2, \ldots, x_n) = \left(\frac{x_2+x_3}{2}, \frac{x_2+x_3}{2}, \frac{x_4+x_5}{2}, \frac{x_4+x_5}{2}, \ldots, \frac{x_n+x_1}{2}, \frac{x_n+x_1}{2}\right).$$

If f^k denotes the kth iterate of f, find

$$\lim_{k\to\infty} f^k(x_1, x_2, \ldots, x_n).$$

[*Edilio Escalona, Caracas, Venezuela.*]

1123. For which positive integers p is the following result true? If $\{a_n\}$ is a sequence of real numbers and Σa_n^p converges, then $\Sigma a_n/n$ must converge. [*Steve Ricci, Boston College.*]

1124. Evaluate the indefinite integral

$$\int \tan^2(x-a_1)\tan^2(x-a_2)\cdots\tan^2(x-a_n)\,dx.$$

[*L. Kuipers, Switzerland.*]

1125. Let $f(x)$ be differentiable on $[0,1]$ with $f(0)=0$ and $f(1)=1$. For each positive integer n and arbitrary given positive numbers $k_1, k_2, \ldots, k_n$, show that there exist distinct $x_1, x_2, \ldots, x_n$ such that

$$\sum_{i=1}^{n} \frac{k_i}{f'(x_i)} = \sum_{i=1}^{n} k_i.$$

[*G. Z. Chang, University of Utah.*]

ASSISTANT EDITORS: DON BONAR, *Denison University;* WILLIAM A. MCWORTER, JR., *The Ohio State University.* *We invite readers to submit problems believed to be new. Proposals should be accompanied by solutions, when available, and by any information that will assist the editors. Solutions to published problems should be submitted on separate, signed sheets. An asterisk (*) will be placed by a problem to indicate that the proposer did not supply a solution. A problem submitted as a Quickie should be one that has an unexpected succinct solution. Readers desiring acknowledgment of their communications should include a self-addressed stamped card. Send all communications to this department to Dan Eustice, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.*

1126. Let G be a connected graph with a nonnegative integer $f(v)$ attached to each of its vertices v. Suppose f has the following properties:

(1) If the vertices v and w are adjacent, then $f(v)$ and $f(w)$ differ by at most 1.
(2) If $f(v)>0$, then v is adjacent to at least one vertex w such that $f(w)<f(v)$.
(3) There is exactly one vertex v_0 such that $f(v_0)=0$.

Prove that $f(v)$ is the distance of v from v_0. [*Peter Ungar, New York University.*]

Quickies

Solutions to Quickies appear at the conclusion of the Problems section.

Q668. Three of the four integers between 100 and 1000 which are equal to the sum of the cubes of their digits are 153, 371, and 407. What is the fourth such integer? [*Unsigned letter from Columbus, Ohio.*]

Solutions

Roll the Dice Again **March 1979**

1071. Player A rolls $n+1$ dice and keeps the highest n. Player B rolls n dice. The higher total wins, with ties awarded to Player B.
(a) For $n=2$, show that Player A wins and find his probability of winning.
(b*) Find the smallest value of n for which Player B wins. [*Joseph Browne, Onondaga Community College.*]

Solution: The solution to part (a) is in this *Magazine*, 53 (1980) page 247.
(b*) 1. (Trivial solution.) Mathematician's answer. You cannot roll a negative number of dice, but you can roll zero dice very easily. Thus the answer is $n=0$.

2. (Nonrigorous solution.) Gambler's answer. Winning ties is worth exactly half a point. When $n=1$, B's average is thus 4, while A's average is easily seen to be 4.5. B's average is always $3.5n+0.5$ and A's average approaches $3.5(n+1)-1$ for large n, an edge for A of just under two points. The conjecture that A's probability of winning approaches .58 is wrong, since for a million dice, 8% of the cases is far more than a two-point spread.

3. Since $n=1$ is easy, we prove that A has the edge if $n>1$. To emphasize symmetry, use dice with faces -5, -3, -1, 1, 3, and 5. Let b be the sum of B's dice and a the sum of the last n of A's dice.

Case (i): $a\neq b$. Each event with $a<b$ can be matched with an event $a>b$ which A wins. A also wins some events when $a<b$. Thus A has the edge in case (i).

Case (ii): $a=b$. A wins just in case $a_0>a_i$ for some $i>0$. The events $a_i=1$ and $a_i=-1$ each have frequency f_1, and define similarly f_3 and f_5, where $2(f_1+f_3+f_5)=1$. Then $a_0>a_i$ with frequency $(5f_5+4f_3+3f_1+2f_1+f_3)/6=5/12$. An easy calculation shows that $a_0>\min(a_1,a_2)$ with frequency greater than $1/2$ and thus A has the edge in case (ii) just when $n>1$.

J. L. Selfridge
Mathematical Reviews

A Congruence **March 1980**

1094. For each positive integer n, show that there is a solution to the congruence

$$32k^2+21k+14\equiv 0\ (\mathrm{mod}\,2^n).$$

[*Tom Moore, Bridgewater State College.*]

Solution: We show more generally that $ak^2+bk+c\equiv 0$ (mod2^n) has a solution for all n whenever b is odd and a or c is even. For $n=1$, take $k=0$ if c is even and $k=1$ if c is odd. Now suppose the claim is true for n. If c is even, then, by assumption, the congruence $2at^2+bt+c/2\equiv 0$ (mod2^n) has some solution t. Letting $k=2t$ we get $ak^2+bk+c=2(2at^2+bt+c/2)\equiv 0$ (mod2^{n+1}). If c is odd, then a is even, so $a+b+c$ is even; hence, by assumption, the congruence $2at^2+(2a+b)t+(a+b+c)/2\equiv 0$ (mod2^n) has some solution t. Letting $k=2t+1$ yields

$$ak^2+bk+c=2\big[2at^2+(2a+b)t+(a+b+c)/2\big]\equiv 0\ (\mathrm{mod}\,2^{n+1}).$$

Thus, whether c is even or odd, the claim is true for $n+1$, and so by induction for all n.

N. M. Rice
Queen's University
Kingston, Ontario, Canada

Also solved by Richard Beigel, Bern Problem Solving Group (Switzerland), Walter Bluger (Canada), Bill Bompart, Michael Carr, Milton P. Eisner, Nick Franceschine III, Gordon Fisher, Tom Haertel, L. S. Kennison, L. Kuipers (Switzerland), Jinku Lee, Graham Lord (Canada), John J. Martinez, Jack McCown, William Myers, Bob Prielipp, Roger Cuculière (France), Harry D. Ruderman, Robert S. Stacy, J. M. Stark, Michael Vowe (Switzerland), Edward T. H. Wang (Canada), Gregory Wulczyn, Ken Yocum, Aleksandras Zujus, and the proposer.

Fundamental Period **March 1980**

1095*. It is mathematical folklore that the fundamental period of a linear combination (with nonzero coefficients) of simple sine and cosine functions having commensurable fundamental periods is the least common multiple of the fundamental periods of the separate functions. Is there a noncalculus proof (or disproof) of this? (The *fundamental period* of a periodic function from R to R is the smallest positive period. The function f is a *simple* sine or cosine function if $f(x)=\sin\alpha x$ for all x, or $f(x)=\cos\alpha x$ for all x, for some positive constant α.) [*S. Yeshurun, Bar-Ilan University, Ramat-Gan, Israel.*]

Solution: Let $f(x)=a\sin\alpha x+b\cos\beta x$, where $ab\neq 0$, $\alpha>0$, $\beta>0$, and α/β is rational. Then the least common multiple of $2\pi/\alpha$ and $2\pi/\beta$ is a period of $f(x)$.

Let T be the fundamental period of $f(x)$. Then T is a period of

$$\frac{1}{2}(f(x+\epsilon)+f(x-\epsilon))=a\sin\alpha x\cos\alpha\epsilon+b\cos\beta x\cos\beta\epsilon$$

and

$$\frac{1}{2}(f(x+\epsilon)-f(x-\epsilon))=a\cos\alpha x\sin\alpha\epsilon-b\sin\beta x\sin\beta\epsilon.$$

Therefore, for all ϵ,δ:

$$b\cos\beta\epsilon=a\sin\alpha T\cos\alpha\epsilon+b\cos\beta T\cos\beta\epsilon;\tag{1}$$

$$a\sin\alpha\epsilon=a\cos\alpha T\sin\alpha\epsilon-b\sin\beta T\sin\beta\epsilon;\tag{2}$$

$$b\cos\beta\delta=a\sin\alpha T\cos\alpha\delta+b\cos\beta T\cos\beta\delta;\tag{3}$$

$$a\sin\alpha\delta=a\cos\alpha T\sin\alpha\delta-b\sin\beta T\sin\beta\delta.\tag{4}$$

Let $0<\epsilon<\delta<\min(\pi/2\alpha,\pi/2\beta)$. Only two cases need to be considered.

Case 1: $\alpha=\beta$. (1) and (2) become
$b=a\sin\alpha T+b\cos\alpha T$;
$a=a\cos\alpha T-b\sin\alpha T$.

Since this system has a nonzero determinant, it has a unique solution, which is $\cos\alpha T=1$ and $\sin\alpha T=0$. Therefore, T is a multiple of $2\pi/\alpha$, which must be the fundamental period of $f(x)$.

Case 2: $\alpha\neq\beta$. The linear system (1)–(4) has a unique solution provided that

$$\cos\alpha\epsilon\cos\beta\delta-\cos\alpha\delta\cos\beta\epsilon\neq 0$$

and

$$\sin\alpha\epsilon\sin\beta\delta-\sin\alpha\delta\sin\beta\epsilon\neq 0.$$

Let $\delta=2\epsilon$. Then

$$\begin{aligned}\cos\alpha\epsilon\cos\beta\delta-\cos\alpha\delta\cos\beta\epsilon&=\cos\alpha\epsilon(2\cos^2\beta\epsilon-1)-(2\cos^2\alpha\epsilon-1)\cos\beta\epsilon\\&=(1+2\cos\alpha\epsilon\cos\beta\epsilon)(\cos\beta\epsilon-\cos\alpha\epsilon)\neq 0.\end{aligned}$$

Also,

$$\begin{aligned}\sin\alpha\epsilon\sin\beta\delta-\sin\alpha\delta\sin\beta\epsilon&=2\sin\alpha\epsilon\sin\beta\epsilon\cos\beta\epsilon-2\sin\alpha\epsilon\cos\alpha\epsilon\sin\beta\epsilon\\&=2\sin\alpha\epsilon\sin\beta\epsilon(\cos\beta\epsilon-\cos\alpha\epsilon)\neq 0.\end{aligned}$$

Therefore, $\cos\alpha T=1$, $\sin\alpha T=0$, $\cos\beta T=1$, $\sin\beta T=0$. Therefore, T is a multiple of the fundamental periods of $\sin\alpha x$ and $\cos\beta x$; hence T is their least common multiple.

The theorem has a simple proof using calculus. If T is the fundamental period of f, then $f(T)=f(0)$ and the first three derivatives of f are the same at T and at 0. Therefore,

$$\begin{aligned}a\sin\alpha T+b\cos\beta T&=b,\\a\alpha\cos\alpha T-b\beta\sin\beta T&=a\alpha,\\-a\alpha^2\sin\alpha T-b\beta^2\cos\beta T&=-b\beta^2,\\-a\alpha^3\cos\alpha T+b\beta^3\sin\beta T&=-a\alpha^3.\end{aligned}$$

Solving these equations in pairs, we find: $\sin\alpha T=0$, $\cos\alpha T=1$, $\sin\beta T=0$, $\cos\beta T=1$, as before.

RICHARD BEIGEL
Stanford, California

The proposer first called attention to the problem in Internat. J. Math. Ed. Sci. Tech., vol. 8 (1977) 411–415. Two solutions using calculus were given for the general case $\Sigma_{k=1}^{n}(a_k\sin\alpha_k x+b_k\cos\alpha_k x)$ by W. C. Waterhouse, ibid., vol. 9 (1978) 375, and one by L. F. Meyers, ibid., vol. 10 (1979) 281–282.

There was one incorrect solution.

Answers

Solutions to the Quickies which appear near the beginning of the Problems section.

Q668. Since 1 is the last digit of 371 and $1^3=1$, the fourth integer must be 370.

REVIEWS

Paul J. Campbell, Editor
Beloit College

Pierre J. Malraison, Jr., Editor
MDSI, Ann Arbor

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.

Kolata, G.B., *Mathematics news in brief*, Science 212 (3 April 1981) 30-31.

Announces a solution to the Littlewood conjecture on integrals of sums of primes, the Van der Warden conjecture on minimum values for permanents, and some new results on integer programming.

Cole, Jonathan R., *Meritocracy and marginality: women in science today and tomorrow*, Association for Women in Mathematics Newsletter 11:1 (Jan.-Feb. 1981) 4-19.

Perceptive essay deals carefully with sensitive issues from a strong empirical basis. The author dispels some myths: that women who enter science are not as "able" as their counterparts; that women are systematically discriminated against. Cole did find differences in "reputational standing" of men and women in the same scientific fields. Inequalities in salaries of men and women scientists "virtually disappear" after a variety of explanatory factors, including research productivity, are considered. However: "Female scientists tend not to publish as much as [comparable] males," yet the data do *not* support the usual speculations as to cause (family obligations, etc.). The author concludes by examining the barriers to full equality for women in science, including the question of how to alter American culture to make science an attractive career for women.

Klarner, David A., ed., The Mathematical Gardner, Wadsworth, 1981; viii + 382 pp, $17.95.

Splendid tribute to Martin Gardner, author of the renowned Mathematical Games column in *Scientific American*, on the occasion of his 65th birthday. His lucid, popular accounts of recent developments in recreational and "serious" mathematics have inspired a generation of amateur and professional mathematicians. Have you ever wondered if it would be possible to devise a fair game of "mental poker?" Or what a bicycle tube looks like when turned inside out? The articles are grouped under games, geometry, 2-dimensional tiling, 3-dimensional tiling, fun and problems, and numbers and coding theory. Other highlights include Schattschneider's "In Praise of Amateurs," which details contributions of amateurs to the problem of tiling the plane with convex pentagons, and Knuth's "Supernatural Numbers," which discusses efficient representations of gigantic numbers.

Fletcher, Colin R., *G.H. Hardy--applied mathematician*, Bulletin of the Institute of Mathematics and Its Applications 16 (1980) 61-67.

Highly interesting and detailed account of the circumstances surrounding Hardy's contribution to genetics (the Hardy-Weinberg law)--the counterexample to his claim "I have never done anything useful." Also includes anecdotes which complement those of Hardy's *A Mathematician's Apology*, and a photo of Hardy in cricket gear.

Noll, Curt and Nickel, Laura, *The 25th and 26th Mersenne primes*, Mathematics of Computation 35 (1980) 1387-1390.

The authors were high school students when they found the primes (via Lucas-Lehmer test) and wrote the paper.

Millen, Jonathan K., *Programming the game of Go*, Byte 6:4 (April 1981) 102-120.

Miniature assembly-language program on a Kim-1, plays beginner-level Go (without Ko rule) on a 15x15 board.

Davidson, James J., *Recurrence in numerical analysis*, Byte 6:4 (April 1981) 20-30.

Considerations in calculations functions by recurrence, applied particularly to Bessel functions, the recursive functions with the greatest engineering utility.

Bronson, Gail, *Turn for the worse: 'simple' little puzzle drives millions mad*, Wall Street Journal (5 March 1981) 1, 20.

Account of the popularity of Rubik's magic cube, of which 4½ million were sold in the U.S. last year.

Hofstadter, D.R., *Metamagical Themas: The magic cube's cubies are twiddled by cubists and solved by cubemeisters*, Scientific American 244:3 (March 1981) 20-39, 194.

Everything you wanted to know about Rubik's cube *except* how to solve it, including the total number of possible positions, relationships to quantum mechanics, "pretty patterns," and where to get the book that does tell you how to solve it.

Singmaster, David, Notes on Rubik's 'Magic Cube,' 5th ed., corrected, Enslow Publishers (Box 777, Hillside, NJ 07205), 1981; iv + 72 pp, $5.95 (P).

Reissue of the famous cubemeister's manual, with errata, new publisher, and new price.

Platonic chemistry: now a dodecahedrane, Science News, 119:6 (7 February 1981) 85.

A 20-faced hydrocarbon structure has been created by researchers at Ohio State University. "It's hard to say what it's going to be good for," says one of the creators--but Plato would have been pleased.

Shenoy, Prakash P., *A two-person non-zero-sum game model of the world oil market*, Applied Mathematical Modelling 4 (1980) 295-300; *A three-person cooperative game formulation of the world oil market*, *ibid.*, 301-307.

Application of many of the prominent concepts of contemporary game theory to a situation of popular interest. Excellent example for a course in game theory. Includes validation for U.S.A. of assumed exponential model for function describing loss in GNP resulting from cuts in oil imports.

Dale, Charles and Workman, Rosemarie, *The arc sine law and the Treasury bill futures market*, Financial Analysts Journal, (Nov.-Dec. 1980) 71-74.

The "arc sine law for last visits," a theorem of mathematical statistics, is here visited upon financial analysts, who are cautioned that it implies that mechanical trading rules (such as use of a moving average) will not be profitable in the long run. This is the same law that says that when two sports teams of equal ability play, one will probably lead most of the game; and that games with a close final score are surprisingly likely to be "last-minute come-from-behind" affairs, in which the winner trailed for most of the game. The article refers the reader to Feller's classic text for the mathematical background.

Gardner, M., *Mathematical Games: How Lavinia finds a room on University Avenue and other geometric problems*, Scientific American 244:4 (April 1981) 18-26.

A collection of problems involving geometry, topology, and game theory, to name a few.

Adams, Merilee, Napkin Folding for the Geometrical Gourmet, Harcourt Brace Jovanovich, 1981; 20 pp, gratis.

A bouquet of 17 ways to fold a napkin, by a former mathematics instructor; "these napkin folds derive from her interest in right-hemisphere mathematics."

Malone, Thomas W., What Makes Things Fun to Learn? A Study of Intrinsically Motivating Computer Games, Cognitive and Instructional Sciences Series CIS-7 (SSL-80-11), Xerox Palo Alto Research Center, August 1980.

Why are computer games so captivating, so addicting? How can we co-opt the essential features to deliver the standard fare of the educational agenda? Should we? The author deals only with the first question, focusing on challenge, fantasy, and curiosity.

Judson, Horace Freeland, The Search for Solutions, Holt, Rinehart and Winston, 1980; x + 211 pp, $16.95.

Magnificent exploration of the methods and basis of scientific knowledge, illustrated with scores of breath-taking photographs. Mathematics is mentioned in connection with patterns and modeling. The unspoken philosophy of the book is *technologism*, *i.e.*, human problems can be solved by enough technological progress.

Gruber, B. and Millman, R.S., Symmetries in Science, Plenum Press, 1980, ix + 495 pp.

Proceedings of a conference held at Southern Illinois University the week of February 29, 1979, in celebration of Einstein's 100th birthday. Topics range from Dirac's "Why We Believe in the Einstein Theory" which is suitable for the general reader, to Bohm's paper on DeSitter Fibers and SO(3,2) Spectrum which is aimed more at the professional theoretical physicist. Thirty papers in all, this book is recommended for any mathematician with an interest in applications of mathematics (algebra, statistics, functional analysis) to physics.

Wang, Hao, Popular Lectures on Mathematical Logic, Van Nostrand Reinhold, 1981; ix + 273 pp, $24.95.

Six extensive lectures given at the Chinese Academy of Science in 1977, covering the modern scene in mathematical logic without bogging the reader down in symbolism. Lectures treat axiomatic method, computers, problems as a driving force, first order logic, computation, and the continuum problem. Extensive appendices cover dominoes and the infinity lemma, algorithms, and abstract machines.

Goldstine, Herman H., A History of the Calculus of Variations from the 17th Through the 19th Century, Springer-Verlag, 1980; xviii + 410 pp.

Excellent survey which takes the reader through the thoughts and calculations of the subject's creators. Accessible to students with background in advanced calculus.

Lavington, Simon, Early British Computers, Digital Press, 1980; iii + 138 pp.

A very entertaining account of the development of British computers and computer companies from the late thirties up to about 1955. Anyone complaining about lack of facility in some high-level languages should read the chapter titled "Programming an Early Computer."

Thomas, A.D. and Wood, G.V., Group Tables, Shiva, 1980; 190 pp, \$12.50, \$6.50 (P).

Multiplication table, subgroup lattice, character table, automorphism group type, and other details on all non-cyclic groups of order ≤ 32. A handy reference source for students and teachers alike. (Errata slip, glued to Contents page, notes six corrections to character tables.)

Warrick, Patricia S., The Cybernetic Imagination in Science Fiction, MIT Press, 1980; xvii + 282 pp, \$15.

Survey, history, analysis, and bibliography of "cybernetic fiction," 255 short stories and novels (1930-1977) involving fictional worlds of computers and robots. The main observation: "Much of the SF [science fiction] is dysutopian, but no such negative attitude prevails in the field of computer science...[M]uch of the fiction written since World War II is reactionary in its attitude toward computers and artificial intelligence. It is often ill-informed about information theory and computer technology and lags behind present developments instead of anticipating the future." If the writers were better informed, would they be any less negative?

Howard, James C., Mathematical Modeling of Diverse Phenomena, U.S. Govt. Printing Office (S/N: 033-000-00777-9), 1979; viii + 394 pp, \$8.

A more accurate title would allude to the fact that this is not a general work in modeling, but an exposition on tensors and their applications in aeronautics, particle dynamics, fluid mechanics, and cosmology. A noteworthy feature is the emphasis on the usefulness of computerized *symbolic* mathematical computation for working out details of models.

Ledermann, Walter and Vajda, Steven, Handbook of Applicable Mathematics, Vol. I: Algebra, Wiley, 1980; xix + 524 pp, \$85.

First of six "core" volumes. Each chapter is, to the greatest extent possible, self-contained. Standard topics on sets, numbers, and matrix algebra are included, plus articles on group theory (including representations), linear programming, integer programming, theory of games, and half a dozen other minor chapters. Subsequent core volumes will treat probability, numerical methods, analysis, geometry and combinatorics, and statistics; in addition, guide books will be published for mathematics users in specific areas.

National Science Foundation & the Department of Education, Science & Engineering Education for the 1980's & Beyond, U.S. Govt. Printing Office (S/N: 038-000-00467-1), October 1980; 82 pp, \$3.75 (P).

Report (to President Carter) of what needs to be done to relieve personnel shortages and strengthen educational capacity, to provide both adequate numbers of highly-trained scientists and engineers and a scientifically literate citizenry. Little that is news to educators, but contains a crisp and well-referenced summary of the problems and their causes.

NEWS & LETTERS

MORE ON MAGIC CIRCLES

In the last paragraph of "Magic Circles" in this issue, two questions were raised about which I now know more than when I wrote the article. It was shown that magic circles of any size $N = p^s + 1$ for p prime can be produced, but whether circles exist when $N \neq p^s + 1$ was unresolved. Unpublished work of V.H. Keiser in fact implies the non-existence for all such $N < 3600$. The other question dealt with a conjecture of Singer that the number of different magic circles of a given size $N = p^s + 1$ is $\Phi(N(N - 1) + 1)/3s$. Using a 1956 result of M. Hall of perfect difference sets, I have recently used the computer to verify this conjecture for all $N < 18$. Finally, I have discovered two nice reference books. T.H. O'Beirne's Puzzles and Paradoxes (Oxford U. Press, 1965), Chapter 6 and pp. 218-220 is great fun, and Cyclic Difference Sets by L.D. Baumert (Springer-Verlag, 1971), especially Chapters IV and V, is a comprehensive and comprehensible source of information on perfect difference sets.

David A. James
University of Michigan
Dearborn, MI 48128

RESPONSIBLE USE OF MATHEMATICS

This note was prompted by Brams' article "Mathematics and Theology..." (this *Magazine*, November 1980, pp. 227-282), but I want to discuss a broader issue. Brams' article is easy to dismiss. He postulates explicitly (p. 279) that God is a liar. No one who takes religion seriously would accept that postulate; the assumptions are false, and so the conclusions are irrelevant.

I am concerned about the questionable use of mathematics in areas relating to morality. The field of game theory (which is the basis of Brams' article) seems to be especially susceptible to abuse.

Although I am a pure mathematician, it seems to me that standards of intellectual rigor are more essential in applied mathematics than in pure. If, for example, someone published a twenty page paper proving that 2 + 2 = 4 (or even 2 + 2 = 5), this would do no harm beyond wasting twenty pages in a journal. Bad applied mathematics can cause airplanes to crash. Bad reasoning in the moral sphere can do even greater harm. It can cast the mathematician in the role of the Pied Piper, seducing students into accepting dubious assumptions because they are clothed in confusing mathematical language.

Naturally I do not wish to deny that there are legitimate uses of applied mathematics, in the social sciences and elsewhere. Suppose we try to formulate a rough criterion as to when mathematics is applicable. The conditions seem to be: 1) that we start with postulates which have a high probability of being true, and 2) that these postulates are of such a nature that, in discussing them, mathematics renders valuable insights.

What, then, should we do about the multitude of human problems which do not lend themselves to mathematical treatment? Isn't it clear that we should return to the venerable tradition of discussing such problems on their historical and philosophical merits. The problems we face are too important for us to indulge the luxury of playing dishonest intellectual games with them.

Ian Richards
University of Minnesota
Minneapolis, MN 55455

REPORT ON CURRICULUM CONFERENCE

New roles for computer science, statistics and applications dominated a conference on the Undergraduate Mathematics Curriculum held on November 14-15, 1980 at St. Olaf College in Northfield, Minnesota. "Mathematical knowledge, both pure and applied, is doubling every ten years," noted speaker William F. Lucas of Cornell Univeristy.

The rapid growth of the newer applied mathematical sciences has severely strained both faculty resources and curriculum standards for undergraduate mathematics, especially in four-year institutions. Student interest in applied courses has skewed enrollment patterns, making it very difficult for small departments to maintain strong offerings in traditional core subjects. Moreover, most departments are constrained by a no-growth budget, making it impossible to add staff to teach new courses in applied areas.

These and other related issues provided a sense of urgency to the conference, which was supported by an NSF CAUSE grant. Seven panelists representing small departments offered advice on how to respond to these new demands, and the participants themselves discussed the issues in simultaneous special topic sessions. The forthcoming CUPM report on a unified undergraduate program in mathematical science served as a basis for much of the discussion.

While the conference made no attempt to formulate recommendations, transcripts of the formal presentations and notes from the discussion groups were assembled into a written Report to record issues and ideas emerging from the conference. Copies of this Report are available for $4.00 (including postage) from Professor Lynn Arthur Steen, CAUSE Project Director, Department of Mathematics, St. Olaf College, Northfield, Minnesota 55057.

CONFERENCE ON ORDERINGS AND QUADRATIC FORMS

An NSF regional conference in mathematics on the topic "Orderings and Quadratic Forms over Fields" will be held at Carleton College from August 10-14, 1981. The main speaker for the conference will be T-Y Lam, of the University of California. For further information write to Professor Steve Galovich, Dept. of Mathematics, Carleton College, Northfield, Minnesota 55057.

CONFERENCE ON EMERGING TRENDS IN MATHEMATICS

The Ninth Annual Mathematics and Statistics Conference at Miami University, Oxford, Ohio, will be held Sept. 25-26, 1981. The theme for this year's conference will be "Emerging Trends in Mathematics and Its Instruction." Featured speakers will include Paul R. Halmos, Indiana Univ.; Alan Tucker, SUNY at Stony Brook; and James W. Wilson, Univ. of Georgia. Contributed papers relating to the general theme and appropriate for a general audience of mathematicians are welcome. Of particular interest will be papers dealing with emerging trends in a particular area of mathematics or mathematics education. Abstracts should be sent by June 1 to David Kullman or Lyman Peck, Dept. of Mathematics and Statistics, Miami Univ., Oxford, OH 45056. Information concerning preregistration, housing, etc., will be available from this address after July 15.

The Ohio Delta Chapter of Pi Mu Epsilon will also hold its annual student conference Sept. 25-26, 1981. Undergraduate and master's mathematics students are invited to contribute papers, and should send abstracts to Milton Cox, Dept. of Mathematics and Statistics, Miami University, Oxford, OH 45056.

David E. Kullman
Lyman C. Peck
Miami University
Oxford, OH 45056

FOCUS ON

FOCUS

"Focus," the national MAA newsletter, made a welcome debut in March. Edited by Marcia Sward, MAA Associate Director, the publication provides a timely forum for news, announcements, and comments of interest to members of the MAA. The Newsletter will be published bi-monthly except for July (in March, June, September, November, and January) and received by all members of the MAA.

PUTNAM WINNERS

A total of 2043 students from 335 colleges and universities in Canada and the U.S. participated in the Putnam Mathematical Competition held on Dec. 6, 1980. There were teams from 237 schools.

The universities with winning teams are listed below in order of rank, together with members of the winning teams.

1 Washington University, St. Louis

Kevin P. Keating, Nathan E. Schroeder, Edwin A. Shpiz

2 Harvard University

Michael Raship, Ehud B. Reiter, Brian F. Sheppard

3 University of Maryland, College Park

Ravi B. Boppanna, Brian R. Hunt, Eric I. Kuritzky

4 University of Chicago

Daniel J. Goldstein, Nicholas F. Reingold, Michael P. Spertus

5 University of California, Berkeley

Randall L. Dougherty, Lin Goldstein, Robin A. Pemantle

The individual students who ranked highest (Putnam Fellows), are listed below in alphabetical order.

Eric D. Carlson	Michigan St. U.
Randall L. Dougherty	U. of Ca., Berkeley
Daniel J. Goldstein	Univ. of Chicago
Laurence E. Penn	Harvard Univ.
Michael Raship	Harvard Univ.

More complete details on the competition and the top participants will appear in the *American Mathematical Monthly*.

AN UNFORGETTABLE SCORE

Stephen Curran, a junior at Beloit College who received honorable mention, is not likely to forget his score or rank for this competition. He placed 41st in the 41st annual Putnam competition with a score of 41.

BIBLIOGRAPHY OF JOURNALS FOR TEACHERS OF MATHEMATICS

As part of its 'documentation' of materials and information project, the Institute für Didaktik der Mathematik at the Universität Bielefeld, W. Germany, has published an International Bibliography of Journals in Mathematical Education (IDM 23/1980). Edited by Gert Schubring and Jutta Richter, the bibliography intends to include all journals with relevance to the teaching of mathematics in its various orientations and levels.

The editors began with the NCTM list of such journals and made some deletions and several additions, and have produced a 108-page bibliography which is extensive both in its number and variety of journals (including some little-known 'local' or 'student' journals). A preliminary version of the publication was circulated at the International Congress on Mathematical Education held in Berkeley last summer; the latest edition contains several corrections and additions.

PROPOSED BUDGET CUTS TO ELIMINATE NSF SCIENCE EDUCATION FUNDS; ANDERSON RESPONDS

The Administration is recommending to the Congress that the Science Education budget at NSF be reduced to zero, except for ongoing, already-committed graduate fellowships, and thus the Science Education Directorate at NSF be effectively terminated. Richard Anderson, president of the MAA, has responded to these proposals by writing letters to Congressional committee members as well as his local representatives; a sample letter is reproduced on the next page. The Council of Scientific Society Presidents, on which Anderson serves, has adopted a position to the same effect and is presently actively working at describing in detail a Presidential Council as mentioned in his letter.

Anderson urges all members of the MAA and others interested to send to their senators and representatives similar letters, one page in length, written on stationery showing their professional affiliation.

Senator Russell B. Long
United States Senate
Washington, DC 20510

Dear Senator Long:

I urge the Congress to restore funding to Science Education at NSF to a level compatible with overall NSF funding, to maintain the Science Education Directorate, and to direct the National Science Board to refocus its science education effort toward pre-college education for science and technology. The potential dollar cost for a properly redirected federal effort is rather small, the potential benefit to our country is very great.

I strongly support the President's decision on the need for economic initiatives including deep budget cuts and, in particular, his thesis that, in the long run, increased productivity is the essential ingredient to control inflation and to make the economy healthy.

To achieve real increased productivity we not only need scientific and technological innovation but a work force able to use and exploit the technological advances already made and to be made. In this connection, we are facing a crisis in pre-college education: the developing computer-calculator age demands many more and better quantitatively-oriented people to enter our work force at all levels, yet there has been a long term pronounced deterioration in average student performances.

As a top priority in education, we must work for real excellence in pre-college education for science and technology: it is at this level that other countries like Japan, Germany, and the Soviet Union appear to be seriously outstripping us. In my opinion, only the federal government can give the needed impetus to a national commitment for excellence. Our nation's educational system is too diverse and too complex for there to be any real or early prospects for needed change without federal intervention. We need a Presidential Council for Excellence in Education for Science and Technology and a refocusing of both Department of Education and NSF efforts in education for science and technology.

Without a major improvement in pre-college education in science and mathematics, we also face very serious future military manpower problems. We need personnel at all levels capable of being trained to maintain and operate increasingly sophisticated military equipment.

There is substantial evidence that both the larger scientific and the larger technological communitites would strongly back the thrusts suggested above.

I would welcome an opportunity to discuss the issues with you further and/or to provide more documentation in support of these statements.

Sincerely,

RD Anderson

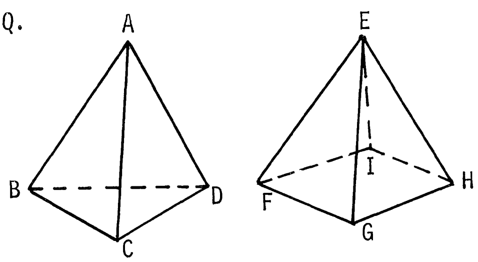
R. D. Anderson

MATHEMATICAL SUNSHINE

Recent "sunshine" decisions by the Educational Testing Service and the 1979 New York Truth-in-Testing Law this year gave students the opportunity to scrutinize (post-test) the questions, their answers, and the test key of the PSAT and SAT. On each test, a student successfully challenged the examiners' choice of answer to a mathematical question. In acknowledging the errors, the ETS raised 250,000 PSAT and 19,000 SAT scores. The embarassing disclosures can be seen as a ray of mathematical sunshine--here is assurance that there are bright students who read problems carefully and are not willing to accept intended, but not stated, assumptions, nor accept answers dictated by authority alone.

The questions challenged appear below; for the ETS answers and the challengers' answers together with their reasoning, see *Time*, 3/31/81, page 51, or *Newsweek*, 4/6/81, page 84.

Q.



In pyramids ABCD and EFGHI shown above, all faces except base FGHI are equilateral triangles of equal size. If face ABC were placed on face EFG so that the vertices of the triangles coincide, how many exposed faces would the resulting solid have?

(A) Five (B) Six (C) Seven
(D) Eight (E) Nine

Q.

Row A	7	2	5	4	6
Row B	3	8	6	9	7
Row C	5	4	3	8	2
Row D	9	5	7	3	6
Row E	5	6	3	7	4

Which row in the list above contains both the square of an integer and the cube of a <u>different</u> integer?

(A) Row A (B) Row B (C) Row C
(D) Row D (E) Row E

A FAULTLESS REJECTION

The following letter is offered as consolation to the many authors who have submitted manuscripts for publication, only to receive a letter with the phrase "we regret...."

After a British writer submitted for publication a paper on the economy to a Chinese journal, he received the following rejection letter (quoted in the *World Business Weekly*):

"We have read your manuscript with boundless delight. If we were to publish your paper it would be impossible for us to publish any work of a lower standard. And as it is unthinkable that, in the next thousand years, we shall see its equal, we are, to our regret, compelled to return your divine composition, and beg you a thousand times to overlook our short sight and timidity."

NATIONAL & INTERNATIONAL OLYMPIADS

On May 5, 1981, more than 100 students who scored high on the Annual High School Mathematics Contest (sponsored in part by the MAA) participated in the U.S.A. Mathematical Olympiad. The next day, in high schools across Canada, over 100 selected students participated in the Canadian Mathematical Olympiad.

This year some winners in each competition will have the chance to be part of a "first." For the Canadians, it will be the first time a team (of 8) will participate in the International Mathematical Olympiad, to be held July 13-14, 1981. The U.S.A. will also select from winners in its annual national Olympiad a team to participate in the IMO, as it has done since 1974. However, this year, for the first time, the United States will be the host country for the IMO, with Washington, D.C. the site of the examination.

Teams from 26 nations are expected to participate in this international competition which consists of six challenging questions, three presented each day, with $4\frac{1}{2}$ hours allotted each day for their solution. Both before and after the exam, sightseeing trips and social events are planned for the participants, who will be the guests of the United States from July 8-20. Award ceremonies will cap the events.

THE TENTH U.S.A. MATHEMATICAL OLYMPIAD

May 5, 1981

1. The measure of a given angle is $180°/n$ where n is a positive integer not divisible by 3. Prove that the angle can be trisected by Euclidean means (straightedge and compass).

2. Every pair of communities in a county are linked directly by exactly one mode of transportation: bus, train, or airplane. All three modes of transportation are used in the county with no community being serviced by all three modes and no three communities being linked pairwise by the same mode. Determine the maximum number of communities in the county.

3. If A, B and C are the measures of the angles of a triangle, prove that

$$-2 \leq \sin 3A + \sin 3B + \sin 3C \leq 3\sqrt{3}/2$$

and determine when equality holds.

4. The sum of the measures of all the face angles of a given convex polyhedral angle is equal to the sum of the measures of all its dihedral angles. Prove that the polyhedral angle is a trihedral angle.

Note: A convex polyhedral angle may be formed by drawing rays from an exterior point to all points of a convex polygon.

5. If x is a positive real number and n is a positive integer, prove that

$$[nx] \geq \frac{[x]}{1} + \frac{[2x]}{2} + \frac{[3x]}{3} + \ldots + \frac{[nx]}{n} ,$$

where $[t]$ denotes the greatest integer less than or equal to t. For example, $[\pi] = 3$ and $[\sqrt{2}] = 1$.

THIRTEENTH CANADIAN MATHEMATICS OLYMPIAD

May 6, 1981

1. For any real number t denote by $[t]$ the greatest integer which is less than or equal to t. For example, $[8] = 8$, $[\pi] = 3$, and $[-5/2] = -3$. Show that the equation $[x] + [2x] + [4x] + [8x] + [16x] + [32x] = 12345$ has no real solution.

2. Given a circle of radius r and a tangent line ℓ to the circle through a given point P on the circle. From a variable point R on the circle, a perpendicular RQ is drawn to ℓ with Q on ℓ. Determine the maximum of the area of triangle PQR.

3. Given a finite collection of lines in a plane P, show that it is possible to draw an arbitrarily large circle in P which does not meet any of them. On the other hand, show that it is possible to arrange an infinite sequence of lines (first line, second line, third line, etc.) in P so that every circle in P meets at least one of the lines. (A point is not considered to be a circle.)

4. $P(x)$ and $Q(x)$ are two polynomials that satisfy the identity $P(Q(x)) \equiv Q(P(x))$ for all real numbers x. If the equation $P(x) = Q(x)$ has no real solution, show that the equation $P(P(x)) = Q(Q(x))$ also has no real solution.

5. Eleven theatrical groups participated in a festival. Each day some of the groups were scheduled to perform while the remaining groups joined the general audience. At the conclusion of the festival each group had seen, during its days off, at least one performance of every other group. At least how many days did the festival last?

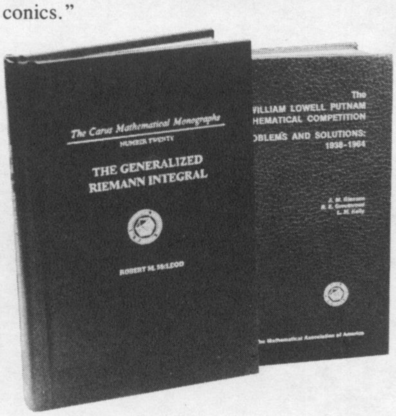
FOUR NEW BOOKS FROM THE MAA...

MAA Studies in Mathematics #20

STUDIES IN ALGEBRAIC GEOMETRY
A. Seidenberg, editor
143 + ix pages hardbound.
List: $16.00 MAA Member: $12.00

STUDIES IN ALGEBRAIC GEOMETRY, a new volume in the MAA Studies in Mathematics Series, contains expository articles which make some of the most interesting results in algebraic geometry accessible to students, as well as teachers and mathematicians who are nonspecialists or scientists working in other fields.

The contents include chapters on "initial results in the theory of linear algebraic groups," "the connectedness theorem and the concept of multiplicity," "space curves as ideal-theoretic complete intersections," and "Chasle's enumerative theory of conics."



Carus Mathematical Monographs #20

THE GENERALIZED RIEMANN INTEGRAL
by Robert M. McLeod
275 + xiii pages hardbound.
List: $18.00 MAA Member: $13.50

THE GENERALIZED RIEMANN INTEGRAL is the first book-length presentation of an exciting new development in integration theory. A new integral, a generalization of the familiar Riemann integral, has been discovered which has all the power and range of the Lebesque integral but which can be readily understood by anyone who has studied calculus through the sophomore level.

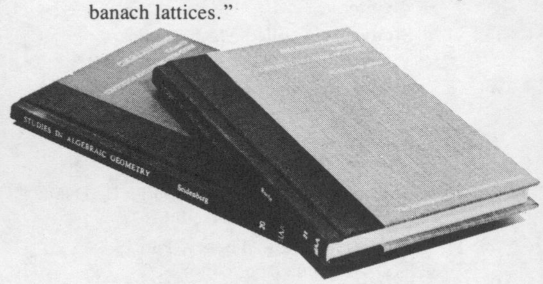
Because of the clarity and organization of the exposition and the inclusion of exercises designed to actively engage the reader in the material, it is eminently suitable for use as a textbook at the advanced undergraduate or beginning graduate level.

MAA Studies in Mathematics #21

STUDIES IN FUNCTIONAL ANALYSIS
Robert Bartle, editor
227 + xi pages hardbound.
List: $19.00 MAA Member: $14.00

STUDIES IN FUNCTIONAL ANALYSIS, a new addition to the MAA's series STUDIES IN MATHEMATICS, presents five expository articles on recent developments in functional analysis. Although these articles cannot provide a full account of the current state of the field, they do give a reader a tantilizing taste of some of the latest ideas and applications of this fascinating area of mathematics.

The contents include chapters on "numerical ranges," "projection operators in approximation theory," "complementably universal separable banach spaces; an application of counter examples to the approximation problem," "integral representation for elements of convex sets," and "aspects of banach lattices."



A New Problem Book from the MAA...

THE WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION: PROBLEMS AND SOLUTIONS: 1938-1964
A. M. Gleason, R. E. Greenwood, and L. M. Kelly
652 + xi pages hardbound.
List: $35.00 MAA Member: $26.00

Problemists the world over, including all past and future Putnam Competitors, will revel in "mastering the difficulties" posed by this collection of problems from the first 25 William Lowell Putnam Competitions. Solutions to all 347 problems are given. In some cases multiple solutions are included, some which contestants could reasonably be expected to find under examination conditions and others which are more elegant or utilize more sophisticated techniques. Valuable references and comments on many of the problems are presented.

ORDER FROM:
The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, D.C. 20036

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, N.W.
Washington, DC 20036

MATHEMATICS MAGAZINE VOL. 54, NO. 3, MAY 1981