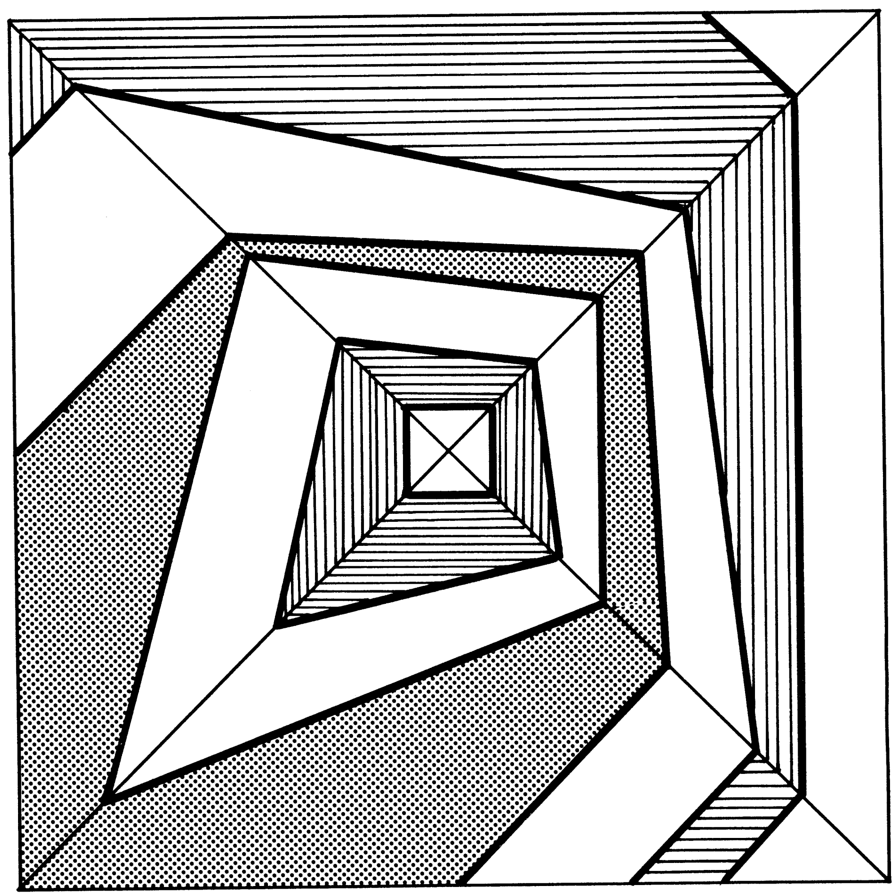
MATHEMATICS MAGAZINE



Vol. 55 No. 4
September, 1982

PYRAMIDAL SECTIONS • INTERRUPTIONS
GEOMETRIC HOMOMORPHISMS • GOAT TALES

MATHEMATICS MAGAZINE

Volume 55, Number 4, September 1982
(ISSN—0025-570X)

COVER: Taxicab ellipses and parabolas as projected sections of a pyramid. See p. 205. Design by the editor.

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

J. W. Lorimer ("Homomorphisms of Affine and Projective Planes") obtained his Ph.D. from McMaster University (Hamilton) in 1971, taught at Loyola College (Montreal) and joined the mathematics faculty of the University of Toronto in 1973 where he is an associate professor. His field of research is Hjelmslev geometry, in which homomorphisms play a fundamental role. His hobby is collecting books, especially in the areas of black literature, children's literature and the supernatural. Co-author **Martin Buchholz** received his B.Sc. in mathematics at Toronto in 1981 and is presently a graduate student there. His main areas of interest are geometry and number theory. He is also interested in chess, in which he has represented Canada several times in international competition.

Richard Laatsch ("Pyramidal Sections in Taxicab Geometry") received his Ph.D. from Oklahoma State University and is currently Professor of mathematics at Miami University in Oxford, Ohio, where he does departmental scheduling because, "I get better hours that way." Author of an abstract algebra text, he has served as vice-chairman of the MAA Oklahoma-Arkansas Section and chairman of the Ohio Section. His article is the outgrowth of a fascination with analytic geometry that he has never outgrown.

ILLUSTRATIONS

Gary Clark sketched the frazzled employee on p. 214.

William McWorter provided the computer-drawn solutions on p. 243.

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. The annual subscription price for Mathematics Magazine to an individual member of the Association is $8, included as part of the annual dues of $40. Students receive a 50% discount. Bulk subscriptions (5 or more copies to a single address) are available to colleges and universities for distribution to undergraduate students at a 35% discount. The library subscription price is $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

Homomorphisms of Affine and Projective Planes

An elementary treatment of homomorphisms between affine and projective planes leads to characterizations of epimorphisms, monomorphisms, and isomorphisms of these geometries.

J. W. LORIMER
M. BUCHHOLZ
University of Toronto
Toronto, Ontario, Canada, M5S 1A1

In most undergraduate mathematics courses a discussion of structure-preserving maps or homomorphisms is customary. Usually surjective homomorphisms or epimorphisms and their associated quotient structures are considered first, then isomorphisms and the standard isomorphism theorems are discussed. In geometry courses, however, only isomorphisms or collineations are studied. The reason for this is rooted historically in Felix Klein's Erlangen program where different geometries were classified via their collineations or transformation group. That is, one was first given the isomorphisms and then had to discover what structure they were preserving. Although it is customary in today's geometry courses to give the geometric structure first, this is not followed by a discussion of homomorphisms. The purpose of this article is to fill this gap and discuss homomorphisms of geometries, expressly those of affine and projective planes.

Geometries and homomorphisms

A **geometry** is a triple $\langle \mathbb{P}, \mathbb{L}, I \rangle$ where $\mathbb{P}$ and $\mathbb{L}$ are disjoint sets and $I \subseteq \mathbb{P} \times \mathbb{L}$. The elements of $\mathbb{P}$, denoted $P, Q, R, \ldots$, are **points**, the elements of $\mathbb{L}$, denoted $a, b, c, \ldots$, are **blocks**, and I is the **incidence relation**. PIl means P "is incident with" or "lies on" the block l. The classical Euclidean, hyperbolic, projective, and conformal planes are all geometries in this sense where blocks are called lines in the first three cases and circles in the last.

It is now apparent how to define a structure-preserving map or homomorphism between two geometries: it must map points to points, blocks to blocks, and **preserve incidence** (if P lies on l, then the image of P lies on the image of l).

To be precise, given two geometries $G = \langle \mathbb{P}, \mathbb{L}, I \rangle$ and $G' = \langle \mathbb{P}', \mathbb{L}', I' \rangle$, a **homomorphism** [8] from G to G' is a pair $\alpha = (\alpha_1, \alpha_2)$ satisfying the following properties:

(H1) $\alpha_1 \colon \mathbb{P} \to \mathbb{P}'$ and $\alpha_2 \colon \mathbb{L} \to \mathbb{L}'$ are functions.
(H2) α preserves incidence; that is, if PIl, then $\alpha_1(P) I' \alpha_2(l)$ for all point-line pairs (P, l) of G.

In general we write I for I' and α for α_1 and α_2. A homomorphism α is a **monomorphism** (**epimorphism**) if α_1 and α_2 are injective (surjective). Finally, α is an **isomorphism** if α_1 and α_2 are bijections and $(\alpha_1^{-1}, \alpha_2^{-1})$ is a homomorphism from G' to G.

For affine and projective planes, an isomorphism can also be described as a point bijection between two planes with the property that whenever P, Q, R lie on a line then so do their images. For each such bijection, α_1, uniquely determines a line bijection α_2 so that (α_1, α_2) is an isomorphism. This, in fact, was the classical way to describe isomorphisms and has led some authors (see [**5**], [**10**], [**18**]) to consider **lineations** or point functions from one plane into another with the above collinearity-preserving property. Clearly, for every homomorphism, the point map is a lineation. However, not every lineation is associated with a unique homomorphism; for example, map all points to one point.

The notion of a homomorphism for geometries appears to have originated simultaneously with the concept of a free projective plane (see [**11**, p. 230] and [**23**, p. 58]). These articles generated a series of papers on homomorphisms of projective planes ([**14**], [**12**], [**7**], [**22**], [**20**], [**15**], [**1**], [**16**]) and later on affine planes ([**6**], [**15**; §2], [**2**; §2]). However, the major results from these papers are either proved using concepts not normally studied in an introductory course—ternary rings, valuation rings, topological planes—or else are immersed in more general settings like projective designs and Hjelmslev planes.

The remainder of this article provides a discussion of homomorphisms of projective and affine planes suitable for an introductory course. Indeed, except for the notion of a free projective plane, we utilize no information other than the axioms of these planes.

Homomorphisms of projective planes

We begin by recalling the definition of a projective plane and introducing some useful notation.

A **projective plane** is a geometry $\mathcal{P} = \langle \mathbb{P}, \mathbb{L}, I \rangle$ satisfying:

(P1) Two points, P, Q, are incident with exactly one line denoted by $P \vee Q$ (P join Q).
(P2) Two lines, l, m, are incident with exactly one point denoted by $l \wedge m$ (l meet m).
(P3) There exists a quadrangle; that is, four points no three of which are collinear.

If a geometry $\mathcal{P}$ satisfies only (P1) and (P2), it is called a **degenerate projective plane**.

It is clear that a homomorphism between projective planes preserves the operations "join" and "meet." In precise terms, if α: $\mathcal{P} \to \mathcal{P}'$ is a homomorphism from a projective plane $\mathcal{P}$ to a projective plane $\mathcal{P}'$, then

(H3) If P, Q are two points with $\alpha(P) \neq \alpha(Q)$, then $\alpha(P \vee Q) = \alpha(P) \vee \alpha(Q)$.
(H4) If l, m are two lines with $\alpha(l) \neq \alpha(m)$, then $\alpha(l \wedge m) = \alpha(l) \wedge \alpha(m)$.

Thus, for homomorphisms of projective planes, the preservation of join and meet replaces the preservation of addition and multiplication for algebraic homomorphisms.

We now present some examples of homomorphisms of projective planes.

EXAMPLE 1. Given a field F we describe the projective plane, $P(F)$ over F, via homogeneous coordinates as follows: the point with coordinates x, y, z is denoted $\langle xyz \rangle$, lines are denoted $[uvw]$, and incidence is defined by

$$\langle xyz \rangle I [uvw] \Leftrightarrow xu + yv + zw = 0.$$

Each nonsingular 3×3 matrix A over F with transpose A^t induces an isomorphism α_A: $P(F) \to P(F)$ where

$$\alpha(\langle xyz \rangle) = \langle (xyz)A^t \rangle \quad \text{and} \quad \alpha([uvw]) = [(uvw)A^{-1}].$$

Our next three examples are proper homomorphisms of projective planes.

EXAMPLE 2. If ψ: $F \to F$ is a nontrivial field homomorphism but not an isomorphism, then it is a proper monomorphism and induces a similar geometric homomorphism

$$\alpha_\psi \begin{cases} \langle xyz \rangle \mapsto \langle \psi(x)\psi(y)\psi(z) \rangle \\ [uvw] \mapsto [\psi(u)\psi(v)\psi(w)]. \end{cases}$$

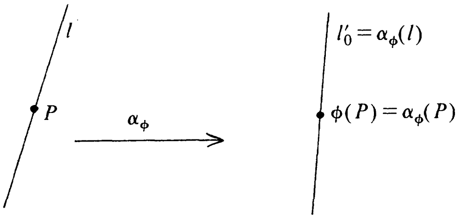


FIGURE 1

EXAMPLE 3. Given two projective planes $\mathcal{P}$ and $\mathcal{P}'$, choose a fixed line l_0' in $\mathcal{P}'$ and any function ϕ: $\mathbb{P} \to l_0'$. Then the map α_ϕ which sends a point P of $\mathbb{P}$ to the point $\phi(P)$ and every line l of $\mathbb{L}$ to l_0' is a homomorphism (see FIGURE 1). In the case where ϕ is injective, α_ϕ gives us an example of a homomorphism which is injective on points, but not on lines.

EXAMPLE 4. Given a projective plane $\mathcal{P}$, choose a fixed point P_0 and a fixed line l_0 with P_0 not on l_0. Then

$$\alpha_1\begin{cases} P \mapsto P_0 & \text{if } P \notin l_0, \quad P \in \mathbb{P} \\ P \mapsto P & \text{if } P \in l_0, \quad P \in \mathbb{P} \end{cases}$$

and

$$\alpha_2\begin{cases} l \mapsto (l_0 \wedge l) \vee P_0 & \text{if } l \neq l_0, \quad l \in \mathbb{L} \\ l \mapsto l_0 & \text{if } l = l_0, \quad l \in \mathbb{L} \end{cases}$$

determine a homomorphism of $\mathcal{P}$ with itself (see FIGURE 2).

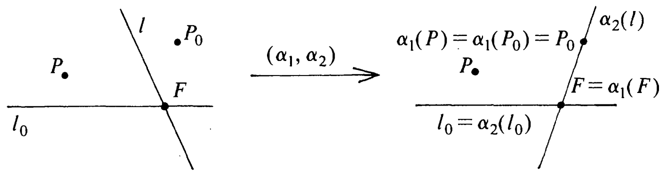


FIGURE 2

Our first question concerning these homomorphisms is naturally, "Is the homomorphic image of a projective plane also a projective plane?" Examples 3 and 4 clearly illustrate that the answer is *no*. In these cases the image plane is a degenerate projective plane, i.e., only (P1) and (P2) are valid. Since properties (H3) and (H4) guarantee that the image plane always satisfies (P1) and (P2) we call a homomorphism **nondegenerate** or **degenerate** depending on whether the image plane satisfies (P3) or not.

One way to ensure that a homomorphism α is nondegenerate is to assume that it **reflects incidence**; that is, $\alpha(P)I\alpha(l)$ implies PIl. We next observe that this property characterizes the monomorphisms.

THEOREM 1. *A homomorphism of projective planes is a monomorphism if and only if it reflects incidence. Hence every monomorphism is nondegenerate.*

Proof. Let α: $\mathcal{P} \to \mathcal{P}'$ be a monomorphism of projective planes, and suppose $\alpha(P)I\alpha(l)$. Choose RIl so that $R \neq P$. Then $\alpha(P) \neq \alpha(R)$ and $\alpha(R)I\alpha(l)$. By (H3), $\alpha(P \vee R) = \alpha(P) \vee \alpha(R) = \alpha(l)$ and so $P \vee R = l$. Hence, PIl.

Conversely, assume α reflects incidence. First we show that α is injective on points. Suppose $\alpha(P) = \alpha(Q)$. Now choose distinct lines l, m through P. Then, $\alpha(Q)$ lies on $\alpha(l)$ and $\alpha(m)$, which implies that Q lies on l and m. Consequently $Q = l \wedge m = P$. A dual argument shows α is injective on lines.

The final claim now follows easily.

Epimorphisms of projective planes

If $\alpha=(\alpha_1,\alpha_2)$ is a homomorphism of projective planes, then Example 3 and its dual show that the injectivity of α_1 does not imply the injectivity of α_2 and conversely. However, we can prove the following result.

THEOREM 2. *A homomorphism of projective planes $\alpha\colon \mathcal{P}\to\mathcal{P}'$ is surjective on points if and only if it is surjective on lines. Moreover, if it is surjective on either points or lines, then it is injective on points if and only if it is injective on lines.*

Proof. Suppose α is surjective on points. Then, if l' is any line of $\mathcal{P}'$ and $l'=P'\vee Q'$ we can choose $P, Q (P\neq Q)$ so that $\alpha(P\vee Q)=l'$ by (H3). By a dual argument, surjectivity on lines implies surjectivity on points.

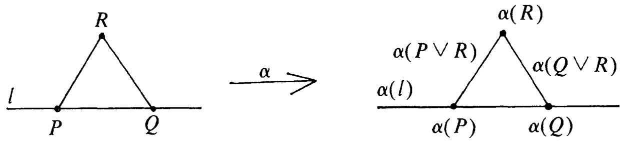


FIGURE 3

Next suppose α is surjective on points and injective on lines. We can then show that α is injective on points. For, if $P\neq Q$ and $l=P\vee Q$, we can choose a point R' not on $\alpha(l)$ such that $R'=\alpha(R)$. Then R does not lie on $P\vee Q$ and $\alpha(R)$ is distinct from $\alpha(P)$ and $\alpha(Q)$. Also $R\vee P\neq R\vee Q$ (see FIGURE 3). Consequently (H3) and the injectivity on $\mathbb{L}$ implies that

$$\alpha(R)\vee\alpha(P)=\alpha(R\vee P)\neq\alpha(R\vee Q)=\alpha(R)\vee\alpha(Q).$$

Hence, $\alpha(P)\neq\alpha(Q)$. A dual argument shows that injectivity on points implies injectivity on lines.

Next we shall observe that epimorphisms of projective planes are either isomorphisms or are very much noninjective in the sense that every point in the image plane has infinitely many preimages. This fact was first demonstrated by Hughes [**12**] in 1960. However, the proof used ternary rings. We shall follow the elementary geometric argument of Mortimer [**17**] given in 1975.

THEOREM 3. *Every epimorphism $\alpha\colon\mathcal{P}\to\mathcal{P}'$ of a projective plane $\mathcal{P}$ onto a projective plane $\mathcal{P}'$ is either an isomorphism or every point of $\mathcal{P}'$ has infinitely many preimages.*

Proof. First we observe that every line of $\mathcal{P}$ has three points incident with it whose images are distinct. Otherwise, since α is surjective, every line of $\mathcal{P}'$ would be incident with one or the other of two points, which cannot occur in a nondegenerate projective plane. Dually, every point has three lines incident with it which have distinct images.

Now, suppose α is not an isomorphism and there is a point $\alpha(A)$ so that $\alpha^{-1}(\alpha(A))$ is finite. Of the lines through A, choose l_1 such that l_1 contains a maximal number of preimages of $\alpha(A)$. Let $Q_1, Q_2,\ldots,Q_n=A$ be all the points of l_1 with $\alpha(Q_i)=\alpha(A)$ for $i=1,2,\ldots,n$. By the above remarks there exists a point B on l_1 with $\alpha(B)\neq\alpha(A)$, and there is a line l_2 through B with $\alpha(l_2)\neq\alpha(l_1)$. Also, there is a point P_1 on l_2 with $\alpha(P_1)\neq\alpha(B)$, and a point O on $P_1\vee Q_1$ with $\alpha(P_1)\neq\alpha(O)$ and $\alpha(O)\neq\alpha(Q_1)$ (see FIGURE 4). For $i=2,\ldots,n$ let $P_i=l_2\wedge(O\vee Q_i)$. Thus $\alpha(l_2)\neq\alpha(Q_i)\vee\alpha(O)$ since equality would imply that $\alpha(Q_i)$ lies on $\alpha(l_2)$, which is distinct from $\alpha(l_1)$, so

$$\alpha(Q_i)=\alpha(l_1)\vee\alpha(l_2)=\alpha(B),$$

contrary to the choice of B. Hence

$$\alpha(P_i)=\alpha(l_2)\vee(\alpha(Q_i)\vee\alpha(O))=\alpha(l_2)\wedge(\alpha(Q_j)\vee\alpha(O))=\alpha(P_j)$$

for $1\leqslant i,j\leqslant n$. So all the P_i's have the same image for $1\leqslant i\leqslant n$.

There is a point C on l_2 with $\alpha(P_i)\neq\alpha(C)$ and $\alpha(C)\neq\alpha(B)$. Let C_i be the intersection of the lines $A\vee C$ and $O\vee P_i$ for $i=1,\ldots,n$. Then, interchanging P_i for Q_i, l_2 for l_1, C for B, and $A\vee C$

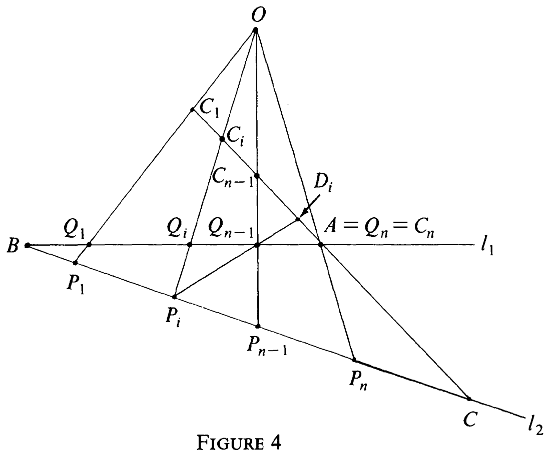


FIGURE 4

for l_2 in the above argument gives $\alpha(A \vee C) \neq \alpha(O) \vee \alpha(P_i)$. Thus $\alpha(C_i) = \alpha(A \vee C) \wedge (\alpha(O) \vee \alpha(P_i)) = \alpha(A)$ for $i = 1, 2, \ldots, n$.

Now, for $i = 1, 2, \ldots, n$, put $D_i = (P_i \vee Q_{n-1}) \wedge (C \vee A)$. We have $\alpha(P_i \vee Q_{n-1}) \neq \alpha(C \vee A)$, since equality would imply that $\alpha(P_i) \vee \alpha(C) = \alpha(C \vee A) = \alpha(l_2)$ which means that

$$\alpha(A) = \alpha(l_1) \wedge \alpha(A \vee C) = \alpha(l_1) \wedge \alpha(l_2) = \alpha(B),$$

contrary to the choice of B. Thus $\alpha(D_i) = \alpha(D_j)$ for $1 \leq i, j \leq n$. But $D_{n-1} = C_{n-1}$ and so $\alpha(D_i) = \alpha(A)$ for $i = 1, \ldots, n$. However, we already have n points on $A \vee C$ whose image is $\alpha(A)$, namely, $C_1, \ldots, C_n$. So each D_i is a C_j for some j. Clearly the D_i's are distinct and $D_i \neq C_n$ for $i = 1, \ldots, n$. Thus there are n points D_i to distribute among $n - 1$ points C_j (since $C_{n-1} = D_{n-1}$), a contradiction. The proof is complete.

As an immediate consequence of the above result we have the following amazing fact!

THEOREM 4. *Every epimorphism of finite projective planes is an isomorphism.*

This fact was first proved by Hughes [**12**] and subsequently other proofs were given by Dembowski [7], André [1], Breitspecher [4], and Drake-Lenz [9]. However, all these proofs involve advanced concepts; for example, projective designs [7], ternary rings [1], [**12**], topological planes [4], and finite Klingenberg planes [9].

The preceding theorems immediately imply the following characterizations of isomorphisms.

THEOREM 5. *A homomorphism α of projective planes is an isomorphism if and only if one of the following conditions holds*:

(1) *α is surjective on points and reflects incidence*;
(2) *α is surjective on points and injective on lines*;
(3) *α is an epimorphism and each point in the image plane has at most finitely many preimages*;
(4) *α is injective on points and surjective on lines*;
(5) *α is bijective on lines*;
(6) *α is bijective on points.*

As mentioned before, homomorphic images of projective planes are not necessarily projective planes. The previous theorems show that images of monomorphisms and epimorphisms are projective planes.

Theorem 4 shows that proper epimorphisms of finite projective planes do not exist. We shall also see later (Theorem 8) that proper epimorphisms between affine planes that map parallel lines to parallel lines do not exist!

It is thus natural to ask whether proper, nondegenerate homomorphisms between projective planes exist. If we choose a proper field monomorphism σ (for instance K is a simple transcendental extension and $x \overset{\sigma}{\mapsto} x^2$) then Example 2 produces a proper monomorphism of projective planes. A proper epimorphism can be constructed using the concept of a **free projective plane** introduced in 1943 by M. Hall in [**11**]. Suppose $G = \langle \mathbb{P}, \mathbb{L}, I \rangle$ is a geometry so that any two points lie on at most one line, every line passes through at least two points, and G contains four points, no three of which are collinear. If G is not a projective plane, then either two points have no joining line or two lines fail to intersect. To make G a projective plane, we must add either the missing lines or the missing points. Such additions may introduce new nonintersecting pairs of lines or noncollinear pairs of points, so further additions may be necessary. If the process is continued indefinitely, the free projective plane generated by G is produced. In the study of such planes it can be shown that

Every projective plane is the homomorphic image of a free projective plane

([**13**, 11.14 corollary] or [**23**, p. 65]).

Since free planes are infinite, the above result shows that there are proper epimorphisms for infinite planes. Other examples involving Desarguesian or cyclic planes may be found in [**16**] or [**19**] respectively.

The existence of proper epimorphisms for (infinite) projective planes enables us to introduce the next concept. If α: $\mathcal{P} \to \mathcal{P}'$ is an epimorphism of projective planes, then α induces an equivalence relation on points (lines) where two points (lines) are defined to be equivalent if they have the same images under α. Let $\bar{P}$ and $\bar{l}$ denote equivalence classes of points and lines in $\mathcal{P}$ respectively. Then the corresponding quotient spaces determine a **quotient geometry** $\bar{\mathcal{P}}$ with incidence defined by: $\bar{P}I\bar{l}$ if and only if there exist $Q \in \bar{P}$ and $m \in \bar{l}$ with QIm. It is then straightforward to verify that

The quotient geometry $\bar{\mathcal{P}}$ is a projective plane and the correspondence $P \to \bar{P}$ and $l \to \bar{l}$ is an epimorphism from $\mathcal{P}$ to $\bar{\mathcal{P}}$.

This result generates, in the obvious manner, a geometric analogue of the first isomorphism theorem for groups and rings.

If α: $\mathcal{P} \to \mathcal{P}'$ is an epimorphism of projective planes, then $\bar{\mathcal{P}}$ is isomorphic to $\mathcal{P}'$.

We next reexamine our results on projective planes for affine planes and obtain two surprising and contrasting results (Theorems 6 and 8).

Homomorphisms of affine planes

If $\mathcal{P}$ is a projective plane and l is any line, then the subgeometry $\mathcal{P}_l = \langle \mathbb{P} \setminus l, \mathbb{L} \setminus \{l\}, I \rangle$ is the **affine plane of $\mathcal{P}$ with l its line at infinity**. Two lines, a and b, of $\mathcal{P}_l$ are **parallel**, denoted $a \| b$, if $a = b$ or their point of intersection lies on l (i.e., $(a \wedge b) \wedge (\mathbb{P} \setminus l) = \varnothing$). If l is a fixed line of $\mathcal{P}$, then any homomorphism α: $\mathcal{P} \to \mathcal{P}'$ of projective planes induces a homomorphism α_l between the affine planes $\mathcal{P}_l$ and $\mathcal{P}'_{\alpha(l)}$ *provided $\alpha(X)I\alpha(l)$ implies XIl*. Then, α_l has the additional property that it **preserves parallelism**; that is, if a and b are parallel lines in $\mathcal{P}_l$, then their images are parallel lines in $\mathcal{P}'_{\alpha(l)}$.

Since parallelism is an intrinsic property of affine planes, it is only natural to study those homomorphisms of affine planes that preserve parallelism.

We recall that a geometry $\mathcal{Q} = \langle \mathbb{P}, \mathbb{L}, I \rangle$ is an **affine plane** if it satisfies the following axioms:

(A1) Two points lie on exactly one line.
We define two lines, a and b, to be parallel if $a = b$ or $a \wedge b = \varnothing$.
(A2) Given a point P and a line l, there exists a unique line incident with P and parallel to l; this line is denoted $L(P, l)$.
(A3) There exists a triangle or three noncollinear points.

A geometry $\mathcal{Q}$ which satisfies only (A1) and (A2) is called a **degenerate affine plane**.

Because of our prior discussion of homomorphisms for affine planes, we call a homomorphism

α: $\mathcal{A} \to \mathcal{A}'$, between affine planes an **affine homomorphism** (**A-homomorphism** for short) if it satisfies the additional property:

(H5) If $a \| b$, then $\alpha(a) \| \alpha(b)$ for all $a, b \in \mathbb{L}$.
Equivalently, for any point-line pair (P, l), $\alpha(L(P, l)) = L(\alpha(P), \alpha(l))$.

It is well known that each affine plane $\mathcal{A}$ has a unique (up to isomorphism) extension to a projective plane, obtained by adding a line l_∞ at infinity. We denote this extension of $\mathcal{A}$ by $\mathcal{P}(\mathcal{A})$. The significance of A-homomorphisms for affine planes is illustrated by the next observation:

A homomorphism α: $\mathcal{A} \to \mathcal{A}'$ of affine planes has a (unique) homomorphic extension $\mathcal{P}(\alpha)$: $\mathcal{P}(\mathcal{A}) \to \mathcal{P}(\mathcal{A}')$ which maps l_∞ to l'_∞ if and only if α is an A-homomorphism.

$\mathcal{P}(\alpha)$ maps l_∞ to l'_∞ and $X_\infty I l_\infty$ to $\alpha(m) \wedge l'_\infty$ where $X_\infty = l_\infty \wedge m$. It is easy to see that there is only one such extension. However Examples 5 and 6 below show that in general, homomorphisms can have more than one projective extension.

EXAMPLE 5. If P_0 is a fixed point of an affine plane $\mathcal{A}$, and ϕ is any map from $\mathbb{L}$ to the pencil of lines incident with P_0, then the function α_ϕ which maps every point to P_0 and every line to $\phi(l)$ is a homomorphism. The function α_ϕ is an A-homomorphism only if parallel lines have the same image under ϕ. Observe that we can extend α_ϕ to a homomorphism of $\mathcal{P}(\mathcal{A})$ fixing l_∞ only if α_ϕ is an A-homomorphism. In any case we can extend α_ϕ to a homomorphism distinct from $\mathcal{P}(\alpha)$ by mapping l_∞ to some fixed line through P_0 and each point at infinity to P_0.

EXAMPLE 6. Suppose that $\mathcal{A}$ and $\mathcal{A}'$ are affine planes and l'_0 is a fixed line of $\mathcal{A}'$. If there exists a bijection ϕ from the points of $\mathcal{A}$ onto l'_0, then the function α_ϕ which maps each point P to $\phi(P)$ and each line l to l'_0 is an A-homomorphism. The obvious projective extension is distinct from $\mathcal{P}(\alpha_\phi)$.

EXAMPLE 7. We are given l_0 and m_0 fixed nonparallel lines of an affine plane $\mathcal{A}$. Then the projection along a parallel pencil, which maps each point P to the intersection of l_0 and the line through P parallel to m_0 and each line l to itself if l is parallel to m_0 and to l_0 otherwise, is an A-homomorphism (see FIGURE 5).

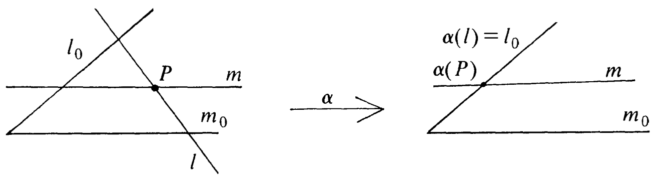


FIGURE 5

Clearly the homomorphic image of an affine plane satisfies (A1) and (A2). However, as the above three examples demonstrate, it need not satisfy (A3). Hence, we call a homomorphism of affine planes **nondegenerate** or **degenerate** depending on whether the image geometry satisfies (A3) or not. Clearly every epimorphism between affine planes is nondegenerate.

We can show that our Examples 5, 6, 7 essentially describe all possible degenerate A-homomorphisms. That is, any degenerate A-homomorphism α: $\mathcal{A} \to \mathcal{A}'$ between affine planes must map all points into a fixed line $\alpha(l_0)$ so that one of the following situations occurs:

(A) *α is constant on points* (Example 5);
(B) *α is not constant on points, but is constant on lines* (Example 6);
(C) *There exists a line m_0 not parallel to l_0 so that $\alpha(l_0) \nparallel \alpha(m_0)$, α is constant on the lines not parallel to m_0 and not constant on the lines parallel to m_0* (Example 7).

To prove this, we argue as in [**2**, page 42]. Suppose $\alpha\colon \mathcal{A} \to \mathcal{A}'$ is degenerate. Then, if α is constant on points, we can choose l_0 to be any line of $\mathcal{A}$ and case (A) holds. If (A) is false, there exists Q so that $\alpha(Q) \neq \alpha(P)$. Put $l_0 = P \vee Q$. Then $\alpha(X) I \alpha(l_0)$ for all points X. If α is constant on lines we have case (B), and if not, there is a line m_0 so that $\alpha(m_0) \neq \alpha(l_0)$. Then YIm_0 implies $\alpha(Y)$ lies on $\alpha(m_0)$ and $\alpha(l_0)$ and so $\alpha(m_0) \nparallel \alpha(l_0)$. Finally, by (H3), α is not constant on the lines parallel to m_0, since $\alpha(L(P, m_0)) \neq \alpha(L(Q, m_0))$. In addition, α is constant on the lines not parallel to m_0, as $m \nparallel m_0$ implies, m meets both $L(P, m_0)$ and $L(Q, m_0)$, yielding $\alpha(m) = \alpha(l_0)$.

We say a homomorphism $\alpha\colon \mathcal{A} \to \mathcal{A}'$ of affine planes **reflects parallelism** if $\alpha(l) \parallel \alpha(m)$ implies $l \parallel m$ for all lines l, m. This is not true in general even for A-homomorphisms (see Example 6 above). But we can show:

(H6) (a) *If the images of two lines, under a homomorphism, are distinct and parallel, then the lines themselves are distinct and parallel.*
(b) *A nondegenerate A-homomorphism reflects parallelism.*

To see (a), suppose $\alpha(l) \neq \alpha(m)$ and $\alpha(l) \parallel \alpha(m)$. If $l \nparallel m$, then $l \wedge m = P$. Hence $\alpha(P)$ lies on both $\alpha(l)$ and $\alpha(m)$ and so $\alpha(l) = \alpha(m)$, a contradiction.

To see (b), suppose α is a nondegenerate A-homomorphism and assume $l \nparallel m$. Then m meets all lines parallel to l and so $\alpha(m)$ meets all the lines $\alpha(x)$ where x is parallel to l. If P_1, P_2, P_3 are three points whose images are noncollinear, then all three lines $\alpha(L(P_i, l))$ $(i = 1, 2, 3)$ cannot be equal. Hence, $\alpha(m)$ meets one of them exactly once, and thus by (H5) meets $\alpha(l)$ once also. Consequently $\alpha(l) \nparallel \alpha(m)$.

Monomorphisms and epimorphisms of affine planes

As with projective planes, we will show that monomorphisms of affine planes are exactly those homomorphisms which reflect incidence (Theorem 1). Hence every monomorphism is a nondegenerate A-homomorphism. We can also show that, unlike the projective case, the converse of the previous statement is also true. Hence, an affine plane has no nonisomorphic (affine plane) images!

THEOREM 6. *A homomorphism α of affine planes is a monomorphism precisely when one of the following conditions holds*:

(1) *α is an A-homomorphism and injective on lines*;
(2) *α reflects incidence*;
(3) *α is a nondegenerate A-homomorphism.*

Proof. Every monomorphism satisfies (1). To see that $(1) \Rightarrow (2)$, suppose $\alpha(P) I \alpha(l)$. Then by (H5), $\alpha(l) = L(\alpha(P), \alpha(l)) = \alpha(L(P, l))$. The injectivity on lines yields $l = L(P, l)$ and so PIl. The implication $(2) \Rightarrow (3)$ is immediate from the comments preceding the theorem.

Finally, we show that every nondegenerate A-homomorphism α is injective on points and lines. First we establish the injectivity on points (see FIGURE 6). Let $P \neq R$ and put $g = P \vee R$. Since α is nondegenerate, we may choose Q so that $\alpha(Q) \not I \alpha(g)$. Thus $Q \not I g$ and $Q \neq P, R$. Hence, $P \vee Q \nparallel R \vee Q$ and so (H6) (b) implies $\alpha(P \vee Q) \nparallel \alpha(R \vee Q)$. Now $\alpha(P)$ and $\alpha(R)$ lie on $\alpha(g)$ and hence $\alpha(Q)$ is distinct from $\alpha(P)$ and $\alpha(R)$. Consequently, $\alpha(P) \vee \alpha(Q) = \alpha(P \vee Q) \neq \alpha(R \vee Q) = \alpha(R) \vee \alpha(Q)$ and thus $\alpha(P) \neq \alpha(R)$. This verifies that α is injective on points.

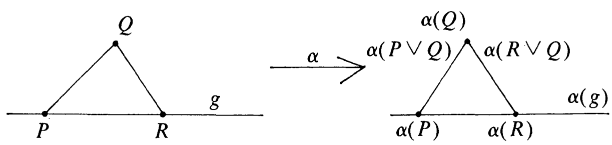


FIGURE 6

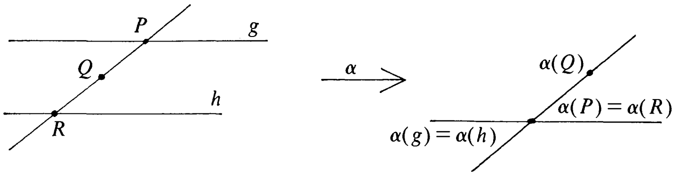


FIGURE 7

We next show that α is injective on lines (see FIGURE 7). Suppose $\alpha(g)=\alpha(h)$. By (H6) (b), $g \| h$. Again, by nondegeneracy, choose Q so that $\alpha(Q) \not I \alpha(g)$. Hence $Q \not I g$. Select $P I g$. Then $P \neq Q$ and $h \nparallel P \vee Q$. Let $R=h \wedge(P \vee Q)$. Since $\alpha(P) \neq \alpha(Q)$ and $\alpha(P) I \alpha(g)$, we conclude that $\alpha(P) \vee \alpha(Q)=\alpha(P \vee Q) \nparallel \alpha(g)=\alpha(h)$. Hence, $\alpha(R)=\alpha(h) \wedge \alpha(P \vee Q)=\alpha(g) \wedge \alpha(P \vee Q)=\alpha(P)$. Since α is injective on points, $R=P$. Hence P lies on g and h and, since $g \| h$, we conclude that $g=h$. The proof is now complete.

As an analogue to Theorem 2 in the projective case we have:

THEOREM 7. *A homomorphism α of affine planes is an epimorphism if it is surjective on points. Moreover, if α is an epimorphism and injective on points then it is also injective on lines.*

The proof of the first statement is the same as in Theorem 2. But the verification of the second assertion is a little trickier: the problem is that we do not know if α preserves parallelism. However, we can utilize the fact (H6) (a) that α "almost" reflects parallelism to show our claim.

Proof. Suppose α is an epimorphism and injective on points. We assume $l \neq m$ and show $\alpha(l) \neq \alpha(m)$. There are two cases to consider.

CASE 1: $l \wedge m=P$. Let $k' \neq \alpha(l)$ and $k' \| \alpha(l)$. For some k, $k'=\alpha(k)$ and so (H6) (a) implies that $k \| l$. Hence $k \nparallel m$ and so k' and $\alpha(m)$ have a point in common, whereas k' does not meet $\alpha(l)$. Consequently $\alpha(m) \neq \alpha(l)$.

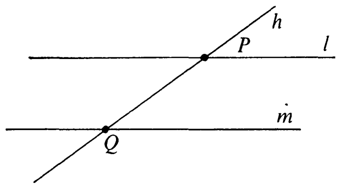


FIGURE 8

CASE 2: $l \| m$ (see FIGURE 8). Choose $h \nparallel m$ and put $P=h \wedge l$, $Q=h \wedge m$. Then $P \neq Q$ since $l \neq m$. By Case 1 and the injectivity of α on points it follows that $\alpha(l) \neq \alpha(h)=\alpha(P) \vee \alpha(Q)$. Since $\alpha(P) I \alpha(l)$, it follows that $\alpha(Q) \not I \alpha(l)$. Because $\alpha(Q) I \alpha(m)$, we conclude $\alpha(l) \neq \alpha(m)$.

Since every epimorphism between affine planes is obviously nondegenerate, Theorem 6 immediately produces the startling fact that *every A-epimorphism is an isomorphism*! This result was first observed by Corbas [**6**], but the proof has an error, as noticed and corrected by Bacon ([**2**], p. 238). We end our discussion by listing this and other characterizations of isomorphisms of affine planes. The proof follows from our previous theorems.

THEOREM 8. *A homomorphism α of affine planes is an isomorphism precisely when one of the following conditions holds*:

(1) *α is an epimorphism which reflects incidence.*
(2) *α is an A-epimorphism.*
(3) *α is an A-homomorphism which is surjective on points.*
(4) *α is a homomorphism which is surjective on points and injective on lines.*
(5) *α is a homomorphism which is bijective on points.*

The following simple consequence of part (6) of Theorem 5 and part (5) of Theorem 8 is useful in constructing isomorphisms of affine and projective planes.

If G and G′ are affine (projective) planes and α_1 is a bijection from the points of G to the points of G′ so that the images of any three collinear points are themselves collinear, then there exists a unique bijection α_2 from the lines of G to the lines of G′ so that (α_1, α_2) is an isomorphism.

The collinearity-preserving property of α_1 ensures that the map $\alpha_2\colon \mathbb{L} \to \mathbb{L}'$ which sends each line l to the line $\alpha_1(P) \vee \alpha_1(Q)$, where P and Q are distinct points on l, is well defined and (α_1, α_2) is an isomorphism. Moreover, if $(\alpha_1, \tilde{\alpha}_2)$ is also an isomorphism, then (H3) implies that $\tilde{\alpha}_2 = \alpha_2$.

In affine and finite projective planes, the absence of proper epimorphisms precludes the possibility of quotient structures. It is perhaps worth noting, however, that in the theory of Hjelmslev and Klingenberg planes (generalizations of ordinary affine and projective planes where two distinct points may lie on no common line or more than one line), proper epimorphisms play a fundamental role (see [**3**], [**24**], and [**25**]).

This research was supported by the Natural Sciences and Engineering Research Council of Canada, and particularly for the second author by a NSERC undergraduate Summer research award.

References

[1] J. André, Über homomorphismen projektiver ebenen, Abh. Math. Sem. Univ. Hamburg, 34 (1969) 98–114.
[2] P. Bacon, Coordinatized Hjelmslev Planes, Ph.D. thesis, University of Florida, 1973.
[3] ______, An introduction to Klingenberg Planes, published by the author, Florida, 1976.
[4] S. Breitspecher, Uniform projektive ebenen, Math. Z., 95 (1967) 139–168.
[5] D. S. Carter and Andrew Vogt, Collinearity-preserving functions between Desarguesian planes, Memoirs of the Amer. Math. Soc., Sept. 1980, vol. 27, 235.
[6] V. Corbas, La non esistenza di amomorfismi propri fra piani affini, Rend. Mat., 24 (1965) 373–376.
[7] P. Dembowski, Homomorphismen von λ-ebenen, Arch. Math. Sem. Univ. Hamburg, 34 (1959) 98–114.
[8] ______, Finite Geometries, Springer-Verlag, New York, 1968.
[9] D. A. Drake and H. Lenz, Finite Klingenberg planes, Abh. Math. Sem. Univ. Hamburg, 44 (1975) 70–83.
[10] L. E. Garner, Fields and projective planes: a category equivalence, Rocky Mountain J. Math., 2 (1972) 605–610.
[11] M. Hall, Projective planes, Trans. Amer. Math. Soc., 54 (1943) 229–277.
[12] D. R. Hughes, On homomorphisms of projective planes, Proc. Symp. Appl. Math., 10 (1960) 45–52.
[13] D. R. Hughes and F. Piper, Projective Planes, Springer-Verlag, New York, 1972.
[14] W. Klingenberg, Projektive geometrien mit homomorphismus, Math. Ann., 132 (1956) 180–200.
[15] J. W. Lorimer and N. D. Lane, Morphisms of affine Hjelmslev planes, Atti Accad. Naz. Lincei Rend., (1974) 880–885.
[16] K. Mathiak, Eine geometrische kennzeichnung von homomorphismen desarguesscher projektiver ebenen, Math. Z., 98 (1967) 257–267.
[17] Brian Mortimer, A geometric proof of a theorem of Hughes on homomorphisms of projective planes, Bull. London Math. Soc., 7 (1975) 267–268.
[18] F. Radó, Non-injective collineations on some sets in Desarguesian projective planes and extensions of non-commutative valuations, Aequationes Math., 4 (1970) 307–321.
[19] D. Row, Homomorphisms of sharply transitive projective planes, Bull. Austral. Math. Soc., 4 (1971) 361–366.
[20] H. Salzmann, Homomorphismen topologischer projektiver ebenen, Arch. Math., 10 (1959) 51–55.
[21] P. Scherk and R. Lingenberg, Rudiments of Affine Planes, Univ. of Toronto Press, 1975.
[22] L. A. Skornyakov, Homomorphisms of projective planes and T-homomorphisms of ternary ring (Russian), Mat. Sb. (N. S.), 43, (85), (1957) 285–294.
[23] ______, Projective planes, Uspehi Mat. Nauk (N. S.), 6, no. 6 (46), (1951) 112–154. (English translation: Amer. Math. Soc., translation #99 (1953).)
[24] G. Törner, Eine klassifizierung von Hjelmslev-ringen und Hjelmslev-ebenen, Mitt. Math. Sem. Giessen, 107 (1974).
[25] ______, (r^{n-1}, r)-Hjelmslev-Ebenen des Tyns n, Math. Z., 154 (1977) 189–201.

Pyramidal Sections in Taxicab Geometry

The eccentricity definition of conics is borrowed and recycled, and Euclid is repaid with interest.

RICHARD LAATSCH
Miami University
Oxford, OH 45056

The formula for the Euclidean distance between two points in the plane leads, early in the study of analytic geometry, to the definitions of the classical conic sections—the circle, ellipse, and hyperbola—in terms of distances to fixed points. Additionally, the parabola can be defined in terms of distances to a fixed point and a fixed line. After modest additional study in the area, the student becomes aware that the ellipse and hyperbola have comparable "point-line" definitions. It is this point-line approach to the conic sections I wish to pursue here—not in the context of Euclidean distance but using, instead, the **taxicab metric** defined in the coordinate plane by

$$d_T(P(x_1, y_1), Q(x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2| .$$

The d_T-distance between points P and Q is the length of a shortest path from P to Q composed of line segments parallel to the coordinate axes.

Previously published studies of the taxicab analogues of the conic sections [**1**], [**4**], [**5**], [**6**], [**7**] consider the plane curves whose definitions use the word "distance," interpret that word to mean the taxicab distance, and determine the shapes thus produced. In examining the ellipse and hyperbola from the standpoint of their "two-fixed-points" definitions, these studies obtain hexagons and octagons for their taxicab ellipses and an intriguing variety of unbounded graphs, some containing entire regions of the plane, for taxicab hyperbolas.

If the "point-line" definition is used instead, then some interesting results can be proved which show a strong analogy between the taxicab figures and the Euclidean conics. First, the taxicab figures obtained by using the focus-directrix definition of ellipse, parabola, and hyperbola can be obtained by projecting plane sections of a square pyramid onto a plane perpendicular to the axis of the pyramid. (The square pyramid is the natural object to section, since it is the taxicab analogue of the right circular cone.) Second, there is a simple way to locate the focus and directrix of these projections in the three-dimensional context of the sectioned pyramid. Third, this way of locating focus and directrix also works for the comparable projections of conic sections in Euclidean space.

Throughout our discussion, unless otherwise noted, points and lines are all contained in a fixed coordinate plane **P**. To begin our investigation, we need a taxicab definition of the distance from a point to a line. Let L be a line in the plane **P** and P a point in **P** not on L. The **taxicab distance** from P to L, denoted $d_T(P, L)$, is defined as

$$d_T(P, L) = \min\{d_T(P, Q): Q \in L\}.$$

A couple of quick sketches show that $d_T(P, L)$ is very easy to compute or measure since, if L is a vertical line or if its slope m satisfies $|m| \geqslant 1$, then $d_T(P, L)$ is the length of the horizontal segment joining P to L; however, if $|m| \leqslant 1$, then $d_T(P, L)$ is the length of the vertical segment joining P to L. (It should be noted that such concepts as line, slope, and equation of a curve are coordinate concepts, dependent only on the coordinates of points, and are unaffected by our change to the taxicab metric.)

We now define the taxicab parabola, ellipse, and hyperbola, proceeding according to the

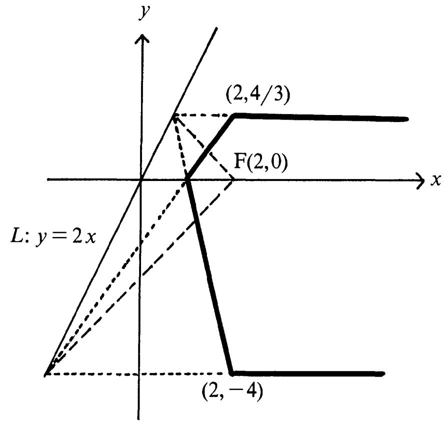


FIGURE 1. **Taxicab parabola (heavy line), general case, with directrix L: $y = 2x$ and focus $F(2,0)$; its equation is (1).**

analogous Euclidean prescription. Let F (the **focus**) be a fixed point in **P** and L (the **directrix**) be a fixed line in **P** which does not contain F. Then the set of all points P in the plane **P** such that $d_T(P,F)/d_T(P,L)=e$ (e is called the **eccentricity**) is called a **projected pyramidal section** (or ***p*-section**). The shape of the curve in **P** is determined by the ratio $d_T(P,F)/d_T(P,L)=e$, and we shall call the resulting plane figure an **ellipse** if $e<1$, a **parabola** if $e=1$, or a **hyperbola** if $e>1$. We will justify the name "projected pyramidal section" by showing that these curves can also be obtained by sectioning a square pyramid with a plane and projecting that section onto another plane. Thus our choice of names for the three classes of curves is consistent with Euclidean geometry since the projection of a Euclidean ellipse, parabola, or hyperbola onto another suitably chosen plane is an ellipse, parabola, or hyperbola, respectively.

There is an inherent "asymmetry" in measurements using the taxicab metric; we have pointed out that the method of measuring distance from a point to a line depends on the slope of the line. For this reason, it is natural to consider three cases for the graphs of each p-section. If the directrix is a horizontal or vertical line in **P**, we shall refer to the graph as the **parallel case**; if the slope of the directrix is ± 1, we shall call the graph the **diagonal case**; the remaining possibilities will be the **general case**. We shall graph only one or two cases of each p-section here and thus leave the interested reader with a few cases to investigate.

We begin with the parabola. FIGURE 1 shows a graph of the general case with directrix L the line $y=2x$ and focus the point $F(2,0)$. The defining relation $d_T(P,F)/d_T(P,L)=1$ yields the equation of the parabola:

$$|x-2|+|y|=|x-y/2|. \tag{1}$$

The parabola can be sketched by dividing the plane into regions using the directrix and the horizontal and vertical lines through the focus as the separating lines, and graphing in each region the linear equation derived from (1) which is appropriate. (For example, in FIGURE 1, $|x-2|=2-x$ for regions to the left of $F(2,0)$ and $|x-y/2|=x-y/2$ for regions to the right of the directrix.) Frequently, however, such a graph can be drawn more quickly using the knowledge that it consists of portions of straight lines and changes in direction only when it crosses the horizontal and vertical lines through the focus.

A parallel case of the parabola with directrix $x=-2$ and focus $F(2,0)$ is graphed in FIGURE 2, and a diagonal case of the parabola with directrix $y=-x-4$ and focus $F(4,4)$ is graphed in FIGURE 3. The symmetry of the graphs in FIGURES 2 and 3 is typical, that is, the parallel and diagonal case will always have graphs symmetric to a line perpendicular to the directrix. Other graphs of taxicab parabolas appear in [**4**, p. 702], [**5**, p. 84], and [**7**, p. 147]; lattice-point graphs of the parallel and diagonal case appear in [**1**, p. 26].

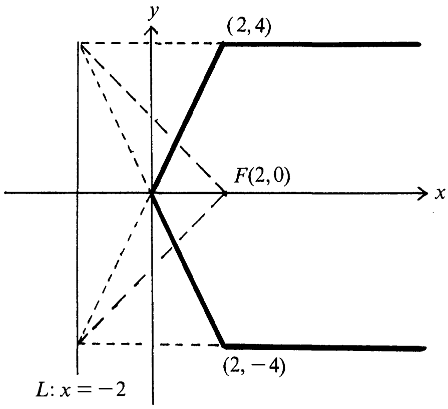


FIGURE 2. Taxicab parabola (heavy line), parallel case, with directrix L: $x = -2$ and focus $F(2,0)$.

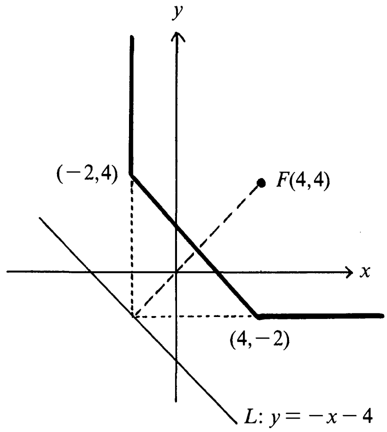


FIGURE 3. Taxicab parabola (heavy line), diagonal case, with focus $F(4,4)$ and directrix L: $y = -x - 4$.

Each taxicab ellipse is a quadrilateral. We show an example of a parallel case having the y-axis as directrix L, $F(3,0)$ as focus, and eccentricity $e = 1/2$. The general relation $d_T(P, F)/d_T(P, L) = 1/2$ gives the equation

$$|x-3| + |y| = \tfrac{1}{2}|x|, \tag{2}$$

which is graphed in FIGURE 4. In the parallel case, it can be proved that the Euclidean ellipse with the same directrix, focus, and eccentricity passes through all four vertices of the taxicab ellipse. Note that the taxicab ellipse in FIGURE 4 has one axis of symmetry, containing the focus and two vertices while the corresponding Euclidean ellipse has two axes of symmetry. The reader is encouraged to explore examples of the diagonal case and general cases, to find the similarities and differences of these taxicab ellipses.

The hyperbola assumes the most varied shapes of the p-sections of taxicab geometry. In the parallel and diagonal cases (as with these cases for the ellipse and parabola), the graph has one axis of symmetry. We leave these cases to be investigated by the reader. The general case is more complex, with the shape of the graph dependent on both the slope m of the directrix and the eccentricity e. There are three subcases. We illustrate the variation in shape of a taxicab hyperbola with focus $F(4,0)$ and eccentricity $e = 3$ for three different choices of directrix L with slope m: $|m| < e$, $|m| = e$, $|m| > e$. For the first subcase, we choose L as the line with equation $y = 2x$; the

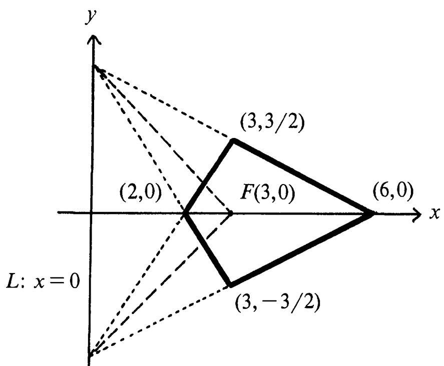


FIGURE 4. Taxicab ellipse (heavy line), parallel case, with directrix L: $x = 0$ and focus $F(3,0)$; its equation is (2).

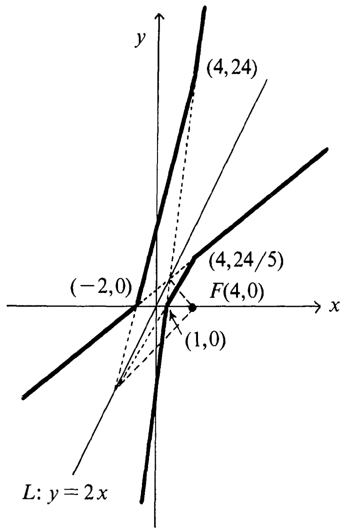


FIGURE 5(a). **Taxicab hyperbola (heavy line), general case with $|m|<e$. Here $e=3$, the focus is $F(4,0)$ and the directrix is L: $y=2x$. The equation of the graph is given by (3).**

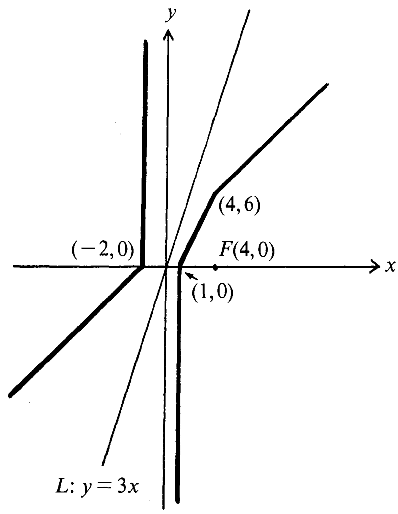


FIGURE 5(b). **Taxicab hyperbola with $e=3$, focus $F(4,0)$ and directrix L: $y=3x$ (case $|m|=e$).**

equation of this hyperbola is

$$|x-4|+|y|=3|x-y/2|, \tag{3}$$

and its graph is shown in FIGURE 5(a). If we increase $|m|$ while holding e constant, we find the lower ray of the right-hand branch of this hyperbola increasing in slope, and becoming vertical when $|m|=e=3$. At that instant the segment and upper ray of the left-hand branch also form a vertical ray. The graph of this hyperbola, whose equation is $|x-4|+|y|=3|x-y/3|$ is shown in FIGURE 5(b). If $|m|$ is increased so that $|m|>e$, then the upper ray on the left-hand branch assumes a negative slope as does the lower portion (no longer a ray) of the right-hand branch. This portion intersects the vertical line $x=4$ and bends, increasing in slope, as it crosses that line. A graph of this subcase, with $m=6$, is shown in FIGURE 5(c).

After examining the equations and graphs of several of the p-sections, it is easy to conjecture (and prove) the following general statement: *each p-section is the union of segments or rays from exactly four distinct lines* (*three for the diagonal parabola*). Far less obvious is the observation in Theorem 1 below. No doubt the curious reader has wondered why dashed and dotted lines occur in the graphs of FIGURES 1–5; this was done to illustrate the following fact.

THEOREM 1. *Two alternate sides of a p-section, if extended, will intersect at a point Q on the directrix of the section. A line of slope $+1$ or -1 through the focus of the section also passes through Q.*

The analytic proof of this theorem is composed of three direct computations, one for each case: parallel, diagonal, and general. Since geometers have been known to become ecstatic over the discovery of three-line incidence properties, this theorem, which asserts the incidence of four lines at the point Q, is superficially spectacular. However, this property turns out to be a routine consequence of the geometry of the sectioning of the square pyramid—to which we now turn our attention.

Consider a square pyramid (of two nappes) in three-dimensional space, the surface of which is the graph of the equation

$$|x|+|y|=c|z|, \quad c>0, \tag{4}$$

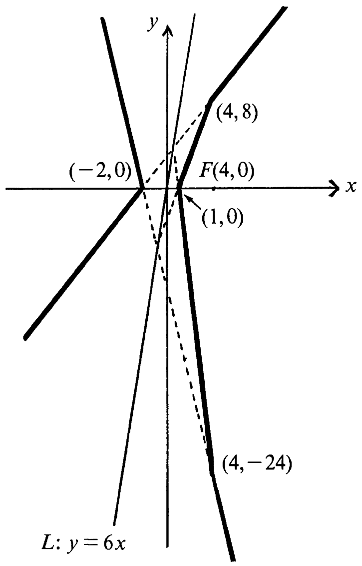


FIGURE 5(c). **Taxicab hyperbola with $e = 3$, focus $F(4,0)$ and directrix L: $y = 6x$ (case $|m| > e$).**

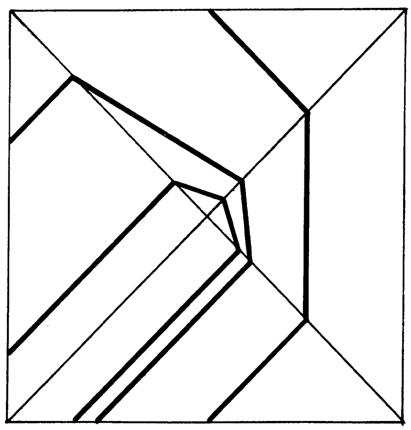


FIGURE 6. **The three cases of taxicab parabolas as projections of pyramidal sections.**

oriented for descriptive purposes with the z-axis vertical. It is not hard to prove that each of the cases (parallel, diagonal, general) of each of the basic shapes (parabola, ellipse, hyperbola) can be obtained from the intersection of a plane with the pyramid by perpendicular projection of that intersection onto the xy-plane. Even without detailed computation this seems plausible, since the equation of a plane is of the form

$$z = ax + by + d. \tag{5}$$

Solving (4) and (5) simultaneously by eliminating z (projecting onto the xy-plane) should produce an equation similar to (1), (2), or (3). It can be shown that, analogous to the prescription of Euclidean conic sections, a sectioning plane which is parallel to an edge of the pyramid will generate a parabola; a plane of lesser inclination, cutting all four edges of one nappe, will generate an ellipse; and one of greater inclination (thus cutting both nappes), a hyperbola.

FIGURE 5 illustrates the variations in shape of a hyperbolic p-section. By cutting the pyramid with a plane in such a way that the plane intersects both nappes *and the axis* of the pyramid, and projecting the section onto the xy-plane, you can achieve this variation in the components of the branches of the section by rotating the plane around the axis of the pyramid. The Euclidean hyperbola can, of course, be obtained by a section of the two nappes of a cone, but since Euclidean distances are unchanged by rotation, sections will not change shape if the sectioning plane is rotated around the axis of a right circular cone.

FIGURE 6 shows the three types of parabola as projections of pyramidal sections. (The reader is encouraged to slice a solid pyramid or draw the intersections of planes with the pyramid on its faces; the view from above the pyramid, looking down at the vertex, appears to the eye as the projections shown in FIGURE 6.) In the diagonal case, the sectioning plane is parallel to one face of the pyramid and in the parallel case the edge of the pyramid is equidistant from the two parallel edges of the section. In general, for all of the p-sections, the direction of the gradient of the sectioning plane σ (the vector in the xy-plane in the direction of the maximum inclination of σ) determines the case of each section. If the gradient is parallel to the x-axis or y-axis, then the pyramidal section will project onto the xy-plane as a parallel case; if the gradient has slope ± 1 in the xy-plane, the section projects as a diagonal case; other sections yield the general case.

We can illustrate these observations by showing how our earlier examples of p-sections are indeed projections of pyramidal sections. For this, we identify our coordinate plane **P** with the xy-plane. A parabolic p-section congruent (via the translation $x \to x+4, y \to y+4$) to that of FIGURE 3 is the projection of the intersection of the pyramid $|x|+|y|=2|z|$ with the plane σ whose equation is $x+y+6=2z$. (More generally, you can replace each "2" in these equations with an arbitrary $c>0$.) To verify this, consider the four equations of the planes $\pm x \pm y = 2z$ which form the faces of the pyramid, and eliminate the z-term between each and the equation of σ. The resulting equations in x and y which are consistent (one of them is not!) will be the equations of the projections of the lines of intersection of the pairs of planes into the xy-plane. Graph the portion of each line in its appropriate quadrant. Similarly, the p-section in FIGURE 1 is congruent (via the translation $x \to x-2$) to the projection onto the xy-plane of the intersection of the pyramid $|x|+|y|=|z|$ with the plane whose equation is $-2x+y+2z=4$. The p-section in FIGURE 4 is congruent (via the translation $x \to x-3$) to the projection of the intersection of the pyramid $|x|+|y|=\frac{1}{2}|z|$ with the plane whose equation is $x+z=-3$.

It should be noted that the p-sections usually are not themselves congruent to sections of a square pyramid. For example, the angle between the segment and a ray of the p-section in FIGURE 3 is 135°, but the angle between the corresponding parts of the pyramidal section whose projection is this p-section, is less than that. Also, in order for the graph in FIGURE 4 to be a pyramidal section would require the existence of a nondegenerate Euclidean right triangle with one leg equal to its hypotenuse.

The focus-directrix definition of a p-section in terms of the taxicab metric is an analytic one, and relies on the use of coordinates in the plane. However, the description of these curves as projections of sections of a square pyramid onto a plane π through the vertex of the pyramid and perpendicular to the axis of the pyramid is coordinate-free. But it is possible to locate the focus and directrix of the p-section in π. The illustration of parabolic p-sections shown in FIGURE 6 suggests an obvious location of the focus of these sections: the vertex of the pyramid.

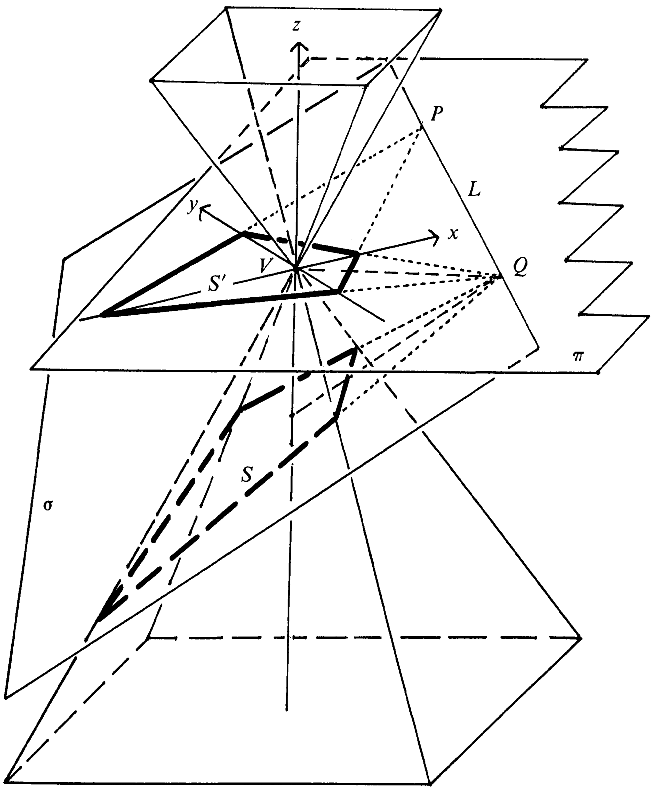


FIGURE 7. **Pyramidal section S and p-section S'.**

THEOREM 2. *Let π be a plane through the vertex V of a square pyramid, perpendicular to the axis of the pyramid. Let σ be a plane not perpendicular to π, which intersects π. Let S be the intersection of the pyramid with σ, and project S parallel to the axis of the pyramid onto the plane π; call this projection S'. Then S' is a p-section with focus V and directrix the line of intersection of the planes π and σ.*

FIGURE 7 illustrates Theorem 2 for the case of a p-section S' which is an ellipse. If we identify π with the xy-coordinate plane as shown in FIGURE 7 (the x and y axes are the projections of the edges of the pyramid onto π), then the fact that V is the focus of S' is immediately apparent, since V is the intersection of the x and y axes, which are the horizontal and vertical lines at which the graph of the p-section changes direction. (Note our earlier remarks about how to graph equation (1) in FIGURE 1.)

Theorem 1 says that the points of intersection of the lines containing alternate sides of a p-section are on its directrix; hence (except in the diagonal parabola case) the two points P and Q, thus determined by the two pairs of alternate sides, determine the directrix. But the lines through the sides of the p-section are just the projections into the xy-plane of the lines in which the sectioning plane σ intersects the face planes of the pyramid. FIGURE 8 shows that lines through alternate edges of the section S of the pyramid intersect in the plane π, so that this point of intersection coincides with the intersection of the projections of these lines in π. (In FIGURE 8,

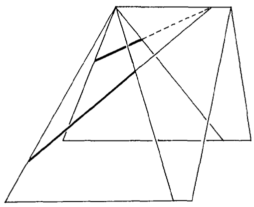


FIGURE 8

opposite faces of the pyramid have been extended to form a prism; the plane σ which cuts the pyramid in S has as its intersection with the prism lines through the edges of S, and clearly these lines intersect in the top edge of the prism, which lies in π.) Since P and Q are both in $\pi \cap \sigma$, and lie on the directrix of S', the line $\pi \cap \sigma$ is the directrix of S'. In the case of a diagonal parabola S', only one point Q on the directrix is determined as the intersection of alternate sides of S', so a different argument is needed to complete the proof that $\pi \cap \sigma$ is the directrix of S'. We leave this to the reader.

Since a search of available literature turned up no result comparable to Theorem 2 for the classic conics, it was natural to consider a comparable, and virtually identical, Euclidean theory of focus and directrix for projections of conic sections. Let π be a plane through the vertex V of a right circular cone K, with π perpendicular to the axis of K and consider any plane σ which is not parallel to π or to the axis of the cone. The intersection of σ and the cone K is a conic section S: an ellipse, parabola, or hyperbola. The projection of S parallel to the axis of the cone onto the plane π is a "fatter" ellipse, parabola, or hyperbola, respectively, which we shall call S'. The following theorem tells us where to find a focus and directrix of S'.

THEOREM 3. *In the above construction, the point V, the vertex of the cone K, is a focus of S', and the line of intersection of π and the sectioning plane σ is the directrix corresponding to this focus V.*

FIGURE 9 illustrates Theorem 3 for the case of an ellipse. The theorem can be proved analytically by direct calculation. Let the cone K have equation $c^2z^2 = x^2 + y^2$ and let the plane σ be given by $z = mx + b$. (The plane π is the xy-plane.) Eliminating z between these equations yields the equation of the projection S' of $S = K \cap \sigma$ into the xy-plane,

$$(1 - c^2m^2)x^2 - 2bc^2mx + y^2 = b^2c^2,$$

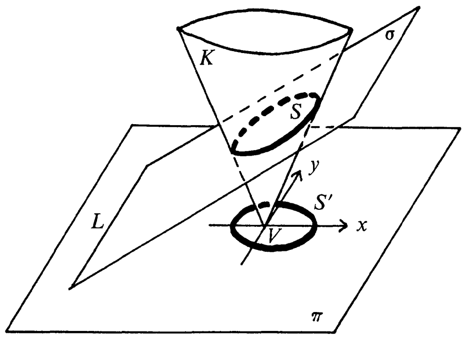


FIGURE 9. **Conic section S and projection S' with focus V and directrix L.**

which can be transformed into the standard equation of the conic. For example, in the case of the ellipse as in FIGURE 9 where $m < 1/c$, we obtain the standard form

$$(x-h)^2/A^2 + (y-k)^2/B^2 = 1$$

with

$$h = bc^2m/(1-c^2m^2), \qquad k=0, \quad A^2 = b^2c^2/(1-c^2m^2)^2,$$

and

$$B^2 = b^2c^2/(1-c^2m^2).$$

Thus the focal distance is C such that

$$C^2 = A^2 - B^2 = b^2c^4m^2/(1-c^2m^2)^2.$$

Since $h = C$, the left-hand focus is (0,0), the origin of the xy-plane (which is the vertex V of the cone K). The equation of the left-hand directrix L is $x = h - A^2/C$, which reduces to $x = -b/m$, the equation of the trace of the plane $z = mx + b$ in the xy-plane. The calculations for the hyperbola ($m > 1/c$) are comparable and those for the parabola ($m = 1/c$) are shorter. (In the cases of the ellipse and hyperbola we have an eccentricity $e = C/A = mc = m/(1/c)$, and we can conclude, additionally, that the eccentricity of S' will always be the ratio between the slope m of the sectioning plane and the slope $1/c$ of an element of the cone.)

There is also an interesting three-dimensional theory of focus and directrix of (unprojected) Euclidean conics, originally described by G. P. Dandelin in 1822 **[2]**, **[3]**, **[8]**, which has no analogue in taxicab geometry. It gives both foci and both directrices for the ellipse and hyperbola using spheres inscribed in the cone; the theory presented in Theorem 3 gives only one focus and its directrix. In the case of the p-sections of taxicab geometry, however, there is only one focus and directrix because the figures lack the symmetry for two. It is unusual to be led to a result in Euclidean geometry by its analogue in a non-Euclidean geometry—particularly when the Euclidean theorem is simpler than the non-Euclidean.

References

[1] M. Gardner, Mathematical games, Sci. Amer., 243 (Nov. 1980) 18, 30.

[2] D. Hilbert and S. Cohn-Vossen, Geometry and the Imagination, Chelsea, New York, 1952, pp. 27–29.

[3] G. B. Huff, On defining conic sections, Selected Papers on Geometry, M.A.A., 1979, 77–78 (reprinted from Amer. Math. Monthly, 62 (1955) 250–251).

[4] E. F. Krause, Taxicab geometry, Math. Teacher, 66 (1973) 695–706.

[5] ———, Taxicab Geometry, Addison-Wesley, Menlo Park, 1975.

[6] L. E. Mehlenbacher, Introduction to Modern Mathematics, Prindle, Weber and Schmidt, Boston, 1968, pp. 140–147.

[7] B. E. Reynolds, Taxicab geometry, Pi Mu Epsilon J., 7 (1979–84) 77–88.

[8] G. Salmon, A Treatise on the Conic Sections, 6th ed., Chelsea, New York, 1954, p. 331.

Planning for Interruptions

JoAnne Simpson Growney
Bloomsburg State College
Bloomsburg, PA 17815

Organizing is what you do before you do something, so that when you do it, it's not all mixed up.
—Christopher Robin in **A. A. Milne's *Winnie the Pooh***

Nothing is particularly hard if you divide it into small jobs.
—**Henry Ford**

It is Tuesday morning. Faculty members Jones and Smith are each in their offices preparing for eleven o'clock classes.

Professor Jones has left his door open. His students who pass by the office stop in with questions and comments. Professor Jones interrupts his lecture preparation to respond to their queries. These students leave the office pleased and satisfied with the help provided by their instructor. But eleven o'clock arrives without Jones' class lecture being fully prepared. His presentation is disorganized, and both he and the class members are dissatisfied with the results.

The door to Professor Smith's office is, on the other hand, closed and locked. Students with questions pass by the door and find their instructor unavailable. They leave dissatisfied. When eleven o'clock arrives, Professor Smith has completed her class preparation. The class material is well organized and she presents it clearly. Both she and her students are pleased and satisfied with the learning that took place in their hour together. Nevertheless, later Professor Smith is dismayed when she learns of student dissatisfaction with her prior unavailability to answer individual questions.

The choices faced by Professors Jones and Smith—either to isolate themselves and concentrate on a task that needs to be thought through carefully without interruption or to permit interruptions that may prevent completion of the desired task—provide a dilemma for many individuals. Teachers, students, managers, parents—to name a few—all face this conflict repeatedly. On the one hand are unpredictable events to which attention should be given as the need arises. On the other hand are tasks that require uninterrupted periods of thought.

The purpose of this article is to develop a mathematical model that confirms the wisdom of planning activities so as to minimize the detrimental effect of interruptions.

The parable of the watchmakers

The value of planning for interruptions is well illustrated by the following tale [**4**]:

> Once upon a time there were two watchmakers, Hora and Tempus. Each produced very fine watches which began to be in great demand. Their workshop phones rang frequently, bringing orders from new customers. Hora prospered but Tempus become poorer and poorer and finally lost his shop. What was the secret of Hora's success?
>
> Both men made watches that consisted of about 1000 parts. Tempus had constructed his so

that if several pieces were assembled and he had to put the assembly down—for example, to answer the phone—all the pieces fell apart and had to be reassembled from scratch. Naturally, the better his customers liked his work, the more they told their friends—who interrupted his work to place orders for more watches. As a result it became more and more difficult for Tempus to find enough uninterrupted time to finish a watch.

Hora, on the other hand, had designed his watches so that he could put together components of about ten parts. As with Tempus, if he was interrupted while working on a particular assembly, the pieces fell apart and had to be reassembled from scratch. But the ten-part components, once completed, were "stable" and could be set aside for later use as a unit. Ten of the components could be assembled into a larger stable subsystem. Later the assembly of the ten larger subsystems constituted the whole watch.

The obvious secret of Hora's success was the use of "stable" subsystems that allowed him to pay only a small penalty in watch-assembly time for each interruption. What the parable fails to tell us, however, is the price that Hora paid to develop his method. That is, we do not know how many hours or days or weeks it took Hora to organize his system of assembly. We see only his success without any history of the size of his initial investment.

In general, individuals are often unwilling to invest time to organize lengthy tasks into "stable" subtasks because the time required to get organized is an initial cost, often great, that appears to lengthen even more the time that the task will require. However, our mathematical model for analysis of interruptions will reveal that extra organizational time is often a wise investment.

For Tempus and Hora, for example, if $p = 0.01$ is the probability of interruption during the addition of any given part to the watch assembly, the formula given below for T can be used to estimate that the time required for Tempus to complete a watch is almost 2000 times as great as that for Hora. Thus, even if organizing his watchmaking into stable subsystems was a very lengthy process for Hora, it proved an essential ingredient in his success and provided substantial improvement over the assembly method of Tempus.

A model for estimating the effect of interruptions

Suppose a certain project is to be performed and that the project is divided into s stable subtasks, each of which requires U consecutive uninterrupted time units for its completion. (For example, in *The Parable of the Watchmakers*, Tempus' method for constructing a watch has $s = 1$ and $U = 1000$; Hora's method has $s = 111$ and $U = 10$.) If no interruptions occur, the total time required for completion of the desired project is $s \cdot U$.

Further, suppose that interruptions occur randomly with probability p during any given time unit. Our model will assume two basic properties concerning the nature of the interruptions. The first is that any interruption has no effect on the pattern of subsequent interruptions—that is, interruptions are *independent*. The second assumption is that each interruption during a stable subtask requires starting over at the beginning of that subtask—that is, interruptions are *devastating*.

The assumptions that interruptions are independent and devastating are not wholly realistic. Yet in cases in which interruptions are troublesome, these are not unreasonable suppositions. For, in such cases, it seems that no matter how many interruptions we deal with we cannot change the likelihood of more later—hence the reasonable assumption of independence. Furthermore, the frustration associated with being diverted from a task that requires sustained concentration makes a later return to the task almost like starting over. Hence the assumption that interruptions are devastating also is reasonable.

If T denotes the expected number of time units required for completion of the desired project, then

$$T = \frac{s}{p}\left[\frac{1}{(1-p)^U} - 1\right] \tag{1}$$

where s, p, and U are as defined above.

By application of l'Hôpital's rule, we can determine that

$$\lim_{p \to 0} T = \lim_{p \to 0} \left\{ \frac{s\left[\frac{1}{(1-p)^U} - 1\right]}{p} \right\} = s \cdot U.$$

This limiting value, $s \cdot U$, is the total time required for completion of the project if no interruptions occur.

We now show how formula (1) for T can be deduced. For background details from the theory of probability, consult references [**1**], [**2**], or [**3**].

The probability that an interruption will *not* occur during a given time interval is $(1-p)$. The probability of no interruption during U consecutive time intervals is $(1-p)^U$—since the interruptions are independent. The quantity $(1-p)^U$ may also be interpreted as the proportion of starts that are completed. Its reciprocal $1/(1-p)^U$ thus gives the expected number of starts per completed subtask. We then have that the expected number of subtask starts required for completion of the entire project is

$$\left.\begin{array}{l}\text{Expected number of}\\ \text{subtask starts required}\\ \text{for project completion}\end{array}\right\} = \frac{1}{(1-p)^U} \cdot s. \tag{2}$$

Since the probability of no interruptions during U consecutive time intervals is $(1-p)^U$, the probability that an interruption will occur during one of U consecutive time intervals is $[1-(1-p)^U]$. This probability may also be interpreted as the proportion of starts that are interrupted—equivalently, as the expected number of interruptions per subtask start. Combining this with (2) gives

$$\left.\begin{array}{l}\text{Expected number of}\\ \text{interruptions during}\\ \text{project completion}\end{array}\right\} = \left[1-(1-p)^U\right]\cdot\frac{1}{(1-p)^U}\cdot s. \tag{3}$$

To obtain the equation for T it remains only to multiply equation (3) by the expected length of time-on-task that elapses between interruptions—that is, the expected number of time units spent per interruption—which is equal to $1/p$. Thus we have

$$\left.\begin{array}{l}\text{Expected number of}\\ \text{time units required}\\ \text{for project completion}\end{array}\right\} = \frac{1}{p}\cdot\left[1-(1-p)^U\right]\cdot\frac{1}{(1-p)^U}\cdot s \tag{4}$$

which simplifies to yield formula (1).

Returning to the watchmaker parable, we have

For Tempus:

$$s=1,$$
$$U=1000,$$
$$p=0.01,$$
$$T=\frac{1}{.01}\left[\frac{1}{(1-.01)^{1000}}-1\right]\approx 2{,}316{,}000 \text{ time units.}$$

For Hora:

$$s=111,$$
$$U=10,$$
$$p=0.01,$$
$$T=\frac{1}{.01}\left[\frac{1}{(1-.01)^{10}}-1\right]\approx 1174 \text{ time units.}$$

Hora, in producing a watch that requires assembly of 1110 parts (including the various stable subsystems), loses an average of 64 time units because of starting over after interruptions. For Tempus, the time loss is so great that he can reasonably expect never to complete a watch under the given circumstances.

Our interruption model, based on the watchmaker parable, has the following moral:

Time invested in organizing a complex project into stable subtasks is amply returned through a reduction in the time lost because of interruptions.

A simple example will illustrate the moral.

EXAMPLE. Two political science students, Jim and Jean, have each been assigned to write a brief paper analyzing and comparing different voting methods. Each estimates that the rough draft will take one hour of intense concentration—that is, it must be done without interruptions. Any interruption will be so devastating that starting over will be required. For both Jim and Jean interruptions occur randomly but, on the average, every twenty minutes.

(a) Jim started writing his rough draft at 6 P.M. last evening. Under the given conditions, what is the expected time at which he is finished?

Analysis. Using the notation of the interruption model, we have

$s=1$ (stable subtasks),

$U=60$ (consecutive uninterrupted minutes),

$$p = \frac{1}{20} = 0.05 \text{ (likelihood of interruption during any given minute).}$$

Substitution of these values into the formula for T yields

$$T = \frac{1}{.05}\left[\frac{1}{(1-.05)^{60}} - 1\right] \approx 414 \text{ minutes.}$$

Thus Jim's rough draft required an estimated seven hours of "time-on-task" and, if his interruptions diverted him for only brief periods, his completion time would be expected to be around 1 A.M. this morning.

(b) Jean was astounded when a sleepy-eyed Jim told her how long it had taken him to complete his rough draft. Based on his experience, she has decided to try a different approach. Jean believes that with fifteen minutes of intense concentration she can outline her paper and divide it into four sections that will each take fifteen minutes to write. If she is interrupted during any fifteen-minute interval, she will need to start over and spend fifteen more minutes on that portion of her task. However, once the outline or any of the subsections is complete, no interruption will affect it.

Jean thus has five fifteen-minute tasks to complete without interruption. How long will it take her? If she begins at 6 P.M. tonight, when can she expect to finish?

Analysis. Using the notation of the interruption model, we have

$$s = 5 \text{ (stable subtasks),}$$

$$U = 15 \text{ (consecutive uninterrupted minutes),}$$

$$p = \frac{1}{20} = 0.05 \text{ (likelihood of interruption during any given minute).}$$

Substitution of these values into the formula for T yields

$$T = \frac{1}{.05}\left[\frac{1}{(1-.05)^{15}} - 1\right] \approx 116 \text{ minutes.}$$

Thus Jean's rough draft will require about two hours for completion. If she starts at 6 P.M. this evening she can expect to complete her task at about 7:56 P.M.$+I$—where I is the number of minutes actually devoted to her interruptions.

When is organizational time worth the investment?

In the example above, Jean's investment of fifteen extra minutes to organize her paper resulted in an expected savings—when compared with Jim's situation—of almost five hours. We easily conclude that her initial investment was worthwhile. However, a general question emerges: *Under what conditions is it likely to be worthwhile to invest time to organize a lengthy task into stable subtasks to reduce the startover time required by interruptions?*

A partial answer results from consideration of a special case. Suppose the probability p of interruption during any given time unit and the number U of consecutive uninterrupted time units required to complete the desired task are related by

$$U = \frac{1}{p}.$$

In this situation, with the task not divided into more than one subtask, we have

$$T = U \cdot \left[\frac{1}{(1-1/U)^{U}} - 1\right].$$

JUST FOR FUN

Paul has recently become aware of the evidences of his own aging and has determined to reverse the process and become physically fit. Being a rational fellow he has consulted his physician before embarking on a fitness program. The doctor has recommended for Paul a daily stint of eight uninterrupted minutes on a treadmill. If Paul is interrupted during his effort then he must start over and complete eight more minutes without interruption.

Sounds simple! However, Paul has another affliction. He also suffers from a chronic skin disorder. Randomly, but on the average of once every two minutes, Paul is interrupted with the necessity to stop and scratch an itch.

How does Paul fare as he tries to carry out his physical fitness program? The truth may never be known but you can simulate Paul's frustrating experience in the following manner:

Repeatedly toss a coin, letting each toss represent a minute for Paul on the treadmill. Let each toss that results in a head designate a minute in which Paul must stop to scratch. A tail permits him to keep jogging. How many tosses altogether are required to simulate a string of eight consecutive uninterrupted minutes of jogging?

From this we obtain

$$\frac{T}{U}=\left[\left(\frac{U}{U-1}\right)^{U}-1\right].$$

TABLE 1 lists a sample of values of U and corresponding values for T/U. More generally, we observe that T/U is a decreasing function of U and, as U becomes large, T/U approaches $(e-1)$. (L'Hôpital's rule is useful in evaluating this limit; for illustration of the method see, for example, [5].)

U	2	3	4	5	7	10	15	25	50	100
T/U	3	2.375	2.160	2.052	1.942	1.868	1.815	1.775	1.746	1.732

TABLE 1

Thus, in the case in which $U=1/p$, the "time-on-task" will be lengthened by over 70% because of interruptions. In such case, an initial investment of organizational time—to divide the task into stable subtasks—is well worth considering. In general, if we value our time, we should never attempt to complete an uninterruptable task unless the time it requires is substantially less than the length of time we expect between interruptions.

The question of how to minimize the sum of organizational time plus time wasted due to interruptions is difficult to solve in general. Nevertheless, an example will illustrate the tractability of the problem in any specific case.

Consider a situation in which a project to be completed requires 100 time units. Suppose that the task is subdivided into s uninterruptable subtasks each requiring U uninterrupted time units so that we have $sU=100$. Suppose further that the probability of interruption during any given time unit is $p=0.05$. TABLE 2 gives the values of T for various subdivisions of the task into s stable subtasks. In this case we have

$$T=20s\cdot\left[\frac{1}{(.95)^{100/s}}-1\right].$$

s	1	2	5	10	20	25	50	100
T	3360	480	179	134	117	114	108	105

TABLE 2

Of course, the time wasted because of interruptions decreases as s increases. But to assume that a large number of subtasks is optimal is naive; it ignores the time required for organizing the project into stable subtasks.

For the project under consideration, let us suppose that the organizational time required is proportional to the number of additional subtasks that must be devised. That is, suppose there is a positive constant C for which $C \cdot (s-1)$ gives the cost in organizational time. TABLE 3 lists values of $C \cdot (s-1)$ for the values $C = 0.1, 1, 10$, and for the s-values of TABLE 2. (We still assume that $s \cdot U = 100$.)

C \ s	1	2	5	10	20	25	50	100
0.1	0	.1	.4	.9	1.9	2.4	4.9	9.9
1	0	1	4	9	19	24	49	99
10	0	10	40	90	190	240	490	990

TABLE 3. **Values of $C \cdot (s-1)$.**

Minimizing the time required for the project requires finding that value of s for which $T + C \cdot (s-1)$ is least. Circled values in TABLE 4 show these minima.

C \ s	1	2	5	10	20	25	50	100
0.1	3360	480.1	179.4	134.9	118.9	116.4	(112.9)	114.9
1	3360	481	183	143	(136)	138	157	204
10	3360	490	(219)	224	307	354	598	1095

TABLE 4. **Values of $T + C \cdot (s-1)$. Circled values are minimal for the given C.**

The obvious conclusion is that when organizational planning is inexpensive, more of it should be done. (For example, when $C = .1$, the project should be divided into $s = 50$ subtasks of duration $U = 2$ time units.) When planning is costly, unless the organizational cost can be distributed over several similar projects, planning time is harder to justify—one may as well suffer some of the effects of interruptions.

In closing, it seems appropriate to remind the reader that we have been engaging in an amusing flight of fancy. Estimation of project completion times and probabilities of interruption requires imaginative speculation. Hence calculations based on them lead to results still more imaginary. Like many flights of fancy our exploration is, however, enlightening. Effective scheduling of time remains pure art. Yet an excursion into the fancy of formulas and calculations can show how to extend the practice of that art.

References

[1] Morris H. DeGroot, Probability and Statistics, Addison-Wesley, Reading, MA, 1975.

[2] B. W. Lindgren and G. W. McElrath, Introduction to Probability and Statistics, 2nd ed., Macmillan, New York, 1966.

[3] Marcel F. Neuts, Probability, Allyn and Bacon, Boston, 1973.

[4] Herbert A. Simon, The Sciences of the Artificial, 2nd ed., MIT Press, Cambridge, MA, 1981, pp. 200–202.

[5] George B. Thomas, Jr. and Ross L. Finney, Calculus and Analytic Geometry, 5th ed., Addison-Wesley, Reading, MA, 1979, Section 16-3.

A Differential Equation in Group Theory

CARROLL F. BLAKEMORE
RIDGLEY LANGE
RALPH TUCCI
University of New Orleans
New Orleans, LA 70148

Seemingly unrelated areas of mathematics may sometimes cross-pollinate to produce interesting results. This phenomenon occurs frequently in research, but is not so common at the undergraduate level. In this note we solve the problem of finding all order-preserving topological group isomorphisms (homeomorphisms which are also group isomorphisms) from a group $(G, *)$ whose underlying set G is the open interval $(-1,1)$ onto the additive group of reals $(\mathbb{R},+)$. We accomplish this by means of a certain differential equation.

Let G be the open interval $(-1,1)$ and for $a, b \in G$ define

$$a*b=\frac{a+b}{1+ab}.$$

First we prove that $(G, *)$ is an abelian group. For a, b in G the inequality $0<(1-a)(1-b)=1-a-b+ab$ leads to $a*b<1$ and, in a similar manner, also to $a*b>-1$. Thus G is closed under $*$, i.e., $*$ is a binary operation on G. Since it is clear that $*$ is commutative, that 0 is the identity, and that the inverse of a is $-a$, we have only to show that $*$ is associative. This straightforward calculation is left to the reader. Next we note that $(G, *)$ is a (linearly) ordered group in the natural order on $(-1,1)$ in the sense that if $a \in G$ and $b>c$ in G, then

$$a*b=\frac{a+b}{1+ab}>\frac{a+c}{1+ac}=a*c,$$

as is easily seen by cross-multiplying.

As topological spaces, $\mathbb{R}=(-\infty,+\infty)$ and $G=(-1,1)$ are homeomorphic, i.e., there is a continuous bijection $f\colon \mathbb{R}\to G$ such that f^{-1} is also continuous. A simple example is given by $f(x)=x/(1+|x|)$; however, this function is not a group isomorphism since it is easy to verify that $f(1+1)\neq f(1)*f(1)$. We now prove that there does exist a homeomorphism from $\mathbb{R}$ onto G which is an order-preserving group isomorphism.

Suppose that $f\colon \mathbb{R}\to G$ is a group isomorphism. Then f must satisfy the functional equation

$$f(x+y)=f(x)*f(y)=\frac{f(x)+f(y)}{1+f(x)f(y)}$$

with $f(x)$ and $f(y)$ taking values in $(-1,1)$ for all real x and y. If f is also continuous, then it can be shown (see the remarks in the final paragraph) that f is differentiable and, in particular, that $f'(0)$ exists. If we recall that $f(0)=0$ (since homomorphisms map identities to identities), a short calculation yields

$$f'(x)=\lim_{y\to 0}\frac{f(x+y)-f(x)}{y}=\lim_{y\to 0}\frac{f(y)}{y}\cdot\frac{1-f^2(x)}{1+f(x)f(y)}=f'(0)\left(1-f^2(x)\right)$$

because f is continuous at $y=0$. We note that f is a solution of the initial value problem

$$f'=f_0(1-f^2),\qquad f(0)=0 \tag{1}$$

where $f_0=f'(0)$. If $f'(0)=0$, then f is identically zero and this is impossible as f is onto. Hence, $f_0=f'(0)\neq 0$. By the method of separation of variables we obtain the solution

$$f(x)=\frac{\exp(2f_0x)-1}{\exp(2f_0x)+1}=\tanh(f_0x), \tag{2}$$

the verification of which is straightforward. It is also easy to verify that (2) defines an isomorphism which is order-preserving when $f_0 > 0$ since $f'(x) > 0$ for all x. Finally, the reader may check that f is also a homeomorphism. The solution (2) shows that there is in fact a distinct solution for each choice of $f_0 = f'(0)$.

There are other approaches to the above problem which we briefly mention. For example, the function k: $\mathbb{R} \to (0, +\infty)$ defined by $k(x) = e^x$ is an order-preserving topological group isomorphism from the additive group of reals $\mathbb{R}$ onto the multiplicative group of positive reals $\mathbb{R}^+ = (0, +\infty)$. Furthermore, h: $\mathbb{R}^+ \to G$ given by $h(y) = (y-1)/(y+1)$ is a homeomorphism such that $h(yz) = h(y) * h(z)$ for $y, z \in \mathbb{R}^+$. Hence, their composition $h \circ k$: $\mathbb{R} \to G$ is the function f in (2) above for $f_0 = \frac{1}{2}$. However, this approach fails to yield the complete family of solutions to the problem, which can be given by $\{\tanh(kx)$: $k \in \mathbb{R}$ with $k \neq 0\}$ where k is a parameter for the family. Finally, we note that if we let $(G, \circ)$ be an arbitrary group and let X be a set having the same cardinality as G, then, if f: $G \to X$ is a bijection, there is a (unique) binary operation $*$ defined on X by $a * b = f(f^{-1}(a) \circ f^{-1}(b))$ for $a, b \in X$ such that $(X, *)$ is a group and f is an isomorphism. Hence, any bijection from the group of additive reals $\mathbb{R}$ onto $(-1, 1)$ can be used to induce a group structure on $(-1, 1)$.

Several comments seem to be in order. First, except for the notion of an ordering (which can be ignored), the foregoing discussion is accessible to any student studying abstract algebra and differential equations. In fact, it is likely that the verification that the open interval $(-1, 1)$ with the operation $*$ is indeed a group has appeared as an exercise in some abstract algebra textbook (though the authors failed to find any mention of this in any of the textbooks they consulted). The discovery of this group operation on $(-1, 1)$ by the second author arose from his investigation of the operation $*$ defined for points in the open unit disc in the complex plane. (Associativity fails for $*$ here.) The group G also provides a nice example of a "bounded" group which is isomorphic to the additive reals. Finally, it is worth observing that the existence of $f'(0)$ may easily be assumed for beginning mathematics students. However, if we seek a group isomorphism from $\mathbb{R}$ onto G which is also assumed to be a homeomorphism and thus *a fortiori* continuous, then, since $\mathbb{R}$ and G are both Lie groups (this requires verification), such a mapping is necessarily differentiable (see [**1**, p. 109]).

The authors are grateful to the editor and referees for their helpful comments and suggestions.

Reference

[1] Frank W. Warner, Foundations of Differentiable Manifolds and Lie Groups, Scott, Foresman and Company, Glenview, Illinois, 1971.

A Tale of Two Goats

Marshall Fraser
Lincoln University
San Francisco, CA 94118

While attending a math conference a friend asked me to solve this problem:

A goat is grazing in a circular field. The goat is tethered to the fence encircling the field by a rope. How long should the rope be so that the goat grazes half the area of the field?

The question had that pleasant quality of being easy to visualize. I felt I should be able to solve it. Also, I wondered if I hadn't seen it before. To stall for time I asked my friend to define "tethered." Of course, to no avail. I was unable to solve the problem.

My ego was soothed when at dinner that evening several friends filled their placemats and the verification of which is straightforward. It is also easy to verify that (2) defines an isomorphism which is order-preserving when $f_0 > 0$ since $f'(x) > 0$ for all x. Finally, the reader may check that f is also a homeomorphism. The solution (2) shows that there is in fact a distinct solution for each choice of $f_0 = f'(0)$.

There are other approaches to the above problem which we briefly mention. For example, the function k: $\mathbb{R} \to (0, +\infty)$ defined by $k(x) = e^x$ is an order-preserving topological group isomorphism from the additive group of reals $\mathbb{R}$ onto the multiplicative group of positive reals $\mathbb{R}^+ = (0, +\infty)$. Furthermore, h: $\mathbb{R}^+ \to G$ given by $h(y) = (y-1)/(y+1)$ is a homeomorphism such that $h(yz) = h(y) * h(z)$ for $y, z \in \mathbb{R}^+$. Hence, their composition $h \circ k$: $\mathbb{R} \to G$ is the function f in (2) above for $f_0 = \frac{1}{2}$. However, this approach fails to yield the complete family of solutions to the problem, which can be given by $\{\tanh(kx)$: $k \in \mathbb{R}$ with $k \neq 0\}$ where k is a parameter for the family. Finally, we note that if we let $(G, \circ)$ be an arbitrary group and let X be a set having the same cardinality as G, then, if f: $G \to X$ is a bijection, there is a (unique) binary operation $*$ defined on X by $a * b = f(f^{-1}(a) \circ f^{-1}(b))$ for $a, b \in X$ such that $(X, *)$ is a group and f is an isomorphism. Hence, any bijection from the group of additive reals $\mathbb{R}$ onto $(-1, 1)$ can be used to induce a group structure on $(-1, 1)$.

Several comments seem to be in order. First, except for the notion of an ordering (which can be ignored), the foregoing discussion is accessible to any student studying abstract algebra and differential equations. In fact, it is likely that the verification that the open interval $(-1, 1)$ with the operation $*$ is indeed a group has appeared as an exercise in some abstract algebra textbook (though the authors failed to find any mention of this in any of the textbooks they consulted). The discovery of this group operation on $(-1, 1)$ by the second author arose from his investigation of the operation $*$ defined for points in the open unit disc in the complex plane. (Associativity fails for $*$ here.) The group G also provides a nice example of a "bounded" group which is isomorphic to the additive reals. Finally, it is worth observing that the existence of $f'(0)$ may easily be assumed for beginning mathematics students. However, if we seek a group isomorphism from $\mathbb{R}$ onto G which is also assumed to be a homeomorphism and thus *a fortiori* continuous, then, since $\mathbb{R}$ and G are both Lie groups (this requires verification), such a mapping is necessarily differentiable (see [**1**, p. 109]).

The authors are grateful to the editor and referees for their helpful comments and suggestions.

Reference

[1] Frank W. Warner, Foundations of Differentiable Manifolds and Lie Groups, Scott, Foresman and Company, Glenview, Illinois, 1971.

A Tale of Two Goats

MARSHALL FRASER
Lincoln University
San Francisco, CA 94118

While attending a math conference a friend asked me to solve this problem:

A goat is grazing in a circular field. The goat is tethered to the fence encircling the field by a rope. How long should the rope be so that the goat grazes half the area of the field?

The question had that pleasant quality of being easy to visualize. I felt I should be able to solve it. Also, I wondered if I hadn't seen it before. To stall for time I asked my friend to define "tethered." Of course, to no avail. I was unable to solve the problem.

My ego was soothed when at dinner that evening several friends filled their placemats and

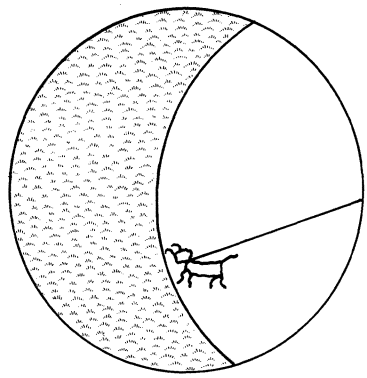


FIGURE 1

napkins with diagrams and integrals. But no solution was actually found. The problem was harder, and more interesting, than I had first thought. I decided to look into it.

Once I was home at my desk I was able to beard the goat, so to speak (not without encountering some difficulties, however). I was still curious about its history. I wrote to Howard Eves who suggested that the problem had been proposed in various old mathematics journals. Sure enough, I started with the very first volume of *The American Mathematical Monthly* (published in 1894), and there it was. It was not stated exactly in the form above but Geometry Problem Number 30 had the essentials:

> A circle containing one acre is cut by another whose center is on the circumference of the given circle, and the area common to both is one-half acre. Find the area of the cutting circle.

The solution published seven months later was quite the same as mine. Five people were credited with the published solution, and four other people were also credited with solutions. I looked through subsequent volumes of the *Monthly* and found that, at least until 1900, the problem was fairly popular. It also has been posed in a variety of puzzle books. The bibliography at the end of this paper makes no claim as to completeness, but gives some references to grazing goat problems.

I will discuss the original grazing goat problem and one of the variations of the problem. In the original the goat grazes inside the fence and in the variation the goat grazes outside the fence. If the goat grazing outside the fence seems unnatural to you, the grazing habits of the goats in some of the other variations are more unnatural still.

The goat inside the fence

We proceed in the most obvious way—draw and label a picture, set up an equation, and then solve the equation. It's easy to draw a picture as in FIGURE 1. But we can save some work if we are clever in labeling the picture. Clearly we will draw two intersecting circles with the center of the larger circle on the circumference of the smaller (FIGURE 2). How do we know which circle is larger? A little experimentation shows that the rope must be at least a little longer than the radius of the field. For if we draw a picture where the length of the rope equals the radius, we see that we do not get a large enough area. So we modify the picture with the cutting circle larger.

If we are lucky we will first draw the line through the centers of the two circles and the line through the two points of intersection as in FIGURE 2. (I was unlucky and first put in Cartesian coordinates, which made the problem harder. A moral is that coordinates are no cure-all in problem solving.) As usual, we label the radius of the field r. I also labeled the length of the rope kr, rather than x or l. This was because my intuition told me I wanted to find how many times longer the rope was than the radius of the field. Watch for r to cancel out of the equations.

Does it make sense to label the angles α and β as in FIGURE 2? We want to find the area that the goat grazes, and this area is the disjoint union of segments of the two circles. But this means that their central angles are important, so we label angles α and β.

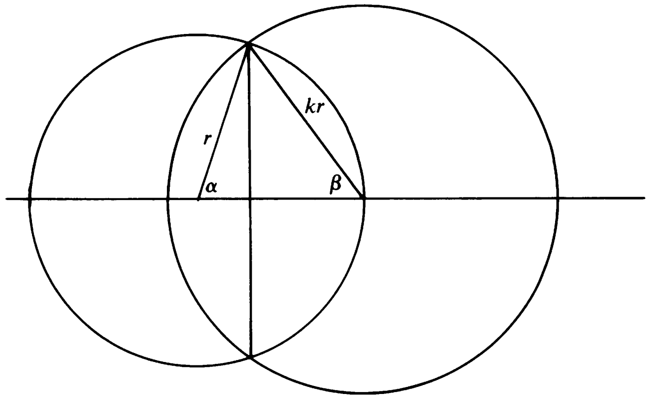


FIGURE 2

In presenting the grazing goat to students I have found that this first stage presents the greatest difficulty. They must be encouraged to draw construction lines and to label a figure. I would say that it is no sin to identify too many variables, because the process of labeling a figure will lead us to think about which lines and angles are important.

If you don't recall the formula for the area of a segment of a circle, it seems to me to be a nice, little trigonometric exercise (see FIGURE 3). The formula is

$$\text{area of segment} = \frac{r^2}{2}(2\theta - \sin 2\theta).$$

Now we can write down a formula for the area the goat grazes:

$$\text{area grazed} = \frac{r^2}{2}(2\alpha - \sin 2\alpha) + \frac{k^2 r^2}{2}(2\beta - \sin 2\beta).$$

Surely we can simplify this formula by finding relationships between α, β and k. You may already have noticed that the triangle is isosceles, and so

$$\alpha + 2\beta = \pi.$$

We also have

$$k = 2\cos\beta.$$

(You can prove this using right triangles, or you can use the law of cosines.)

Now we use trigonometric identities to simplify the formula for the area grazed. We can express the area as a function of either α or β. We get a slightly simpler function by using β:

$$\text{area grazed} = r^2[\pi + 2\beta\cos 2\beta - \sin 2\beta].$$

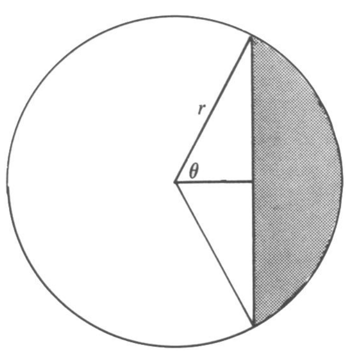


FIGURE 3. **The segment is shaded; the central angle is 2θ.**

It's time to set up the required equation:

$$\text{area grazed} = \tfrac{1}{2}\ \text{area of field}$$

$$r^2[\pi + 2\beta\cos 2\beta - \sin 2\beta] = \tfrac{1}{2}\pi r^2.$$

Cancel r^2 and rearrange a bit:

$$\sin 2\beta - 2\beta\cos 2\beta = \pi/2.$$

If we prefer we can let $\gamma = 2\beta$ and write

$$\sin\gamma - \gamma\cos\gamma = \pi/2.$$

A nice looking equation, but we see that it is transcendental! At first this surprised me. More experience with grazing goats, however, showed that it was not uncommon to obtain nonalgebraic equations. Even goats have their transcendental moments when they graze.

Surely everyone now reaches for a calculator. This equation is easy to solve by trial and error or is a nice example to be solved by Newton's method. I obtained the approximate solution

$$\begin{aligned}\gamma &\approx 1.90569573 \text{ radians},\\ \beta &\approx 0.952847865 \text{ radians}\\ &\approx 54^\circ\ 35'\ 39'',\\ k = 2\cos\beta &\approx 1.158728472.\end{aligned}$$

Therefore, the rope should be approximately 1.16 times as long as the radius of the field.

Actually, I came upon this solution in a roundabout way. First I used calculus to set up an equation and obtained the formidable-looking

$$\frac{k\sqrt{4-k^2}}{2} + k^2\arcsin\frac{k}{2} + \arcsin\frac{2-k^2}{2} = \frac{k^2\pi}{2}.$$

Then I realized that this equation could be simplified using $k = 2\cos\beta$. It's a pleasant exercise and shows that trigonometry is easier than calculus.

A natural variation I did not find in the *Monthly* is to have the goat graze in a square field with the rope tethered to a midpoint on the fence. I will let you set up and solve this problem which is the same level of difficulty as the original.

You may also wish to investigate the function

$$A(\beta) = \pi + 2\beta\cos 2\beta - \sin 2\beta.$$

This function gives the area grazed in the unit circle and is strictly decreasing between $\beta = 0$ and $\beta = \pi/2$. It has an inflection point at $\beta \approx 58^\circ\ 7'\ 11''$. What is an interpretation of this inflection point in terms of the goat grazing?

The goat outside the fence

This variation is posed as follows:

A goat is grazing outside a circular field. The goat is tethered by a rope to the outside of the fence encircling the field. How long should the rope be so that the goat grazes half the area of the field?

We begin with a sketch like FIGURE 4. In this case (for reasons which will soon be clear) we call the radius of the field a and the length of the rope ka.

The border of the region that the goat can graze is partly a semicircle, but the rest of the border is not part of any circle. This is because the rope starts winding about the fence, continually shortening the goat's reach. So this portion of the border is an involute of the circle.

With some reluctance we introduce polar coordinates to compute the area grazed. The polar axis is the line through the centers of the circles, with the origin at the center of the field. The points along the 'bad' portion of the border are represented as (r, θ) in FIGURE 4. The length of the rope, ka, breaks up naturally into the sum $l_1 + l_2$, where l_1 is the length that has wound around part of the circle, and l_2 is a straight line segment. The angle θ is broken up accordingly,

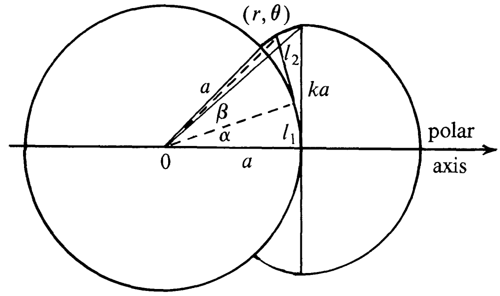


FIGURE 4

into $\alpha+\beta$.

We can obtain parametric equations for (r,θ) using β as a parameter. Clearly we have $r=a\sec\beta$. Now for θ, notice that $ka=l_1+l_2=a\alpha+a\tan\beta$ implies $k=\alpha+\tan\beta$. So $\theta=\alpha+\beta$ $=k+\beta-\tan\beta$. Therefore we have the parametric equations

$$r=a\sec\beta,$$
$$\theta=k+\beta-\tan\beta.$$

To find the area of the bad part of the region grazed we will integrate in polar coordinates. Notice that the bad area is the area of a bad wedge plus the area of a triangle minus the area of a sector. To find the area of the bad wedge we integrate:

$$\text{area bad wedge}=\frac{1}{2}\int_{\arctan k}^{k} r^2\,d\theta=\frac{a^2k^3}{6}.$$

Now we piece all these areas together. We have

$$\text{bad area}=\text{area bad wedge}+\text{area triangle}-\text{area sector}=\frac{a^2k^3}{6}.$$

The bad area is not so bad after all. Obtaining such a simple expression for the area suggests that we look for a simpler way to derive it. We like to think that simple answers possess simple derivations. Once again, the introduction of coordinates, which comes to most of us quite instinctively, obscures an easier approach.

Look at FIGURE 5 and think of the bad area as being approximated by a Riemann sum. The elements of area can be approximated by the area of a sector of a circle of radius PT. Thus

$$\Delta A\approx\tfrac{1}{2}\left(PT^2\Delta\phi\right)=\tfrac{1}{2}a^2\phi^2\Delta\phi,$$

so

$$\text{bad area}=\frac{1}{2}a^2\int_0^k\phi^2\,d\phi=\frac{a^2k^3}{6}.$$

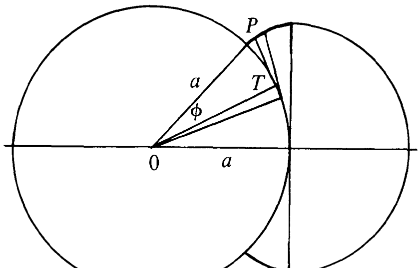


FIGURE 5

This coordinate-free method for finding the area under an involute was used in several solutions listed in the bibliography. It is presently a rather neglected topic of undergraduate teaching, but arises naturally in this setting if you wish to revive it.

We now have the total area the goat grazes:

$$\text{area grazed} = 2 \text{ bad areas} + \text{area of semicircle}$$
$$= \frac{a^2k^2}{6}(2k+3\pi).$$

We can set up and solve the required equation:

$$\text{area grazed} = \tfrac{1}{2} \text{ area field}$$
$$\frac{a^2k^2}{6}(2k+3\pi) = \tfrac{1}{2}\pi a^2$$
$$2k^3 + 3\pi k^2 = 3\pi.$$

At least this goat provides us with a polynomial equation, and a cubic at that. I found it interesting to find the exact solution. We eliminate the quadratic term with the substitution $k = x - \pi/2$. After a little simplification we obtain

$$4x^3 - 3\pi^2 x = 6\pi - \pi^3.$$

Now we can make the substitution $x = \pi\cos\phi$, a nice trick, which yields

$$\cos 3\phi = \frac{6}{\pi^2} - 1$$
$$\phi = \frac{1}{3}\arccos\left(\frac{6}{\pi^2} - 1\right).$$

Tracing back our steps, we obtain

$$k = \pi\cos\phi - \frac{\pi}{2},$$

with the value of ϕ given above. An approximate solution is

$$k \approx 0.9150896408.$$

In this variation the rope should be slightly shorter than the radius of the field—about 0.915 of the radius.

My experience with these two goat problems was to approach them using a "standard," "tried-and-true" method by introducing coordinates. In each case I later discovered a simpler, more satisfying solution, thus increasing my appreciation of the problems.

I find both of these problems interesting. Each is easy to comprehend. Each involves some difficulties, but the difficulties are surmountable using some fine elementary mathematics. Also, the problems invite further questions. I'm presently looking into an n-dimensional generalization of the grazing goat which I hope to write up soon.

I wish to thank Cathi Colin, Fred Safier, and Peter Renz for discussions and suggestions concerning grazing goats, Howard Eves for his informative letter, and a referee for many suggestions.

References

Here are some grazing goat problems from early issues of *The American Mathematical Monthly*. The animal, when there is one, is a horse. A mathematician who has herded goats tells me that they are much too independent to submit to tethering.

[1] Geometry Problem 30; proposed in 1 (1894) 132; solution in 1 (1894) 395–96.

[2] Arithmetic Problem 32; proposed in 1 (1894) 266; solution in 1 (1894) 431.
Horse tethered to a corner of a square building.

[3] Geometry Problem 34; proposed as Problem 33 in 1 (1894) 317; solution as Problem 34 in 2 (1895) 48-49.
Both problems—a horse can graze inside the fence and then outside the fence. The solution published for outside the fence seems obviously incorrect.
[4] Calculus Problem 37; proposed in 2 (1895) 52; solution in 2 (1895) 277–78.
Two mules graze together, but on opposite sides of the fence.
[5] Calculus Problem 55; proposed in 3 (1896) 148; solutions in 4 (1897) 17–18.
Both problems again—a horse can graze outside the fence and then inside the fence. Involutes enter into the solutions.
[6] Calculus Problem 69; proposed in 5 (1898) 29; solution in 5 (1898) 111; note in 5 (1898) 177.
Horse is tethered to a sliding ring outside an elliptical field, perhaps the most exotic problem. The note points out that the published solution is incorrect, but a correct solution is not offered.
[7] Arithmetic Problem 93; proposed as Problem 91 in 5 (1898) 60; solution as Problem 93 in 5 (1898) 170–71.
Another horse is tethered to the corner of a barn. But six different answers were received.
[8] Geometry Problem 103; proposed in 5 (1898) 217; solutions in 5 (1898) 295–96.
A horse is grazing on the edge of a circular pond. Three solutions reflect different assumptions about whether the rope can stretch across the pond. Curiously, the answer is almost the same in each case.
[9] Calculus Problem 103; proposed in 6 (1899) 289; remark in 7 (1900) 267–68.
Essentially the same as Calculus Problem 69—a horse is grazing outside an elliptical field tethered to a sliding ring. No solution was received and, so far as I know, these two problems remain unsolved.

Grazing goat problems and some variants from recent sources.

[10] V. W. B., An iterative process: the goat's share revisited, Math. Gaz., 65 (1981) 137–39.
The classical goat; discussion of iterative calculator solution of the formidable-looking arcsine equation.
[11] Howard P. Dinesman, Superior Mathematical Puzzles, Simon and Schuster, New York, 1968; Puzzle 8 on p. 20 and Puzzle 53 on p. 71.
Puzzle 8 is very easy; in Puzzle 53 a goat is tethered to a circular silo.
[12] Jordi Dou, Solution to Problem S19 (proposed by Harley Flanders), Amer. Math. Monthly, 88 (1981) 147.
[13] Henry Earnest Dudeney, Amusements in Mathematics, Dover, New York, 1958 (reprint of 1917 edition), Problem 196 on p. 53.
Goat grazing in an equilateral triangle; easy.
[14] L. A. Graham, Ingenious Mathematical Problems, Dover, New York, 1959, p. 6.
This has a literary twist, and is called "Mrs. Miniver's Problem."
[15] S. I. Jones, Mathematical Nuts, S. I. Jones Co., 1932.
Note the quick calculation of the area of the involute.
[16] P. M. H. Kendall and G. M. Thomas, Mathematical Puzzles for the Connoisseur, Thomas Y. Crowell, New York, 1962, pp. 24–25.
A macabre goat is tethered to a mausoleum in a circular field. A nice involute problem.
[17] L. H. Longley-Cook, Work This One Out, Fawcett, 1960, Problem 69.
[18] James E. Schultz and Bert K. Waits, A new look at some old problems in light of the hand calculator, TYCMJ, 10 (1979) 20–27.
Discusses calculator solutions of three classical problems—one of which is the grazing goat. A simple derivation using sectors of a complicated-looking equation.

Triangle Constructions with Three Located Points

William Wernick
3986 Hillman Avenue
Bronx, NY 10463

Given a triangle ABC we can construct its medians, AM_a, BM_b, and CM_c which are concurrent at the centroid G (Figure 1). Now suppose we erase almost all of this figure, leaving only the points A, B, and M_a *in position*. Can we reconstruct the original figure? Yes, very easily, since if we double the segment BM_a we get the point C.

[3] Geometry Problem 34; proposed as Problem 33 in 1 (1894) 317; solution as Problem 34 in 2 (1895) 48-49.
Both problems—a horse can graze inside the fence and then outside the fence. The solution published for outside the fence seems obviously incorrect.
[4] Calculus Problem 37; proposed in 2 (1895) 52; solution in 2 (1895) 277–78.
Two mules graze together, but on opposite sides of the fence.
[5] Calculus Problem 55; proposed in 3 (1896) 148; solutions in 4 (1897) 17–18.
Both problems again—a horse can graze outside the fence and then inside the fence. Involutes enter into the solutions.
[6] Calculus Problem 69; proposed in 5 (1898) 29; solution in 5 (1898) 111; note in 5 (1898) 177.
Horse is tethered to a sliding ring outside an elliptical field, perhaps the most exotic problem. The note points out that the published solution is incorrect, but a correct solution is not offered.
[7] Arithmetic Problem 93; proposed as Problem 91 in 5 (1898) 60; solution as Problem 93 in 5 (1898) 170–71.
Another horse is tethered to the corner of a barn. But six different answers were received.
[8] Geometry Problem 103; proposed in 5 (1898) 217; solutions in 5 (1898) 295–96.
A horse is grazing on the edge of a circular pond. Three solutions reflect different assumptions about whether the rope can stretch across the pond. Curiously, the answer is almost the same in each case.
[9] Calculus Problem 103; proposed in 6 (1899) 289; remark in 7 (1900) 267–68.
Essentially the same as Calculus Problem 69—a horse is grazing outside an elliptical field tethered to a sliding ring. No solution was received and, so far as I know, these two problems remain unsolved.

Grazing goat problems and some variants from recent sources.

[10] V. W. B., An iterative process: the goat's share revisited, Math. Gaz., 65 (1981) 137–39.
The classical goat; discussion of iterative calculator solution of the formidable-looking arcsine equation.
[11] Howard P. Dinesman, Superior Mathematical Puzzles, Simon and Schuster, New York, 1968; Puzzle 8 on p. 20 and Puzzle 53 on p. 71.
Puzzle 8 is very easy; in Puzzle 53 a goat is tethered to a circular silo.
[12] Jordi Dou, Solution to Problem S19 (proposed by Harley Flanders), Amer. Math. Monthly, 88 (1981) 147.
[13] Henry Earnest Dudeney, Amusements in Mathematics, Dover, New York, 1958 (reprint of 1917 edition), Problem 196 on p. 53.
Goat grazing in an equilateral triangle; easy.
[14] L. A. Graham, Ingenious Mathematical Problems, Dover, New York, 1959, p. 6.
This has a literary twist, and is called "Mrs. Miniver's Problem."
[15] S. I. Jones, Mathematical Nuts, S. I. Jones Co., 1932.
Note the quick calculation of the area of the involute.
[16] P. M. H. Kendall and G. M. Thomas, Mathematical Puzzles for the Connoisseur, Thomas Y. Crowell, New York, 1962, pp. 24–25.
A macabre goat is tethered to a mausoleum in a circular field. A nice involute problem.
[17] L. H. Longley-Cook, Work This One Out, Fawcett, 1960, Problem 69.
[18] James E. Schultz and Bert K. Waits, A new look at some old problems in light of the hand calculator, TYCMJ, 10 (1979) 20–27.
Discusses calculator solutions of three classical problems—one of which is the grazing goat. A simple derivation using sectors of a complicated-looking equation.

Triangle Constructions with Three Located Points

William Wernick
3986 Hillman Avenue
Bronx, NY 10463

Given a triangle ABC we can construct its medians, AM_a, BM_b, and CM_c which are concurrent at the centroid G (Figure 1). Now suppose we erase almost all of this figure, leaving only the points A, B, and M_a *in position*. Can we reconstruct the original figure? Yes, very easily, since if we double the segment BM_a we get the point C.

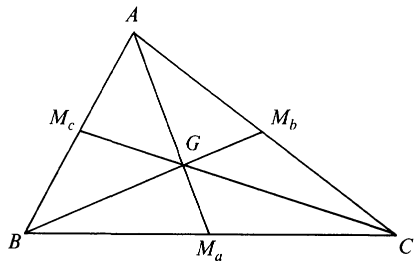


FIGURE 1

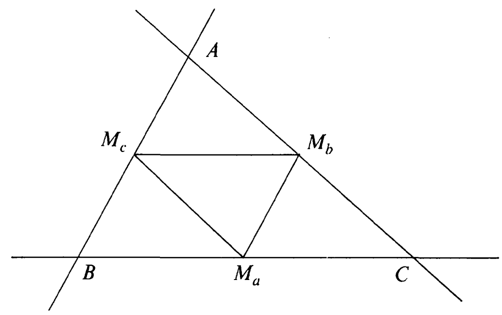


FIGURE 2

Suppose we erase all but the three midpoints M_a, M_b, M_c, *in location*; can we again reconstruct the original figure? One solution is to start by drawing a line through M_c parallel to the line M_aM_b, then another line through M_a parallel to the line M_bM_c. These two constructed lines will intersect in point B, a vertex of the original triangle, and the rest of the construction follows easily (FIGURE 2). We indicate this solution as follows:

$$//(M_c, M_aM_b) \cap //(M_a, M_bM_c) = B; \qquad 2(BM_c) = A, \qquad 2(BM_a) = C.$$

It is interesting to investigate the construction of a triangle ABC, given certain triples of located points selected from the following list of sixteen points (see FIGURES 1, 3, 4, 5):

A, B, C, O	Three vertices and circumcenter
M_a, M_b, M_c, G	Three feet of the medians, and centroid
H_a, H_b, H_c, H	Three feet of the altitudes, and orthocenter
T_a, T_b, T_c, I	Three feet of the internal angle bisectors, and incenter.

In these problems we may take two approaches: (1) we assume that a triangle has been erased, except for three located points, and we try to recover that original triangle; or (2) we choose any three distinct points of the plane and designate these as three particular points among the list of sixteen, then try to construct a triangle to fit. It is clear that the second approach includes the first and is a little more general, raising questions of constructibility and redundancy. Since it is more interesting, it is the approach we shall use.

The list in TABLE 1 is a careful compilation of exactly 139 such problems, all significantly distinct. For example, the selection of the triple of points to be two vertices and the centroid of a triangle could be listed as A, B, G; or B, C, G; or A, C, G, which are surely not significantly distinct. Only the first choice is listed (problem 4). Some selections of triples are redundant, that is, one of the points can be determined from the other two. For example, the triple A, B, M_c is redundant since M_c is the midpoint of segment AB (clearly any two of these three points determine the third). Such redundant triples in TABLE 1 are noted with the letter **R**; I have found three such triples in this list.

Note also that some selections, while not redundant, are still restricted as to choice of the points; the triple A, B, O is such a selection, since the circumcenter O must lie on the perpendicu-

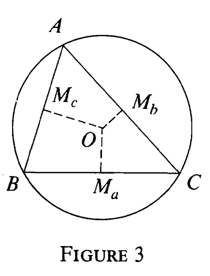


FIGURE 3

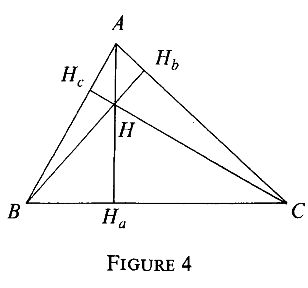


FIGURE 4

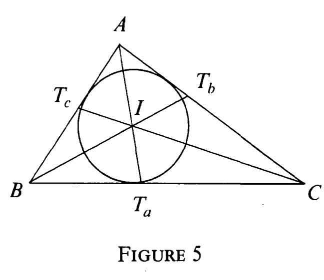


FIGURE 5

lar bisector of the side AB. If the point designated O does lie on the perpendicular bisector of the segment AB, then the third vertex C of the triangle may lie anywhere on the circumcircle with center O and radius OA. If the point designated as O does not lie on the perpendicular bisector of segment AB, then no solution is possible. In such a situation, the locus restriction gives us either infinitely many or no solutions to the problem. Twenty-two of the problems in TABLE 1 have been identified as being of this type; they are designated with the symbol **L**.

Of the remaining 114 problems, I have found solutions for only 65 of these, which are indicated on the list with the symbol **S**. Many of these 65 are quite easy and straightforward to solve. Some are harder and would challenge any student of geometry. Problems 69, 76, and 101 in particular are real challenges; they appear as Proposal 1149 in the Problems section of this issue of this *Magazine*, p. 236.

Fortunately, some problems are related, in the sense that if you solve one of them, you have ready solutions to some others. Consider, for example, problems 26, 42, and 95, whose triples of given points are as follows:

$$\begin{aligned} 26&: \quad A, M_a, T_b \\ 42&: \quad A, G, T_b \\ 95&: \quad M_a, G, T_b. \end{aligned}$$

Since the triple A, M_a, G is redundant (problem 21), it follows that a solution to any one of the problems 26, 42, 95 leads to immediate solutions of the other two problems. Other sets of such related problems are 27, 43, 96; 72, 79, 120; and 73, 80, 121.

Our list in TABLE 1 contains 41 seemingly independent, presently unsolved problems. It may be that some of these lead to provably impossible constructions, in which case such a proof would

1. A,B,O	L	29. A,M_b,G	S	57. A,H,I		85. M_a,M_b,H_a	S	113. M_a,T_b,T_c	
2. A,B,M_a	S	30. A,M_b,H_a	L	58. A,T_a,T_b		86. M_a,M_b,H_c	S	114. M_a,T_b,I	
3. A,B,M_c	R	31. A,M_b,H_b	L	59. A,T_a,I	L	87. M_a,M_b,H		115. G,H_a,H_b	
4. A,B,G	S	32. A,M_b,H_c	L	60. A,T_b,T_c	S	88. M_a,M_b,T_a		116. G,H_a,H	S
5. A,B,H_a	L	33. A,M_b,H	S	61. A,T_b,I	S	89. M_a,M_b,T_c		117. G,H_a,T_a	S
6. A,B,H_c	L	34. A,M_b,T_a	S	62. O,M_a,M_b	S	90. M_a,M_b,I		118. G,H_a,T_b	
7. A,B,H	S	35. A,M_b,T_b	L	63. O,M_a,G	S	91. M_a,G,H_a	L	119. G,H_a,I	
8. A,B,T_a	S	36. A,M_b,T_c	S	64. O,M_a,H_a	L	92. M_a,G,H_b	S	120. G,H,T_a	
9. A,B,T_c	L	37. A,M_b,I	S	65. O,M_a,H_b	S	93. M_a,G,H	S	121. G,H,I	
10. A,B,I	S	38. A,G,H_a	L	66. O,M_a,H	S	94. M_a,G,T_a	S	122. G,T_a,T_b	
11. A,O,M_a	S	39. A,G,H_b	S	67. O,M_a,T_a	L	95. M_a,G,T_b		123. G,T_a,I	
12. A,O,M_b	L	40. A,G,H	S	68. O,M_a,T_b		96. M_a,G,I		124. H_a,H_b,H_c	S
13. A,O,G	S	41. A,G,T_a	S	69. O,M_a,I	S	97. M_a,H_a,H_b	S	125. H_a,H_b,H	S
14. A,O,H_a	S	42. A,G,T_b		70. O,G,H_a	S	98. M_a,H_a,H	L	126. H_a,H_b,T_a	S
15. A,O,H_b	S	43. A,G,I		71. O,G,H	R	99. M_a,H_a,T_a	L	127. H_a,H_b,T_c	
16. A,O,H	S	44. A,H_a,H_b	S	72. O,G,T_a		100. M_a,H_a,T_b		128. H_a,H_b,I	
17. A,O,T_a	S	45. A,H_a,H	L	73. O,G,I		101. M_a,H_a,I	S	129. H_a,H,T_a	L
18. A,O,T_b	S	46. A,H_a,T_a	L	74. O,H_a,H_b		102. M_a,H_b,H_c	S	130. H_a,H,T_b	
19. A,O,I	S	47. A,H_a,T_b	S	75. O,H_a,H	S	103. M_a,H_b,H	S	131. H_a,H,I	
20. A,M_a,M_b	S	48. A,H_a,I	S	76. O,H_a,T_a	S	104. M_a,H_b,T_a	S	132. H_a,T_a,T_b	
21. A,M_a,G	R	49. A,H_b,H_c	S	77. O,H_a,T_b		105. M_a,H_b,T_b	S	133. H_a,T_a,I	S
22. A,M_a,H_a	L	50. A,H_b,H	L	78. O,H_a,I		106. M_a,H_b,T_c		134. H_a,T_b,T_c	
23. A,M_a,H_b	S	51. A,H_b,T_a	S	79. O,H,T_a		107. M_a,H_b,I		135. H_a,T_b,I	
24. A,M_a,H	S	52. A,H_b,T_b	L	80. O,H,I		108. M_a,H,T_a		136. H,T_a,T_b	
25. A,M_a,T_a	S	53. A,H_b,T_c	S	81. O,T_a,T_b		109. M_a,H,T_b		137. H,T_a,I	
26. A,M_a,T_b		54. A,H_b,I	S	82. O,T_a,I		110. M_a,H,I		138. T_a,T_b,T_c	
27. A,M_a,I		55. A,H,T_a	S	83. M_a,M_b,M_c	S	111. M_a,T_a,T_b		139. T_a,T_b,I	S
28. A,M_b,M_c	S	56. A,H,T_b		84. M_a,M_b,G	S	112. M_a,T_a,I	S		

TABLE 1. **For each of the 139 triples of points listed, construct the corresponding triangle ABC. Problems solved by the author are noted with an S, L, or R to designate a solution triangle ABC, a locus-dependent solution, or a redundant triple, respectively.**

constitute a solution. However, I have the firm conviction that they can all be "done," and that eventually every one of the 139 problems will have the appropriate **R** or **L** or **S** designation. I welcome any companionship and friendly competition on the last 41 (which incidentally is a prime number with pleasant associations).

I close with a few solutions to whet your appetite.

PROBLEM 18. Given points A, O, T_b.

Solution. $\perp(O, AT_b) \cap \text{Cir}(O, OA) = K$; $KT_b \cap \text{Cir}(O, OA) = B$; $AT_b \cap \text{Cir}(O, OA) = C$. In words, the perpendicular from O to the line AT_b will meet the circle with center O and radius OA at the point K; the line KT_b will meet the circle with center O and radius OA at the point B; (now it's your turn) $AT_b \cap \text{Cir}(O, OA) = C$.

PROBLEM 49. Given points A, H_b, H_c.

Solution. $\text{Cir}(A, H_b, H_c) = \text{Cir}(K)$; $\text{Diam}\,\text{Cir}(K) = AH$; $H_bH \cap AH_c = B$; $H_cH \cap AH_b = C$. That is, the circle which goes through points A, H_b, and H_c (the circumcircle of triangle AH_bH_c) we call circle K; the diameter of circle K which has one end at A has the other end at the point H; the line H_bH will meet the line AH_c at the point B; and now $H_cH \cap AH_b = C$.

PROBLEM 65. Given points O, M_a, H_b.

Solution. $\perp(M_a, M_aO) = xM_ay$; $\text{Cir}(M_a, M_aH_b) \cap xM_ay = B, C$; $\text{Cir}(O, OB) = \text{Cir}(K)$; $\text{Cir}(K) \cap CH_b = A$. The perpendicular at the point M_a to the line M_aO we call the line xM_ay; the circle with center M_a and radius M_aH_b will meet the line xM_ay at points B and C; the circle with center O and radius OB is called circle K; circle K will meet the line CH_b at point A.

PROBLEM 103. Given points M_a, H_b, H.

Solution. $\perp(M_a, H_bH) \cap H_bH = Q$; $2(H_bQ) = B$; $2(BM_a) = C$; $\perp(B, CH) \cap CH_b = A$. The perpendicular from M_a to the line H_bH will meet the line H_bH at the point Q (that is, Q is the orthogonal projection of M_a to line H_bH); extend segment H_bQ to double its length to get point B; double the segment BM_a to get point C; the perpendicular from point B to line CH will meet line CH_b at point A.

A Papal Conclave: Testing the Plausibility of a Historical Account

ANTHONY LO BELLO
Allegheny College
Meadville, PA 16335

Among the stories to be found in Valérie Pirie's *The Triple Crown*, a gossipy and unreliable account of the Papal conclaves since 1458, is one having to do with the election of 1513 that appears particularly suspicious to a reader with some mathematical abilities.

> Twenty-five cardinals entered the conclave. The absence of the French element left practically only two contending parties—the young and the old. The former had secretly settled on Giovanni de' Medici; the second openly supported S. Giorgio, England's candidate.... The Sacred College had been assembled almost a week before the first serious scrutiny took place. Many of the cardinals, wishing to temporise and conceal their real intentions, had voted for the man they considered least likely to have any supporters. As luck would have it, thirteen prelates had selected the same outsider, with the result that they all but elected Arborense, the

constitute a solution. However, I have the firm conviction that they can all be "done," and that eventually every one of the 139 problems will have the appropriate **R** or **L** or **S** designation. I welcome any companionship and friendly competition on the last 41 (which incidentally is a prime number with pleasant associations).

I close with a few solutions to whet your appetite.

PROBLEM 18. Given points A, O, T_b.

Solution. $\perp(O, AT_b) \cap \mathrm{Cir}(O, OA) = K$; $KT_b \cap \mathrm{Cir}(O, OA) = B$; $AT_b \cap \mathrm{Cir}(O, OA) = C$. In words, the perpendicular from O to the line AT_b will meet the circle with center O and radius OA at the point K; the line KT_b will meet the circle with center O and radius OA at the point B; (now it's your turn) $AT_b \cap \mathrm{Cir}(O, OA) = C$.

PROBLEM 49. Given points A, H_b, H_c.

Solution. $\mathrm{Cir}(A, H_b, H_c) = \mathrm{Cir}(K)$; $\mathrm{Diam}\,\mathrm{Cir}(K) = AH$; $H_bH \cap AH_c = B$; $H_cH \cap AH_b = C$. That is, the circle which goes through points A, H_b, and H_c (the circumcircle of triangle AH_bH_c) we call circle K; the diameter of circle K which has one end at A has the other end at the point H; the line H_bH will meet the line AH_c at the point B; and now $H_cH \cap AH_b = C$.

PROBLEM 65. Given points O, M_a, H_b.

Solution. $\perp(M_a, M_aO) = xM_ay$; $\mathrm{Cir}(M_a, M_aH_b) \cap xM_ay = B, C$; $\mathrm{Cir}(O, OB) = \mathrm{Cir}(K)$; $\mathrm{Cir}(K) \cap CH_b = A$. The perpendicular at the point M_a to the line M_aO we call the line xM_ay; the circle with center M_a and radius M_aH_b will meet the line xM_ay at points B and C; the circle with center O and radius OB is called circle K; circle K will meet the line CH_b at point A.

PROBLEM 103. Given points M_a, H_b, H.

Solution. $\perp(M_a, H_bH) \cap H_bH = Q$; $2(H_bQ) = B$; $2(BM_a) = C$; $\perp(B, CH) \cap CH_b = A$. The perpendicular from M_a to the line H_bH will meet the line H_bH at the point Q (that is, Q is the orthogonal projection of M_a to line H_bH); extend segment H_bQ to double its length to get point B; double the segment BM_a to get point C; the perpendicular from point B to line CH will meet line CH_b at point A.

A Papal Conclave: Testing the Plausibility of a Historical Account

ANTHONY LO BELLO
Allegheny College
Meadville, PA 16335

Among the stories to be found in Valérie Pirie's *The Triple Crown*, a gossipy and unreliable account of the Papal conclaves since 1458, is one having to do with the election of 1513 that appears particularly suspicious to a reader with some mathematical abilities.

> Twenty-five cardinals entered the conclave. The absence of the French element left practically only two contending parties—the young and the old. The former had secretly settled on Giovanni de' Medici; the second openly supported S. Giorgio, England's candidate.... The Sacred College had been assembled almost a week before the first serious scrutiny took place. Many of the cardinals, wishing to temporise and conceal their real intentions, had voted for the man they considered least likely to have any supporters. As luck would have it, thirteen prelates had selected the same outsider, with the result that they all but elected Arborense, the

most worthless nonentity present. This narrow shave gave the Sacred College such a shock that its members determined to come to some agreement which would put matters on a more satisfactory basis for both parties.

[**5**, p. 49]

We may assume that all but a handful of the twenty-five cardinals were recognized as having no chance whatsoever of winning seventeen votes, the two-thirds majority required for an election. Of these, a smaller number m would have been considered to have absolutely no supporters; this number cannot have been too small, or there would not have been so great a surprise when one man among these "least likely to have any supporters" won thirteen votes. If we further assume that r cardinals, $17 \leqslant r \leqslant 25$, participated in the strategy, and that each voted at random on that first ballot for one of the m "dark horses," *what was the probability that one of the m would receive thirteen votes*? More importantly, since their eminences were terrified when this event actually occurred and determined to change their ways, we may wonder what was the probability of anyone actually being elected by this method. We shall presume that the $25-r$ prelates not participating in the maneuver each voted for one of the serious candidates.

It must be pointed out that it is quite possible that the m cardinals who were considered to have no supporters were ranked according to undesirability, in which case the voting would not have been at random. In our analysis which follows we assume that the m undesirable candidates were deemed "equally undesirable."

The number of different ways in which r votes can be randomly cast among m candidates is the same as the number of different ways r indistinguishable objects can be distributed among m cells, that is, $\binom{m+r-1}{r}$ (see [**2**, p. 31], and take $n=m-1$). The number of different ways in which a particular one of the m candidates can get thirteen votes is the same as the number of ways in which the remaining $r-13$ votes can be divided among the remaining $m-1$ candidates, i.e., $\binom{m+r-15}{r-13}$. The probability of the event described by Pirie is therefore

$$\frac{m\binom{m+r-15}{r-13}}{\binom{m+r-1}{r}}, \qquad m>1.$$

These probabilities are listed in TABLE 1. In compiling this table we have adopted the praiseworthy axiom of Emile Borel, presented in his discussion of the Petersburg Paradox in [**1**], that, as a working rule, events of probability less than .001 should be considered impossible. The table displays only those columns for which $17 \leqslant r \leqslant 20$, since a small number of pious cardinals would

m \ r	17	18	19	20
1	0	0	0	0
2	.1111	.1053	.1000	.0952
3	.0877	.0948	.1000	.1039
4	.0526	.0632	.0727	.0813
5	.0292	.0383	.0474	.0565
6	.0159	.0225	.0296	.0373
7	.0087	.0131	.0182	.0241
8	.0048	.0077	.0112	.0155
9	.0027	.0046	.0069	.0100
10	.0016	.0028	.0043	.0065
11	0	.0017	.0027	.0042
12	0	0	.0017	.0027
13	0	0	.0011	.0018
14	0	0	0	.0012
15	0	0	0	0

TABLE 1. **Probability of one of the "least likely" candidates receiving exactly 13 votes.**

not have taken part in what must have appeared to them to be a blasphemous procedure. However, it would seem from the last sentence quoted from Pirie that there were enough participants to raise the fear that one of the undesirables might be elected Pontiff. We do not have enough information to determine what values are most likely for m, though we might suppose that in those times there were great numbers of unworthy prelates.

The probability that one of the m candidates will receive at least thirteen votes under our hypotheses is

$$\frac{m}{\binom{m+r-1}{r}}\sum_{i=1}^{r-13}\binom{m+r-15-i}{r-13-i}=\frac{m\binom{m+r-14}{r-13}}{\binom{m+r-1}{r}}=\frac{m(m+r-14)!r!}{(r-13)!(m+r-1)!},\tag{1}$$

where we have used the identity [2, p. 30]

$$\binom{n}{r}=\sum_{i=1}^{r+1}\binom{n-i}{r-i+1}$$

to establish the first equality. For all values of r, the probability that one of the m candidates will get at least thirteen votes decreases as m increases. (See TABLE 2.)

We may next inquire whether the cardinals were mathematically justified in fearing that their strategy might actually produce a Pope. *What is the probability that some dark horse would get seventeen or more votes under the conditions imposed above?* By reasoning as we did to establish (1), we find that this probability is given by

$$\frac{m}{\binom{m+r-1}{r}}\sum_{i=0}^{r-17}\binom{m+r-19-i}{r-17-i}=\frac{m\binom{m+r-18}{r-17}}{\binom{m+r-1}{r}}=\frac{m(m+r-18)!r!}{(m+r-1)!(r-17)!}.$$

It would appear, upon examination of TABLE 3, that their eminences were wrong to abandon their strategem in fright, unless m was small (in which case it was a poor strategy to begin with). We may, finally, point out that if we ignore the trivial case of $m=1$, the probability of producing a winner is greatest when $m=2$ if $r\leqslant 23$ but greatest when $m=3$ for $r=24$ or 25.

A simpler problem, also arising from an election, is the "Imperial Election Problem": if the seven electors of the Holy Roman Emperor determined to elect one of themselves Emperor by random voting, what is the probability that an election would be accomplished on the first ballot?

m \ r	17	18	19	20
1	1	1	1	1
2	.5556	.6316	.7000	.7619
3	.2632	.3316	.4000	.4675
4	.1228	.1684	.2182	.2710
5	.0585	.0861	.1186	.1553
6	.0287	.0449	.0652	.0895
7	.0146	.0241	.0365	.0522
8	.0076	.0132	.0209	.0309
9	.0041	.0074	.0122	.0186
10	.0023	.0043	.0073	.0114
11	.0013	.0025	.0044	.0071
12	0	.0015	.0027	.0045
13	0	.0010	.0017	.0029
14	0	0	.0011	.0019
15	0	0	0	.0013
16	0	0	0	0

TABLE 2. **Probability of one of the "least likely" candidates receiving at least 13 votes.**

m \ r	17	18	19	20
1	1	1	1	1
2	.0556	.1053	.1500	.1905
3	.0176	.0474	.0857	.1299
4	.0035	.0120	.0260	.0452
5	.0008	.0034	.0085	.0165
6	0	0	0	.0063
7	0	0	0	.0026
8	0	0	0	.0011
9	0	0	0	0

TABLE 3. **Probability of a "dark horse" receiving 17 or more votes under the stated conditions.**

(A simple majority was required for election.) Another fable presented as fact in *The Triple Crown* is discussed in [3].

For a true account of the 1513 conclave, see [4]. The final outcome was that Giovanni de' Medici succeeded in emerging as Pope and reigned until 1521 under the title Leo X.

References

[1] E. Borel, Probability and Certainty, trans. by Douglas Scott, Walker and Co., 1963.
[2] J. Freund, Mathematical Statistics, 2nd ed., Prentice-Hall, Englewood Cliffs, 1971.
[3] A. Lo Bello, When Clio sleeps, The Tablet, London, (August 25, 1979) 821–824.
[4] L. Fr. von Pastor, History of the Popes, vol. 7, trans. by R. F. Kerr, Herder, St. Louis, 1908.
[5] V. Pirie, The Triple Crown, G. P. Putnam's Sons, New York, 1936.

An Application of Desargues' Theorem

JOHN MCCLEARY
Vassar College
Poughkeepsie, NY 12601

Among the highlights of projective geometry is

DESARGUES' THEOREM. *If two triangles are in perspective from a point, the three intersections of the extensions of their corresponding sides lie on one line.*

The purpose of this note is to apply Desargues' theorem to prove the result in Euclidean geometry known as the *Three-circle theorem* [3, p. 115]. The relation of projective geometry to Euclidean geometry makes Desargues' theorem, when properly interpreted, a useful tool in Euclidean proofs. This fact was exploited, for example, by the French geometer Poncelet [4].

In [1], Coxeter gives a proof of the Three-circle theorem using inversive geometry techniques. He also quotes the philosopher Herbert Spencer on his fascination with "a truth which I never contemplate without being struck with its beauty..." [5]. Though Spencer's public interest in mathematics is noteworthy, his understanding lagged far behind his enthusiasm. In [2], MacKay critiques Spencer's statements on mathematics and he describes the philosopher's mathematical knowledge as "both slender and scrappy."

m \ r	17	18	19	20
1	1	1	1	1
2	.0556	.1053	.1500	.1905
3	.0176	.0474	.0857	.1299
4	.0035	.0120	.0260	.0452
5	.0008	.0034	.0085	.0165
6	0	0	0	.0063
7	0	0	0	.0026
8	0	0	0	.0011
9	0	0	0	0

TABLE 3. **Probability of a "dark horse" receiving 17 or more votes under the stated conditions.**

(A simple majority was required for election.) Another fable presented as fact in *The Triple Crown* is discussed in [3].

For a true account of the 1513 conclave, see [4]. The final outcome was that Giovanni de' Medici succeeded in emerging as Pope and reigned until 1521 under the title Leo X.

References

[1] E. Borel, Probability and Certainty, trans. by Douglas Scott, Walker and Co., 1963.
[2] J. Freund, Mathematical Statistics, 2nd ed., Prentice-Hall, Englewood Cliffs, 1971.
[3] A. Lo Bello, When Clio sleeps, The Tablet, London, (August 25, 1979) 821–824.
[4] L. Fr. von Pastor, History of the Popes, vol. 7, trans. by R. F. Kerr, Herder, St. Louis, 1908.
[5] V. Pirie, The Triple Crown, G. P. Putnam's Sons, New York, 1936.

An Application of Desargues' Theorem

JOHN MCCLEARY
Vassar College
Poughkeepsie, NY 12601

Among the highlights of projective geometry is

DESARGUES' THEOREM. *If two triangles are in perspective from a point, the three intersections of the extensions of their corresponding sides lie on one line.*

The purpose of this note is to apply Desargues' theorem to prove the result in Euclidean geometry known as the *Three-circle theorem* [3, p. 115]. The relation of projective geometry to Euclidean geometry makes Desargues' theorem, when properly interpreted, a useful tool in Euclidean proofs. This fact was exploited, for example, by the French geometer Poncelet [4].

In [1], Coxeter gives a proof of the Three-circle theorem using inversive geometry techniques. He also quotes the philosopher Herbert Spencer on his fascination with "a truth which I never contemplate without being struck with its beauty..." [5]. Though Spencer's public interest in mathematics is noteworthy, his understanding lagged far behind his enthusiasm. In [2], MacKay critiques Spencer's statements on mathematics and he describes the philosopher's mathematical knowledge as "both slender and scrappy."

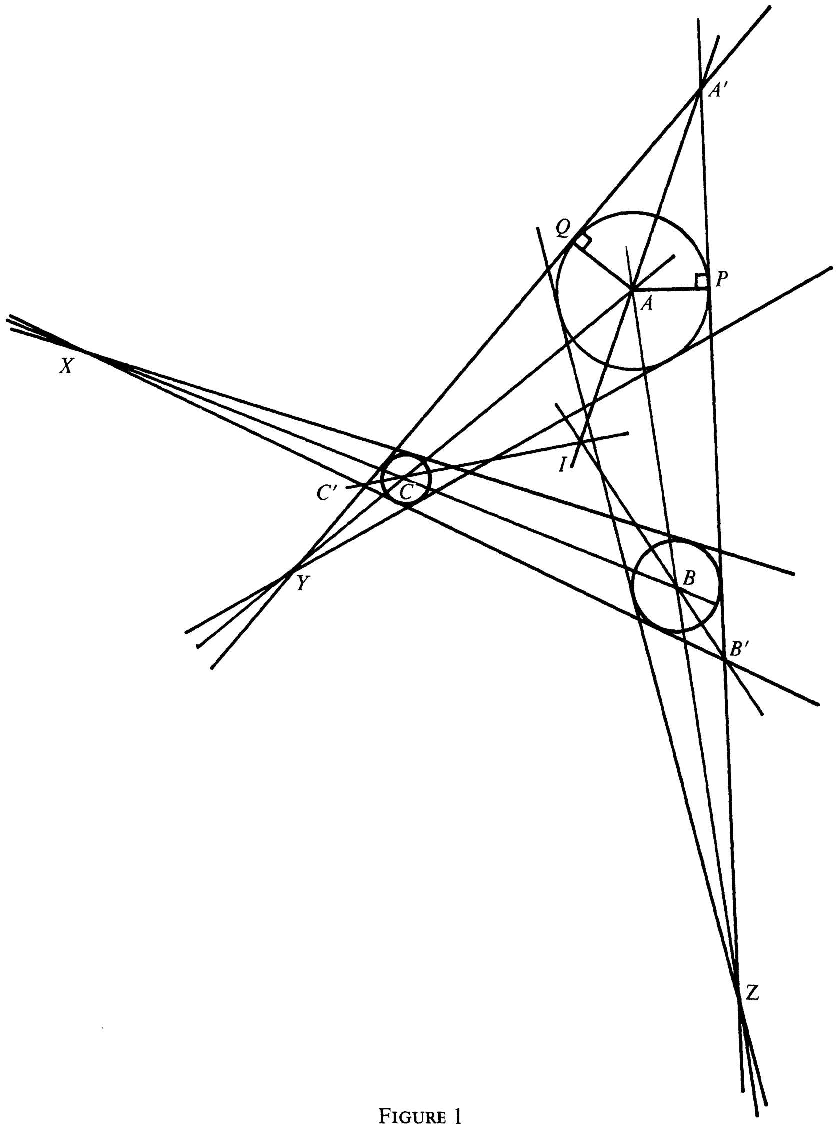


FIGURE 1

THREE-CIRCLE THEOREM. *Given three circles, nonintersecting, mutually external, and with distinct radii, then the three points obtained by taking the intersections of external common tangents of pairs of circles lie on one line.*

Proof. In the proof refer to FIGURE 1. We shall suppose that A, B, and C are the centers of the circles. First observe that the lines joining the centers of the circles also pass through the intersection points of the common tangents.

Consider the triangle $A'B'C'$ obtained from the intersections of tangents external to the triangle ABC. If P and Q denote the points of tangency to the circle centered at A of the lines $A'C'$ and $A'B'$, then triangles $AA'P$ and $AA'Q$ are seen to be congruent right triangles and so AA' bisects angle $B'A'C'$.

Similarly BB' and CC' are angle bisectors of the triangle $A'B'C'$. But the angle bisectors of a triangle are concurrent so AA', BB', and CC' meet at a point I. It follows immediately that triangles ABC and $A'B'C'$ are in perspective from the point I and hence, by Desargues' theorem, X, Y, and Z are collinear.

The reader will want to repeat the proof using the triangle obtained from the intersections of tangents interior to the triangle *ABC*. Using this interior triangle establishes the theorem in cases when the exterior triangle is degenerate. Furthermore, by introducing internal rather than external common tangents, new collinearities can be found, for example by taking the points of intersection of two pairs of common internal tangents and one pair of common external tangents. By a judicious choice of the intersections of tangent lines, the reasoning in the proof above carries through to prove the collinearity of the corresponding points.

The author would like to thank the referees for many helpful comments and further historical information. The author also thanks Kate Turner for asking for a proof of the three-circle theorem.

References

[1] H. S. M. Coxeter, The problem of Apollonius, Amer. Math. Monthly, 75 (1968) 5–15.
[2] J. S. MacKay, Herbert Spencer and mathematics, Proc. Edinburgh Math. Soc., 25 (1907) 95–106.
[3] C. Stanley Ogilvy, Excursions in Geometry, Oxford, New York, 1969.
[4] J.-V. Poncelet, Applications d'Analyse et de Géometrie, Mallet-Bachelier, Paris, 1862.
[5] Herbert Spencer, Autobiography, vol. 1, New York, 1904.

Mathematician

In midair somewhere
he lays an axiomatic floor.
On it he sets a hypothetical plank
on which he raises a logical ladder
which he proceeds to climb.
    There is risk, suspense and drama:
    any loose rung, any misstep fatal.
At the proper confluence of space and time
he steps off onto a higher platform
with a broader panorama.

The whole thing is fabrication.
But so was Creation.

—KATHARINE O'BRIEN

A version of this poem was read at a meeting of the Poetry Society of America in New York City in February 1981; the program was devoted to poems related to science.

The reader will want to repeat the proof using the triangle obtained from the intersections of tangents interior to the triangle ABC. Using this interior triangle establishes the theorem in cases when the exterior triangle is degenerate. Furthermore, by introducing internal rather than external common tangents, new collinearities can be found, for example by taking the points of intersection of two pairs of common internal tangents and one pair of common external tangents. By a judicious choice of the intersections of tangent lines, the reasoning in the proof above carries through to prove the collinearity of the corresponding points.

The author would like to thank the referees for many helpful comments and further historical information. The author also thanks Kate Turner for asking for a proof of the three-circle theorem.

References

[1] H. S. M. Coxeter, The problem of Apollonius, Amer. Math. Monthly, 75 (1968) 5–15.
[2] J. S. MacKay, Herbert Spencer and mathematics, Proc. Edinburgh Math. Soc., 25 (1907) 95–106.
[3] C. Stanley Ogilvy, Excursions in Geometry, Oxford, New York, 1969.
[4] J.-V. Poncelet, Applications d'Analyse et de Géometrie, Mallet-Bachelier, Paris, 1862.
[5] Herbert Spencer, Autobiography, vol. 1, New York, 1904.

Mathematician

In midair somewhere
he lays an axiomatic floor.
On it he sets a hypothetical plank
on which he raises a logical ladder
which he proceeds to climb.
    There is risk, suspense and drama:
    any loose rung, any misstep fatal.
At the proper confluence of space and time
he steps off onto a higher platform
with a broader panorama.

The whole thing is fabrication.
But so was Creation.

—KATHARINE O'BRIEN

A version of this poem was read at a meeting of the Poetry Society of America in New York City in February 1981; the program was devoted to poems related to science.

PROBLEMS

Leroy F. Meyers, Editor
G. A. Edgar, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before February 1, 1983.

1149. For each of the parts (a), (b), (c), separately, construct a triangle ABC, given in position the three points:

(a) O, M_a, I;
(b) O, H_a, T_a;
(c) M_a, H_a, I;

where O and I are the circumcenter and incenter of triangle ABC and M_a, H_a, and T_a are the points in which the median, the altitude, and the angle bisector to side a meet this side. (See the Note, "Triangle constructions with three located points", this issue, pp. 227–230.) [*William Wernick, The Bronx, New York.*]

1150. To aid those uncertain of their fitness to join Mensa (a social organization for people of high intelligence), Mensa publishes tests, tests not infrequently flawed. A recent question was: "If you left your house and walked one mile east, three miles north, one mile west, and two miles south, how far would you then be from home?" What further conditions are needed to make the given answer, "One mile," correct? [*Marlow Sholander, Case Western Reserve University.*]

1151. Three points P, Q, R move on curves in the plane. At each instant, the normal at P to the curve on which P is moving coincides with the bisector of the angle RPQ. Corresponding conditions hold for the points Q and R.

(a) Show that the perimeter of the triangle PQR is constant.

(b) Find examples other than that of an equilateral triangle whose vertices move around a fixed circle. (If, say, P is fixed, then find examples where the bisector of the angle RPQ is a fixed line.)

(c)* Does the result in (a) have a dynamic interpretation in terms of three heavy masses moving on a smooth horizontal table with a light inextensible string looped over them? [*Peter J. Giblin, University of North Carolina at Chapel Hill and University of Liverpool.*]

ASSISTANT EDITORS: Daniel B. Shapiro *and* 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 number 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 acknowledgement of their communications should include a self-addressed stamped card. Send all communications to this department to Leroy F. Meyers, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.*

1152. An elevator starts on the top floor of a 100-floor building and in its descent to the bottom (first) floor stops at at least 40 floors, counting both the top and bottom floors as stops. Show that somewhere in its travel the elevator had to stop at two floors that were exactly 9, 10, or 19 floors apart. [*Michael Gilpin and Robert Shelton, Michigan Technological University.*]

1153. Let ϕ be Euler's totient function. Show that:

(a) if $n>79$, then there are a and b such that $a+b=n$, $a>1$, $b>1$, and $\phi(a)/a+\phi(b)/b<1$;
(b) if $n>4$, then there are a and b such that $a+b=n$, $a>1$, $b>1$, and $\phi(a)/a+\phi(b)/b>1$.
[*Charles R. Wall, Trident Technical College.*]

Quickies

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

Q674. What quantity is measured in units of dollars squared? [*Robert Messer, Albion College.*]

Q675. Prove that if the series $\Sigma_{n=0}^{\infty} a_n$ of nonnegative terms is convergent, then for every partition of $N=\{0,1,\dots\}$ into countably many disjoint finite sets I_n (for $n\in N$) there is an $m\in N$ such that $a_m \geqslant \Sigma_{n\in I_m} a_n$. [*Warren Page, New York City Technical College.*]

Q676. Prove that $e+\ln 4>4$. [*A. Wilansky, Lehigh University.*]

Solutions

Fred, not Branch **March 1981**

1121. After getting three hits in four times at bat, a baseball player's average changes from .233 to .252. Can one determine how many times the player has been at bat during the season? [*V. Frederick Rickey, Bowling Green State University.*]

Solution: Let x and h be the number of times at bat and the number of hits, respectively, when the player is batting .233. Then the two given batting averages yield the inequalities

$$.2325 \leqslant \frac{h}{x} < .2335 \tag{1}$$

and

$$.2515 \leqslant \frac{h+3}{x+4} < .2525. \tag{2}$$

Rewriting (1) and (2) gives

$$\frac{h}{.2335} < x \leqslant \frac{h}{.2325} \tag{3}$$

and

$$.2515(x+4) \leqslant h+3 < .2525(x+4). \tag{4}$$

Applying (3) to (4) and using transitivity yields

$$23.13 < h < 25.87.$$

Since h must be an integer, we have $h = 24$ or $h = 25$. But (3) with $h = 25$ gives $107.06 < x < 107.53$, and there would be no integral solution for x. However, (3) with $h = 24$ gives (rounded) $102.78 < x < 103.23$, whence $x = 103$. So the player is 24 for 103 while batting .233, and 27 for 107 while batting .252.

LEE O. HAGGLUND
Wofford College

Also solved using inequalities directly by James Bowe, Chico Problem Group, Wayne M. Delia & Michael Fox, David Del Sesto, Thomas P. Dence, Clayton W. Dodge, Philip M. Dunson, Michael W. Ecker, Milton P. Eisner, Thomas E. Elsner, Gordon Fisher, Nick Franceschine III, Dean Furbish, Anton Glaser, Clifford H. Gordon, Dale E. Grussing, Steven Gustafson, Joel K. Haack, C. V. Heuer, G. A. Heuer, J. R. Hilditch (England), Steven Izen, John Kahila, Kent Mac Dougall, Peter M. Makus, Robert A. Meyer, Roger B. Nelsen, John C. Nichols, Charles F. Pinzka, Daniel A. Rawsthorne, Steve Ricci, Marlow Sholander, Harris Shultz, Nat Smith, Robert S. Stacy, Daniel F. Symanczyk, Jim Tattersall, Michael Vowe (Switzerland), Eugene Wermer, G. Willikers, Ken Yocom, Harry Zaremba, Gene Zirkel, and the proposer.

Also solved mentioning continued fractions by Walter Bluger (Canada), M. S. Krishnamoorthy, Michael I. Ratliff, and the proposer.

The proposer found the problem mentioned in a review by Peter Hilton and Jean Pedersen in *The American Mathematical Monthly*, 87 (1980) 144, of Sheila Tobias's Overcoming Math Anxiety, esp. p. 91. A related problem is found in Donald E. Knuth, The Art of Programming, vol. 2, 2nd ed., p. 363, §4.5.3, problem 39; solution, p. 606.

Clustering Means — May 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.*]

Solution: We use the notation x for $(x_1, x_2, \ldots, x_n)$ and $\|x\|$ for $(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$. Direct computation using the inequality $2ab \leqslant a^2 + b^2$ shows that f is a norm-decreasing function (i.e., $\|f(x)\| \leqslant \|x\|$ for all x) and that $\|f(x)\| = \|x\|$ if and only if $x_2 = x_3$, $x_4 = x_5, \ldots,$ and $x_n = x_1$. Thus the sequence $(f^k(x))_{k=1}^{\infty}$ in R^n is bounded and so has a cluster point a. Since f is continuous, $f^k(a)$ will also be a cluster point for each k. Since the norm of any cluster point is the limit of the nonincreasing sequence $(\|f^k(x)\|)_{k=1}^{\infty}$, we see that

$$\|a\| = \|f(a)\| = \|f^2(a)\|.$$

Hence $a_2 = a_3$, $a_4 = a_5, \ldots, a_n = a_1$, and

$$a_2 + a_3 = a_4 + a_5 = \cdots = a_n + a_1,$$

so that $a_1 = a_2 = \cdots = a_n$. Since the sum of the coordinates of $f(x)$ is the same as that of x, we see that there is only one cluster point, namely $(c, c, \ldots, c)$, where c is the mean of the coordinates of x. This cluster point must be the limit of the sequence $(\|f^k(x)\|)_{k=1}^{\infty}$, since the unique cluster point of a bounded sequence must be its limit.

Note that $x_1, \ldots, x_n$ are arbitrary real numbers, not necessarily positive.

C. RAY ROSENTRATER
Westmont College

Also solved by Robert E. Bernstein, Chico Problem Group, Victor Hernandez (Spain), L. Kuipers (Switzerland), Roger B. Nelsen, Morris Newman & Marvin Marcus, Edward Schmeichel, Paul J. Zwier, and the proposer. There were two incorrect solutions.

Most of the solvers observed that the matrix of f (considered as a linear transformation on R^n) is a (doubly stochastic) transition matrix in the theory of Markov chains. Other solvers found explicit bounds on the coordinates of $f^k(x_1,\dots,x_n)$.

Powerful Convergence **May 1981**

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.*]

Solution: The result is true if p is 1 or even, and false if p is odd and greater than 1.

Dirichlet's test (see R. Johnsonbaugh and W. E. Pfaffenberger, *Foundations of Mathematical Analysis*, Marcel Dekker, 1981, p. 90) states that if $\Sigma_1^n a_k$ is bounded and $\{b_n\}$ decreases to zero, then $\Sigma a_n b_n$ converges. Taking $b_n = 1/n$, we see that the result is true for $p = 1$.

Hölder's inequality (see E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer, 1965, p. 190) states that

$$\sum_1^n |a_k b_k| \leq \left(\sum_1^n |a_k|^p\right)^{1/p} \left(\sum_1^n |b_k|^q\right)^{1/q},$$

where $p > 1$ and $q = p/(p-1)$. If p is even and $b_k = 1/k$, it follows that the result is true.

Now suppose p is odd and greater than 1. Let Σa_n be the series

$$0 + \frac{1}{(\log 2)^{1/p}} + \frac{1}{(\log 3)^{1/p}} - \frac{2^{1/p}}{(\log 4)^{1/p}} + \frac{1}{(\log 5)^{1/p}} + \frac{1}{(\log 6)^{1/p}} - \frac{2^{1/p}}{(\log 7)^{1/p}} + \cdots.$$

By Dirichlet's test, Σa_n^p converges. If $\Sigma a_n/n$ converges, we may subtract from it the convergent series

$$0 + \frac{1}{2(\log 2)^{1/p}} + \frac{1}{3(\log 3)^{1/p}} - \frac{2}{4(\log 4)^{1/p}} + \frac{1}{5(\log 5)^{1/p}} + \frac{1}{6(\log 6)^{1/p}} - \frac{2}{7(\log 7)^{1/p}} + \cdots$$

to obtain the divergent series

$$\frac{2-2^{1/p}}{4(\log 4)^{1/p}} + \frac{2-2^{1/p}}{7(\log 7)^{1/p}} + \cdots$$

which is a contradiction. Thus, the result is false if p is odd.

RICHARD JOHNSONBAUGH
Chicago State University

Also solved by R. P. Boas, M. S. Klamkin (Canada), David S. Liang, and the proposer. Partial solutions by David A. Rawsthorne, Edward Schmeichel, and David H. Vetterlein. There were two incorrect solutions.

Integrating a Square Tangent Product **May 1981**

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.*]

Solution: The basic idea used in the solution is that for all α, β,

$$\tan(\alpha-\beta)\tan\alpha\tan\beta = \tan\alpha - \tan\beta - \tan(\alpha-\beta). \tag{1}$$

Let $d_1 = 1$, and for $n > 1$, let

$$d_n = \prod_{1 \leqslant i < j \leqslant n} \tan(a_j - a_i).$$

Also, let

$$P_n(x) = \prod_{i=1}^{n} \tan(x - a_i).$$

LEMMA. *For each $n \geqslant 1$ there exist constants b_n and C_{in} (for $i = 1, 2, \ldots, n$) such that*

$$d_n P_n(x) = b_n + \sum_{i=1}^{n} C_{in} \tan(x - a_i). \tag{2}$$

Proof. We proceed by induction. Take $b_1 = 0$ and $C_{11} = 1$ for the case $n = 1$. Assume that b_n and $C_{1n}, C_{2n}, \ldots, C_{nn}$ have been defined so that (2) holds. Let

$$q_{n+1} = \prod_{i=1}^{n} \tan(a_{n+1} - a_i).$$

Then

$$\begin{aligned} d_{n+1} P_{n+1}(x) &= d_n q_{n+1} P_n(x) \tan(x - a_{n+1}) \\ &= \left(b_n + \sum_{i=1}^{n} C_{in} \tan(x - a_i) \right) q_{n+1} \tan(x - a_{n+1}) \\ &= b_n q_{n+1} \tan(x - a_{n+1}) \\ &\quad + \sum_{i=1}^{n} \frac{C_{in} q_{n+1}}{\tan(a_{n+1} - a_i)} \tan(a_{n+1} - a_i) \tan(x - a_i) \tan(x - a_{n+1}) \\ &= b_n q_{n+1} \tan(x - a_{n+1}) + \sum_{i=1}^{n} \frac{C_{in} q_{n+1}}{\tan(a_{n+1} - a_i)} \tan(x - a_i) \\ &\quad - \sum_{i=1}^{n} \frac{C_{in} q_{n+1}}{\tan(a_{n+1} - a_i)} \tan(x - a_{n+1}) - \sum_{i=1}^{n} C_{in} q_{n+1}. \end{aligned}$$

Hence with the definitions

$$b_{n+1} = -\sum_{i=1}^{n} C_{in} q_{n+1},$$

$$C_{i,n+1} = \frac{C_{in} q_{n+1}}{\tan(a_{n+1} - a_i)} \quad \text{for } i = 1, 2, \ldots, n,$$

$$C_{n+1,n+1} = b_n q_{n+1} - \sum_{i=1}^{n} \frac{C_{in} q_{n+1}}{\tan(a_{n+1} - a_i)},$$

the induction is complete.

A straightforward calculation using (2) and (1) shows that

$$\begin{aligned} d_n^2 P_n^2(x) &= b_n^2 + 2 b_n \sum_{i=1}^{n} C_{in} \tan(x - a_i) + \sum_{i=1}^{n} C_{in}^2 \tan^2(x - a_i) \\ &\quad + 2 \sum_{1 \leqslant i < j \leqslant n} \frac{C_{in} C_{jn}}{\tan(a_j - a_i)} \big(\tan(x - a_i) - \tan(x - a_j) - \tan(a_j - a_i) \big), \end{aligned}$$

so that

$$d_n^2 \int P_n^2(x)\,dx = b_n^2 x - \left(\sum_{i=1}^{n} C_{in}\right)^2 x + \sum_{i=1}^{n} C_{in}^2 \tan(x-a_i) - 2b_n \sum_{i=1}^{n} C_{in} \ln|\cos(x-a_i)|$$

$$+ \sum_{1 \leq i<j \leq n} \frac{2C_{in}C_{jn}}{\tan(a_j-a_i)} \ln\left|\frac{\cos(x-a_j)}{\cos(x-a_i)}\right| + C. \qquad (3)$$

We can use (3) to calculate $\int P_n^2(x)\,dx$ if no two of the a_i for $i=1,2,\dots,n$ differ by an integral multiple of π. If this is not so, we can take limits (using l'Hôpital's rule) after dividing by d_n^2.

PAUL J. ZWIER
Calvin College

Also solved by the proposer. There was one incorrect solution.

Meaner Intermediate Value Properties **May 1981**

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 $k_1, k_2, \dots, k_n$, show that there exist distinct $x_1, x_2, \dots, x_n$ such that

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

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

Solution: Set $a_0=0$, $a_n=1$, and $\sum_{i=1}^{n} k_i = M$. Since f is differentiable on $[0,1]$, it is also continuous there, and by the Intermediate Value Theorem there is for each i, $1 \leq i \leq n$, an a_i such that $0 \leq a_{i-1} < a_i \leq 1$ with

$$f(a_i) = \sum_{j=1}^{i} k_j/M, \qquad i \geq 1.$$

By the Mean Value Theorem there is an x_i, $a_{i-1} < x_i < a_i$, such that

$$f'(x_i) = \frac{f(a_i) - f(a_{i-1})}{a_i - a_{i-1}} = \frac{k_i/M}{a_i - a_{i-1}}.$$

Therefore,

$$\sum_{i=1}^{n} \frac{k_i}{f'(x_i)} = M \sum_{i=1}^{n} (a_i - a_{i-1}) = M(a_n - a_0) = M.$$

(This generalizes Solution I of problem 1053, this MAGAZINE, 53 (1980) 51.)

MICHAEL VOWE
Therwil, Switzerland

Also solved by Chico Problem Group, Michael W. Ecker, Robert Heller, Johnny Henderson, Victor Hernandez (Spain), M. S. Klamkin (Canada), Mou-Liang Kung, Steve Ricci, Paul J. Zwier, and the proposer.

Ricci proved the following generalization. Let F be a function defined on a subset of $[a,b]$ such that either (a) $F(x)=1$ has infinitely many solutions x in $[a,b]$ or (b) there exists an $\varepsilon>0$ such that for each y in $[1-\varepsilon, 1+\varepsilon]$ there is an x in $[a,b]$ with $F(x)=y$. Let k_i be positive constants for $1 \leq i \leq n$. Then there are distinct x_i in $[a,b]$ such that $\sum_{i=1}^{n} k_i F(x_i) = \sum_{i=1}^{n} k_i$. (If $n>1$, then there are infinitely many choices for the x_i.) In particular, $F(x)$ may be $(f'(x))^{\alpha}$, where α is any real number and f is any differentiable function such that $f(b)-f(a)=b-a>0$.

Axiomatic Characterization of Distance — May 1981

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.*]

Solution: It is shown that for each nonnegative integer n and each vertex v, $d(v_0, v)=n$ if and only if $f(v)=n$. This is clear for $n=0$ since that is condition (3); assume it is true for all $k \leqslant n$.

If $d(v, v_0)=n+1$, there exists v_1 so that $d(v_0, v_1)=n$ and $d(v_1, v)=1$. By the induction hypothesis $f(v) \geqslant n+1$ and by condition (1), $f(v) \leqslant n+1$, so $f(v)=n+1$.

If $f(v)=n+1$, by condition (2) there exists a vertex v_1 so that $f(v_1)<f(v)$ and $d(v, v_1)=1$. Then $f(v_1) \leqslant n$ forces $f(v_1)=d(v_0, v_1)$. By the induction hypothesis, $d(v_0, v) \geqslant n+1$. However, $d(v_0, v) \leqslant d(v_0, v_1)+d(v_1, v) \leqslant n+1$; thus, $d(v_0, v)=n+1$.

Note: This proves that G must be connected, so that the hypothesis is redundant. [The hypothesis of connectedness was mistakenly inserted by the editors.—*Ed. note.*]

S. F. Barger
Youngstown State University

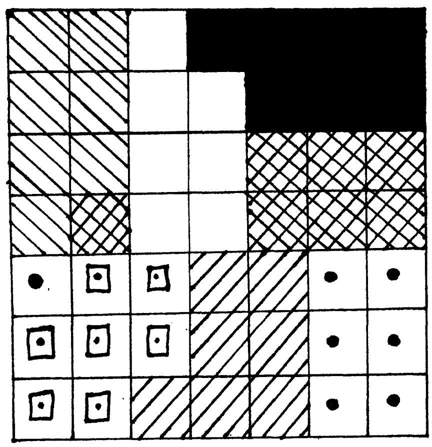
Also solved by Robert Cabane (France), Chico Problem Group, Miklós Csúri (Hungary), Michael W. Ecker, Victor Hernandez (Spain), Steve Kahn, Graham Lord (Canada), John Mitchem & Edward Schmeichel, Klára Pintér (Hungary), C. Ray Rosentrater, Gerald Thompson, Gregory Wene, and the proposer.

Colorful Squares — September 1981

1127. (a) Color a 7×7 board with seven colors so that there are exactly seven squares of each color, with no more than three colors in each row and column.

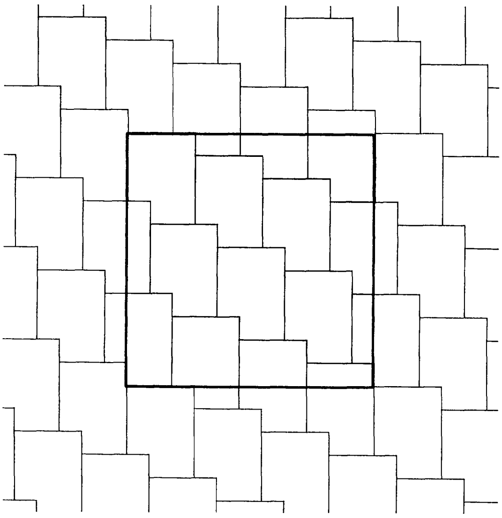
(b)* Color an $n \times n$ board with n colors so that there are exactly n squares of each color, with fewer than $\sqrt{n}+1$ colors in each row and column. [*J. L. Selfridge, Mathematical Reviews.*]

Solution: (a)

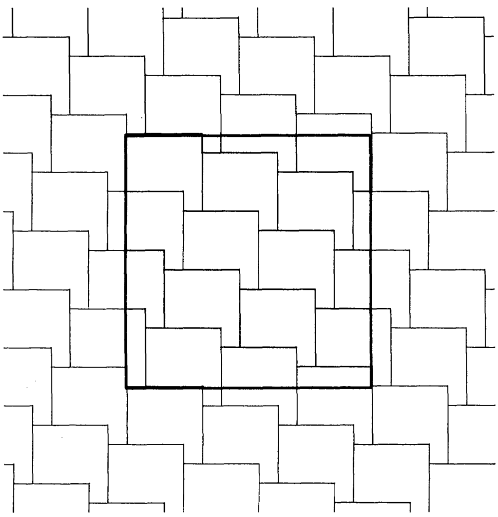


Andrea Paffrath, age 16
Mädchengymnasium Essen-Borbeck
Essen, West Germany

Solution: (b) We shall construct a solution for every size n. Select k so that $(k-1)^2 < n \leq k^2$, and set w to be the least integer not less than n/k, i.e., $w = k-1$ or k. Observe that $n = (k-1)w + r$ with $1 \leq r \leq w$. Label the positions on the board (a, b) with $0 \leq a, b < n$ and entries interpreted modulo n. For each i, $0 \leq i \leq n-1$, we place the n cells of color i in a block, B_i, formed from a $(k-1) \times w$ rectangle with an additional partial row of r cells placed below the lower right end of the rectangle. The upper left-hand cell of the block is positioned at (i, wi).



$n = 11$



$n = 13$

We shall demonstrate that each cell has been assigned to precisely one block. First, we examine row k. For $1 \leq i \leq k-1$, block B_i contains the cells (k, j) provided $wi \leq j < w(i+1)$, while the remaining cells (k, j) have $w - r \leq j < w$ and lie in B_0. Thus, row k is covered. Now the bijection $\theta(a, b) = (a+1, b+w)$ permutes the cells of the board, mapping block B_i onto B_{i+1} and sending row k to row $k+1$. Thus, successive applications of θ assure us that every row has been covered.

Since each block can cover only n cells, it follows that each cell is covered precisely once. Since k colors appear in row k, there must be k colors appearing in each row. Similarly, every block appearing in a specified column must cover $k-1$ or k entries in that column. We find that precisely r blocks cover k entries and $w-r$ blocks cover $k-1$ entries, accounting for all $rk+(w-r)(k-1)=n$ entries. Thus $w=k-1$ or k colors appear in each column. Since $k<\sqrt{n}+1$, we have found a solution for every n.

Notice that when n happens to fall in the range $(k-1)^2<n\leqslant k(k-1)$ we have actually done a touch better since we escape with $w=k-1$ colors appearing in each column. However, we can never achieve $k-1$ colors simultaneously for rows and columns. Suppose we had done so. Then the average number of occurrences of each color actually appearing in a row or column is at least $n/(k-1)>\sqrt{n}$. Select a color i which attains at least this average value for the r rows and c columns containing it. Since each of the n cells of color i appears in a row and in a column, we have

$$\frac{2n}{r+c}\geqslant\frac{n}{k-1}>\sqrt{n}.$$

But the n cells of color i must appear among the rc entries in the prescribed rows and columns, so $n\leqslant rc$. It follows that

$$\frac{2n}{r+c}\leqslant\frac{2n}{r+n/r}=\frac{2\sqrt{n}}{r/\sqrt{n}+\sqrt{n}/r}\leqslant\sqrt{n}.$$

This contradiction implies that the number of colors appearing in each row and column cannot be reduced further.

ALLEN J. SCHWENK
University of Waterloo, Canada

Part (a) also solved by Anders Bager (Denmark), Samuel Chort†, Milton P. Eisner, Krishnamoorthy, Ronald A. Luther*, Roger B. Nelsen*, John P. Robertson, Steve Schonberger*, J. Suck* (West Germany), Zalman Usiskin, the following students at the Mädchengymnasium, Essen-Borbeck, West Germany: Simone Buschmann, Sandra Dombrowski, Heike Koller, Beate Küpper, Andrea Paffrath (second solution), Nicole Pukies & Uschi Roberts, Heike Schmitz, and Kerstin Winkler, as well as by the proposer. An asterisk and a dagger indicate a partial and a complete solution of part (b), respectively.*

Answers

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

Q674. The variance of a monetary quantity (i.e., the variance in income of workers in the United States).

Q675. Otherwise there would be a partition $\{I_n\colon n\in N\}$ of N such that $a_m<\sum_{n\in I_m}a_n$ for every $m\in N$, in which case

$$\sum_{m=0}^{\infty}a_m<\sum_{m=0}^{\infty}\left(\sum_{n\in I_m}a_n\right)=\sum_{n=0}^{\infty}a_n,$$

which is impossible if $\sum_{n=0}^{\infty}a_n<\infty$.

Q676. $0<\int_{\ln 2}^{1}(e^x-2)\,dx=e-4+\ln 4$. (This uses only the monotonicity of the exponential function.) A similar proof yields $e+\ln 27>6$.

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, Gina, *New code is broken*, Science 216 (28 May 1982) 971-972.

One of the first public key cryptosystems, devised in 1976, has been broken by Ali Shamir (Weizmann Institute), an inventory of another scheme. The broken scheme, a trapdoor code based on the knapsack problem, involves scrambling a superincreasing sequence ($a_{n+1} > a_1 + \ldots + a_n$). The breakthrough involved converting the decoding problem to an integer programming problem and solving the latter efficiently (Lenstra's new algorithm, or alternatively continued fractions). Other trapdoor codes not based on the knapsack problem appear safe--for the time being.

Kolata, Gina, *Perfect shuffles and their relation to math*, Science 216 (30 April 1982) 505-506.

A perfect riffle shuffle alternately interlaces cards from two piles. Persi Diaconis (Stanford), a statistician who is also a magician, asked: Which rearrangements of the deck can occur after repeated perfect shuffles? The solution required help from several quarters, including the programming efforts of an undergraduate; the answer involves M_{12}, one of the Mathieu simple groups. The same mathematics of shuffling is also used by engineers to interconnect computers in networks for parallel processing.

Bernstein, Jeremy, *Profiles: A. I. (Marvin Minsky)*, New Yorker (14 December 1981) 50-126.

Long interview with Marvin Minsky, tracing his biography and the development of the field of artificial intelligence.

Richards, Ian, *The invisible prime factor*, American Scientist 70 (1982) 176-179.

How--and even more important, why--is it possible to know that a number is not prime without knowing any of its divisors? Using the proof of Euclid's lemma (p prime never divides ab if $a, b < p$) as an example, the author shows that by contradiction one can prove a "breakdown" occurs in a chain of reasoning that may be so long that it is quite hopeless to pinpoint where the breakdown occurred. (There is no mention of intuitionists or constructivists, who find this notion absurd.) The article then considers Fermat's little theorem, its use in showing $2^{257} - 1$ is not prime, and its application to public key codes.

Peterson, Ivars, *Quickening the pursuit of primes*, Science News 121 (6 March 1982) 158.

Notes development of a new algorithm for testing primality that permits efficient testing of 100-digit numbers. Like other relatively fast algorithms for this purpose, this one begins from Fermat's little theorem ($a^{p-1} \equiv 1 \bmod p$ for p prime). Succinct summary of the state of the art: "factoring is hard, and primality testing is easy."

Rothman, Tony, *The short life of Evariste Galois*, Scientific American 246:4 (April 1982) 136-149, 162.

A more popular version of the author's reexamination of the myths about Galois (*Amer. Math. Monthly*, 8(1982) 84-106), including exposition on the Galois theory of equations.

Hofstadter, Douglas R., *Metamagical Themas: Number numbness, or why innumeracy may be just as dangerous as illiteracy*, Scientific American 246:5 (May 1982) 20-34, 178.

"The fact is that when numbers get too large, people's imaginations balk....an inability to relate to large numbers is clearly bad for society. It leads people to ignore big issues on the grounds that they are incomprehensible." Hofstadter goes on to illustrate how to estimate large numbers and how to think logarithmically--essential arithmetic skills we would do well to substitute in school for pencil-and-paper computation practice in this calculator age. (A member of the White House science staff underscores Hofstadter's point with an anecdote of a 6 degree-of-magnitude error (247:1 (July 1982) 8).)

Parsons, J.E., *Social forces shape math attitudes and performance*, Association for Women in Mathematics Newsletter, 12:3 (May-June 1982) 4-10.

What factors most strongly affect the course-taking plans and achievement of junior-high-school students? Conclusions from a two-year longitudinal study of 250 students, their parents and teachers: not math aptitude, but social factors, particularly mothers' beliefs regarding how difficult math is for their child. The data suggest that "exposure to mothers' sex stereotyped beliefs" is a viable explanation for observed sex differences in math achievement.

Solow, Daniel, How to Read and Do Proofs: An Introduction to Mathematical Thought Processes, Wiley, 1982; xiv + 172 pp, (P).

This book may become the fundamental guide to proof mathematics that Pólya's *How to Solve It* is for problem solving. Solow emphasizes a systematic "forward-backward" approach to building a proof, then devotes a chapter each to specific techniques, corresponding to the logical form of the proposition (the construction method, the choose method, induction, specialization), followed by chapters on the contradiction method, the contrapositive method, and how to negate a statement. Solutions to exercises are included (a student version without solutions might be desirable). *Every* serious mathematics student at calculus level or above can benefit from this book. Examples are chosen from geometry, elementary number theory, college algebra, and calculus.

Brewer, James W., and Smith, Martha K. (eds.), Emmy Noether: A Tribute to Her Life and Work, Dekker, 1981; x + 180 pp, $19.75.

Progressing from biography to mathematics, this book contains notable essays on Noether and her influence, mathematics at Göttingen, personal recollections, and individual articles on Noether's contributions to Galois theory, calculus of variations, commutative ring theory, representation theory, and algebraic number theory. There are also photographs, reprints of obituaries, a list of Noether's publications, and the text of her address to the 1932 International Congress.

Berlekamp, Elwyn, Conway, John H., and Guy, Richard K., Winning Ways for Your Mathematical Plays, Vol. 1: Games in General; Vol. 2: Games in Particular; Academic Pr, 1982, xxxii + 850 + xix pp, $64.50, $22.50 (P).

Dedicated to Martin Gardner, this long-awaited book is undoubtedly the preeminent recreational mathematics work of the decade. From Hackenbush, through Nim and Sprague-Grundy theory, to Kayles, Toads and Frogs, Cutcake, Hotcakes, and hundreds of other games and variations, the authors systematically treat in Volume 1 theory and examples of two-person games of complete information and strategy (except for chess, checkers, and Go). Volume 2 is devoted to case studies of games to which their theories do (and don't) apply, grouped according to how the games are played: with coins, with pencil and paper, or on a board. Solitaire games are also included (Soma cube, Chinese rings, Fifteen puzzle, Rubik's cube, peg solitaire), and so are even no-player games (Life). This is a marvelous compendium, with lots of open questions, that takes the reader to the frontiers in combinatorial game theory. It is splendidly printed, with attractive figures in red, blue and green. Altogether, it's a book to savor and treasure.

Balinski, Michel L., and Young, Peyton, Fair Representation: Meeting the Ideal of One Man, One Vote, Yale U Pr, 1982; xi + 191 pp, $27.50.

Full treatment of the various methods of apportionment of representatives, with prose in the body of the book and mathematics (62 pp) in the first appendix. The authors "show how the choice among methods can be reduced to a choice among principles." They recommend for the U.S. the Webster divisor system as the only method which avoids the paradoxes, is unbiased, and in practice stays within quota. Since the Webster system was abandoned by the U.S. in 1941, one must wonder: will this book provoke a return to it?

Gould, Stephen Jay, The Mismeasure of Man, Norton, 1981; 352 pp, $14.95.

Brilliant and masterful indictment of 19th-century craniometry, biological identification of criminals, and the current 20th-century hereditarian theory of IQ, with its error of reifying "intelligence". Gould offers a readable nontechnical explanation of the plethora of statistical fallacies indulged by researchers in these areas, together with a disillusioning look at self-serving failures in scientific objectivity.

Gardner, Martin, Science Fiction Puzzle Tales, Potter, 1981; xii + 148 pp, $4.95 (P).

Collection of 36 puzzle tales, each a page or so in length, which originally appeared in *Isaac Asimov's Science Fiction Magazine*. Some puzzles are mathematical, others involve wordplay, a few even deal with a little physics. All are typical Gardner, and readers of his *Scientific American* column will welcome this new trove. A unique feature is three levels of answers to the puzzles!

Niven, Ivan, Maxima and Minima Without Calculus, Dolciani Mathematical Expositions No. 6, MAA, 1981; xv + 303 pp, $24.50.

Are alternatives to the calculus max/min methods just "trick procedures of limited usefulness"? Ivan Niven answers in the negative by offering unifications of those procedures into general methods that can solve even problems impossible by elementary calculus. Niven prefers to solve geometric problems by reformulating them algebraically, and he deliberately omits consideration of linear programming as well as calculus. (The book is written for an audience "at or near the maturity level of second- and third-year students in North American universities and colleges", but both level and material make it attractive to bright high-school juniors and seniors.)

Ascher, Marcia, and Ascher, Robert, Code of the Quipu: A Study in Media, Mathematics, and Culture, U Michigan Pr, 1981; vii + 166 pp, $18.95, $8.95 (P).

The Inca accounting system was done on knotted colored cords called quipus. The authors, a mathematician and an anthropologist who are the world experts in quipus, have written a book which acquaints the general reader with the cultural context, interpretation problems and mathematical ideas of the quipu.

Winfree, Arthur T., The Geometry of Biological Time, Springer-Verlag, 1980; xiii + 530 pp.

An innovative and extraordinary book! The topic is "temporal morphology", and treated are all manner of biological and biochemical processes that repeat themselves regularly. "Phase singularities" are a central feature of the phenomena, which include circadian rhythms, excitable kinetics, fireflies, cell metabolism, aggregation of slime mold amoebae, and pattern formation in fungi. True to the title, the mathematics used is mostly geometric, especially in the second half, but all manner of continuum mathematics, calculus and differential equations, plus some differential geometry, is employed to some degree. The topics are exciting, the account is at the frontiers of research, and the graphic layout is outstanding. A fine book, rewarding both in conceptual and modeling aspects.

Lumsden, Charles J., and Wilson, Edward O., Genes, Mind and Culture: The Coevolutionary Process, Harvard U Pr, 1981; xii + 428 pp, $20.

"This book contains the first attempt to trace development all the way from genes through the mind to culture." The considerable mathematics used (vector calculus, linear algebra, probability and statistics, continuous and discrete modeling) will tax most of the intended audience of biologists and social scientists. Whether the arguments here can serve as a quantitative basis for Wilson's theories of sociobiology will require detailed examination by mathematicians as well as biologists.

Griffiths, H.B., Surfaces, 2nd ed., Cambridge U Pr, 1981; xii + 128 pp, $29.95, $12.95 (P).

Classification of two-dimensional surfaces, invariance of the Euler characteristic, and Morse theory; all presented in an informal fashion, with notation as simple as possible and great regard for the pedagogy of the subject. Intended for teachers, the book would be a welcome spatial component of an undergraduate--or even high school!--geometry (or topology) course, since it presumes only comfort with simple algebra and the ability (and curiosity!) to follow an argument.

Eves, Howard, Great Moments in Mathematics (After 1650), Dolciani Mathematical Expositions No. 7, MAA, 1981; xii + 263 pp, $23.50.

Twenty more lectures on outstanding events in mathematics, from the birth of probability to the resolution of the four-color conjecture, together with a list of 20 "regrets" that could not be included. Like the several encore volumes to the author's *In Mathematical Circles*, can we perhaps hope yet to see the "regrets" come to print?

Page, Warren (ed.), Two-Year College Mathematics Readings, MAA, 1981; vii + 304 pp, $19.50 (P).

Collection of 45 articles, all new and *not* previously published in the *Two-Year College Mathematics Journal*. Selections are divided evenly among headings of Algebra, Geometry, Number Theory, Calculus, Probability & Statistics, Calculators & Computers, Mathematics Education, and Recreational Mathematics.

NEWS & LETTERS

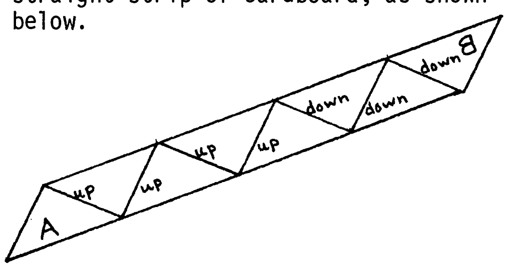
DE BRUIJN SEQUENCES

A very nice chapter entitled "Memory Wheels" in Sherman Stein's book Mathematics, The Man-made Universe provides some interesting history and application of the sequences discussed by Ralston in "DeBruijn Sequences..." (this *Magazine*, May 1982, 131-143).

Joe Malkevitch
Associate Editor

SOLUTION TO PROBLEM 1110

In the solution to Problem No. 1110 (this *Magazine*, January 1982, 47-48) the case of $n = 7$ can be illustrated by a model of seven hinged isosceles triangles, which is easily made from a straight strip of cardboard, as shown below.



Here eight triangles are used, and the first, A, is pasted to the last, B. The hinge edges are scored and folded to make movable hinges. The short edges of the triangular plates represent the seven pipes of equal length.

Michael Goldberg
5823 Potomac Ave., N.W.
Washington, DC 20016

MATHEMATICS IN LITERATURE

The recent paper of D.O. Koehler ("Mathematics and Literature", this *Magazine*, March 1982) prompts me to call attention to some entertaining occurrences of mathematics and mathematicians in literature.

Sometimes mathematics aids in literary humor.

"The king reigning over that part of the country now called Lincolnshire and Cambridge in A.D. 609 was called Eghedde the Bald. The people of his kingdom were middle angles--that is, they were neither right nor straight--and were known as the Treacherous Forty-Fives." [3]

Mathematics sometimes fares well in books.

"He was an arithmetician rather than a mathematician. None of the humour, the music or the mysticism of higher mathematics ever entered his head." [4]

Other writers think less well of mathematics.

"I am a mathematician, sir. I never permit myself to think." [1]

The mystery story writer John Dickson Carr (just cited) confessed,

"Mathematics was the business that ruined me--mathematics, that last refuge of the half-wit. I have no word low enough for this miserable, stupid, utterly inexcusable insult to the intelligence called mathematics, which so neatly excludes all possible attempts to use the brain. (This peroration is partly inspired by the fact that I could never figure out how long it took those delightful asses A, B and C to fill the cistern with wall-paper and didn't much care anyway.)" [2]

Ernest Hemingway used mathematics metaphorically in an interview [5] (with Harvey Breit of the *N.Y. Times*),

"In writing I have moved through arithmetic, through plane geometry and algebra, and now I am in Calculus. If they (the critics) don't understand that, to hell with them."

My own collection of references to mathematics started as a youth upon reading Robert Browning's "A Grammarian's Funeral" where I was delighted to encounter the line "Calculus racked him." That this reference is to stones (such as gall stones) did not lessen my comfort from this line as I was learning limits.

At about the same time I discovered Edna St. Vincent Millay's sonnet "Euclid Alone Has Looked On Beauty Bare." The poetic excess of this sonnet cancels out the attitude of John Dickson Carr.

Since 1917 the *Mathematical Gazette* has often published brief quotations from literary works or their authors.

[1] John Dickson Carr, The Hollow Man (1935).
[2] ———, in Books of the Month (1929).
[3] J.N. Chance, Murder in Oils (1945).
[4] John Steinbeck, The Moon Is Down (1942).
[5] *Time*, September 25, 1950.

W. R. Utz
University of Missouri
Columbia, MO 65211

(*Editor's note:* We could not publish Utz's lengthy list of references to the *Gazette*; interested readers should contact him.)

THE BOOKWORM

I would like to add to the description of your most attractive cover on the March 1982 issue of *Mathematics Magazine*. The print is obviously an adaptation of the painting "Der Bücherwurm" by Carl Spitzweg, 1808-1855, a German painter of genre scenes.

Alan Wayne
Pasco-Hernando Community College
New Port Richey, FL 33552

(*Ed. note:* Leary's Old Bookstore, Philadelphia, had a stained glass window of "The Bookworm" as well as the print we used, which was their logo.)

OOPS ON PUTNAM A-3

Several readers responded to an error in the solution to Putnam Problem A-3 as presented in the May issue: the final series should be an alternating series rather than a series of positive terms.

One way to see that $G'(t)/e^t$ diverges is to note that for sufficiently large t,

$$\frac{G'(t)}{2e^t} = \int_0^t \frac{e^{x-t} - 1}{x-t}\,dx =$$

$$\int_0^t \frac{1 - e^{-y}}{y}\,dy > \int_1^t \frac{1 - e^{-y}}{y}\,dy$$

$$> (1 - e^{-1}) \log t.$$

Loren C. Larson
St. Olaf College
Northfield, MN 55057

WINNERS OF 1981 PUTNAM

The winning teams of the 42nd annual William Lowell Putnam Competition held on December 5, 1981 are as follows (listed in descending rank):

Washington University, St. Louis
Kevin P. Keating, Edward A. Shpiz, Richard A. Strong.

Princeton University
Gregg N. Patruno, David P. Roberts, Charles H. Walter.

Harvard University
Michael J. Larsen, Laurence E. Penn, Michael Raship.

Stanford University
Richard J. Beigel, Thomas C. Hales, Kenneth O. Olum.

Univ. of Maryland, College Park
Ravi B. Boppana, Andrew E. Gelman, Brian R. Hunt.

The Putnam Fellows, the five highest ranking individuals, who each received a $500 award, are:

David W. Ash, U. of Waterloo;
Scott R. Fluhrer, Case Western Res. U.;
Michael J. Larsen, Harvard University;
Robin A. Pemantle, U. of Ca., Berkeley;
Adam Stephanides, U. of Chicago.

PÓLYA AND ALLENDOERFER AWARDS

Four authors were honored at the business meeting of the Mathematical Association of America on Tuesday, August 24, 1982 at the University of Toronto. The awards of $200 each recognize excellence in expository writing for articles published in *Mathematics Magazine* and the *Two-Year College Mathematics Journal*.

Recipients of the Carl B. Allendoerfer awards, for articles which appeared in 1981 in *Mathematics Magazine* are:

Ian Richards, "Continued Fractions Without Tears", *Math. Magazine*, 54 (1981) 163-171.

Marjorie Senechal, "Which Tetrahedra Fill Space?", *Math. Magazine*, 54 (1981) 227-243.

Recipients of the George Pólya awards, for articles which appeared in 1981 in the *Two-Year College Mathematics Journal* are:

John Mitchem, "On the History and Solution of the Four-Color Map Problem", TYCMJ, 12 (1981) 108-116.

Peter Renz, "Mathematical Proof: What It Is and What It Ought to Be", TYCMJ, 12 (1981) 83-103.

CRYPTOLOGIA UNDERGRADUATE PAPER COMPETITION

The journal *Cryptologia* announces its second annual undergraduate paper competition to encourage the study of cryptology in the undergraduate curriculum. The contest is open to any undergraduate student. Closing date is January 1, 1983, and the prize is $300, with the winning paper to be published in *Cryptologia*. The papers will be judged by the editors of *Cryptologia* and the winner will be announced on April 1, 1983.

Papers are to be original works which have not been previously published. They are to be no more than 20 typewritten pages, double-spaced, and fully referenced on any topic related to cryptology, including technical, historical and literary subjects. Five copies of the paper must be submitted.

For inquiries and submission information, write *Cryptologia*, Editorial Office, Rose-Hulman Institute of Technology, Terre Haute, IN 47803.

POLYHEDRA FANS TAKE NOTE

We hope to hold a conference on polyhedra at Smith College sometime during the 83-84 academic year. Topics would include current research, applications, teaching of 3-dimensional geometry, model-making, and visualizations of polyhedra. An exhibit on the history of polyhedra is also planned. We welcome expressions of interest and suggestions.

Marjorie Senechal
Dept. of Mathematics
Smith College
Northampton, MA 01063

Robert Whorf
Symmetrics, Inc.
P.O. Box 26791
Philadelphia, PA 19117

OHIO DEVELOPMENTAL EDUCATION CONFERENCE

The IX Ohio Developmental Education Conference entitled "ABC's of D.E.: Articulation, Burnout and Competencies" will be held November 5-7, 1982 at the Harley Hotel, Cincinnati, Ohio. The program is focused to meet the needs of professionals at both the secondary and post-secondary level. National participation is invited. For information and registration forms contact Dr. Tanya Ludutsky or Dr. Phyllis Sherwood, Raymond Walters General and Technical College, 9555 Plainfield Road, Cincinnati, Ohio 45236 (513-745-4202).

USA MATH OLYMPIAD WINNERS

The 100 top performers in the Annual High School Mathematics Contest competed on May 4, 1982 in the Eleventh USA Mathematical Olympiad. From the winners listed below, four were selected (noted by *) to be the USA team to compete at the 1982 International Mathematical Olympiad, held in Budapest, Hungary, on July 9-10, 1982. The U.S. team placed third in the IMO, with a team score of 136. More news on the IMO, including the contest problems, will appear in our November issue. USAMO winners:

*Noam D. Elkies, New York, NY;
*Douglas S. Jungreis, N. Woodmere, NY;
*Brian R. Hunt, Silver Spring, MD;
Tsz Mei KO, Corona, NY;
*Washington Taylor IV, Cambridge, MA;
Vance Maverick, Los Angeles, CA;
W. David Vinke, Alvinston, Ontario;
Edith N. Starr, Philadelphia, PA.

Match your wits with these bright high school students--try the questions from both the USA and Canadian Olympiads.

11TH USA MATHEMATICAL OLYMPIAD
MAY 4, 1982

1. In a party with 1982 persons, among any group of four there is at least one person who knows each of the other three. What is the minimum number of people in the party who know everyone else?

2. Let $S_r = x^r + y^r + z^r$ with x, y and z real. It is known that if $S_1 = 0$,

$$\frac{S_{m+n}}{m+n} = \frac{S_m}{m} \cdot \frac{S_n}{n} \qquad (*)$$

for $(m,n) = (2,3),(3,2),(2,5)$ or $(5,2)$. Determine *all* other pairs of integers (m,n), if any, satisfying (*).

3. If point A_1 is in the interior of an equilateral triangle ABC and point A_2 is in the interior of triangle A_1BC, prove that

$$I.Q.(A_1BC) > I.Q.(A_2BC)$$

where the *isoperimetric quotient* of a figure F is defined by

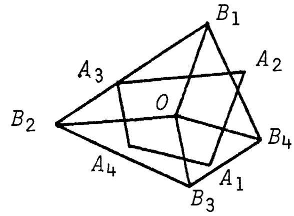
$$I.Q. = \frac{\text{Area}(F)}{[\text{Perimeter}(F)]^2} .$$

4. Prove that there exists a positive integer k such that $k2^n+1$ is composite for every positive integer n.

5. Given that A, B and C are three interior points of a sphere S such that AB and AC are perpendicular to the diameter of S through A. Through A, B and C, two spheres can be constructed which are both tangent to S. Prove that the sum of their radii is equal to the radius of S.

CANADIAN MATHEMATICAL SOCIETY
14TH MATHEMATICS OLYMPIAD
MAY 5, 1982

1.



In the diagram, OB_i is parallel and equal in length to A_iA_{i+1} for $i = 1,2,3$ and 4 ($A_5 = A_1$). Show that the area of $B_1B_2B_3B_4$ is twice that of $A_1A_2A_3A_4$.

2. If a, b and c are the roots of the equation $x^3 - x^2 - x - 1 = 0$,

(i) show that a, b and c are distinct;

(ii) show that

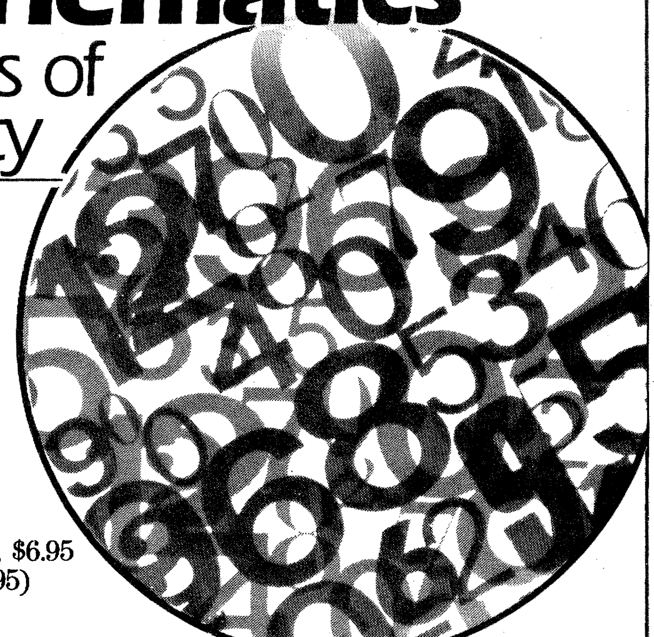
$$\frac{a^{1982} - b^{1982}}{a - b} + \frac{b^{1982} - c^{1982}}{b - c} + \frac{c^{1982} - a^{1982}}{c - a}$$

is an integer.

3. Let R^n be n-dimensional Euclidean space. Determine the smallest number $g(n)$ of points of a set in R^n such that every point in R^n is at an irrational distance from at least one point in that set.

4. Let p be a permutation of the set $S_n = \{1,2,\ldots,n\}$. An element $j \in S_n$ is called a fixed point of p if $p(j) = j$. Let f_n be the number of permutations of S_n having no fixed points, and g_n be the number with exactly one fixed point. Show that $|f_n - g_n| = 1$.

5. The altitudes of a tetrahedron $ABCD$ are extended externally to points A', B', C' and D' respectively, where $AA' = k/h_a$, $BB' = k/h_b$, $CC' = k/h_c$ and $DD' = k/h_d$. Here k is a constant and h_a denotes the length of the altitude of $ABCD$ from vertex A, etc. Prove that the centroid of the tetrahedron $A'B'C'D'$ coincides with the centroid of $ABCD$.



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

MATHEMATICS MAGAZINE VOL. 55, NO. 4, SEPTEMBER 1982