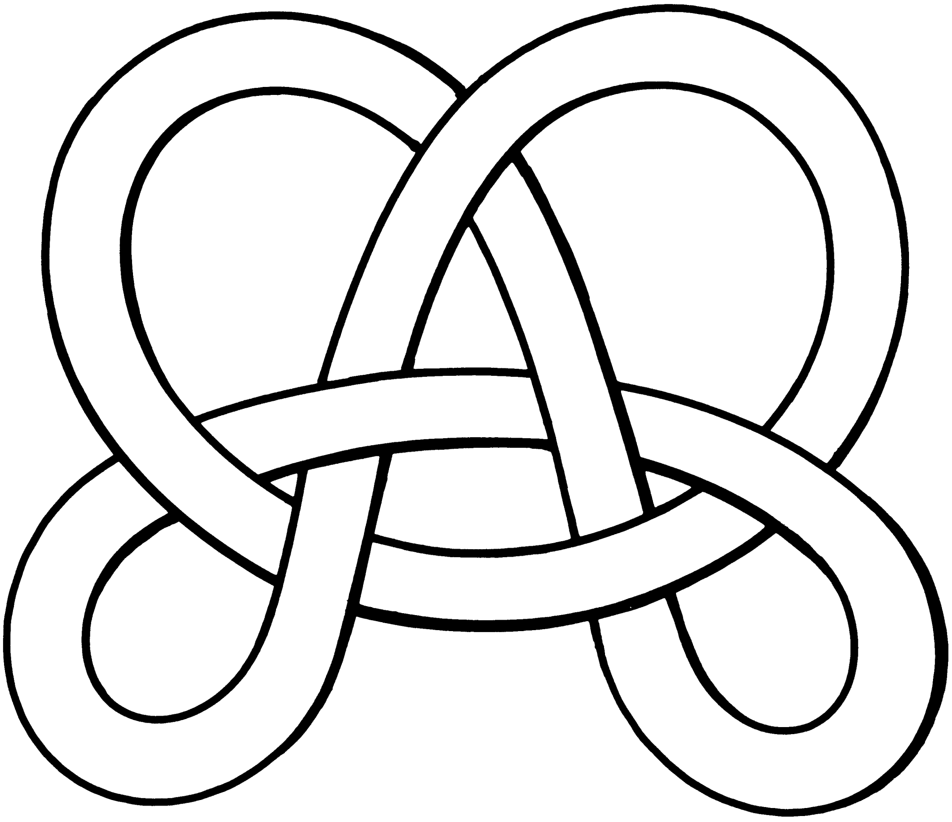
Vol. 62, No. 4 October 1989

MATHEMATICS MAGAZINE



- Connectivity and Smoke-Rings
- Taxicab Geometry
- Tennis Ball Paradox

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

The aim of *Mathematics Magazine* is to provide lively and appealing mathematical exposition. This is not a research journal and, in general, the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for an article for the *Magazine*. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Articles on pedagogy alone, unaccompanied by interesting mathematics, are not suitable. Neither are articles consisting mainly of computer programs unless these are essential to the presentation of some good mathematics. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

The full statement of editorial policy appears in this *Magazine*, Vol. 54, pp. 44–45, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, nor published by another journal or publisher.

Send new manuscripts to: G. L. Alexanderson, Editor, Mathematics Magazine, Santa Clara University, Santa Clara, CA 95053. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and one copy and keep one copy. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added.

AUTHORS

Thomas Archibald obtained a Ph.D. from the University of Toronto in 1987, and is currently teaching in the Department of Mathematics of Acadia University, Wolfville, Nova Scotia. His research interests include the history of analysis in the nineteenth century, with particular emphasis on the relationship between developments in pure mathematics and those in mathematical physics.

Editorial note. After careful searches through standard sources and consultations with Green's biographers, we have concluded that a portrait of George Green does not exist. If any reader has information to the contrary we would be very pleased to learn of it.

Vol. 62 No. 4 October 1989

MATHEMATICS MAGAZINE

The *MATHEMATICS MAGAZINE* (ISSN 0025–570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.
The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $13 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $29. Student, unemployed and emeritus members receive a 50% discount; new members receive a 30% dues discount for the first two years of membership.) The nonmember/library subscription price is $45 per year. Bulk subscriptions (5 or more copies) are available to colleges and universities for classroom distribution to undergraduate students at $5.40 per copy (minimum order 5 copies).
Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.
Advertising correspondence should be addressed to Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

Second class postage paid at Washington, D.C. and additional mailing offices.
Postmaster: Send address changes to Mathematics Magazine Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20077-9564.
PRINTED IN THE UNITED STATES OF AMERICA

ARTICLES

Connectivity and Smoke-Rings: Green's Second Identity in Its First Fifty Years

THOMAS ARCHIBALD
Acadia University,
Wolfville, Nova Scotia B0P 1X0

Introduction

James Clerk Maxwell, in his review of Thomson and Tait's *Treatise on Natural Philosophy*, noted an important innovation in the authors' approach to mathematics:

> The first thing which we observe in the arrangement of the work is the prominence given to kinematics,... and the large space devoted under this heading to what has been hitherto considered part of pure geometry. The theory of curvature of lines and surfaces, for example, has long been recognized as an important branch of geometry, but in treatises on motion it was regarded as lying as much outside of the subject as the four rules of arithmetic or the binomial theorem.
>
> The guiding idea, however... is that geometry itself is part of the science of motion, and that it treats, not of the relations between figures already existing in space, but of the process by which these figures are generated by the motion of a point or a line. [1]

This "guiding idea," which treats geometric entities as physical objects in some sense, had been influential with mathematicians for many years. Countless mathematical problems have their origin in the investigation of the natural world. However, it also happens that the solutions of some problems may be facilitated by attributing physical properties to the mathematical objects under study. In addition, mathematical constructs usually thought of as "purely geometric" may be created by considering such mathematico-physical entities.

It is my purpose in this article to illustrate some aspects of the cross-fertilization of mathematics and physics by examining the development of Green's second identity (known to physicists as Green's theorem) and its generalizations over a fifty-year period, from 1828 to 1878. During this period, despite an increased emphasis on logical rigour in some circles, many mathematicians continued to accept physical proofs of analytic theorems as valid. Such proofs used hypothetical physical properties such as incompressibility to characterize the regions in space; this trend was found most strongly in nineteenth-century British mathematics, though it was not unknown elsewhere.

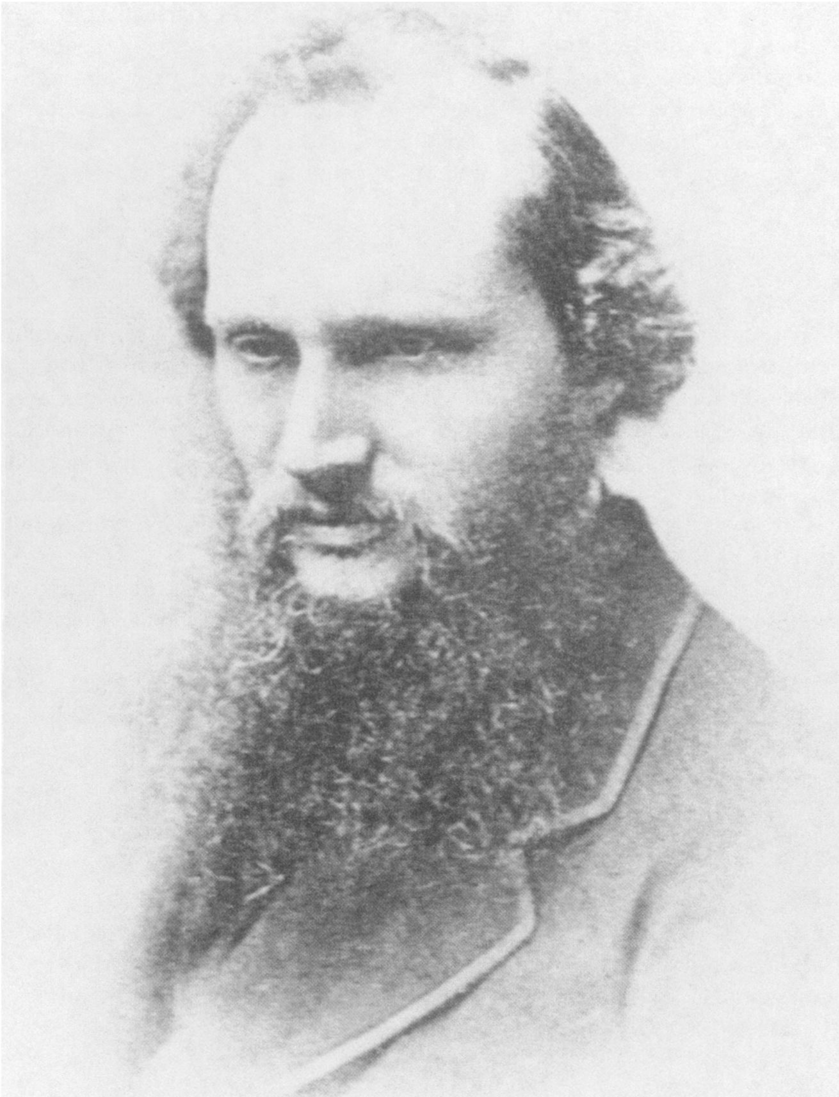
Green's second identity, well known from vector calculus, states that

Peter Guthrie Tait

$$\iiint(\varphi\nabla^2\psi-\psi\nabla^2\varphi)\,dx\,dy\,dz=\oint\left(\varphi\frac{\partial\psi}{\partial n}-\psi\frac{\partial\varphi}{\partial n}\right)da.$$

The integration on the left is performed over a region bounded by a closed surface S. The integral on the right is then a surface integral over S, and n is an outward normal to S. Finally, φ and ψ are continuously differentiable real-valued functions (scalar fields) on R^3. This theorem first appeared in a paper by George Green published in 1828, along with a number of other lemmas which Green employed in his study of electrostatics and magnetism. Green's results remained virtually unknown, however, until William Thomson (later Lord Kelvin) obtained two copies of Green's pamphlet in 1845. Green's results subsequently became widely known, and were central to the mathematical theory of potential, one of the most important tools of mathematical physics in the following decades.

Potential theory had originated as a body of results which arose in connection with the efforts of French mathematical physicists (notably Poisson, Laplace, and Biot), to



William Thomson, later Lord Kelvin

extend the methods of Newton. Laplace attempted to explain many natural phenomena as the result of forces proportional to the inverse square of the distance between the interacting objects. To achieve this, it was necessary to determine the integrals of vector forces. Laplace showed that such forces could be treated as what we now term the gradient of a scalar function, and hence was able to simplify the calculations greatly. Such a function, the gradient of which is a force, is known as a potential for that force. (We will also see velocity potentials in the course of this article, which are functions the gradient of which gives a velocity.) Like that of Laplace, Green's work was a contribution both to mathematical physics and to potential theory, since it expresses relationships between potentials and their integrals as well as applying the results to physical problems [2].

At its beginning, the idea of potential was a mathematical convenience. However, by the 1850s it had acquired physical interpretations. In particular, if a vector function has a potential, the integral of that potential along a curve depends only on the endpoints of the integration, and integrals around closed paths are zero. This expresses the fact that the vector function is an exact differential. Physically, this implies that the force described by the function is conservative, so that potential functions are closely associated with potential energy.

Green's 1828 Essay

Green's paper, which was published privately in Nottingham in 1828, was called *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism* [3]. George Green (1793–1841), a miller's son, had been given access to the library of a local aristocrat interested in science. This opportunity, and Green's ability, permitted him to master basic works by Laplace, Lagrange, and Poisson. Inspired by Laplace's work on gravitation and Poisson's on electrostatics and magnetism, Green set forth to investigate electrostatics using similar hypotheses but new methods.

Of particular interest to us are Green's general mathematical theorems, presented at the beginning of the paper, which he later applied to particular electrical and magnetic calculations. Green's second identity is the key theorem in this section. It is the essential tool in solving the Laplace equation and the Poisson equation by the method which Green introduced. A detailed discussion of this method, today known as the method of Green's functions, would take us too far afield.

In modern notation, the identity Green proved was

$$\int U\nabla^2 V d^3x + \oint U\frac{\partial V}{\partial n}\,d\sigma = \int V\nabla^2 U d^3x + \oint V\frac{\partial U}{\partial n}\,d\sigma. \tag{1}$$

Here U and V are any two functions which are continuously differentiable in the region of differentiation, and n is now the inward normal from the surface σ. Green used the symbol δ to express what we have denoted by ∇^2. Green's proof of this identity rests on applying integration by parts to the expression

$$\iiint\left\{\frac{\partial V}{\partial x}\cdot\frac{\partial U}{\partial x}+\frac{\partial V}{\partial y}\cdot\frac{\partial U}{\partial y}+\frac{\partial V}{\partial z}\cdot\frac{\partial U}{\partial z}\right\}dx\,dy\,dz = \int(\nabla V)\cdot(\nabla U). \tag{2}$$

Assuming U, V are sufficiently differentiable, we can integrate by parts in each variable. For example let

$$u = \frac{\partial U}{\partial x}\,dx \quad \text{and} \quad v = \frac{\partial V}{\partial x}\,dx.$$

Then substitution in (2) yields

$$\iint dy\,dz\left(\int\frac{dV}{dx}\frac{dU}{dx}\,dx\right)$$

$$= \iint V(x_1)\left.\frac{dU}{dx}\right|_{x=x_1} dy\,dz - \iint V(x_0)\left.\frac{dU}{dx}\right|_{x=x_0} dy\,dz - \int V\frac{\partial^2 U}{\partial x^2}\,dx\,dy\,dz.$$

Green then argued that, if σ is a surface element and n an inward normal, we have

$$\iint dy\,dz\left(V(x_1)\frac{dU}{dx}\bigg|_{x_1} - V(x_0)\frac{dU}{dx}\bigg|_{x_0}\right) = -\int d\sigma\frac{\partial x}{\partial n}V\frac{dU}{dx}.$$

Hence the partial integral becomes

$$\iint dy\,dz\left(\int\frac{\partial V}{\partial x}\frac{\partial U}{\partial x}dx\right) = -\int d\sigma\frac{\partial x}{\partial n}V\frac{\partial U}{\partial x} - \int V\frac{\partial^2 U}{\partial x^2}.$$

Consequently the result of integration with respect to all three variables gives

$$\iint dx\,dy\,dz(\nabla V)\cdot(\nabla U) = -\int d\sigma V\frac{\partial U}{\partial n} - \int V\nabla^2 U. \tag{3}$$

This is often known as Green's first identity. By symmetry, we may interchange U and V in (3) to obtain the second identity:

$$\int d\sigma V\frac{\partial U}{\partial n} - \int V\nabla^2 U = -\int d\sigma\, U\frac{\partial V}{\partial n} - \int V\nabla^2 U.$$

Many present-day niceties in the proof of this identity were not considered by Green. A full proof involves dealing properly with the relationship between the infinitesimals and the finite, and we must use the equivalence of multiple and interated integrals, which Green did not distinguish. The advances in rigorous analysis due to Cauchy may well have been unknown to Green at this time, since he mentions his limited access to the latest work. Instead his arguments rely on the geometry of infinitesimals. In this, his work resembles that of most of his contemporaries, even in France.

Green's work went almost entirely unnoticed for many years. None of the private subscribers who purchased the pamphlet appears to have been capable of appreciating its worth, and his results and methods remained little known [4]. Green's work might have been forgotten had it not been mentioned by the Irish electrician Robert Murphy. Murphy referred to Green as the originator of the term potential, though Murphy's own definition of potential was erroneous, indicating that he had not actually seen Green's work [5].

Green himself did not revive interest in his earlier work. His efforts in the interim were devoted to further research, and to an education at Cambridge. His other papers met a happier immediate reception; several were published in the *Transactions of the Cambridge Philosophical Society*, where they attracted the interest of the British scientific community. Green thus made a name for himself before his death in 1841, though his reputation was considerably enhanced by the rediscovery of the 1828 *Essay*.

Thomson Rediscovers Green

It was William Thomson (later Lord Kelvin) who first drew the attention of the international scientific world to Green's results. Sometime in 1842 Thomson had read a reference by Murphy to Green's paper; his interest was piqued for several reasons. Thomson was himself then engaged in research on the theory of attraction, and published papers on the subject in 1842 and 1843. Murphy had referred to Green's use of the term potential, a notion which, as Thomson states, was also employed by Gauss with great success in his 1839 paper on inverse-square forces. Thomson doubtless wondered how Green, whose name he knew well, had employed the notion of potential, and was curious about the exact nature of his results.

Thomson was unable to see a copy of Green's work until January 25, 1845, shortly before he was about to embark on a trip to France following the completion of his studies at Cambridge. By chance, Thomson's tutor, Hopkins, had two copies which he had apparently never examined, and sent them with Thomson. Thomson was very impressed with the generality of Green's results, and was soon endeavouring to apply them in his own research. On his arrival in France, Thomson showed the paper to Liouville, Sturm, and Chasles, among others. Soon the Paris mathematical community was well aware of Green's work: for example, Liouville gave Green and Gauss equal credit for the introduction of the term potential in an 1847 paper [6]. Thomson sent the other copy of Green's paper to Germany with Cayley, who delivered it to August Crelle, editor of the *Journal für die reine und angewandte Mathematik*. Crelle published a translation of Green's paper in three installments between 1850 and 1854, hence it became well known to interested researchers in Germany. Thus, 25 years after Green's original publication, his methods began to find their way into the scientific literature and textbooks of Europe [7]. One of the first to make use of Green's work was Bernhard Riemann (1826–1866), who was then writing his doctoral dissertation at Göttingen.

Riemann and Multiply-Connected Regions

Riemann's principal interest in Green's work was in the method of Green's functions, which Green had used to solve boundary-value problems involving functions satisfying Laplace's equation

$$\nabla^2\varphi = 0.$$

Here φ is to be interpreted as the potential function of the electrostatic force due to a charge density on a conductor. Riemann, however, noticed that Green's methods could be useful in the study of functions of a complex variable, since the real and complex parts of such functions must satisfy Laplace's equation. Employing this insight, Riemann developed methods that enabled him to specify a complex function by its boundary values and discontinuities. In so doing, Riemann presented the idea of *multiply-connected regions* of the plane: a region is simply connected if a cross-cut divides it in two, and has connectivity equal to the number of cuts taken to separate it. (See FIGURE 1.) This notion was published in Riemann's dissertation (1851) and found wider circulation with the appearance of his paper on abelian integrals (1857) [8]. It was here that it was seen by Hermann von Helmholtz, who was attempting to employ Green's ideas in a different way.

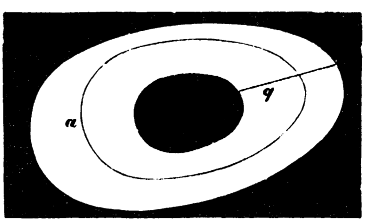


FIGURE 1.
Riemann's illustration of a doubly connected surface. (From B. Riemann, *Gesammelte Math. Werke*, 2nd edition, Leipzig, Teubner, 1892.)

Helmholtz and Vortices

Hermann von Helmholtz's 1858 paper *On Integrals of Hydrodynamic Equations Which Yield Vortex Motion* was also deeply influenced by Green's work [**9**]. Helmholtz (1821–1894) had become interested in the solution of boundary-value problems in fluid mechanics in connection with his investigation of the physiology of the ear. (He was at that time a professor of anatomy in Bonn [**10**].) Furthermore, Helmholtz saw a parallel between certain problems in hydrodynamics and problems in electromagnetic theory, a longstanding interest of his. Attempts to provide a detailed theoretical treatment of the analogy between electromagnetic theory and fluid dynamics may have been sparked by the superficial resemblance between electrical and hydrodynamical phenomena. The electric current was widely viewed in the mid-nineteenth century as the flow of one or two "electric fluids" along a conductor. The motion of this fluid produces a magnetic effect. André-Marie Ampère demonstrated in the 1820s that magnetism may be explained as the result of hypothetical microscopic electric currents in a body, and hence the existence of the electromagnetic phenomenon should mean that a current gives rise to other currents. If the original current flows in a straight line, the currents responsible for magnetic effects must be helical, forming microscopic vortices.

The researches of Helmholtz and others in this area aimed to make this rather vague picture precise. In the 1858 paper, Helmholtz examined the following question: suppose we are given a closed container filled with a frictionless incompressible fluid. How does action on the boundary of the container affect the motion inside?

Helmholtz apparently saw the value of Green's theorems in such an investigation soon after reading Green's paper, but was kept from working out his ideas because of other academic obligations. However, Helmholtz had also recently read Riemann's paper of 1857, which made it clear to him that Green's theorem could only be used when the regions involved were simply connected. This is because functions with potentials—what we would now term conservative vector fields—may in fact be multiple-valued in multiply-connected regions.

Let us discuss how Helmholtz used Green's theorem. He began with Euler's equation of fluid dynamics, which we may write in vector notation as

$$\vec{F} = \frac{1}{\rho}\nabla\vec{p} + \left(\frac{\partial}{\partial t} + \vec{v}\cdot\nabla\right)\vec{v}, \tag{4}$$

$$\nabla\cdot\vec{v} = 0. \tag{5}$$

Here ρ is the density of the fluid, $\vec{p}$ the pressure, $\vec{v}$ the velocity. Helmholtz supposed that

$$\vec{F} = \nabla V \quad \text{(where } V \text{ is a force-potential)}$$

and

$$\vec{v} = \nabla\varphi \quad (\varphi \text{ is a velocity-potential}).$$

From (5) we have that φ satisfies Laplace's equation, since $\nabla\cdot\vec{v} = \nabla^2\varphi = 0$. Helmholtz then noted that this implies

$$\nabla\times\vec{v} = \nabla\times(\nabla\varphi) = 0.$$

Bernhard Riemann

If the walls of the container are rigid, this means that the component of velocity perpendicular to the boundary is zero. Hence, if n is an outward normal, $\partial\varphi/\partial n$ is equal to zero everywhere. But by Green's first identity,

$$\int_R (\nabla U \cdot \nabla V)\, dx^3 = \iint V \frac{\partial U}{\partial n}\, ds - \int_R V \nabla^2 U\, dx^3.$$

If $U = V = \varphi$ we have

$$\int_R (\nabla\varphi)^2\, dx^3 = \iint \varphi \frac{\partial\varphi}{\partial n}\, ds = 0, \qquad \text{(Remember } \nabla^2\varphi = 0.\text{)}$$

so that $\nabla\varphi = 0$, i.e., no motion is induced in the fluid. Hence, any motion of a fluid in a closed vessel (with simply-connected interior) which has a velocity potential must depend exclusively on a motion of the boundary. Helmholtz went on to show that a motion of the boundary uniquely determines such a motion in the fluid. A further important conclusion stated that vortices can only be produced by a motion which has no velocity potential. More important still, if vortices do exist initially, they are stable

Hermann von Helmholtz

under the action of conservative forces. Either they must be closed tubes, or else they extend from the boundary to the boundary.

When vortices do exist one can consider the portion of fluid without vortices as a multiply-connected region, the vortices being the "holes." Thus to solve boundary value problems in such a region, one would ideally have an extension of Green's theorem to deal with such cases. Helmholtz stressed the desirability of such a generalization, and noted that Riemann's notion of connectivity could readily be extended to three dimensions for this purpose.

In Helmholtz's work the geometric entities immediately become physical. For one thing, they are three dimensional. Also the points of space become associated with the molecules of a fluid, and holes in the space correspond to vortices. In this instance, the geometric entities may be given clear physical interpretation, and physical questions (the solution of specific boundary-value problems, for example) dictate the mathematical problems which are important.

Helmholtz's research was received with greater sympathy by British mathematical physicists, especially Thomson, than by his German colleagues. This occurred in part because of the shared interests of Helmholtz and Thomson in hydrodynamic models for electromagnetic theory, an interest that arose because of their attitude toward the then-prevailing thought on electromagnetic theory in Germany. This theory, based on work by Gauss's collaborator Wilhelm Weber, explained electrical phenomena on the basis of a velocity-dependent force law. Both Helmholtz and Thomson felt that such a force could not satisfy the energy conservation principle. Helmholtz apparently felt as well that his mathematical skills were dimly regarded by his German contemporaries, because he had not been formally trained as a mathematician. Thus it was among British mathematical physicists that Helmholtz's papers were read with greatest interest and understanding.

Tait, Thomson, Smoke-Rings and Atoms

Helmholtz's approach found an enthusiastic admirer in Peter Guthrie Tait (1837–1901), a Cambridge-educated Scot who was teaching in Belfast in 1858. Tait was attempting at that time to master William Rowan Hamilton's method of quaternions, and to demonstrate the physical usefulness of the method by obtaining significant applications. In this respect Helmholtz's work interested him, and he made an English translation for his own use [**11**].

A parenthetic note about quaternions: Nowadays, this set of objects is most likely to show up in algebra courses or proofs in algebraic number theory which can make use of its properties as a noncommutative division ring. This is quite remote from their original intended use in geometry and analysis. Hamilton invented quaternions in 1843, and introduced with them the idea of operators. Particularly important was the del or nabla operator, our ∇. For Hamilton and Tait, a quaternion described a quotient of what we would term vectors; such a quotient consists of a 4-tuple which describes the stretch and the three rotations which bring an arbitrary pair of vectors into coincidence [**12**]. Later on we shall see how Tait used this approach to obtain what he called "physical proofs" of analytic statements.

Tait moved to Edinburgh in 1860, where he began a collaboration with William Thomson, then at Glasgow. In 1866 and 1867 their collaboration was at its peak, as they prepared their *Treatise on Natural Philosophy* (which was to become the standard introductory physics text in Britain for decades). Early in 1867, Tait showed Thomson an experimental demonstration of the stability properties of vortices by means of smoke-rings, as well as Helmholtz's mathematical treatment of the problem. Thomson described this event to Helmholtz in a letter:

> Just now, however, vortex motions have displaced everything else, since a few days ago Tait showed me in Edinburgh a magnificent way of producing them. Take one side (or a lid) off a box (any old packing box will serve) and cut a large hole in the opposite side. Stop the open side *AB* loosely with a piece of cloth, and strike the middle of the cloth with your hand. If you leave anything smoking in the box, you will see a magnificent ring shot out by every blow.

Thomson then went on to describe what he found particularly interesting about the phenomenon and the theory.

> The absolute permanence of the rotation, and the unchangeable relation you have proved between it and the portion of the fluid once acquiring such motion in a perfect fluid, shows that if there is a perfect fluid all through space, constituting the substance of all matter, a vortex-ring would be as permanent as the solid hard atoms assumed by Lucretius and his followers (and predecessors) to account for the permanent properties of bodies... thus if two vortex rings were once created in a perfect fluid, passing through one

another like links of a chain, they could never come into collision, or break one another, they would form an indestructible atom. [**13**]

Thomson embarked on the mathematical theory of these apparently indestructible vortex atoms at once, and his results were read before the Royal Society of Edinburgh a little over three weeks later. His paper *On Vortex Motion*, much augmented, appeared in 1878 [**14**]. Here he encountered and solved the problem posed by Helmholtz of extending Green's theorem to multiply-connected regions. For in order to investigate the properties of vortex atoms, it was necessary to solve boundary value problems where complicated vortices (such as those pictured) formed part of the boundary. (See FIGURE 2.)

In this paper, Thomson wrote the original version of Green's theorem thus:

$$\int_R \nabla\varphi\cdot\nabla\varphi'\,dV = \oint \varphi\frac{\partial\varphi'}{\partial n}\,da - \int_R \varphi\nabla^2\varphi'\,dV$$
$$= \oint \varphi'\frac{\partial\varphi}{\partial n}\,da - \int_R \varphi'\nabla^2\varphi\,dV.$$

Here φ and φ' must be single-valued. Thomson then investigated what happens if φ' is multivalued, that is to say, if we consider R to be a multiply-connected region, and then considered how to set the problem up in general. A multiply-connected space can be made simply connected by making cuts, or by inserting what Thomson calls stopping barriers. The integral around an (almost) closed path from one side of the barrier to the other has a constant value k_i, which is the same for all such paths; such constants k_i exist for all stopping barriers and the integral in question becomes

$$\int_R \nabla\varphi\cdot\nabla\varphi'\,dV = \oint \varphi\frac{\partial\varphi'}{\partial n}\,d\sigma + \sum_i k_i \iint \frac{\partial\varphi'}{\partial n}\,d\sigma' - \int_R \varphi\nabla^2\varphi'\,dV$$
$$= \oint \varphi'\frac{\partial\varphi}{\partial n}\,d\sigma + \sum_i k_i \iint \frac{\partial\varphi}{\partial n}\,d\sigma' - \int_R \varphi'\nabla^2\varphi\,dV.$$

Here $d\sigma'$ represents a surface element of the barrier surface.

Thus armed, Thomson was able to examine fluid motion in multiply-connected regions, concluding that the normal component of velocity of a fluid at every point of

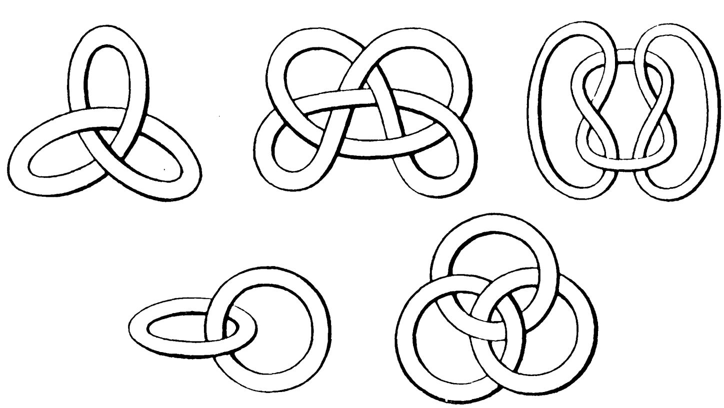


FIGURE 2.
William Thomson's Knots. (From W. Thomson, On vortex motion, *Trans. Royal Soc. Edinburgh*, 25, 1869.)

the boundary determines the motion inside a multiply-connected region (provided we know the circulation of the fluid in each region). He then considered how best to define the order of connectivity, noting that for some of the surfaces shown the stopping barriers must be self-intersecting and difficult to distinguish. He therefore proposed a definition using what he called "irreconcilable paths"—which we would now call homotopy classes of closed paths with base point. He selected a point on the surface, and noted that the connectivity of the surface is determined if we see how many mutually irreconcilable paths can be drawn on its surface. For a simply-connected region, for example, all closed paths on the surface are homotopic. Although this affords an unambiguous definition of connectivity, it is not easily possible using this method to get the generalization of Green's theorem. For that the stopping barriers are required.

Thomson's interest in vortex atoms thus led him directly to a generalization of Green's theorem, and to the question of the proper definition of connectivity. His proof technique is in essence the same as that of Green, augmented by the stopping barriers, and it is this method that is usually taught today in courses on vector calculus.

Tait's Quaternion Version of Green's Theorem

By the late 1860's, Tait's interest in quaternions had turned into a crusade. To the British association in 1871, he said:

> comparing a Cartesian investigation...with the equivalent quaternion one...one can hardly help making the remark that they contrast even more strongly than the decimal notation with the binary scale or with the old Greek arithmetic, or than the well-ordered subdivision of the metrical system with the preposterous non-systems of Great Britain.

In the same address, Tait pointed out that from the quaternion point of view:

> Green's celebrated theorem is at once seen to be merely the well-known equation of continuity expressed for a heterogeneous fluid, whose density at every point is proportional to one electric potential, and its displacement or velocity proportional to and in the direction of the electric force due to another potential. [**15**]

Let us see exactly what he means. In his 1870 paper *On Green's and Other Allied Theorems*, Tait supposed a spatial region R to be uniformly filled with points [**16**]. If points inside and outside the regions are displaced by a vector then we may have a net decrease or increase of the volume—that is, of the number of points—in the region R. This can be calculated in two ways:

1. We can find the total increase in density throughout R

$$\int_R \operatorname{div}\sigma\, dV \left(\text{in Tait's notation } \iint S\cdot\nabla\sigma\, ds\right). \tag{6}$$

2. We can estimate the excess of those that pass inwards through the surface over those that pass outwards:

$$\iint \sigma\cdot\vec{n}\,da \left(\text{in Tait's notation } \iint s\cdot\sigma UV ds\right). \tag{7}$$

The expressions (6) and (7) must be equal, yielding what we now call the divergence theorem from the equation of continuity. If we consider that the density—for example, of electric fluid— is given by a potential P_1, and the displacement is proportional to a force σ with potential P, then we have:

$$\nabla(PP_1) = P\nabla P_1 + P_1\nabla P$$

and

$$\nabla^2(PP_1) = P\nabla^2 P_1 + P_1\nabla^2 P + 2(\nabla P \cdot \nabla P_1). \tag{8}$$

But by the divergence theorem

$$\int_R \nabla^2(PP_1)\,dV = \int_R \operatorname{div}(\nabla PP_1)\,dV = \int_{\partial R}(\nabla PP_1)\cdot\vec{n}\,da$$

$$= \int_{\partial R}(P\nabla P_1 + P_1\nabla P)\cdot\vec{n}\,da.$$

Hence from (8)

$$\int_{\partial R}(P\nabla P_1 + P_1\nabla P)\cdot\vec{n}\,da = \int_R (P\nabla^2 P_1\nabla^2 P)\,dV + 2\int_R \nabla P\cdot\nabla P_1\,dV.$$

But the left side here, by the divergence theorem, is

$$= \int_R (P\nabla^2 P_1 - P_1\nabla^2 P)\,dV.$$

Combining these two yields Green's theorem in the form

$$\int_R (\nabla P\cdot\nabla P_1)\,dV = -\int_R P_1\nabla^2 P + \int_{\partial R} P\frac{\partial P}{\partial n}\,da$$

$$= -\int_R P\nabla^2 P_1 + \int_{\partial R} P\frac{\partial P_1}{\partial n}\,da.$$

Notice that the argument depends on treating geometric points as mobile physical entities, with continuity properties like those of a fluid.

We find Tait's views nicely summarized in his 1892 review of Poincaré's *Thermodynamique*:

> Some forty years ago, in a certain mathematical circle at Cambridge, men were wont to deplore the necessity of introducing words at all in a physico-mathematical textbook: the unattainable, though closely approachable Ideal being regarded as a world devoid of aught but formulae! But one learns something in forty years, and accordingly the surviving members of that circle now take a very different view of the matter. They have been taught alike by experience and by example to regard mathematics, so far at least as physical enquiries are concerned, as a mere auxiliary to thought... this is one of the great truths which were enforced by Faraday's splendid career. **[17]**

Conclusion

Our excursion from Green to Tait has taken us from electrostatics and potential theory, via complex analysis and fluid dynamics, to homotopy classes of maps and vector analysis. While I have only touched on a few of the interesting problems associated with these developments, I hope that I have shown that physical thinking is important, not only in posing mathematical problems, but also in solving them. In particular, physical thinking may lead to the creation of certain mathematical notions, such as connectivity, which are of interest in their own right, for example in the classification of the knots described by Thomson. Tait undertook this classification problem around 1870, achieving the first basic results of knot theory.

Acknowledgements The research for this paper was supported in part by a grant from the Social Sciences and Humanities Research Council of Canada. I am also grateful to two anonymous referees for helpful comments, and to J. Marontate for assistance at all stages of preparation of this paper.

REFERENCES

1. J. Clerk Maxwell, Review of Thomson and Tait in *Nature*, vol. xx, reprinted in *Scientific Papers*, vol. 2., 777.
2. O. D. Kellogg, *Foundations of Potential Theory*, Springer, Berlin, 1929.
3. George Green, *Mathematical Papers*, ed. N. Ferrers, Macmillan, London, 1871; reprint, Chelsea, New York, 1970.
4. H. Gwynedd Green, A Biography of George Green, in A. Montagu ed., *Studies and Essays in the History of Science and Learning*, Schumann, New York, 1956.
5. For information on Murphy, see W. Thomson, *Reprint of Papers on Electrostatics and Magnetism*, Macmillan, London, 1872, 1–2 and 126.
6. J. Liouville, reprinted in [5] 174.
7. S. P. Thompson, *Life of Lord Kelvin*, Chelsea, New York, 1976, vol. 1, 113–119.
8. G. F. B. Riemann, *Gesammelte Mathematische Werke*, Teubner, Leipzig, 1892, 92–96.
9. H. Helmholtz, Ueber Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen, in *Journal für Mathematik* 55 (1858), 25–55.
10. Leo Königsberger, *Hermann von Helmholtz*, Vieweg, Braunschweig, 1902, vol. 1, 307–312.
11. S. P. Thompson, *Life of Lord Kelvin*, Chelsea, New York, 1976, vol. 1, 511.
12. P. G. Tait, and P. Kelland, *Introduction to Quaternions*, Macmillan, London, 1904, 37–38.
13. S. P. Thompson, *Life of Lord Kelvin*, Chelsea, New York, 1976, vol. 1, 514–515.
14. W. Thomson, On vortex motion, *Trans. Royal Soc. Edinburgh* 25 (1869), 217–260.
15. C. G. Knott, *Life and Scientific Work of P. G. Tait*, Cambridge University Press, 1911.
16. P. G. Tait, *Scientific Papers*, Cambridge University Press, 1911, vol. 1, 136–150.
17. C. G. Knott, *Life and Scientific Work of P. G. Tait*, Cambridge University Press, 1911, p. 273.

NOTES

An Application of Set Theory to Coding Theory

NOGA ALON
Tel Aviv University
Ramat-Aviv, Tel Aviv, Israel

ANDY LIU
University of Alberta
Edmonton, T6G 2G1, Canada

"As high-speed electronic communication becomes commonplace, there is a tremendous need for better transmission schemes—ones that minimize the effect of inevitable transmission errors, ones that protect confidential or secret messages, ones that route messages most efficiently. Many of the best schemes are based on patterns or properties of classical algebraic and geometric objects, originally studied for their intrinsic interest. Mathematically, these are the subjects of information theory, coding and encryption" [**5**].

In this note, we will confine our attention to one aspect of information theory: error-correcting codes. We first state our assumptions about the setting in which these codes come into play. Messages are in the form of sequences of 0's and 1's. In transmission, the only errors that may occur are in the form of digit-reversals; that is, an error may change a 0 to a 1, or vice versa. The transmission device handles blocks of l digits at a time, and we know the maximum number of errors per transmission.

If the original message is broken up into blocks of length l, and transmitted as is, potential transmission errors will compromise the reliability of the received message. To trade off economy for accuracy, the original message is broken up into words of length $m < l$, and each word is augmented with $l - m$ digits in such a way that the correct word can be deciphered despite possible errors.

An error-correcting code may be defined as a pair of companion procedures. The first, that of determining how the $l - m$ additional digits are to be chosen, is called encoding. The second, that of recovering the correct word from the received block, is called decoding. The ratio m/l is a measurement of the efficiency of the code.

The subject of error-correcting codes is of immense scope and depth, as detailed in the monumental treatise by MacWilliams and Sloane [**6**]. An excellent exposition by Thompson [**10**] shows the interrelationship between codes and many other mathematical structures. Our primary purpose is to give another example along this line. We will show how a recent set-theoretic result of Frankl and Pach [**2**] provides an alternative justification for a family of known codes.

Error-correcting codes may be based on very simple ideas (see for example [**1**] and [**8**]), but these tend to suffer in efficiency. The extended Hamming codes (see [**3**] and [**4**]), discovered early in the history of information theory, enjoy the best of both worlds. We will begin by describing such a code in set-theoretic language.

Suppose we have a transmission device which handles 15 digits at a time, with at most 1 error per transmission. We break up the message into words of length 11.

TABLE I

a	a	a	a		a	a	a				a			
b	b	b		b	b			b	b			b		
c	c		c	c		c		c		c			c	
d		d	d	d			d		d	d				d
1	0	1	1	0	0	0	1	1	0	0	0	1	1	0

TABLE I shows how the word 10110001100 is to be encoded. All 15 nonempty subsets of $\{a, b, c, d\}$ are listed, each occupying a column over the horizontal line. A vertical line separates the 4 one-element subsets from the other 11. Beneath the horizontal line under these 11 subsets, the word is copied with one digit in each column.

The extra digit to be placed under $\{a\}$ is obtained as follows. Consider all other subsets which *include* $\{a\}$, namely, $\{a, b, c, d\}$, $\{a, b, c\}$, $\{a, b, d\}$, $\{a, c, d\}$, $\{a, b\}$, $\{a, c\}$ and $\{a, d\}$. (We could have referred to these as subsets which contain the element "a." However, anticipating our main result, we prefer to describe them using the relation of set inclusion.) Under the corresponding columns, there are four 1's. We append a 0 under $\{a\}$ so that the total number of 1's under all subsets which include $\{a\}$ is even. A similar process is used to determine the remaining three appended digits.

TABLE II

a	a	a	a		a	a	a				a			
b	b	b		b	b			b	b			b		
c	c		c	c		c		c		c			c	
d		d	d	d			d		d	d				d
1	0	1	1	0	0	1	1	1	0	0	0	1	1	0

Suppose the received block is 101100111000110 as shown on the bottom line of TABLE II. Because the total number of 1's under the columns corresponding to subsets which include $\{a\}$, $\{b\}$, $\{c\}$ and $\{d\}$ are 5, 4, 5 and 4 respectively, it is clear that an error has occurred. The parities for $\{a\}$ and $\{c\}$ are disturbed. Since there is at most one digit-reversal, it must have occurred under $\{a, c\}$. Thus the message can be decoded correctly.

More generally, suppose the blocks are of length $l \geqslant 3$. Let n be the greatest integer such that $2^n - 1 \leqslant l$. The word length is chosen to be the greatest integer m such that $m \leqslant 2^n - \binom{n}{0} - \binom{n}{1}$. We use an n-element set instead of $\{a, b, c, d\}$. The word is copied under the columns corresponding to the subsets of size at least two. If $m < 2^n - \binom{n}{0} - \binom{n}{1}$, the appropriate number of these columns are omitted. The digits under the columns corresponding to the subsets of size one are determined by the parity condition as in TABLE I.

The extended Hamming codes cannot correct multiple transmission errors. For instance, a double-error on the digits under $\{a, b\}$ and $\{a\}$ has the same net effect as a single error on the digit under $\{b\}$. However, our set-theoretic approach suggests a generalization immediately. Again, we use an example.

Suppose we have a transmission device which handles 15 digits at a time, with up to 3 errors per transmission. We break up the messages into words of length 5.

TABLE III is set up just like TABLE I, but with an additional vertical line separating the two-element subsets from the rest. The word 10110 is copied beneath the

TABLE III

a	a	a	a		a	a	a				a			
b	b	b		b	b			b	b			b		
c	c		c	c		c		c		c			c	
d		d	d	d			d		d	d				d
1	0	1	1	0	0	0	1	1	0	0	0	1	1	0

horizontal line under the columns corresponding to subsets of size at least three.

The digit under $\{a, b\}$ is obtained by ensuring that the total number of 1's under the subsets which include $\{a, b\}$ is even. It will be 0 since the number of 1's under the relevant subsets other than $\{a, b\}$ itself is even (2, under $\{a, b, c, d\}$ and $\{a, b, d\}$). The other five digits between the vertical lines are similarly determined. The table is then completed as in the extended Hamming code.

TABLE IV

a	a	a	a		a	a	a				a			
b	b	b		b	b			b	b			b		
c	c		c	c		c		c		c			c	
d		d	d	d			d		d	d				d
1	0	1	0	0	1	0	1	1	0	0	0	1	1	1

Suppose the block 101001011000111 is received as shown on the bottom line of TABLE IV. Parity checks reveal that violations occur with respect to $\{a, b\}$, $\{a, c\}$, $\{a, d\}$, $\{c, d\}$, $\{b\}$ and $\{c\}$. We use P, Q and R to denote the (at most) three subsets of $\{a, b, c, d\}$ such that the digits beneath them are reversed.

Note that each of $\{b\}$, $\{c\}$ and $\{a, d\}$ is a subset of an odd number (1 or 3) or P, Q and R, while each of $\{a\}$, $\{d\}$ and $\{b, c\}$ is included in an even number (0 or 2) of them. It follows that $\{a, d\}$ is a subset of exactly 1 of P, Q and R, say P. Moreover, one of Q and R, say Q, includes $\{a\}$ while the other, R, includes $\{d\}$.

Similarly, $\{b, c\}$ is not a subset of any of P, Q and R, and that each of $\{b\}$ and $\{c\}$ is a subset of exactly 1 of them. Violations with respect to $\{a, c\}$ and $\{c, d\}$ imply that the element "c" is in P, while the absence of a violation with respect to $\{b, d\}$ means that "b" cannot be in R. Hence "b" is in Q, which is consistent with the violation with respect to $\{a, b\}$.

Thus the errors occur under $P = \{a, c, d\}$, $Q = \{a, b\}$ and $R = \{d\}$, and decoding is completed. This type of code is known as the punctured Reed-Muller codes (see [7] and [9]).

More generally, suppose the blocks are of length $l \geqslant 3$. Let n be the greatest integer such that $2^n - 1 \leqslant l$. The maximum number of errors per transmission that can be accommodated is $2^{n-1} - 1$. Let t be the least integer such that $2^t - 1$ is greater than or equal to the maximum number of transmission errors. The word length is chosen to be the greatest integer such that $m \leqslant 2^n - \binom{n}{0} - \binom{n}{1} - \cdots - \binom{n}{t}$. We use an n-element set instead of $\{a, b, c, d\}$. Additional vertical lines separate subsets of size up to t. The word is copied under the columns corresponding to the subsets of sizes at least $t + 1$. The digits under the columns corresponding to the subsets of sizes $t, t-1, \dots, 1$ are determined recursively by the parity condition.

If $m < 2^n - \binom{n}{0} - \binom{n}{1} - \cdots - \binom{n}{t}$, the appropriate number of subsets of sizes at least $t + 1$ are omitted. It should be pointed out that in such cases, the codes can no longer be called the punctured Reed-Muller codes. It seems natural to refer to them as the extended Reed-Muller codes. In particular, the extended Hamming codes are special cases of the extended Reed-Muller codes.

Although the decoding schemes for both are clearly related, that for the extended

Reed-Muller codes, as illustrated by decoding TABLE IV, does seem rather ad hoc. We now come to the set-theoretic result which guarantees that, in fact, the errors can always be uniquely identified. For practical applications, decoding tables can be constructed.

To facilitate the description of parity-disturbances, we introduce a notation. For any set S and any positive integer t, let S^t denote the collection of nonempty subsets of S of sizes at most t. If $S=\{a,b,c,d\}$, then $S^2=\{\{a\},\{b\},\{c\},\{d\},\{a,b\},\{a,c\},\{a,d\},\{b,c\},\{b,d\},\{c,d\}\}$.

We also remind the reader of the associative operation of symmetric difference between two sets. $A\triangle B$ yields the set of elements belonging to exactly one of A and B. In general, $A_1\triangle A_2\triangle\cdots\triangle A_k$ yields the set of elements which belong to an odd number of the A_i.

To see how this notation helps describe parity-disturbances, we return to the example in TABLE IV. We have

$$\begin{aligned}&\{a,c,d\}^2\triangle\{a,b\}^2\triangle\{d\}^2\\&\quad=\{\{a\},\{c\},\{d\},\{a,c\},\{a,d\},\{c,d\}\}\\&\quad\quad\triangle\{\{a\},\{b\},\{a,b\}\}\triangle\{\{d\}\}\\&\quad=\{\{b\},\{c\},\{a,b\},\{a,c\},\{a,d\},\{c,d\}\}.\end{aligned}$$

In the general case of the extended Reed-Muller codes, if the transmission errors occur under the subsets $A_1, A_2,\ldots, A_r$, this will result in some parity-disturbances. Specifically, if D is a subset of the n-element set such that the size of D is at most t, then the parity with respect to D will be disturbed if and only if D is a subset of an odd number of the A_i, that is, if and only if D belongs to $A_1^t\triangle A_2^t\triangle\cdots\triangle A_r^t$.

The decoding problem is to determine the A_i, given the subsets of the same type as D. Thus it is crucial to know that the subsets A_i may be determined uniquely.

Suppose to the contrary that there are two distinct collections of subsets, $\{A_1, A_2,\ldots, A_r\}$ and $\{B_1, B_2,\ldots, B_s\}$, which give rise to the same pattern of parity-disturbances. Then we have $A_1^t\triangle A_2^t\triangle\cdots\triangle A_r^t=B_1^t\triangle B_2^t\triangle\cdots\triangle B_s^t$ or equivalently $A_1^t\triangle\cdots\triangle A_r^t\triangle B_1^t\triangle\cdots\triangle B_s^t=\varnothing$. If $A_i=B_j$ for some i and j, we may remove them. Since each of r and s is at most 2^t-1 (the maximum number of errors per block), $r+s\leqslant 2^{t+1}-2$. This contradicts the following result.

THEOREM. *Let $S_1, S_2,\ldots, S_r$ be distinct nonempty finite sets. If $S_1^t\triangle S_2^t\triangle\cdots\triangle S_r^t=\varnothing$, then $r\geqslant 2^{t+1}-1$ and this is best possible.*

Proof. If $S_1, S_2,\ldots, S_r$ are the $2^{t+1}-1$ nonempty subsets of a $(t+1)$-element set, then we have $S_1^t\triangle S_2^t\triangle\cdots\triangle S_r^t=\varnothing$. This shows that $r\geqslant 2^{t+1}-1$ is best possible. We must now show that this inequality always holds.

We begin by noting that because the S_i are distinct, the largest one among them is a subset of itself but of no other S_i. That is, there is certainly *some* set which is a subset of an odd number of the S_i. Thus it will be possible to choose a nonempty set B of minimal cardinality k which is a subset of an odd number of the S_i. If $k\leqslant t$, then B belongs to $S_1^t\triangle S_2^t\triangle\cdots\triangle S_r^t$. Since this collection is empty, we have $k\geqslant t+1$.

Next we claim that if C is any one of the 2^k-1 nonempty subsets of B, then there will be an odd number (and hence at least one) of the S_i having the property that $S_i\cap B=C$. Of course, if $C_1\neq C_2$, then the sets S_i with $S_i\cap B=C_1$ must be distinct from the sets S_j with $S_j\cap B=C_2$. That is, if this claim is true, there must be at least 2^k-1 different sets S_i. In other words, $r\geqslant 2^k-1\geqslant 2^{t+1}-1$, and the theorem will be proved.

We shall prove this claim by a "descending" inductive argument on $|C|$. When $|C| = k$, we have $C = B$, and $S_i \cap B = B$ for an odd number of S_i by the definition of B. Suppose then that the claim is true for all subsets of B having cardinality *greater* than j, $1 \leqslant j < k$. We will show that this implies that it must also be true for any subset of B of cardinality j.

We consider such a subset C, restricting our attention to those S_i which include C. We expect that some of these S_i will contain elements of $B - C$. Indeed, if D is any one of the 2^{k-j} subsets of $B - C$, we can form the class of S_i such that $S_i \cap B = C \cup D$.

It is our goal to show that the class corresponding to $D = \varnothing$ has an odd number of members. We establish this via three observations.

(1) *The total number of classes is even.*

More explicitly, there are exactly 2^{k-j} classes.

(2) *The total number of S_i under consideration is even.*

This follows from the definition of B (in particular, the minimality condition) and the fact that $|C| < |B|$.

(3) *All but perhaps one class contain an odd number of the S_i.*

This is the induction hypothesis, which applies to all classes except that corresponding to $D = \varnothing$.

It follows from these observations that the final class, corresponding to $D = \varnothing$, must necessarily contain an odd number of the S_i. The claim is thus proved, and the theorem follows.

Our theorem is essentially a paraphrase, in set-theoretic language, of a result in [2]. Conversely, this result could have been derived from the knowledge that the extended Reed-Muller codes do work.

The authors thank the referees for helpful suggestions which improved the presentation of this paper.

REFERENCES

1. F. L. Alt, A Bell Telephone Laboratories' computing machine (I), *Math. Tables and Aids to Comput.* (*Math. Comput.*) 3 (1948/49), 1–13.
2. P. Frankl and J. Pach, On the number of sets in a null t-design, *Europ. J. Comb.* 4 (1983), 21–23.
3. M. J. E. Golay, Notes on digital computing, *Proc. IRE* (*IEEE*) 37 (1949), 657.
4. R. W. Hamming, Error-detecting and error-correcting codes, *Bell System. Tech. J.* 29 (1950), 147–160.
5. A. Jaffe, Ordering the universe, *Renewing U.S. Mathematics*, National Academy Press, Washington, DC, 1984, p. 148.
6. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977, pp. 372–377.
7. D. E. Muller, Application of Boolean algebra to switching circuit design and to error detection, *IRE Trans.* EC-3 (1954), 6–12.
8. M. Rabenstein, An example of an error-correcting code, this MAGAZINE 58 (1985), 225–226.
9. I. S. Reed, A class of multiple-error-correcting codes and the decoding scheme, *IRE Trans.* IT-4 (1954), 38–49.
10. T. M. Thompson, *From Error-Correcting Codes through Sphere Packings to Simple Groups*, MAA, Washington, DC, 1983.

Taxicab Geometry—A New Slant

KATYE O. SOWELL
East Carolina University
Greenville, NC 27858

As anyone who has played with a bit of screen wire knows, a square is not a rigid geometrical figure. A slight push to the right or left will deform a square into a nonsquare rhombus. If such a transformation is performed on the points of a plane which has been coordinatized by a square grid so that the positive y-axis makes an angle of 60° with the positive x-axis and the shorter diagonals of each rhombus are drawn, an isometric grid results. Points will still be named by ordered pairs of real numbers with respect to the x-axis and the transformed y-axis (FIGURE 1a).

There are only three regular polygons which will tessellate the plane: the equilateral triangle, the square, and the regular hexagon. The first two of these can be subdivided into smaller similar polygons (yielding the isometric grid and the square grid). Work has already been done on taxicab geometry using the square grid; this note considers taxicab geometry using the isometric grid.

Square-taxi geometry arises because, for the two points $A=(x_1, y_1)$ and $B=(x_2, y_2)$, a new distance function is chosen:

$$d_S(A, B)=|x_1-x_2|+|y_1-y_2|.$$

Thus square-taxi geometry is not Euclidean because there the distance function, derived from the Pythagorean Theorem, is

$$d_E(A, B)=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}.$$

In iso-taxi geometry three distance functions arise depending upon the relative positions of the points A and B. At the origin three axes occur: the x-axis, the y-axis and the y'-axis. This latter axis, one of the lines formed by the shorter diagonals of the rhombi mentioned earlier, forms an angle of 60° with the y-axis and with the x-axis. The three axes separate the plane into six regions called hextants. These hextants will be numbered I–VI in a counterclockwise direction beginning with the hextant where the coordinates of the points are both positive (FIGURE 1a). At any point in the plane three lines may be drawn parallel to the axes which, in turn, separate the plane into six regions. Two points could then have a I–IV orientation, a II–V orientation or a III–VI orientation to one another.

The three distance functions for iso-taxi geometry are

i) if the two points have a I–IV orientation (FIGURE 1c) then $d_I(A, B)=|x_1-x_2|+|y_1-y_2|$;
ii) if the two points have a II–V orientation (FIGURE 1b) then $d_I(A, B)=|y_1-y_2|$;
iii) if the two points have a III–VI orientation (FIGURE 1b) then $d_I(A, B)=|x_1-x_2|$.

If the two points lie on a line parallel to the x-axis, then formula iii) is used; if the two points lie on a line parallel to the y-axis or to the y'-axis then formula ii) is used.

The reader may observe that the length of the segment $\overline{AB}$ can be quickly found by counting the number of units along the two adjacent sides of the parallelogram ($\square$ $AQBP$ in FIGURE 1c), having $\overline{AB}$ as its longer diagonal, formed by lines parallel to the axes drawn at each endpoint of the segment.

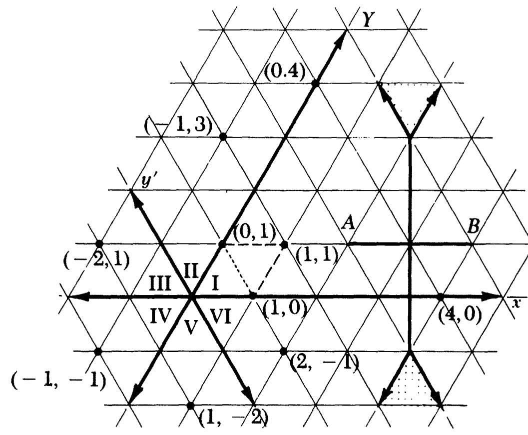


FIGURE 1a

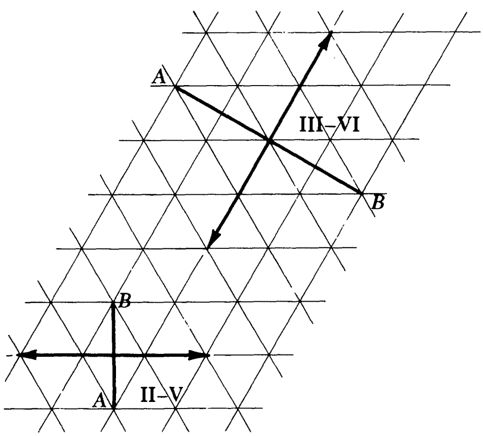


FIGURE 1b

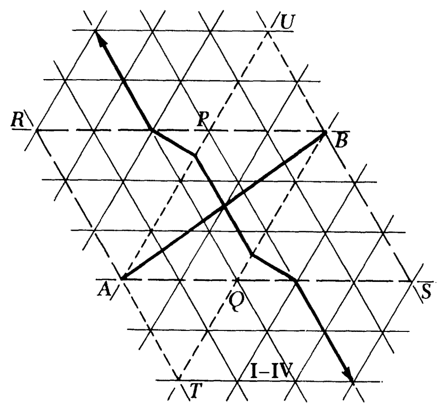


FIGURE 1c

As in Euclidean geometry and in square-taxi geometry, it can be shown that for any two points A and B

a) $d_I(A, B) \geqslant 0$;
 $d_I(A, B) = 0$ if and only if $A = B$;
b) $d_I(A, B) = d_I(B, A)$;
c) if X is any point in the plane,
 $d_I(A, X) + d_I(X, B) \geqslant d_I(A, B)$;
 if equality holds and A, X and B are distinct and collinear, then X is "between" A and B, and conversely.

In iso-taxi geometry two segments are defined to be congruent if and only if they have the same iso-taxi length. The transformations of the plane which preserve iso-taxi distances are reflections about any line such that the angle it makes with the horizontal is a multiple of 30°, all translations; and rotations about any point such that the angle of rotation is a multiple of 60°.

Because angle measure is not dependent upon the distance function, angles may be measured as they are in both Euclidean and in square-taxi geometry. Only Euclidean plane geometry has a rotation invariant metric; therefore one should expect properties involving angle measurement to be altered in taxicab geometries. Specifically, the side-angle-side axiom may be assumed in neither iso-taxi geometry nor in square-taxi geometry. Of course, any theorem which relies on the side-angle-side axiom is also invalid in each of these taxicab geometries. Among these are the angle-side-angle theorem, the angle-angle-side theorem and the side-side-side theorem. The reader may wish to find counterexamples in iso-taxi geometry for the axiom and for each of the theorems. Also, an isosceles triangle may no longer have its base angles congruent and triangles exist where the sum of the lengths of two of the sides equals the length of the third side.

Special geometric sets The *iso-taxi circle* is a hexagon. Because the circumference of the iso-taxi circle is six times the radius, $pi_I = 3$. The square-taxi circle is a square and $pi_S = 4$. It should be observed that each of these values for *pi* is the extreme value obtainable when an affinely regular hexagon is used (iso-taxi) or when a parallelogram is used (square-taxi); other values must fall between these extremes [**6**].

Using an equilateral triangle (whose sides are parallel to the axes and are of one unit in length) as a unit of area, we can find areas of figures in iso-taxi geometry by tessellating the region to be measured with these "equi-tri" units and counting the number required. The formula for the area of an iso-taxi circle is $A = 2 * pi_I * r^2$ "equi-tri" units, whereas the formula for the area of a square-taxi circle is $A = .5 * pi_S * r^2$ square units.

The set $\{X | d_I(A, X) + d_I(X, B) = d_I(A, B)\}$ might be called the *iso-taxi minimum distance set*. It consists of all points in the interior and on the parallelogram ($\square$ $AQBP$ in FIGURE 1c) determined by lines parallel to the axes drawn at the points A and B which has $\overline{AB}$ as its longer diagonal. The minimum distance set in square-taxi geometry consists of all points in the interior and on the rectangle having sides parallel to the axes and having the segment as a diagonal. In each case, if the two points are on a line parallel to an axis then the minimum distance set is a segment.

The set of points in a plane equidistant from two fixed points may be called a mid-set of the segment determined by the two fixed points. Points of the mid-set may be found by locating points of intersection of circles of equal radii having A and B as centers.

The *iso-taxi mid-set* $\{X | d_I(A, X) = d_I(X, B)\}$, as well as the square-taxi mid-set, may assume different shapes depending on the relative positions of the endpoints of the segment (FIGURE 1). The three cases are

a) The points A and B lie on a line parallel to an axis.
   i) In iso-taxi geometry the mid-set is a perpendicular segment and two regions (FIGURE 1a).
   ii) In square-taxi geometry the mid-set is the perpendicular bisector of the segment.
b) The points A and B lie on a line which forms an angle with the x-axis whose measure is
   i) 30° or 90° (iso-taxi). The mid-set is the perpendicular bisector of the segment (FIGURE 1b).
   ii) 45° (square-taxi). The mid-set is a perpendicular segment and two regions.
c) The points A and B lie on a line different from those of a) or b).
   i) In iso-taxi geometry the mid-set consists of three segments and two rays (FIGURE 1c). The "breaks" occur on the edges of the two parallelograms, $ASBR$ and $ATBU$, determined by lines parallel to the axes drawn at the points A and

B, which have $\overline{AB}$ as their shorter diagonal.

ii) In square-taxi geometry the mid-set consists of one segment and two rays. The "breaks" occur on the edges of the minimum distance set.

An ellipse is the set of points in a plane such that the sum of its distances from two fixed points (foci) is a constant. Points of the ellipse may be found by locating points of intersection of circles having A and B as centers such that the sum of the lengths of their radii equals the constant.

The *iso-taxi ellipse*, $\{X | d_I(A, X) + d_I(X, B) = k\}$ where $k > d_I(A, B)$, also assumes different shapes for each of the three cases enumerated above and for different values of the constant sum, k. For example, by considering sets of confocal ellipses, one observes (FIGURE 2) that

in case a) the iso-taxi ellipse is a decagon except when k equals twice the distance between the foci, then it becomes an octagon;

in case b), it is a dodecagon except when k equals one and one-half the distance between the foci; then it is an octagon;

in case c), it is a dodecagon except in two cases in which it becomes a decagon.

It is interesting to observe that the exceptional cases occur at points of intersection of lines parallel to the axes, drawn at each foci. Also when these lines are drawn, the plane is separated into regions and, in each case, segments of the confocal ellipses are parallel in each, or all but one, of these regions. In cases a) and b) there exist two lines of symmetry; in case c) there exists only point symmetry.

In square-taxi geometry ellipses are octagons in cases b) and c) but they are hexagons in case a). Different values of the constant do not affect the shape of the curves. If lines parallel to the axes are drawn at each vertex, parallel segments of the ellipse occur in each region in case a) and in all but one of the regions in the latter two cases. In each case there exist two lines of symmetry—one vertical, one horizontal.

A hyperbola is the set of points in a plane such that the difference of its distances from two fixed points is a constant. Points of the hyperbola may be found by locating points of intersection of circles having A and B as centers such that the difference of the lengths of their radii equals the constant.

The *iso-taxi hyperbola*, $\{X | \, |d_I(A, X) - d_I(X, B)| = k\}$ where $0 < k < d_I(A, B)$, also assumes different shapes for each of the three cases and for different values of k

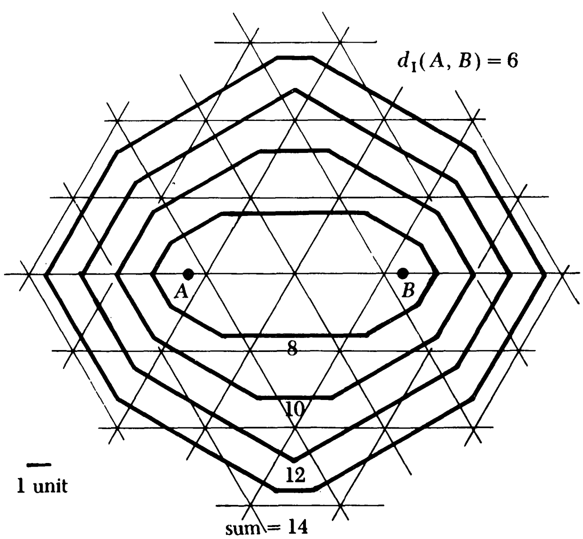


FIGURE 2a

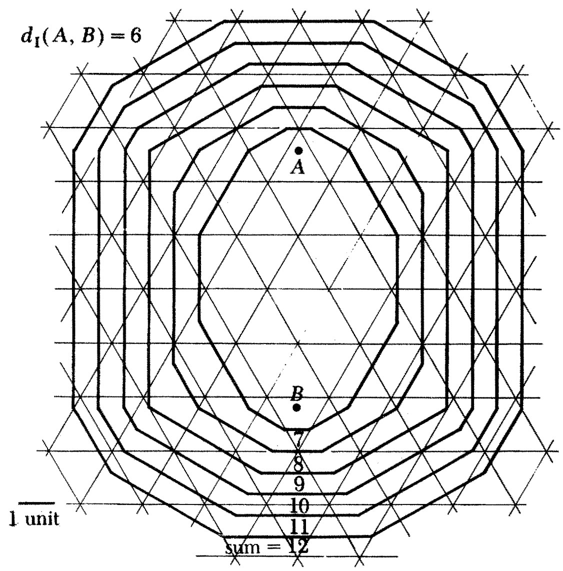


FIGURE 2b

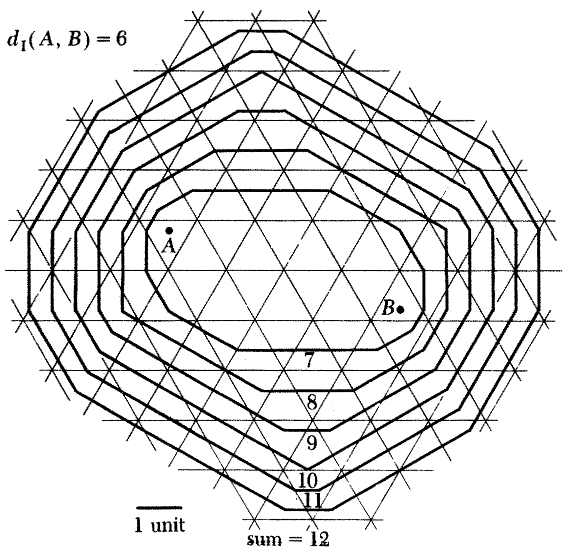


FIGURE 2c

in two of them. For example, by considering confocal hyperbolas, one observes (FIGURE 3) that

in case a) each wing of an iso-taxi hyperbola consists of a segment and two rays;

in case b), for $k < .5 * d_I(A, B)$ each wing consists of three segments and two collinear rays, for $k > .5 * d_I(A, B)$ each wing consists of three segments and two noncollinear rays, and for $k = .5 * d_I(A, B)$ each wing consists of three segments and two regions;

in case c) many different configurations occur—a wing of the iso-taxi hyperbola may consist of three segments and two parallel rays, or of a region, a segment and a ray, or of three segments and two nonparallel rays, or of a region, three segments and a ray.

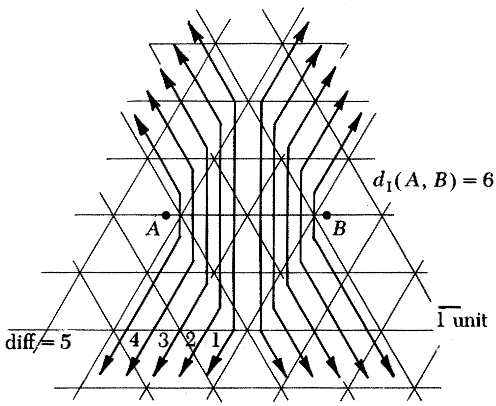


FIGURE 3a

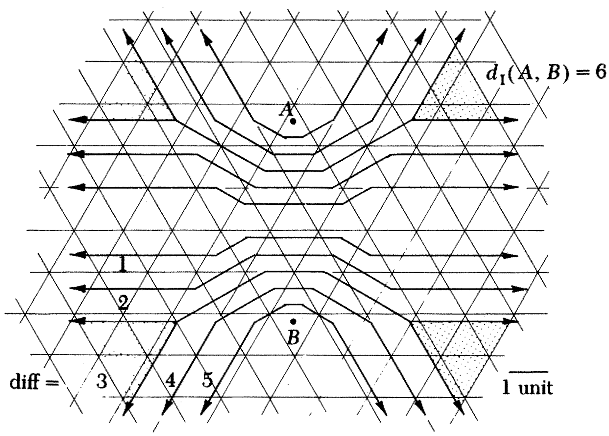


FIGURE 3b

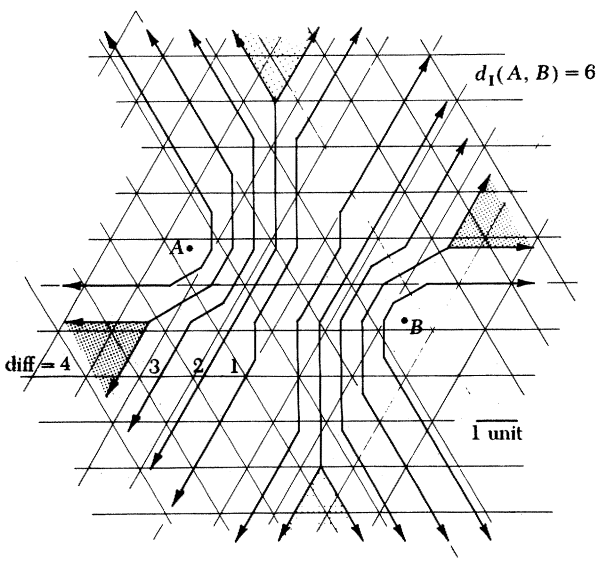


FIGURE 3c

Again, if at each focus lines parallel to the three axes are drawn then parallelism of segments or of rays is observed in all but six of the regions in case a) and in all but two of the regions in cases b) and c)—or the entire region becomes a part of the graph of an iso-taxi hyperbola. Two lines of symmetry occur in the first two cases; a point of symmetry occurs in the last.

For a square-taxi hyperbola one observes that

in case a) each wing is a straight line;

in case b) each wing consists of a segment and two rays;

in case c) each wing may consist of a segment and two parallel rays, or a region, a segment and a ray, or a segment and two nonparallel rays.

If lines parallel to the axes are drawn at each vertex, parallel segments occur, parallel rays occur, or a region of the hyperbola occurs in all but two of the regions in case c) and in all but four of the regions in the former two cases. Symmetry patterns for the three cases are the same as those of their iso-taxi counterparts.

The distance from a point to a line is the length of the radius of a circle, having the point as its center, which is tangent to the line. Thus, in iso-taxi geometry (Figure 4), or in square-taxi geometry, this distance is the length of the shortest segment to the line in a direction parallel to an axis.

The set of points in a plane equidistant from a line consists of two lines parallel to the given line. The use of parallel lines is important in locating points of parabolas and points of angle mid-sets.

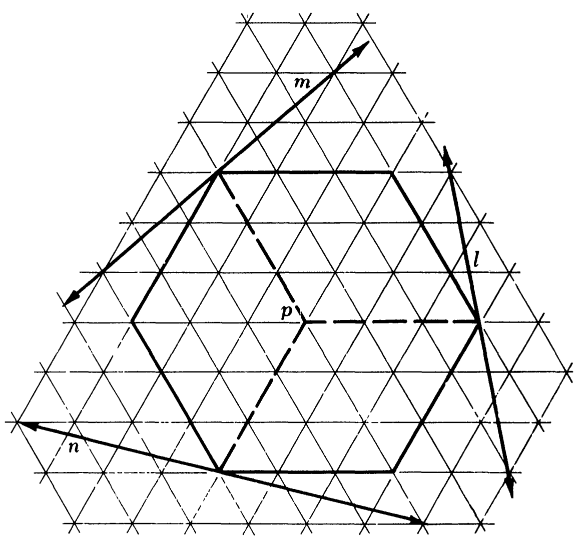


FIGURE 4

The set of points in a plane equidistant from a fixed point (focus) and a fixed line (directrix) is a parabola. Points of a parabola may be found by locating points of intersections of circles having their center at the focus and radius k, and lines parallel to the directrix and k units distant from it, $k \geqslant .5 * d$ where d is the distance from the focus to the directrix. The set where $k = .5 * d$ is called the vertex of the parabola.

The *iso-taxi parabola*, $\{X | d_I(F, X) = d_I(X, \overleftrightarrow{AB})\}$ where F is the focus and the directrix is determined by the points A and B, assumes different shapes depending on the relative positions of the points A and B. Families of parabolas having the same directrix and different foci are considered in each of the three cases (FIGURE 5). Each iso-taxi parabola in case a) consists of three segments and two nonparallel rays; the vertex, in this case, is one of these segments. In cases b) and c) the vertex is a point and the parabola consists of four segments and two parallel rays. In the first two cases there is a symmetry about the line containing the focus that is perpendicular to the directrix. In case c) all symmetry is lost. The "break" points of the parabola in each case may be observed to lie on lines parallel to the axes drawn at the focus.

In square-taxi geometry a parabola consists of a segment (the vertex) and two nonparallel rays in case b) whereas in the other two cases it consists of two segments and parallel rays, and the vertex is a point. Symmetry and "break" points occur as they do for iso-taxi parabolas.

The set of points in a plane equidistant from two noncollinear rays having a common endpoint may be called the angle mid-set. Points of the angle mid-set may be found by locating points of intersection of lines equidistant from the lines containing

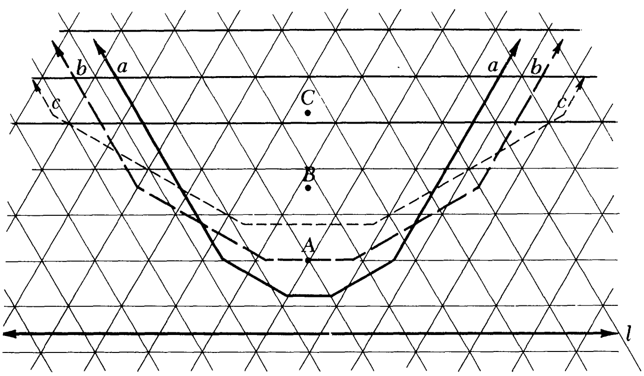


FIGURE 5a

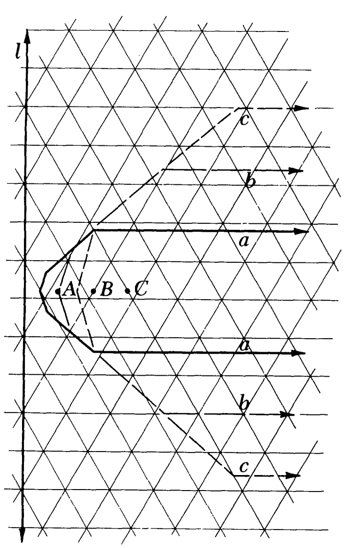


FIGURE 5b

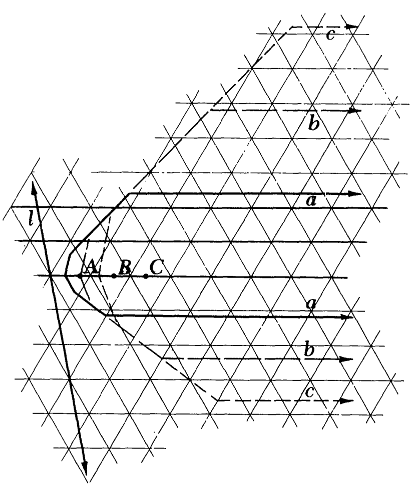


FIGURE 5c

the rays of the angle; these points must also lie in the interior of the angle. The angle mid-set is, therefore, a half-line. Points of the angle mid-set are also the centers of circles which are tangent to each of the two rays.

The *iso-taxi angle mid-set*, $\{X \mid d_I(X,\overrightarrow{BA}) = d_I(X,\overrightarrow{BC})\}$, is the same as the Euclidean angle bisector only for especially placed angles (for example, angles whose rays are symmetric to a line forming an angle with the x-axis such that its measure is a multiple of 30°). The iso-taxi angle mid-set is unique for any given angle; however, if the angle is especially placed then the rays may be tangent to the iso-taxi circles only at the vertex of the angle (for example, if the angle has a ray parallel to an axis and the measure of the angle is greater than or equal to 120°). Consequently, different angles having a common ray may have the same iso-taxi angle mid-set. In FIGURE 6 each of the angles $\angle ABC$, $\angle ABD$ and $\angle ABE$ has $\overrightarrow{BF}$ as its angle mid-set. Also iso-taxi angle mid-sets of two pairs of vertical angles may not be perpendicular.

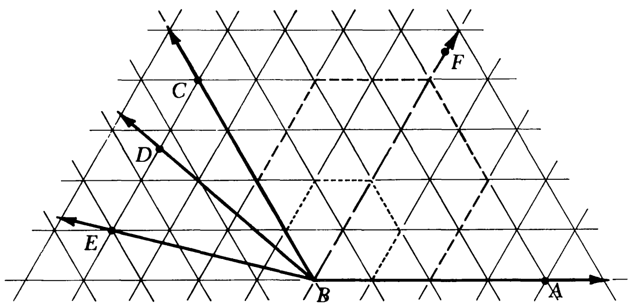


FIGURE 6

The same conclusions are true for square-taxi angle mid-sets except in the first case, the measure of the angle is a multiple of 45°, and in the second case, the angle has a ray which is contained in a line which forms an angle with the x-axis whose measure is 45° and the angle, itself, has measure greater than or equal to 90°.

Because of the uniqueness of the iso-taxi angle mid-set, the incenter of a triangle and its corresponding inscribed circle are also unique. However, because of the observations noted above, the inscribed circle may be tangent to a side of the triangle at a vertex of the triangle (FIGURE 7). Similar results arise in square-taxi geometry.

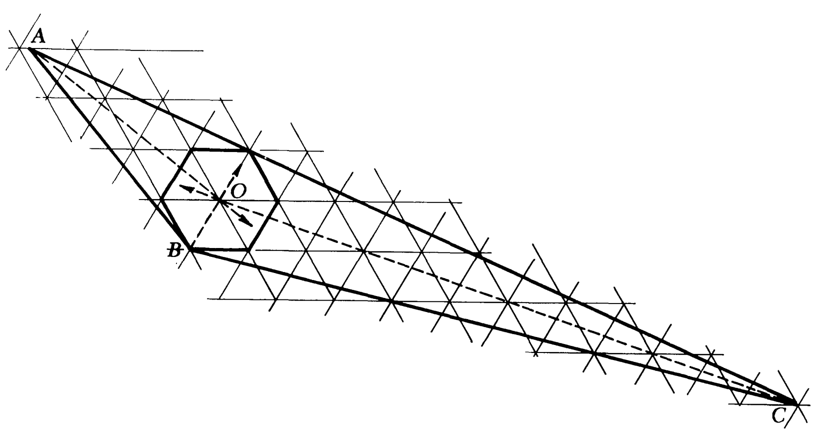


FIGURE 7

The circumcenter of a triangle and its corresponding circumcircle may not be unique in taxicab geometry. In fact the set of circumcenters may consist of a point, a segment, a segment and a ray, or a ray. In iso-taxi geometry a segment occurs if one side is parallel to an axis and the other sides are congruent and more than twice the length of the first side (FIGURE 8a); a segment and a ray occur if two sides are parallel to the axes and the length of the third side equals the sum of the lengths of the other two sides (FIGURE 8b); and a ray occurs if two congruent sides are parallel to the axes and the length of the third side equals the sum of the lengths of the other two sides (FIGURE 8c). In square-taxi geometry a segment occurs if exactly one side is a segment of a line making an angle with the x-axis whose measure is 45° and the other two sides are congruent and greater in length than the first side; a segment and a ray occur if two sides are segments of such 45° lines and the triangle is isosceles but not equilateral; and a ray occurs if at least one side is a segment of a 45° line and another side is parallel to an axis.

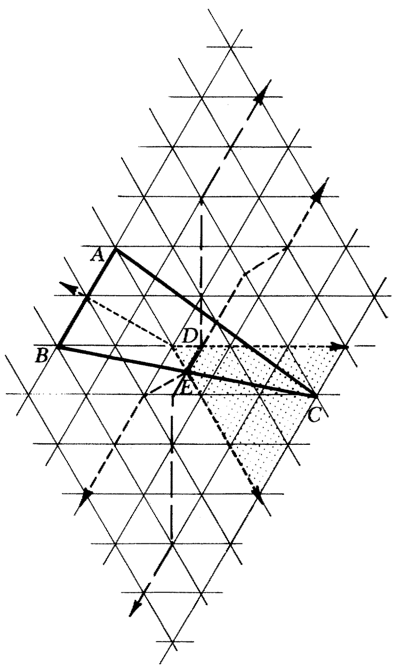


FIGURE 8a

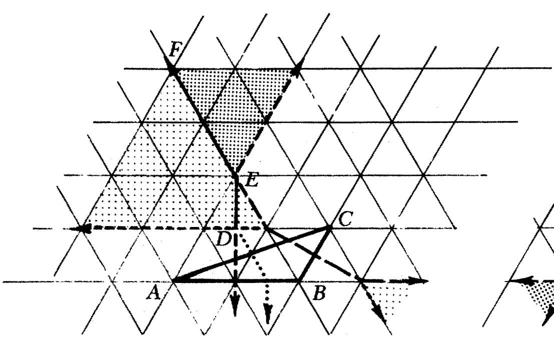


FIGURE 8b

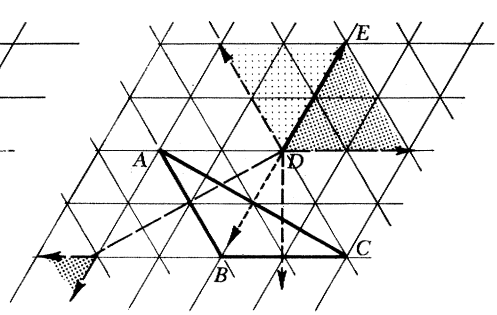


FIGURE 8c

Other taxicab geometric sets might also be explored and additional theorems established as valid or invalid for this geometry. Surely changing the distance function and changing the coordinate grid from a square configuration to a triangular configuration leads to geometries that behave very differently from Euclidean geometry.

The author expresses her appreciation to Mr. Ron Spangler, ECU, for his assistance in the preparation of the illustrations.

REFERENCES

1. Martin Gardner, Mathematical games, *Scientific American* 243 (Nov. 1980), 18–30.
2. Eugene F. Krause, Taxicab geometry, *Mathematics Teacher* 66 (1973), 695–706.
3. ——, *Taxicab Geometry*, Addison-Wesley, Menlo Park, NJ, 1975.
4. Richard Laatsch, Pyramidal sections in taxicab geometry, this MAGAZINE 55 (1982), 205–212.
5. Barbara E. Reynolds, Taxicab geometry, *Pi Mu Epsilon J.* 7 (1980), 77–88.
6. J. J. Schaffer, *Geometry of Spheres in Normed Spaces*, Marcel Dekker, Inc., New York, 1976, pp. 17–18.
7. Doris J. Schattschneider, The taxicab group, *Amer. Math. Monthly* 91 (1984), 423–428.

The Axis of a Rotation: Analysis, Algebra, Geometry

DAN KALMAN
5332 Ironwood St.
Rancho Palos Verdes, CA 90274

Introduction Recently, while working on some problems related to coordinate transformations, I happened on the following discovery:

If the 3 by 3 matrix A represents a rotation (i.e., A is orthogonal with determinant 1), and the trace of A is $\mathrm{tr}(A)$, then for any vector $\mathbf{x}$

$$A\mathbf{x} + A^T\mathbf{x} + [1 - \mathrm{tr}(A)]\mathbf{x}$$

lies on the axis of the rotation.

I recognized immediately that this result must be well known (in certain circles). However, it seems to me that my route of discovery illustrates some important principles of problem solving and mathematical discovery. I present this account with the idea in mind that a suitable modification might be presented in a linear algebra course. In addition to serving as a case study in discovery, the topic is a natural application of eigenvalues and eigenvectors, and the result has an attractive simplicity. As suggested by the title, analysis, algebra, and geometry each play a role in the development to follow.

Background Before proceeding further, it will be useful to review some facts about rotation matrices and to establish the notation and nomenclature to be used. A

Other taxicab geometric sets might also be explored and additional theorems established as valid or invalid for this geometry. Surely changing the distance function and changing the coordinate grid from a square configuration to a triangular configuration leads to geometries that behave very differently from Euclidean geometry.

The author expresses her appreciation to Mr. Ron Spangler, ECU, for his assistance in the preparation of the illustrations.

REFERENCES

1. Martin Gardner, Mathematical games, *Scientific American* 243 (Nov. 1980), 18–30.
2. Eugene F. Krause, Taxicab geometry, *Mathematics Teacher* 66 (1973), 695–706.
3. ——, *Taxicab Geometry*, Addison-Wesley, Menlo Park, NJ, 1975.
4. Richard Laatsch, Pyramidal sections in taxicab geometry, this MAGAZINE 55 (1982), 205–212.
5. Barbara E. Reynolds, Taxicab geometry, *Pi Mu Epsilon J.* 7 (1980), 77–88.
6. J. J. Schaffer, *Geometry of Spheres in Normed Spaces*, Marcel Dekker, Inc., New York, 1976, pp. 17–18.
7. Doris J. Schattschneider, The taxicab group, *Amer. Math. Monthly* 91 (1984), 423–428.

The Axis of a Rotation: Analysis, Algebra, Geometry

DAN KALMAN
5332 Ironwood St.
Rancho Palos Verdes, CA 90274

Introduction Recently, while working on some problems related to coordinate transformations, I happened on the following discovery:

> If the 3 by 3 matrix A represents a rotation (i.e., A is orthogonal with determinant 1), and the trace of A is $\mathrm{tr}(A)$, then for any vector $\mathbf{x}$
>
> $$A\mathbf{x}+A^T\mathbf{x}+[1-\mathrm{tr}(A)]\mathbf{x}$$
>
> lies on the axis of the rotation.

I recognized immediately that this result must be well known (in certain circles). However, it seems to me that my route of discovery illustrates some important principles of problem solving and mathematical discovery. I present this account with the idea in mind that a suitable modification might be presented in a linear algebra course. In addition to serving as a case study in discovery, the topic is a natural application of eigenvalues and eigenvectors, and the result has an attractive simplicity. As suggested by the title, analysis, algebra, and geometry each play a role in the development to follow.

Background Before proceeding further, it will be useful to review some facts about rotation matrices and to establish the notation and nomenclature to be used. A *rotation* is a rigid transformation of real 3 dimensional space leaving the origin fixed. Such a transformation is necessarily linear, and is represented with respect to the standard basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$ by a *rotation matrix*, A. The columns of A are the images of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ under the rigid motion, and so comprise a right-handed triple of orthogonal unit vectors. Therefore, $A^TA = I$ and $\det(A) = 1$.

In general, a square matrix A that satisfies $A^TA = I$ is called orthogonal. Using the matrix product notation $\mathbf{x}^T\mathbf{y}$ and the inner product notation $\mathbf{x} \cdot \mathbf{y}$ interchangeably (for vectors $\mathbf{x}$ and $\mathbf{y}$), observe that $A\mathbf{x} \cdot A\mathbf{y} = (A\mathbf{x})^TA\mathbf{y} = \mathbf{x}^TA^TA\mathbf{y} = \mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$. Thus, orthogonal matrices preserve inner products, and in particular, preserve angles and lengths.

For the special case of a 3×3 orthogonal matrix with unit determinant, it can be shown that 1 is an eigenvalue, and that the corresponding eigenspace is one dimensional, as follows. Since the matrix preserves lengths, its eigenvalues must all have magnitude 1. The product of the eigenvalues is the determinant, and must therefore equal 1, as well. Thus, aside from the trivial case of the identity matrix, there must be a unique eigenvalue equal to 1, and a pair of complex conjugates $r = \cos\theta + i\sin\theta$ and $s = \cos\theta - i\sin\theta$. As an unrepeated eigenvalue, 1 has a one-dimensional eigenspace, as asserted.

The eigenspace for the eigenvalue 1 is a line of fixed points for the transformation. It will now be shown that the transformation is geometrically a rotation about this fixed line. Every vector may be resolved into orthogonal components parallel to the fixed line and in the plane perpendicular to the fixed line. By virtue of linearity, it suffices to show that the transformation acts as a rotation on the perpendicular plane. Accordingly, consider the special case of a vector perpendicular to the fixed line. Its image has equal length, and must also be perpendicular to the fixed line. Therefore, it is possible to rotate the vector about the fixed line to obtain the image vector. Moreover, preservation of the angle between vectors implies that any two vectors in the perpendicular plane must be rotated by the same amount. Thus, the transformation acts on the perpendicular plane as a rotation about the fixed line, as desired.

To summarize the preceding paragraphs, a 3×3 matrix represents a rotation if and only if it is orthogonal with unit determinant. A nontrivial rotation matrix possesses a one-dimensional eigenspace corresponding to the eigenvalue 1, and this eigenspace is, in fact, the axis of the rotation. With this background established, the discussion proceeds to the main topic of the paper.

Analysis Permit me to set the stage. I was interested in developing a computer program to generate the rotation matrix linking two right-handed coordinate systems in $\mathbf{R}^3$, given some information about their relative orientations. As a side topic, I wished to find the axis of the rotation, that is, to find one eigenvector corresponding to the eigenvalue 1. Let the rotation matrix A have entries a_{ij}. A solution of $(A - I)\mathbf{x} = 0$ must be orthogonal to the first two rows of $A - I$. Denoting the mth row of A by (row m), the first two rows of $A - I$ are (row 1) $- \mathbf{i}$ and (row 2) $- \mathbf{j}$. A vector orthogonal to both is obtained by taking the vector product $\mathbf{c}$. This results in

$$\mathbf{c} = (\text{row } 1) \times (\text{row } 2) - \mathbf{i} \times (\text{row } 2) + \mathbf{j} \times (\text{row } 1) + \mathbf{i} \times j.$$

Now, since A is a rotation matrix, its columns form a right handed triple. The same may be said of A^T, so the rows of A also form a right handed triple. In particular, (row 1) $\times$ (row 2) = (row 3). Applying this result and simplifying the remaining three cross products leads to

$$\mathbf{c} = \begin{bmatrix} a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} - \begin{bmatrix} 0 \\ -a_{23} \\ a_{22} \end{bmatrix} + \begin{bmatrix} a_{13} \\ 0 \\ -a_{11} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{31} + a_{13} \\ a_{32} + a_{23} \\ 1 + a_{33} - a_{22} - a_{11} \end{bmatrix}.$$

Of course, the vector $\mathbf{c}$ might equal zero (if the first two rows of $A - I$ are dependent) but in this case a cross product of a different pair can be calculated. This provides enough information for a computer program, and concludes the analysis phase of discovery.

Algebra The formula for $\mathbf{c}$ derived above has too much symmetry to be left alone. Among possible rearrangements, the following pleases the eye:

$$\mathbf{c} = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} + \begin{bmatrix} a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 - a_{11} - a_{22} - a_{33} \end{bmatrix}$$

$$= \text{row } 3 + \text{column } 3 + [1 - \operatorname{tr}(A)]\mathbf{k}.$$

Next, observe that column 3 is just $A\mathbf{k}$, and similarly, row 3 is $A^T\mathbf{k}$. Thus, we have

$$\mathbf{c} = A\mathbf{k} + A^T\mathbf{k} + [1 - \operatorname{tr}(A)]\mathbf{k}$$

$$= (A + A^T + [1 - \operatorname{tr}(A)]I)\mathbf{k}.$$

Is there any reason for the vector $\mathbf{k}$ to be singled out in this fashion? Surely a similar formula involving $\mathbf{i}$ or $\mathbf{j}$ must exist. It is even tempting to believe that replacing $\mathbf{k}$ with *any* vector produces a vector $\mathbf{c}$ in the eigenspace corresponding to the eigenvalue 1. How might such an assertion be proved?

The conjecture is this: for any vector $\mathbf{v}$, $(A + A^T + [1 - \operatorname{tr}(A)]I)\mathbf{v}$ is an eigenvector with eigenvalue 1. That is,

$$A(A + A^T + [1 - \operatorname{tr}(A)]I)\mathbf{v} = (A + A^T + [1 - \operatorname{tr}(A)]I)\mathbf{v}.$$

To establish this for all $\mathbf{v}$ requires showing that the matrices multiplying $\mathbf{v}$ on each side of the equation are equal. Rearranging the necessary identity yields

$$A^2 + I + [1 - \operatorname{tr}(A)]A = A + A^T + [1 - \operatorname{tr}(A)]I$$

and hence

$$A^2 - \operatorname{tr}(A)A + \operatorname{tr}(A)I - A^T = 0.$$

Finally, since A is nonsingular, we may multiply both sides by A to obtain

$$A^3 - \operatorname{tr}(A)A^2 + \operatorname{tr}(A)A - I = 0.$$

Thus, the conjecture at hand is equivalent to a certain polynomial identity for A. This immediately suggests consideration of the characteristic polynomial of A.

Let $p(x)$ be the characteristic polynomial of A. As mentioned earlier, p has roots 1, $r = \cos\theta + i\sin\theta$, and $s = \cos\theta - i\sin\theta$. Moreover, $rs = 1$ and $r + s = 2\cos\theta$. Then, the factored form $p(x) = (x-1)(x-r)(x-s)$ may be multiplied out to give

$$\begin{aligned} p(x) &= x^3 - (1+r+s)x^2 + (1+r+s)x - 1 \\ &= x^3 - (1+2\cos\theta)x^2 + (1+2\cos\theta)x - 1 \\ &= x^3 - \mathrm{tr}(A)x^2 + \mathrm{tr}(A)x - 1. \end{aligned}$$

Now we note that $p(A) = 0$, and the desired identity is established. What was discovered through analysis has been more generally supported by algebra.

Geometry In this section a geometric explanation will be presented for the result established algebraically above. Paraphrased, this result states that for any vector v, $A\mathrm{v} + A^T\mathrm{v} + [1 - \mathrm{tr}(A)]\mathrm{v}$ lies on the axis of rotation A. Assume that A represents a rotation of space through an angle ϕ about a fixed axis. (Note here that no connection has been established between ϕ and θ at this point.) To simplify notation, identify vectors with points in space in the usual way, and perform vector operations on points accordingly. Thus, given a point R, we apply the rotation A to find $S = A(R)$ and the inverse rotation A^T to find $T = A^T(R)$. The points R, S, and T all lie on a cone whose axis is the axis of rotation, and with vertex at the origin. This situation is illustrated in FIGURE 1 as a perspective view, and in FIGURES 2 and 3 as top and side views, respectively.

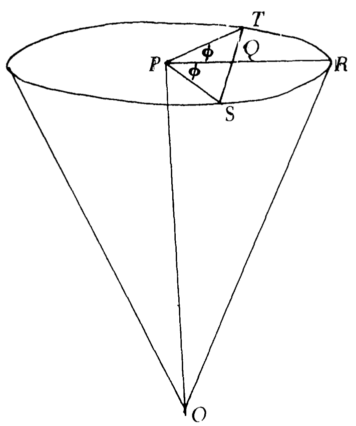


FIGURE 1

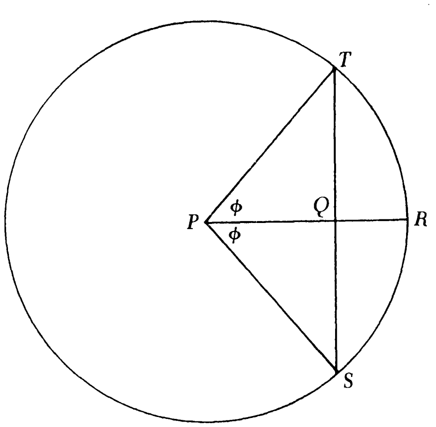


FIGURE 2

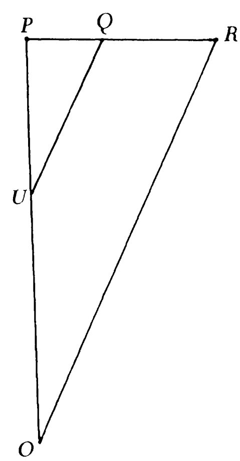


FIGURE 3

Let $Q = .5(S + T)$, the midpoint of segment ST. Clearly, some vector parallel to R can be drawn at Q so that it terminates at a point U on the axis of rotation. In fact, this vector must have length UQ and be parallel to unit vector $-R/(OR)$, hence it is given by $-(UQ/OR)R$. From FIGURE 3, the ratio UQ/OR is equal to PQ/PR. Now the vector from point Q to point U is given by the vector difference $U - Q$. Thus, with $f = PQ/PR$, we have $U - Q = -fR$ or $U = Q - fR$. To relate f to the angle ϕ, observe in FIGURE 2 that $PR = PS$ giving $f = PQ/PS = \cos\phi$. Combining these results produces $U = .5(S + T) - \cos\phi R$ as a point on the axis of rotation. Furthermore, $2U$ is also on the axis of rotation, and is given by $2U = S + T - 2\cos\phi R = (A + A^T - 2\cos\phi)R$. It remains but to show that $-2\cos\phi = 1 - \mathrm{tr}(A)$ and the geometric construction will reestablish the result of the preceding section.

Since $\mathrm{tr}(A)$ is invariant under similarity transformations, we may choose to represent the rotation relative to an orthonormal basis in which the third element lies on the axis of rotation. The corresponding matrix is easily seen to be

$$\begin{bmatrix} \cos\phi & \pm\sin\phi & 0 \\ \mp\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where the ambiguous signs depend on the direction of the rotation. Regardless, the trace is evidently $1 + 2\cos\phi$, as required.

Classroom presentations In describing this material to you, the reader, quite a bit of background has been presented or assumed: characterization of rotations as unit determinant orthogonal matrices, the Cayley-Hamilton theorem, invariance of $\mathrm{tr}(A)$ under changes of basis, etc. Depending on the background knowledge of the students involved, some modifications may be required for classroom presentation. One possible approach is to assume from the outset that A is a geometric rotation about a fixed axis. The existence of a unique one-dimensional eigenspace is then evident. If need be, the algebraic part of the discussion can be omitted in favor of passing directly from the cross product argument to the geometric construction. Indeed, one may even leave the connection between $\cos\phi$ and $\mathrm{tr}(A)$ unproved and use the geometric discussion as a plausibility argument. The most general version of the result could then be established for vectors $\mathbf{i}$ and $\mathbf{j}$ by using cross products, and extended to all vectors by linearity. This approach can also be assigned as an exercise. At the other extreme, with sufficient background, the students should be able to follow the development presented here. For these students especially, this topic provides a simple example of the interplay between various approaches to a problem, and illustrates one way that mathematical discoveries are propagated.

An Iterative Method for Approximating Square Roots

KENNETH R. JOHNSON
North Dakota State University
Fargo, ND 58105

Introduction Mathematicians have been approximating the square roots of non-square integers with rational numbers for thousands of years. Many ingenious schemes have been used to generate these approximations and a good deal of interesting number theory has been discovered in the process. After reviewing some of these historical attempts and some of the number theory involved, we present an iterative procedure for approximating square roots which is based on an observation of M. A. Grant [3].

Some early approximations The Babylonians may have used the approximation formula

$$(a^2+h)^{1/2} \approx a + \frac{h}{2a}, \qquad 0 < h < a^2 \quad [2, \text{p. } 33].$$

The fact that this is a reasonable approximation is easily seen when we use the first two terms the binomial series

$$(1+x)^{1/2} = 1 + \sum_{n=1}^{\infty} \frac{1/2(1/2-1)\cdots(1/2-n+1)}{n!} x^n, \qquad |x| < 1,$$

with $x = h/a^2$, which gives us

$$\left(1 + \frac{h}{a^2}\right)^{1/2} \approx 1 + \frac{1}{2}\frac{h}{a^2}.$$

Multiplying both sides of this equation by a gives us the Babylonian approximation. Taking $a = 4$ and $h = 1$, we see that $33/8$ is a Babylonian approximation of $\sqrt{17}$.

Heron of Alexandria (perhaps A.D. 50–100, [2, p. 157]) took an approximation a to $\sqrt{d}$ and then improved it by computing

$$\frac{a + \frac{d}{a}}{2}.$$

Notice that both a and d/a are approximations to $\sqrt{d}$. Since one of a or d/a is larger than $\sqrt{d}$ while the other is smaller than $\sqrt{d}$, the average of the two will be a better estimate. For example, with $a = 4$, the next approximation, to $\sqrt{17}$ is $33/8$. This method is easily iterated and is used in some computers even today.

The Renaissance algebraist Rafael Bombelli is generally credited as the first to study continued fractions and the first mathematician to employ continued fractions to approximate $\sqrt{d}$.

We say that a rational number p/q is a "good approximation" of an irrational number ξ if

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{q^2}.$$

For irrational numbers there are infinitely many such pairs of rational numbers. The set of all such good rational approximations for a given irrational number ξ is intimately related to the continued fraction expansion of ξ. In the next section we will develop the interplay between the "good" approximations of ξ and the continued fraction expansion of ξ.

More recently, M. A. Grant [3] demonstrated a method of approximating $\sqrt{d}$ by evaluating a sequence of rational functions at a certain rational point. His method is computationally simple and leads to very good approximations of $\sqrt{d}$. These approximations, in some cases, are good enough that one wonders if there is a connection between Grant's approximations and continued fractions. Happily, it turns out that an interesting connection exists between Grant's approximations and the theory of continued fractions. By exploiting this connection, we find an algorithm for producing approximations that is iterative and produces exceedingly good approximations to $\sqrt{d}$.

Continued fractions By a finite continued fraction, we mean a fraction of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}}.$$

Since this notation is cumbersome, this expression is denoted by $[a_0; a_1, \dots, a_n]$. If all of the numbers $a_0, \dots, a_n$ are integers, the continued fraction is called simple. Clearly a finite simple continued fraction is always rational. By the infinite simple continued fraction $[a_0; a_1, a_2, \dots]$ we mean the limit of the sequence of rational numbers $a_0, [a_0; a_1], [a_0; a_1, a_2], \dots$. That is,

$$[a_0; a_1, \dots] = \lim_{n \to \infty} [a_0; a_1, \dots, a_n].$$

The fraction $[a_0; a_1, \dots, a_j]$ is called the jth convergent to $[a_0; a_1, \dots]$ and is denoted by c_j. When we write $[a_0; a_1, \dots, \overline{a_{k+1}, \dots, a_{k+n}}]$, the line over $a_{k+1}, \dots, a_{k+n}$ indicates that this block of integers is repeated over and over. Such an infinite continued fraction is called periodic, and we say that the period of the continued fraction has length n. To illustrate all of this notation, let us look at $[1; \overline{1, 2}]$, which represents the number

$$\alpha = 1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2}}}} \tag{1}$$
$$\ddots$$

What is the value of α? If we look carefully at (1), we see that

$$\alpha = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \left(1 + \cfrac{1}{2 + \cdots}\right)}}$$
$$= 1 + \cfrac{1}{1 + \cfrac{1}{1 + \alpha}}.$$

After a little bit of algebra, one obtains $\alpha^2 = 3$ and since α is positive, $\alpha = \sqrt{3}$. Thus $\sqrt{3}$ is represented by the periodic continued fraction $[1; \overline{1,2}]$. The second convergent to $\sqrt{3}$ is

$$c_2 = [1;1,2] = 1 + \frac{1}{1 + 1/2} = \frac{5}{3}$$

while the third convergent is

$$c_3 = [1;1,2,1] = \frac{7}{4}.$$

As it turns out, all infinite simple continued fractions are irrational numbers. Also, every periodic simple continued fraction is an irrational root of a quadratic equation and conversely, every such irrational root has an infinite periodic simple continued fraction expansion.

In fact it is easy to find the continued fraction expansion of any irrational number. Let [ ] represent the greatest integer function. Then if x_0 is irrational, successively compute

$$x_1 = \frac{1}{x_0 - [x_0]}$$

$$\vdots$$

$$x_{j+1} = \frac{1}{x_j - [x_j]}, \quad \text{and}$$

take $a_n = [x_n]$. Then $x_0 = [a_0; a_1, \ldots]$.

For example using this algorithm, we can compute the first three convergents to $\sqrt{17}$. We have $a_0 = 4$, $a_1 = 8$ and $a_2 = 8$ and so the first three convergents to $\sqrt{17}$ are

$$\frac{4}{1}, \frac{33}{8}, \quad \text{and} \quad \frac{268}{65}.$$

Two classical theorems [e.g. 4, theorems 171 and 184] relate irrational numbers to the convergents of continued fractions. One theorem tells us that if ξ is an irrational number, then the convergents to ξ are good approximations to ξ ($|\xi - p/q| < 1/q^2$). While the other theorem says that if p/q is a sufficiently good approximation to ξ ($|\xi - p/q| < 1/2q^2$) then p/q is a convergent to ξ.

Pells equation The connection between Grant's approximations and continued fractions involves "Pell's equation." Given a positive integer d, Pell's equation is satisfied by finding integers x and y such that

$$x^2 - dy^2 = 1. \tag{2}$$

If we take $x = \pm 1$ and $y = 0$ we get a trivial solution of this equation regardless of the value of d. If d is the square of an integer, say $d = n^2$, then

$$1 = x^2 - dy^2 = (x + ny)(x - ny),$$

which means that both $x + ny$ and $x - ny$ have to be ± 1 and since

$$x = \frac{(x + ny) + (x - ny)}{2},$$

we see that $x = \pm 1$ so $y = 0$ and so this is again the trivial solution of (2). Thus the only interesting cases of Pell's equation are when d is a nonsquare integer. Henceforth we restrict our attention to the case when d is a nonsquare integer.

Fermat was the first to state that for such d, the Pell equation has an infinite number of solutions but he did not provide a proof. This assertion was probably quite surprising to mathematicians of the time since there are cases with small d when the smallest solutions (x, y) are quite large. For example when $d = 13$, the smallest solution (x, y) is the pair (649, 180) while for $d = 29$ the smallest solution is $(x, y) = (9801, 1820)$.

We will require the following lemmas which relate the study of continued fractions and Pell's equation. These lemmas may be found in any of several standard number theory texts [e.g., 1]. Remember that in each of these lemmas, d is a nonsquare positive integer.

LEMMA 1. *If (p, q) is a positive solution to $p^2 - dq^2 = 1$, then p/q is a convergent of the continued fraction expansion of $\sqrt{d}$.*

LEMMA 2. *Let p_k/q_k be the convergents of the continued fraction expansions of $\sqrt{d}$, and let n be the period of the continued fraction expansion for $\sqrt{d}$.*

1) *If n is even, then all positive solutions of $x^2 - dy^2 = 1$ are given by $x = p_{kn-1}$ and $y = q_{kn-1}$ $(k = 1, 2, 3, \ldots)$.*

2) *If n is odd, then all positive solutions of $x^2 - dy^2 = 1$ are given by $x = p_{2kn-1}$ and $y = q_{2kn-1}$ $(k = 1, 2, 3, \ldots)$.*

LEMMA 3. *If (x_1, y_1) is the smallest positive solution of $x^2 - dy^2 = 1$, then every positive solution of the equation is given by (x_j, y_j), where x_j and y_j are the integers defined by $x_j + y_j\sqrt{d} = (x_1 + y_1\sqrt{d})^j$.*

Before considering examples illustrating these ideas, we observe that the last two lemmas provide different descriptions of the solutions to Pell's equation. Since the solution sequences are both increasing and exhaustive, these two descriptions are, in fact, the same. Consequently, using the notation of Lemmas 2 and 3, we have

COROLLARY 1. *If n is even, then $x_k = p_{nk-1}$ and $y_k = q_{nk-1}$. If n is odd, then $x_k = p_{2nk-1}$ and $y_k = q_{2nk-1}$.*

Let us illustrate all of this with a pair of examples. Using the continued fraction algorithm described earlier, one can quickly find that $\sqrt{7} = [2; \overline{1, 1, 1, 4}]$ and $\sqrt{17} = [4; \overline{8}]$. Lemma 1 tells us that since $8^2 - 7 \cdot 3^2 = 1$, $8/3$ is a convergent to $\sqrt{7}$ and indeed $8/3 = [2; 1, 1, 1] = c_3$. Since the period in the partial fraction expansion of $\sqrt{7}$ is 4, Lemma 2 (part one) tells us that the convergents of $\sqrt{7}$ whose numerators and denominators give solution to Pells equation are $c_3, c_7, c_{11}, \ldots$. Similarly Lemma 2 (part 2) says that the convergents of $\sqrt{17}$ which give rise to solutions of Pell's equation are $c_1, c_3, c_5, \ldots$ since the period of $[4; \overline{8}]$ is one. Finally, Lemma 3 tells us that all of the convergents of $\sqrt{7}$ which give solutions to Pell's equation are defined by collecting terms in

$$x_j + y_j\sqrt{d} = (8 + 3\sqrt{7})^j.$$

And so the second smallest solution to the Pellian equation $x^2 - 7y^2 = 1$, which we already know to be c_7 by Lemma 2, is found by collecting terms of $(8 + 3\sqrt{7})^2$, thus

$$x_2 + y_2\sqrt{7} = (127 + 48\sqrt{7}),$$

i.e. $x_2 = 127$ and $y_2 = 48$ so we know that $c_7 = 127/48$. A quick calculation verifies this. Similarly, the convergent c_3 to $\sqrt{17}$ can be found by collecting the terms of $(33 + 8\sqrt{17})^2$ since $c_1 = 33/8$ gives the smallest positive solution of Pell's equation.

Approximation to $\sqrt{d}$ Grant makes a first approximation $x = p/q$ to $\sqrt{d}$ under the restriction $[\sqrt{d}] < x < [\sqrt{d}] + 1$, where [ ] again denotes the greatest integer function. With this choice of x, $0 < (x - \sqrt{d}) < 1$, and so $(x - \sqrt{d})^n \to 0$ as $n \to \infty$. If we expand $(x - \sqrt{d})^n \approx 0$ using the binomial theorem, we can solve for $\sqrt{d}$. For example,

$$(x - \sqrt{d})^4 \approx 0 \quad \text{so}$$

$$\sqrt{d} \approx \frac{x^4 + 6dx^2 + d^2}{4x^3 + 4dx}. \tag{3}$$

After choosing an approximation to $\sqrt{d}$, Grant improves his approximation by evaluating expressions like (3) for successively higher powers of $(x - \sqrt{d})$.

The examples of approximations that Grant provides are sometimes convergents to $\sqrt{d}$ and sometimes not. As we shall see, if we start with a convergent that gives a positive solution to Pell's equation, we may iterate this procedure and obtain a convergent to $\sqrt{d}$ after each iteration.

Grant uses only even powers $(x - \sqrt{d})^{2n}$ for his expansions. Although this is not necessary, it simplifies his exposition, and for the same reason we also adopt this simplification. Expanding $(x - \sqrt{d})^{2j} \approx 0$ and solving for $\sqrt{d}$, we obtain,

$$\sqrt{d} \approx \frac{\displaystyle\sum_{i=0}^{j} \binom{2j}{2i} x^{2i} d^{j-i}}{\displaystyle\sum_{i=0}^{j-1} \binom{2j}{2i+1} x^{2i+1} d^{j-i-1}}. \tag{4}$$

Let us denote the rational function obtained by manipulating $(x - \sqrt{d})^k$ in the above manner by $f_k(x)$. The right-hand side of the example in (3) above is thus denoted by $f_4(x)$ and Grant's method of improving the first approximation x to $\sqrt{d}$ involves successively computing the values of $f_2(x)$, $f_4(x)$, etc.

As it happens, the values $f_2(x)$, $f_4(x)$, etc. are not necessarily convergents of $\sqrt{d}$ and each successive approximation requires the computation of a new polynomial. At this point, we take a different tack. We fix a value of k, take a convergent x that gives a positive solution to Pell's equation, and then successively evaluate $f_k(x)$, $f_k(f_k(x))$, etc. Given this choice of x, the successive values are always convergents to $\sqrt{d}$.

The following theorem makes this statement precise.

THEOREM 1. *Let d be a nonsquare positive integer, k an even integer, and p/q a convergent to $\sqrt{d}$ such that (p, q) is a solution (x, y) of $x^2 - dy^2 = \pm 1$. Denote $f_k(p/q)$ by P/Q, then P/Q is also a convergent to $\sqrt{d}$ with (P, Q) a solution of Pell's equation.*

Proof. It is straightforward to show that

$$P^2 - dQ^2 = (p^2 - dq^2)^k \tag{5}$$

if we see the identity

$$\sum_{i=1}^{2j}\binom{2v}{i}\binom{2v}{2j-i}(-1)^i=(-1)^j\binom{2v}{j}.$$

We can see that the identity is true by comparing the coefficients of x^{2j} in $(1-x^2)^{2v}$ and $(1-x)^{2v}(1+x)^{2v}$. Returning to (5), we see that the value of P^2-dQ^2 must be 1 since $p^2-dq^2=\pm 1$ and k is even. Thus (P,Q) satisfies Pell's equation and so by Lemma 1, P/Q is a convergent to $\sqrt{d}$.

Notice that if k were odd, we would not be able to start with solutions (p,q) to $x^2-dy^2=-1$. In any case, since (P,Q) is a solution to Pell's equation, the process may be iterated with each step producing convergents to $\sqrt{d}$.

The next natural question to ask is, of course, "if p/q is the jth convergent of $\sqrt{d}$, then which convergent is $f_k(p/q)=P/Q$?" The answer is,

THEOREM 2. *Let d be a nonsquare positive integer and suppose that p/q is the jth convergent to $\sqrt{d}$ and that (p,q) is a positive solution to Pell's equation. Then $f_k(p/q)$ is the $[k(j+1)-1]$st convergent to $\sqrt{d}$.*

Proof. Again denote $f_k(p/q)$ by P/Q and the period of the continued fraction expansion by n. If p/q is the jth convergent to $\sqrt{d}$ and (p,q) is a positive solution of Pell's equation, then $j=mn-1$ for some m (if n is even) or $j=2mn-1$ for some m (if n is odd) by Lemma 2. In either case, by Lemma 3

$$p+q\sqrt{d}=\left(x_1+y_1\sqrt{d}\right)^m \tag{6}$$

since (p,q) is the mth positive solution of Pell's equation. By the previous theorem P/Q is also a convergent to $\sqrt{d}$ and a positive solution of Pell's equation. Say $P/Q=c_r$, then as before $r=an-1$ for some a (if n is even) or $r=2an-1$ for some a (if n is odd). In either case,

$$P+Q\sqrt{d}=\left(x_1+y_1\sqrt{d}\right)^a. \tag{7}$$

Now notice in equation (4) that $k=2j$ and if we let $x=p/q$ and multiply numerator and denominator by q^k, the resulting quotient is P/Q. Thus an alternative way to define P and Q would be to use the equation

$$P+Q\sqrt{d}=\left(p+q\sqrt{d}\right)^k.$$

Plugging equation (6) into this expression and equating with (7), we see that $a=mk$. Thus, for even n

$$\begin{aligned} r &= mkn-1\\ &= k(mn-1)+k-1\\ &= kj+k-1\\ &= k(j+1)-1. \end{aligned}$$

A similar calculation for odd n produces the same result and so P/Q is the $k[(j+1)-1]$st convergent to $\sqrt{d}$.

To illustrate this, we begin with Heron's approximation of $33/8$ for $\sqrt{17}$. As we have seen, $\sqrt{17}=[4;\overline{8}]$. Since $33/8$ is the first convergent to $\sqrt{17}$, Theorem 2 tells us that $f_4(33/8)$ is c_7, the seventh convergent to $\sqrt{17}$. With the aid of the computer, we have

$$f_4(c_1) = c_7 = \frac{9478657}{2298192}.$$

Iterating once produces

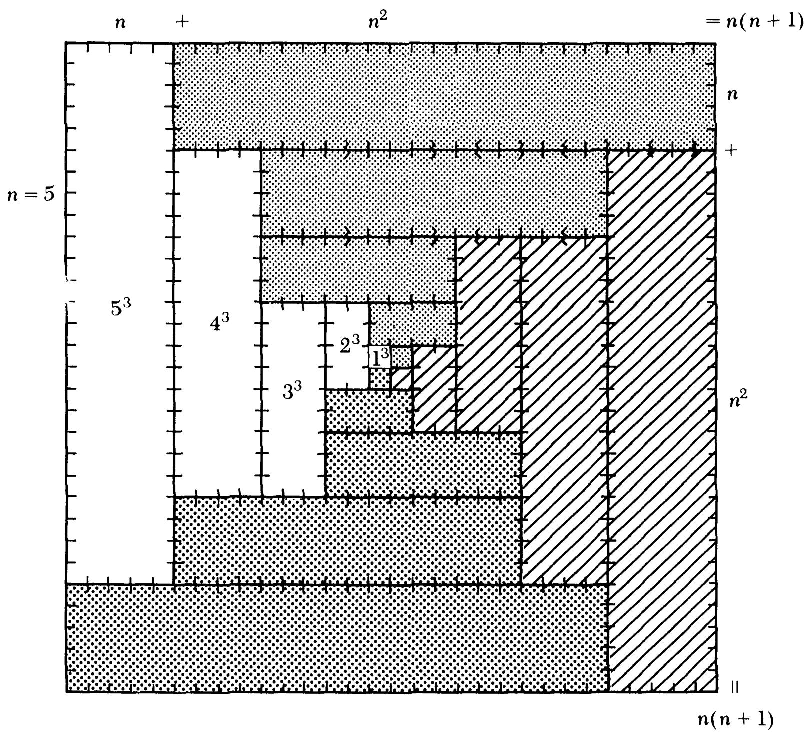
$$f_4(c_7) = c_{31} = \frac{64576903826545426454350012417}{15662199732482357532660158592}.$$

REFERENCES

1. D. M. Burton, *Elementary Number Theory*, Allyn and Bacon, 1980.
2. H. Eves, *An Introduction to the History of Mathematics*, Holt, Rinehart and Winston, 1969.
3. M. A. Grant, Approximation square roots, Note 66.36, *Math. Gaz.* (October 1982), 230–231.
4. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, 1960.

Proof without Words:

$$1^3 + 2^3 + 3^3 + 4^3 + \cdots + n^3 = \frac{\{n(n+1)\}^2}{4}$$



—Antonella Cupillari
Pennsylvania State University
Behrend College
Erie, PA 16563

$$f_4(c_1) = c_7 = \frac{9478657}{2298192}.$$

Iterating once produces

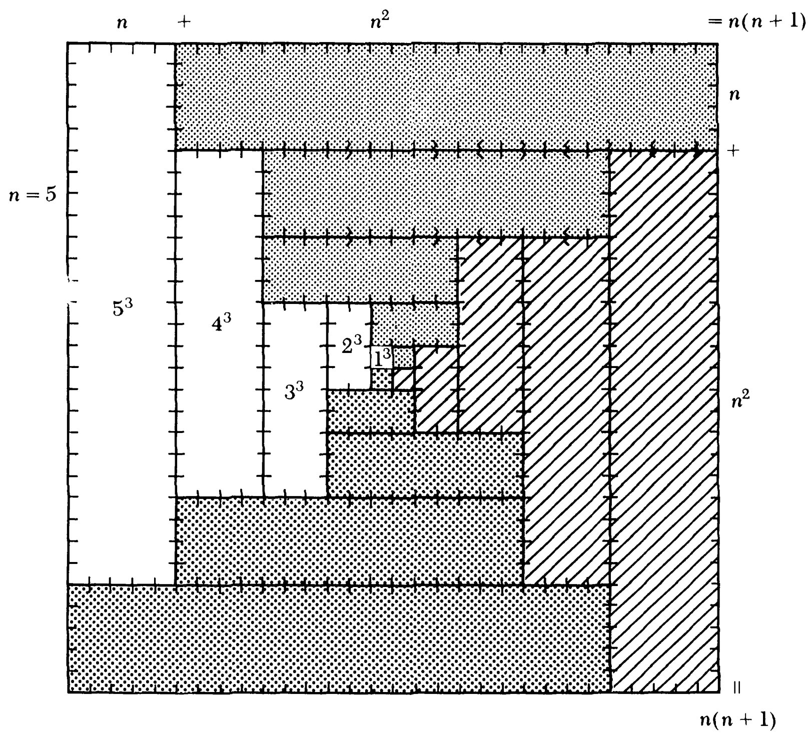
$$f_4(c_7) = c_{31} = \frac{64576903826545426454350012417}{15662199732482357532660158592}.$$

REFERENCES

1. D. M. Burton, *Elementary Number Theory*, Allyn and Bacon, 1980.
2. H. Eves, *An Introduction to the History of Mathematics*, Holt, Rinehart and Winston, 1969.
3. M. A. Grant, Approximation square roots, Note 66.36, *Math. Gaz.* (October 1982), 230–231.
4. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, 1960.

Proof without Words:

$$1^3 + 2^3 + 3^3 + 4^3 + \cdots + n^3 = \frac{\{n(n+1)\}^2}{4}$$



—ANTONELLA CUPILLARI
Pennsylvania State University
Behrend College
Erie, PA 16563

Just an Average Integral

JOHN FROHLIGER
RICK POSS
St. Norbert College
De Pere, WI 54115

What's so special about $\int \sec^3 x\,dx$? Is it merely that many students find the integral fairly difficult to calculate? Well, it does require integration by parts, the use of trigonometric identities, and a trick to obtain

$$\int \sec^3 x\,dx = (1/2)(\sec x \tan x + \ln|\sec x + \tan x| + C).$$

(See, for example, [2].)

This integral is also special because it is useful in solving seemingly unrelated problems. For example, if you wanted to find the arc length of a parabolic curve you would use an integral of the form $\int \sqrt{u^2+a^2}\,du$ and the solution of this involves the integral $\int \sec^3 x\,dx$.

Indeed, there are many useful integrals whose solutions require some ingenuity. However, $\int \sec^3 x\,dx$ has another especially interesting property: *it is precisely the average (arithmetic mean) of the derivative and the antiderivative of* $\sec x$.

Recognizing this might make the solution of $\int \sec^3 x\,dx$ easy to remember, but it also brings up a very good question: what other functions have this property? (An obvious answer is $\csc x$.) We can rephrase the question. What are the solutions of

$$\int y^3\,dx = (1/2)\left(y' + \int y\,dx\right)?$$

We can solve this mixed integral-differential equation by first differentiating both sides. This gives us

$$y^3 = (1/2)y'' + (1/2)y \text{ or } y'' + y - 2y^3 = 0. \qquad (*)$$

Some solutions of $(*)$ are the constant functions $y=0$ and $y = \pm\sqrt{2}/2$. We can find a more general solution by first assuming that $y' \neq 0$ in some interval. Then in that interval we can think of y' as a function of y and make the substitution (see [1])

$$u = y' = \frac{dy}{dx}.$$

Then

$$y'' = \frac{du}{dx} = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy}.$$

Now $(*)$ becomes $u(du/dy) + y - 2y^3 = 0$ or $u\,du = (2y^3 - y)\,dy$. Integrating both sides yields $u^2/2 = (y^4 - y^2 + K)/2$. This can be written as

$$\left(\frac{dy}{dx}\right)^2 = y^4 - y^2 + K \quad \text{or}$$

$$\frac{dy}{dx} = \pm\sqrt{y^4 - y^2 + K}$$

which is separable. After separating the variables and integrating, we finally obtain the solution

$$x = \pm\int \frac{1}{\sqrt{y^4 - y^2 + K}}\, dy.$$

The general closed-form solution of this equation is not obvious (it is not listed in standard integral tables), but several interesting solutions can be obtained by looking at special cases. For example, when $K = 0$

$$x = \pm\int \frac{1}{\sqrt{y^4 - y^2}}\, dy = \pm\int \frac{1}{|y|\sqrt{y^2 - 1}}\, dy = \pm\sec^{-1} y + C$$

or $y = \sec(x + C)$. (Note: If C is replaced by $C_1 - \pi/2$, we obtain $y = \csc(x + C_1)$.)

For another special case, use $K = 1/4$. Here one solution is

$$\begin{aligned} x &= \int \frac{1}{\sqrt{y^4 - y^2 + 1/4}}\, dy \\ &= \int \frac{1}{\sqrt{(y^2 - 1/2)^2}}\, dy \\ &= \int \frac{1}{y^2 - 1/2}\, dy. \end{aligned}$$

Using the method of partial fractions, we obtain

$$x = (1/\sqrt{2})(\ln|y - 1/\sqrt{2}| - \ln|y + 1/\sqrt{2}|) + C.$$

Solving for y yields

$$\begin{aligned} y &= \frac{1 + C_0 e^{\sqrt{2}x}}{\sqrt{2}\left(1 - C_0 e^{\sqrt{2}x}\right)} \\ &= \frac{\sqrt{2}}{1 - C_0 e^{\sqrt{2}x}} - \frac{1}{\sqrt{2}}. \end{aligned}$$

Other interesting problems can be developed by looking for a function the antiderivative of whose *square* is the average of its derivative and its antiderivative. Indeed, any power can be used and the method of solution will be basically the same. So, here we have a situation in which recognizing a pattern in the solution of a problem leads not only to an easy way to remember the solution, but also to an entirely new class of problems to investigate.

REFERENCES

1. Michael Golomb and Merrill Shanks, *Elements of Ordinary Differential Equations*, 2nd ed., McGraw-Hill, New York, 1965.
2. Dennis G. Zill, *Calculus with Analytic Geometry*, PWS Publishers, Boston, 1985.

Bilinear Basics

T. HOY BOOKER
Gallaudet University
Washington, DC 20002

Conformal mapping quite often captivates a group of students in a complex analysis class. Complete with space folds, a point zooming off to infinity, and distortions like a variable fun-house mirror, it combines mathematical modeling and logic with the lure of video games. The first step in this process is the introduction of the bilinear transformation.

In this paper, an analysis of two common problems in this area is presented. They are to find the most general bilinear transformation of the right half-plane onto itself and of the unit disk onto itself. This analysis differs from the conventional approach in that it is based on the internal characteristics of the transformation.

The bilinear transformation is given by

$$w = T(z) = \frac{az+b}{cz+d}, \qquad ad - bc \neq 0.$$

This maps the z-complex plane into the w-complex plane. In most textbooks, this transformation is broken down in the following way.

Case 1. $c = 0$.

For this case, the transformation reduces to $T(z) = (a/d)z + b/d$. This is interpreted as a rotation and a stretching/shrinking by a factor of a/d followed by a translation of b/d.

Case 2. $c \neq 0$. For this case, the transformation is considered in four steps.

1. Translation:

$$T_1(z) = z + \frac{d}{c}.$$

2. Inversion:

$$T_2(z) = \frac{1}{z}.$$

3. Rotation and stretching/shrinking:

$$T_3(z) = \frac{bc - ad}{c^2} z.$$

4. Translation:

$$T_4(z) = z + \frac{a}{c}.$$

This composes to form

$$w = T(z) = T_4 \circ T_3 \circ T_2 \circ T_1(z) = \frac{bc-ad}{c^2} \cdot \frac{1}{z + (d/c)} + \frac{a}{c}.$$

Unfortunately, this decomposition is not used for anything except a fairly mechanical proof of the fact that lines and circles map into lines or circles under this transformation.

The purpose of this note is to present an analysis of two problems in terms of this decomposition and to suggest that you present these problems to your students. With the development of these analyses, the student should better understand the processes involved in this basic transformation and will have a more receptive mind for the mappings to come.

A common approach in this type of problem is to develop the solution to one of the typical problems (unit disk onto unit disk is common) [1]–[4]. Then, this solution is used as the critical intermediate step in solving other problems of this type by funneling these secondary problems through the first solution. The first problem is generally solved by developing the fact that there is just one bilinear transformation that maps three given distinct points z_1, z_2, and z_3 into three specified distinct points w_1, w_2, w_3, respectively. This fact and the fact that, for a bilinear transformation, inverse points map to inverse points about the transformed circle are used to constrain the coefficients of the bilinear transformation suitably. This approach is highly algebraic and it is the opinion of the author that this approach does not promote an understanding of the transformation.

Problem # 1. Determine the most general bilinear mapping on the complex plane such that the right half-plane, $\mathrm{Re}(z) > 0$, maps onto itself.

Analysis

Case 1. $c = 0$.

For $c = 0$, $T(z) = (a/d)z + b/d$. No rotation can be allowed. Hence a/d is a positive real number. Considering the coefficients as vectors on the complex plane, this means that a and d are collinear and pointing in the same direction. Only vertical translation is allowable, hence b/d is purely imaginary, and b is normal to a and d. Hence, the transformation has the form $T(z) = rz + is$ for real r and s, $r > 0$.

Case 2. $c \neq 0$.

Step 1—Translation:

$$T_1(z) = z + \frac{d}{c}.$$

For this step, no horizontal translation can be allowed because a conflict will arise when inverting in step 2. To show this, consider the ε neighborhood of the origin. If horizontal translation to the left is allowed in step 1, then for sufficiently small $\varepsilon > 0$, the ε neighborhood is included in the image of step 1 and this will invert onto the region $|w| > 1/\varepsilon$. The rotation, stretching/shrinking, and translation in steps 3 and 4 cannot map this entire exterior region back to the right half-plane. If horizontal translation to the right is allowed in step 1, then the ε neighborhood is not included in the image of step 1. This ε neighborhood will invert onto the region $|w| > 1/\varepsilon$. This external region will not be in the image of step 2 and subsequent operations in steps 3 and 4 cannot fill the space of the right half-plane. It follows that d/c is purely imaginary.

Step 2—Inversion:

$$T_2(z) = \frac{1}{z}.$$

It is clear that the right half-plane maps onto itself under inversion and no problem is encountered in this step.

Step 3—Rotation, stretching/shrinking:

$$T_3(z) = \frac{bc - ad}{c^2} z.$$

The only conclusion necessary in this step is that $(bc - ad)/c^2$ must be a positive real number. Otherwise, a rotation is involved and the translation in the next step cannot correct this.

Step 4—Translation:

$$T_4(z) = z + \frac{a}{c}.$$

$\mathrm{Re}(a/c) = 0$. Otherwise, some portion of the left half-plane is included or some portion of the right half-plane is not.

Conclusion Interpreting the coefficients as vectors, we conclude from step 1 that d is normal to c and from step 4 that a is normal to c. A relationship between b and c can be shown by observing that since

$$\frac{bc - ad}{c^2} = x \quad \text{for real } x > 0,$$

we have

$$\frac{b}{c} = x + (a/c)(d/c) = x + (r_1 i)(r_2 i) \quad \text{for real } r_1, r_2,$$

and hence

$$\frac{b}{c} \quad \text{is real,} \quad \text{subject to } \frac{b}{c} > \frac{ad}{c^2}.$$

Hence, for $c \neq 0$, the coefficients of the transformation must be constrained as shown below. By forming

$$T(z) = \frac{az + b}{cz + d} = \frac{(a/c)z + b/c}{z + d/c},$$

we conclude that

$$T(z) = \frac{irz + s}{z + it}$$

for real r, s, and t, $s > rt$.

For each step considered separately, the analysis shows that a bilinear transformation maps the right half-plane onto itself if and only if this is true for each of the four mappings involved in the above decomposition of the transformation.

Problem # 2. Determine the most general bilinear mapping on the complex plane of the unit disk onto itself.

Analysis

Case 1. $c=0$.

For $c=0$, $T(z)=(a/d)z+b/d$. It is clear that $b=0$, $|a/d|=1$ and this results in a pure rotation.

Case 2. $c\neq 0$. Use FIGURE 1 as a guide.

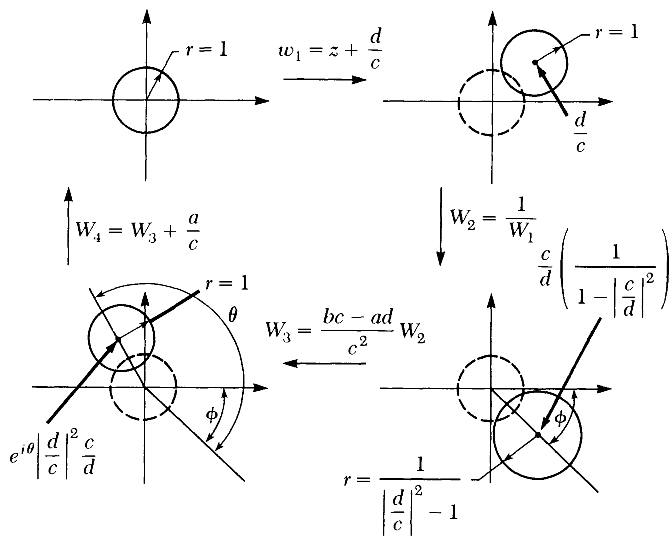


FIGURE 1

Step 1—Translation:

$$T_1(z)=z+d/c.$$

The only restriction necessary here is that $|d/c|>1$. Otherwise, the origin would be inside the translated circle and when inverted, would produce an image including infinity.

Step 2—Inversion:

$$T_2(z)=\frac{1}{z}.$$

To calculate the radius of the inverted circle, consider that the points nearest to and farthest from the origin would invert to the points farthest from and nearest to the origin. Hence, the image of $d/c+(d/c)/|d/c|$ and $d/c-(d/c)/|d/c|$ forms a diameter which extends through the origin with the same direction as c/d. Hence, the radius after inversion is

$$\frac{1}{2}\left|\frac{1}{\frac{d}{c}+\frac{d/c}{|d/c|}}-\frac{1}{\frac{d}{c}-\frac{d/c}{|d/c|}}\right|=\frac{1}{|d/c|^2-1},$$

and the center is located at

$$\frac{1}{\frac{d}{c}+\frac{d/c}{|d/c|}}+\frac{1}{|d/c|^2-1}\cdot\frac{c/d}{|c/d|}=\frac{c}{d}\cdot\frac{1}{1-|c/d|^2}.$$

Step 3—Rotation, stretching/shrinking:

$$T_3(z)=\frac{bc-ad}{c^2}z.$$

The rotation in this step is unrestricted. It is necessary, however, to return the radius to unity to set up the next translation. Hence, for real θ,

$$\frac{bc-ad}{c^2}=e^{i\theta}(|d/c|^2-1).$$

The center of the circle is now located at

$$e^{i\theta}(|d/c|^2-1)(c/d)\left(\frac{1}{1-|c/d|^2}\right)=e^{i\theta}(|d/c|^2)(c/d).$$

Step 4—Translation:

$$T_4(z)=z+a/c.$$

This step must bring the center of the circle in step 3 back to the origin. Hence,

$$\frac{a}{c}=-e^{i\theta}\left|\frac{d}{c}\right|^2(c/d).$$

Conclusion At this point, the problem is essentially finished but we are not left with a clear picture of the mapping as we were in the first problem. To achieve this, we combine the pieces to form

$$\begin{aligned}w&=e^{i\theta}\left(\left|\frac{d}{c}\right|^2-1\right)\left(\frac{1}{z+d/c}\right)-e^{i\theta}\left|\frac{d}{c}\right|^2\frac{c}{d}\\&=e^{i\theta}\frac{|d/c|^2-1-|d/c|^2(c/d)z-|d/c|^2}{z+d/c}\\&=e^{i\theta}\frac{-|d/c|^2(c/d)\left(z-\left(-|c/d|^2(d/c)\right)\right)}{(-d/c)((-c/d)z-1)}.\end{aligned}$$

Noting that $(|d/c|^2(c/d))/(d/c)$ is a pure rotation which can be absorbed by the $e^{i\theta}$ term and letting $z_0=-|c/d|^2(d/c)$ we achieve the familiar result

$$w=e^{i\theta}\frac{z-z_0}{z_0z-1}.$$

where, by step 1, $|z_0|<1$.

Thus, to map the unit disk onto itself, the transformation is either a simple rotation or a sequence of steps beginning with a translation bounded away from the origin, followed by an inversion, an unrestricted rotation about the origin, a multiplicative factor to adjust the radius to unity, and a translation of the center back to the origin.

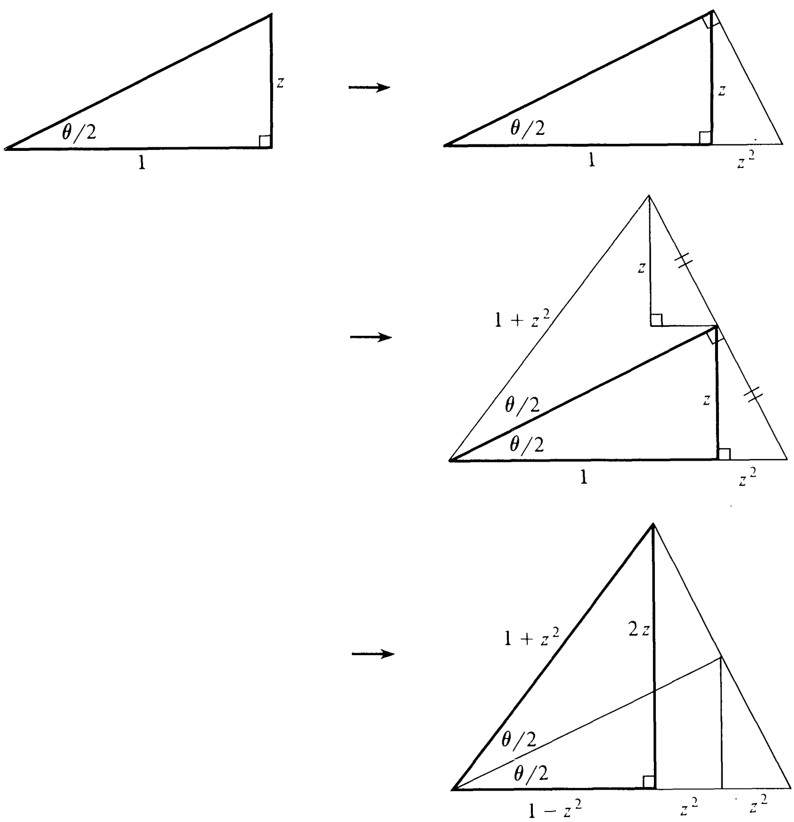
Recommendations for further study Most textbooks on complex analysis have examples of these types of problems and each solution seems to require its own particular analysis. It is suggested that the student find the appropriate constraints for the following:

1. upper half-plane onto itself;
2. real or imaginary axis onto itself;
3. unit circle onto itself;
4. unit disk onto upper or right half-plane;
5. upper or right half-plane onto the unit disk.

REFERENCES

1. L. Ahlfors, *Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1966.
2. R. Churchill, *Complex Variables and Applications*, 2nd ed., McGraw-Hill, New York, 1960.
3. W. Derrick, *Complex Analysis and Applications*, 2nd ed., Wadsworth, Belmont, CA, 1984.
4. E. Hille, *Analytic Function Theory*, Vol. 1, 2nd ed., Chelsea, New York, 1976.

Proof without Words: The substitution to make a rational function of the sine and cosine



$$z = \tan\frac{\theta}{2} \Rightarrow \sin\theta = \frac{2z}{1+z^2} \quad \text{and} \quad \cos\theta = \frac{1-z^2}{1+z^2}$$

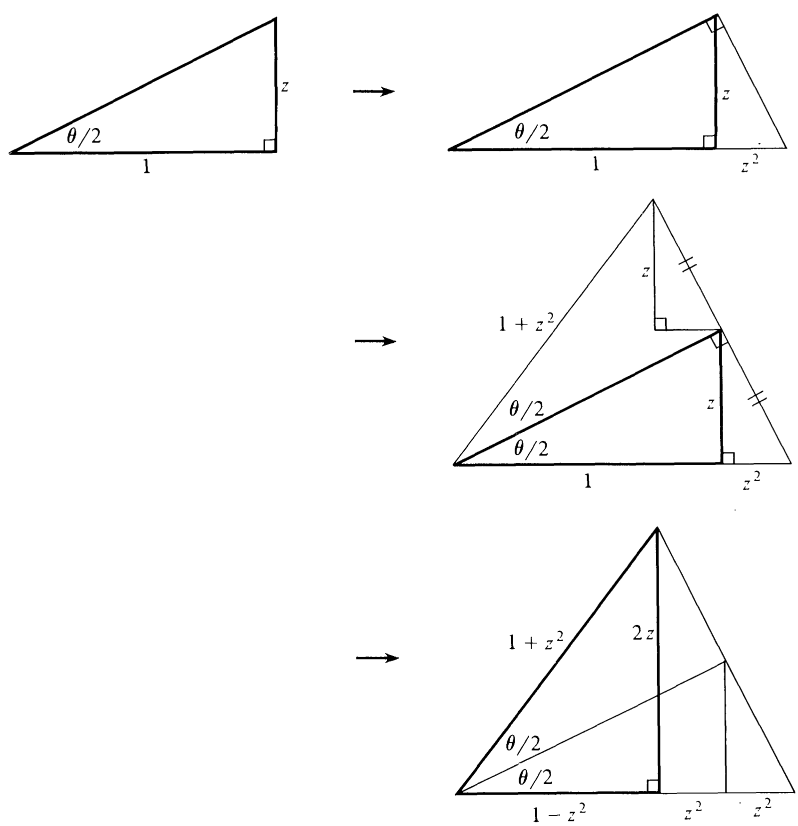
—Roger B. Nelsen
Lewis and Clark College

1. upper half-plane onto itself;
2. real or imaginary axis onto itself;
3. unit circle onto itself;
4. unit disk onto upper or right half-plane;
5. upper or right half-plane onto the unit disk.

REFERENCES

1. L. Ahlfors, *Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1966.
2. R. Churchill, *Complex Variables and Applications*, 2nd ed., McGraw-Hill, New York, 1960.
3. W. Derrick, *Complex Analysis and Applications*, 2nd ed., Wadsworth, Belmont, CA, 1984.
4. E. Hille, *Analytic Function Theory*, Vol. 1, 2nd ed., Chelsea, New York, 1976.

Proof without Words: The substitution to make a rational function of the sine and cosine



$$z = \tan\frac{\theta}{2} \Rightarrow \sin\theta = \frac{2z}{1+z^2} \quad \text{and} \quad \cos\theta = \frac{1-z^2}{1+z^2}$$

—Roger B. Nelsen
Lewis and Clark College

The Tennis Ball Paradox

R. W. HAMMING
Naval Postgraduate School
Monterey, CA 93943

On several occasions I have been told the paradox about a room and some tennis balls. The story goes that in the first half minute two tennis balls, #1 and #2, are tossed into the room and #1 is thrown out. In the next quarter minute balls #3 and #4 are tossed in and #2 is tossed out. In the next 1/8 minute #5 and #6 are tossed in and #3 is tossed out. Etc.

1. I was told that in the limit there are no tennis balls in the room because "if you think that there are then when you name the ball I can tell you when it was thrown out."

2. However, I said to myself that if I look at the number of balls in the room, then at stage N there are N balls left and I am being asked to believe that the limit of N as N goes to infinity is 0.

To get a different view on what is essentially the same situation we examine a well-known mathematical result. The constant γ is defined by

$$\lim_{N\to\infty}\left\{\sum_{n=1}^{N}\frac{1}{n}-\log N\right\}=\gamma=0.5772\ 15664\dots.$$

Define

$$\gamma_N=\sum_{n=1}^{N}\frac{1}{n}-\log N \qquad (\gamma_N\to\gamma)$$

and

$$S_N=\gamma_{2N}-\gamma_N=\sum_{n=N+1}^{2N}\frac{1}{n}-\log\frac{2N}{N}$$
$$=\left[\frac{1}{N+1}+\frac{1}{N+2}+\cdots+\frac{1}{2N}\right]-\log 2.$$

Since γ_{2n} and γ_n both approach γ,

$$\lim_{N\to\infty}\left[\frac{1}{N+1}+\frac{1}{N+2}+\cdots+\frac{1}{2N}\right]=\log 2.$$

We see that, as in the tennis ball paradox, at each stage (as N goes to $N+1$) two terms of the series are added and one term is removed. The first argument would assert that in the limit there are no terms in the sum since if you name the particular term, then I can tell you when it was removed; hence the sum, so the argument goes, has no terms. Yet the sum is apparently $\log 2$. The difference between the two examples is that in the series we have the terms $1/n$, and in the tennis balls we count the number by integers and each corresponding term is 1. Also, the limit of the tennis balls is infinite while for the second illustration the limit is finite.

The trouble arises from the introduction of dynamics into mathematics—which is common in thinking about limits! If we use the terminology of the standard textbook for handling limits then we have the two situations:

1. For any given term N there exists an n_0 such that for all $n \geqslant n_0$ the ball #N is not there.

2. For any given N there exists an n_0 such that for all $n \geqslant n_0$ the number of balls in the room is greater than N.

It is the translation into colorful words that produces the paradox; one doubts the wisdom of the first translation to the assertion that in the limit there are no balls in the room. The careful textbook methods were developed for clarity in thinking, not just to befuddle the student.

Unexpected Occurrences of the Number e

HARRIS S. SHULTZ
BILL LEONARD
California State University
Fullerton, CA 92634

Martin Gardner once observed [**2**, p. 34] that of the three most famous irrational numbers, π, the golden ratio, and e, the third is least familiar to students early in their study of mathematics. The number e, named by Leonhard Euler (1707–1783), is usually encountered for the first time during the second course in calculus, either through the equation

$$\int_1^e \frac{dt}{t} = 1,$$

introduced in connection with natural logarithms, or else through the equation

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n.$$

However, neither of these definitions of e provides immediate insight into this important number. As a result, students rarely come away with a solid grasp of this number, beyond, perhaps, rote memory of the phrase, "It is the base of the natural logarithms." Given the importance of e, students' first meeting with this number should be a more memorable experience than the standard introductions provide.

In this note we present an expository catalog of occurrences of e in probability which might be used to increase students' appreciation of this number. Most likely, some, but not all, of these examples will be familiar to teachers of mathematics.

Example 1. Each of two people is given a shuffled deck of playing cards. Simultaneously they expose their first cards. If these cards do not match (for example, two "four of clubs" would be considered a "match"), they proceed to expose their second cards and so forth through the decks. What is the probability of getting through the decks without a single match? In [**4**, p. 281] it is shown that the answer is given by the sum,

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{1}{52!},$$

which is the initial portion of a series for $1/e$ (based on the Maclaurin series

1. For any given term N there exists an n_0 such that for all $n \geqslant n_0$ the ball #N is not there.

2. For any given N there exists an n_0 such that for all $n \geqslant n_0$ the number of balls in the room is greater than N.

It is the translation into colorful words that produces the paradox; one doubts the wisdom of the first translation to the assertion that in the limit there are no balls in the room. The careful textbook methods were developed for clarity in thinking, not just to befuddle the student.

Unexpected Occurrences of the Number e

HARRIS S. SHULTZ
BILL LEONARD
California State University
Fullerton, CA 92634

Martin Gardner once observed [2, p. 34] that of the three most famous irrational numbers, π, the golden ratio, and e, the third is least familiar to students early in their study of mathematics. The number e, named by Leonhard Euler (1707–1783), is usually encountered for the first time during the second course in calculus, either through the equation

$$\int_1^e \frac{dt}{t} = 1,$$

introduced in connection with natural logarithms, or else through the equation

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n.$$

However, neither of these definitions of e provides immediate insight into this important number. As a result, students rarely come away with a solid grasp of this number, beyond, perhaps, rote memory of the phrase, "It is the base of the natural logarithms." Given the importance of e, students' first meeting with this number should be a more memorable experience than the standard introductions provide.

In this note we present an expository catalog of occurrences of e in probability which might be used to increase students' appreciation of this number. Most likely, some, but not all, of these examples will be familiar to teachers of mathematics.

Example 1. Each of two people is given a shuffled deck of playing cards. Simultaneously they expose their first cards. If these cards do not match (for example, two "four of clubs" would be considered a "match"), they proceed to expose their second cards and so forth through the decks. What is the probability of getting through the decks without a single match? In [4, p. 281] it is shown that the answer is given by the sum,

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{1}{52!},$$

which is the initial portion of a series for $1/e$ (based on the Maclaurin series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \Big).$$

This problem has appeared in the literature under the name "Hat-Check Problem," wherein the following question is asked: If n men have their hats randomly returned, what is the probability that none of the men winds up with his own hat?

Example 2. The following appeared as a Putnam examination problem [**1**]: If numbers are randomly selected from the interval $[0, 1]$, what is the expected number of selections necessary until the sum of the chosen numbers first exceeds 1? The answer is e. An elementary proof of this is given in [**6**].

Example 3. The "Secretary Problem" concerns an employer who is about to interview n applicants for a secretarial position. At the end of each interview he must decide whether or not this is the applicant he wishes to hire. Should he pass over an interviewee, this person cannot be hired thereafter. If he gets to the last applicant, this person gets the job by default. The goal is to maximize the probability that the person hired is the one most qualified. His strategy will be to decide upon a number $k < n$, to interview the first k applicants, and then to continue interviewing until an applicant more qualified than each of those first k is found. As seen in [3], the probability of hiring the most qualified applicant is greatest when k/n is approximately $1/e$. Moreover, this number, $1/e$, is in fact the approximate maximum probability. For example, if there are $n = 30$ applicants, the employer should interview 11 (which is approximately $30/e$) and then select the first thereafter who is more qualified than all of the first eleven. The probability of obtaining the most qualified applicant is approximately $1/e$.

Example 4. With each purchase, a certain fast-food restaurant chain gives away a coin with a picture of a state capitol on it. The object is to collect the entire set of 50 coins. Question: After 50 purchases, what fraction of the set of 50 coins would one expect to have accumulated? In [**5**] it is shown that this fraction is $1 - (1 - \frac{1}{50})^{50}$, which is approximately $1 - (1/e)$.

Example 5. A sequence of numbers, $x_1, x_2, x_3, \ldots$, is generated randomly from the interval $[0, 1]$. The process is continued as long as the sequence is monotonically increasing or monotonically decreasing. What is the expected length of the monotonic sequence? For example, for the sequence beginning .91, .7896, .20132, .41, the length of the monotonic sequence is three. For the sequence beginning .134, .15, .3546, .75, .895, .276, the length of the monotonic sequence is five.

The probability that the length L of the monotonic sequence is greater than k is given by

$$\begin{aligned} P(L > k) &= P(x_1 < x_2 < \cdots < x_{k+1}) + P(x_1 > x_2 > \cdots > x_{k+1}) \\ &= \frac{1}{(k+1)!} + \frac{1}{(k+1)!} \\ &= \frac{2}{(k+1)!}. \end{aligned}$$

If we denote $P(L = k)$ by p_k, the expected length of the monotonic sequence is

$$E(L) = 2p_2 + 3p_3 + 4p_4 + 5p_5 + \cdots.$$

We rewrite this in the form

$$\begin{aligned} E(L) = \quad & p_2 + p_3 + p_4 + p_5 + p_6 + \cdots \\ + & p_2 + p_3 + p_4 + p_5 + p_6 + \cdots \\ + & \qquad p_3 + p_4 + p_5 + p_6 + \cdots \\ + & \qquad\qquad p_4 + p_5 + p_6 + \cdots \\ + & \qquad\qquad\qquad p_5 + p_6 + \cdots . \end{aligned}$$

Adding this triangular array row by row, we obtain

$$\begin{aligned} E(L) &= 1 + 1 + P(L > 2) + P(L > 3) + P(L > 4) + \cdots \\ &= 1 + 1 + 2\left(\frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \cdots\right) \\ &= 1 + 1 + 2\left(e - \tfrac{5}{2}\right) \\ &= 2e - 3 \approx 2.4366. \end{aligned}$$

Example 6. A slight revision of the previous example gives a more pleasing answer. First, we require the sequence to be monotonically increasing and, second, in computing the length of the sequence we include the first number that reverses the increasing direction of the sequence. Hence, the sequence beginning .154, .3245, .58, .432 is assigned a score of four and the sequence beginning .6754, .239 is assigned a score of two. Using an argument similar to that of the preceding example, it can be shown that the expected score is that ubiquitous and fascinating number, e.

REFERENCES

1. L. E. Bush, The William Lowell Putnam Mathematical Competition, *Amer. Math. Monthly* 68 (1961), 18–33.
2. M. Gardner, *The Unexpected Hanging*, Simon & Schuster, New York, 1969.
3. J. P. Gilbert and F. Mosteller, Recognizing the maximum of a sequence, *Amer. Statist. Assoc. J.* 61 (1966), 35–73.
4. A. P. Hillman, G. L. Alexanderson and R. M. Grassl, *Discrete and Combinatorial Mathematics*, Dellen Publishing, San Francisco, 1986.
5. R. C. Shiflett and H. S. Shultz, Can I expect a full set?, *Math. Gaz.* 64 (1980), 262–266.
6. H. S. Shultz, An expected value problem, *Two-Year College Math. J.* 10 (1979), 179.

Decompositions of *U*-Groups

YUNGCHEN CHENG
Southwest Missouri State University
Springfield, MO 65804

In a recently published undergraduate abstract algebra textbook [**1**], the group of units, $U(n)$, of the ring of integers modulo n is used as a concrete way to illustrate the concepts of external direct products of groups and internal direct products of subgroups. In particular, when m and n are relatively prime, the group $U(mn)$ is isomorphic to the external direct product $U(m) \oplus U(n)$ and the internal direct product $U_m(mn) \times U_n(mn)$ where $U_s(st) = \{x \in U(st) | x \equiv 1 \bmod s\}$ (see [**2**] for the

We rewrite this in the form

$$\begin{aligned} E(L) = \quad & p_2 + p_3 + p_4 + p_5 + p_6 + \cdots \\ + & p_2 + p_3 + p_4 + p_5 + p_6 + \cdots \\ + & \qquad p_3 + p_4 + p_5 + p_6 + \cdots \\ + & \qquad\qquad p_4 + p_5 + p_6 + \cdots \\ + & \qquad\qquad\qquad p_5 + p_6 + \cdots . \end{aligned}$$

Adding this triangular array row by row, we obtain

$$\begin{aligned} E(L) &= 1 + 1 + P(L > 2) + P(L > 3) + P(L > 4) + \cdots \\ &= 1 + 1 + 2\left(\frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \cdots\right) \\ &= 1 + 1 + 2\left(e - \tfrac{5}{2}\right) \\ &= 2e - 3 \approx 2.4366. \end{aligned}$$

Example 6. A slight revision of the previous example gives a more pleasing answer. First, we require the sequence to be monotonically increasing and, second, in computing the length of the sequence we include the first number that reverses the increasing direction of the sequence. Hence, the sequence beginning .154, .3245, .58, .432 is assigned a score of four and the sequence beginning .6754, .239 is assigned a score of two. Using an argument similar to that of the preceding example, it can be shown that the expected score is that ubiquitous and fascinating number, e.

REFERENCES

1. L. E. Bush, The William Lowell Putnam Mathematical Competition, *Amer. Math. Monthly* 68 (1961), 18–33.
2. M. Gardner, *The Unexpected Hanging*, Simon & Schuster, New York, 1969.
3. J. P. Gilbert and F. Mosteller, Recognizing the maximum of a sequence, *Amer. Statist. Assoc. J.* 61 (1966), 35–73.
4. A. P. Hillman, G. L. Alexanderson and R. M. Grassl, *Discrete and Combinatorial Mathematics*, Dellen Publishing, San Francisco, 1986.
5. R. C. Shiflett and H. S. Shultz, Can I expect a full set?, *Math. Gaz.* 64 (1980), 262–266.
6. H. S. Shultz, An expected value problem, *Two-Year College Math. J.* 10 (1979), 179.

Decompositions of *U*-Groups

YUNGCHEN CHENG
Southwest Missouri State University
Springfield, MO 65804

In a recently published undergraduate abstract algebra textbook [**1**], the group of units, $U(n)$, of the ring of integers modulo n is used as a concrete way to illustrate the concepts of external direct products of groups and internal direct products of subgroups. In particular, when m and n are relatively prime, the group $U(mn)$ is isomorphic to the external direct product $U(m) \oplus U(n)$ and the internal direct product $U_m(mn) \times U_n(mn)$ where $U_s(st) = \{x \in U(st) \mid x \equiv 1 \bmod s\}$ (see [**2**] for the proofs). The approach taken in [1] is a global one, dealing only with groups involved and not specifically indicating how the factorization of any particular element of $U(mn)$ is achieved. For instance, in the factorization of $U(105)$ as $U(7) \oplus U(15)$ or as $U_7(105) \times U_{15}(105)$ how is the element 38 (cf. Exercise 10 in [1, p. 115]), say, represented? Even more complicated is the factorization of 38 using $U(3) \oplus U(5) \oplus U(7)$ or $U_{35}(105) \times U_{21}(105) \times U_{15}(105)$. Conversely, given $U(n)$ as an external direct product of two or more U-groups it is natural to ask which element of $U(n)$ corresponds to any particular ordered tuple from the direct product under the isomorphism given in [2]. In this note we give explicit formulas that answer these questions.

Let n be a positive integer > 1. If x is an integer, let $[x]_n$ denote the remainder of x upon division by n. One can find the following results in [2]:

(1) If $(m, n) = 1$, the mapping φ from $U_m(mn)$ to $U(n)$ given by $\varphi(x) = [x]_n$ is an isomorphism and the inverse isomorphism φ^{-1} is given by

$$\varphi^{-1}(y) = [sm(y-1)+1]_{mn},$$

where $sm + tn = 1$, $s, t \in Z$.

(2) If $(m, n) = 1$, then $U(mn) = U_m(mn) \times U_n(mn)$ (internal direct product) $\simeq U(m) \oplus U(n)$ (external direct product).

(3) If $m = n_1 n_2 \cdots n_k$ where $(n_i, n_j) = 1$ for $i \neq j$, then $U(m) = U_{m/n_1}(m) \times U_{m/n_2}(m) \times \cdots \times U_{m/n_k}(m)$ (internal direct product) $\simeq U(n_1) \oplus U(n_2) \oplus \cdots \oplus U(n_k)$ (external direct product).

Now, suppose $(m, n) = 1$ and $sm + tn = 1$, $s, t \in Z$. For each $a \in U(mn)$, we know $[a]_n \in U(n)$ and $sm(a-1)+1 \equiv sm([a]_n - 1) + 1 (\text{mod } mn)$. Therefore $x = [sm(a-1)+1]_{mn} \in U_m(mn)$ by (1).

Similarly, $y = [tn(a-1)+1]_{mn} \in U_n(mn)$. Furthermore, $xy \equiv sm(a-1) + tn(a-1) + 1 = a (\text{mod } mn)$. So,

(4) The decomposition $U(mn) = U_m(mn) \times U_n(mn)$, $(m, n) = 1$ is accomplished by

$$a = ([sm(a-1)+1]_{mn})([tn(a-1)+1]_{mn}),$$

where $sm + tn = 1$.

Example. Let's illustrate (4) with the element 38 from the group $U(105) = U_7(105) \times U_{15}(105)$. Since $(-2) \cdot 7 + 1 \cdot 15 = 1$ we have $s = -2$ and $t = 1$ and we obtain $38 = [(-2)(7)(37) + 1]_{105} \cdot [(1)(15)(37) + 1]_{105} = 8 \cdot 31$ in the decomposition of $U(105) = U_7(105) \times U_{15}(105)$.

To establish the isomorphism $U(mn) \simeq U(m) \oplus U(n)$ we collect the isomorphisms $U(mn) = U_m(mn) \times U_n(mn) \simeq U_m(mn) \oplus U_n(mn) \simeq U(n) \oplus U(m) \simeq U(m) \oplus U(n)$ and observe that $[[x]_{mn}]_n = [x]_n$ for any m, n, x. To find the inverse isomorphism we reverse the direction and multiply at the end in $U(mn)$. Therefore, combining (1) and (4), we have the following:

(5) The isomorphism Φ: $U(mn) \to U(m) \oplus U(n)$, $(m, n) = 1$ is given by

$$\Phi(x) = ([tn(x-1)+1]_m, [sm(x-1)+1]_n),$$

where $sm + tn = 1$. The inverse isomorphism Φ^{-1}: $U(m) \oplus U(n) \to U(mn)$ is given by $\Phi^{-1}(y, z) = [sm(z-1) + tn(y-1) + 1]_{mn}$.

Example. Let's illustrate (5) with the element 38 from the group $U(105)$ and the element $(3, 8)$ from the group $U(7) \oplus U(15)$. Again, we have $s = -2$ and $t = 1$ and we obtain $38 \to ([(1)(15)(37)+1]_7, [(-2)(7)(37)+1]_{15}) = (3,8)$ via the isomorphism Φ: $U(105) \to U(7) \oplus U(15)$ and $(3,8) \to [(-2)(7)(7)+(1)(15)(2)+1]_{105} = 38$ via the inverse isomorphism Φ^{-1}: $U(7) \oplus U(15) \to U(105)$.

What can we do when $m = n_1 n_2 \cdots n_k$, $(n_i, n_j) = 1$ for $i \neq j$? Assume that s_i and t_i are integers such that $s_i n_i + t_i(m/n_i) = 1$, $1 \leqslant i \leqslant k$. For each i, we already know that each element x in $U(m)$ has a unique factorization $x = x_i y_i$ in $U(m) = U_{m/n_i}(m) \times U_{n_i}(m)$. On the other hand, x also has a unique factorization $x = z_1 z_2 \cdots z_k$ in $U(m) = U_{m/n_1}(m) \times U_{m/n_2}(m) \times \cdots \times U_{m/n_k}(m)$. For $j \neq i$, we observe that $U_{m/n_j}(m)$ is a subgroup of $U_{n_i}(m)$ because n_i is a divisor of m/n_j. Therefore, $x = x_i y_i = z_i(z_1 \cdots z_{i-1} z_{i+1} \cdots z_k)$ are two factorings of x in $U(m) = U_{m/n_i}(m) \times U_{n_i}(m)$. Hence, $x_i = z_i$. By (4) we conclude:

(6) If $m = n_1 n_2 \cdots n_k$, $(n_i, n_j) = 1$ for $i \neq j$, then the decomposition $U(m) = U_{m/n_1}(m) \times U_{m/n_2}(m) \times \cdots \times U_{m/n_k}(m)$ is accomplished by

$$x = \prod_{i=1}^{k} [t_i(m/n_i)(x-1)+1]_m \pmod m,$$

where $s_i n_i + t_i(m/n_i) = 1$, $1 \leqslant i \leqslant k$.

EXAMPLE. Let's illustrate (6) with the element 38 from the group $U(105) = U_{35}(105) \times U_{21}(105) \times U_{15}(105)$. Since $12 \cdot 3 + (-1) \cdot 35 = 1$, $(-4) \cdot 5 + 1 \cdot 21 = 1$, and $(-2) \cdot 7 + 1 \cdot 15 = 1$, we have $m = 105$, $n_1 = 3$, $n_2 = 5$, $n_3 = 7$; $t_1 = -1$, $s_1 = 12$, $t_2 = 1$, $s_2 = -4$, $t_3 = 1$, $s_3 = -2$. Thus, we obtain $38 = [(-1)(35)(37)+1]_{105} \cdot [(1)(21)(37)+1]_{105} \cdot [(1)(15)(37)+1]_{105} = 71 \cdot 43 \cdot 31$ in the decomposition $U(105) = U_{35}(105) \times U_{21}(105) \times U_{15}(105)$.

Finally, using the same argument as in (5), we obtain the following:

(7) If $m = n_1 n_2 \cdots n_k$, $(n_i, n_j) = 1$ for $i \neq j$, and $s_i n_i + t_i(m/n_i) = 1$, $1 \leqslant i \leqslant k$, then the isomorphism Φ: $U(m) \to U(n_1) \oplus U(n_2) \oplus \cdots \oplus U(n_k)$ is given by $\Phi(x) = ([t_1(m/n_1)(x-1)+1]_{n_1}, \ldots, [t_k(m/n_k)(x-1)+1]_{n_k})$. The inverse isomorphism Φ^{-1}: $U(n_1) \oplus \cdots \oplus U(n_k) \to U(m)$ is given by

$$\Phi^{-1}(x_1, \ldots, x_k) = \prod_{i=1}^{k} [t_i(m/n_i)(x_i - 1) + 1]_m$$
$$= 1 + \sum_{i=1}^{k} [t_i(m/n_i)(x_i - 1)]_m \pmod m.$$

EXAMPLE. Let's illustrate (7) with the element 38 from the group $U(105)$ and the element $(2,3,3)$ from the group $U(3) \oplus U(5) \oplus U(7)$. Again, we have $m = 105$, $n_1 = 3$, $n_2 = 5$, $n_3 = 7$; $t_1 = -1$, $t_2 = 1$, $t_3 = 1$. Thus we have $38 \to ([71]_3, [43]_5, [31]_7) = (2,3,3)$ via the isomorphism Φ: $U(105) \to U(3) \oplus U(5) \oplus U(7)$ and $(2,3,3) \to 1 + [(-1)(35)(1)]_{105} + [(1)(21)(2)]_{105} + [(1)(15)(2)]_{105} = 1 + 70 + 42 + 30 = 143 \equiv 38 \pmod{105}$ via the inverse isomorphism Φ^{-1}: $U(3) \oplus U(5) \oplus U(7) \to U(105)$.

The author is grateful to the editor and the referee for valuable suggestions that led to improvements in the original draft.

REFERENCES

1. J. A. Gallian, *Contemporary Abstract Algebra*, D. C. Heath, Lexington, MA, 1986.
2. J. A. Gallian and D. Rusin, Factoring groups of integers modulo n, this MAGAZINE 53 (1980), 33–36.

PROBLEMS

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

GEORGE GILBERT, *associate editor*
St. Olaf College

Proposals

To be considered for publication, solutions should be received by March 1, 1990.

1327. *Proposed by Hüseyin Demir, Middle East Technical University, Ankara, Turkey.*

Let the sides PQ, QR, RS, SP of a convex quadrangle $PQRS$ touch an inscribed circle at A, B, C, D and let the midpoints of the sides AB, BC, CD, DA be E, F, G, H. Show that the angle between the diagonals PR, QS is equal to the angle between the bimedians EG, FH.

1328. *Proposed by Ronald E. Ruemmler, Middlesex County College, Edison, New Jersey.*

Find the largest even number which cannot be expressed as the sum of two composite odd numbers.

1329. *Proposed by Joe Flowers, Northeast Missouri State University, Kirksville, Missouri.*

For real-valued functions f and g defined on the set of positive integers, let $f * g$ denote the Dirichlet product defined by

$$(f * g)(n) = \sum_{d|n} f(d) \cdot g(n/d).$$

Let $\sigma(n)$ denote the sum of all positive factors of n and let σ^{-1} denote the inverse of σ with respect to Dirichlet multiplication. (Note: the identity element is I, where $I(1) = 1$ and $I(n) = 0$ for all $n > 1$.)

ASSISTANT EDITORS: Clifton Corzat and Theodore Vessey, *St. Olaf College. We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) next to a problem number indicates that neither the proposer nor the editors supplied a solution.*

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

Solutions and new proposals should be mailed in duplicate to Loren C. Larson, Department of Mathematics, St. Olaf College, Northfield, MN 55057.

a. Evaluate $\sigma^{-1}(p^a)$, where p is a prime.
b. Call n an "inperfect" (inverse-perfect) number if $\sigma^{-1}(n) = 2n$. Determine all inperfect numbers.

1330. *Proposed by K. L. McAvaney, Deakin University, Geelong, Victoria, Australia.*

Show that for all integers $p \geq 2$ and $n = 2, 4, \ldots, 2p-2$,

$$\sum_{k=1}^{p} (-1)^k \binom{2p}{p+k} k^n = 0.$$

1331. *Proposed by Daniel Shapiro, The Ohio State University, Columbus, Ohio.*

Suppose **A** and **B** are $n \times n$ matrices (with complex entries, say). If there exists another matrix **P** with the property that $\mathbf{A} = f(\mathbf{P})$ and $\mathbf{B} = g(\mathbf{P})$ for some polynomials $f(x)$, $g(x)$, then clearly $\mathbf{AB} = \mathbf{BA}$. Is the converse true? That is, if $\mathbf{AB} = \mathbf{BA}$ for two matrices **A**, **B**, does it follow that **A** and **B** are expressible as polynomials in some matrix **P**?

Quickies

Answers to the Quickies are on page 281.

Q752. *Proposed by Russell Jay Hendel, Dowling College, Oakdale, New York.*

Show that the "little" Fermat theorem, $a^p - a \equiv 0 \pmod{p}$, p a prime, is a consequence of the binomial theorem.

Q753. *Proposed by Norman Schaumberger, Bronx Community College, Bronx, New York.*

Show that

$$\sum_{k=1}^{n} \frac{1}{k+1} \ln(k^2+k) < \ln^2(n+1) < \sum_{k=1}^{n} \frac{1}{k} \ln(k^2+k).$$

Solutions

Zeta Function Identity **October 1988**

1302. *Proposed by W. E. Briggs, University of Colorado, Boulder, Colorado.*

Show, for integral $n \geq 3$, that

$$\zeta(n) \equiv \sum_{r=1}^{\infty} \frac{1}{r^n} = \sum_{i=1}^{n-2} \sum_{p,q=1}^{\infty} \frac{1}{p^i (p+q)^{n-i}}.$$

Solution by Nick Franceschine, Sebastopol, California.

$$\sum_{i=1}^{n-2} \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \frac{1}{p^i (p+q)^{n-i}}$$

$$= \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} \left(\sum_{i=1}^{n-2} \frac{1}{p^i (p+q)^{n-i}} \right)$$

$$= \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} \frac{1}{p(p+q)^{n-1}}\left(1+\left(\frac{p+q}{p}\right)+\cdots+\left(\frac{p+q}{p}\right)^{n-3}\right)$$

$$= \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} \frac{(p+q)^{n-2}-p^{n-2}}{p^{n-2}q(p+q)^{n-1}}$$

$$= \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} \frac{1}{p^{n-2}q(p+q)} - \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} \frac{1}{q(p+q)^{n-1}}.$$

Letting $m = p + q$ in the second summation and summing on the diagonals, this is

$$= \sum_{p=1}^{\infty} \frac{1}{p^{n-1}}\left(\sum_{q=1}^{\infty}\left(\frac{1}{q}-\frac{1}{p+q}\right)\right) - \sum_{m=2}^{\infty}\frac{1}{m^{n-1}}\sum_{q=1}^{m-1}\frac{1}{q}$$

$$= \sum_{p=1}^{\infty} \frac{1}{p^{n-1}}\left(1+\frac{1}{2}+\cdots+\frac{1}{p}\right) - \sum_{p=2}^{\infty}\frac{1}{p^{n-1}}\left(1+\frac{1}{2}+\cdots+\frac{1}{p-1}\right)$$

$$= \sum_{p=1}^{\infty}\frac{1}{p^n},$$

as desired.

Also solved by Duane Broline, The Oxford Running Club, Michael Vowe (Switzerland), Morton Zweiback, and the proposer. There was one incorrect solution.

Integral Equation **October 1988**

1303. *Proposed by George T. Gilbert, St. Olaf College, Northfield, Minnesota.*

Find all continuous functions f on $(0, \infty)$ such that

$$\int_x^{x^2} f(t)\,dt = \int_1^x f(t)\,dt \quad \text{for all } x > 0.$$

Solution by Amir Akbary Majdabad No (student), Tehran, Iran.

Differentiate each side of the equality to get

$$2xf(x^2) - f(x) = f(x),$$

$$f(x^2) = \frac{f(x)}{x},$$

$$f(x) = \frac{f(\sqrt{x})}{\sqrt{x}}.$$

A simple induction yields,

$$f(x) = \frac{f\left(\sqrt[2^n]{x}\right)}{x^{\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^n}}} = \frac{f\left(\sqrt[2^n]{x}\right)}{x^{1-\frac{1}{2^n}}}$$

and as $n \to \infty$, using the continuity of f, we have

$$f(x) = \frac{f(1)}{x} = \frac{c}{x}.$$

It is easy to check that $f(x)=\frac{c}{x}$ satisfies the conditions of the problem.

Also solved by Phillip Abbott, AIG Financial Products Problem Group, Donald F. Bailey, Seung-Jin Bang (Korea), Katalin Bencsath, Dipendra Bhattacharya, Irl C. Bivens, Anna Boettcher and Václav Konečný, Duane M. Broline, David C. Brooks, Nicholas Buck (Canada), David Callan, Calvin College Problem Solvers, Centre College Mathematical Problem Solving Group, Onn Chan (student), Chico Problem Group, Jeffrey W. Clark, Bruce Dearden, Jim Delany, Fred Dodd, David Doster, Robert Doucette, Drake University Problem Solving Group, François Dubeau (Canada), Alberto Facchini (Italy), Michael V. Finn, William G. Frederick, E. S. Freidkin (South Africa), Jayanthi Ganapathy, Tom Goebeler (student), Michael B. Gregory, Russell Jay Hendel, G. A. Heuer, Christopher Hunter, Hans Kappus (Switzerland), Roger B. Kirchner, Benjamin G. Klein, Y. H. Harris Kwong, Kee-Wai Lau (Hong Kong), Rebecca Lee, J. C. Linders (The Netherlands), Kim McInturff, Leroy F. Meyers, Andreas Müller (Switzerland), Roger B. Nelsen, Stephen Noltie, Northern Kentucky University Problem Group, Maria Luisa Oliver (Argentina), Gene M. Ortner, Edward Osypowski and Larry Olson, The Oxford Running Club, Richard E. Pfiefer, Robert L. Raymond, Edgar N. Reyes, Laurel Rogers and David Ruch, Hyman Rosen, Giuseppe Russo (Canada), P. K. Sahoo, Volkhard Schindler (East Germany), Harry Sedinger and Charles Diminnie, M. A. Selby (Canada), Raul A. Simon (Chile), Michiel Smid (The Netherlands), Claude C. Thompson, Nora S. Thornber and Karvel K. Thornber, Siu Ming Tong, San Vo, Julius Vogel, Michael Vowe (Switzerland), Walter O. Walker, Gary L. Walls, Edward T. H. Wang (Canada), Abraham Wender, Joseph Wiener, Bettina Zoeller, A. Zulauf (New Zealand), and the proposer.

Binomial Identity **October 1988**

1304. *Proposed by William Moser, McGill University, Montreal, Canada.*

For nonnegative integers, k, n, establish the identity

$$\sum_{i\geq 0}(-1)^i\binom{n-k+1}{i}\binom{n-3i}{n-k}=\sum_{i\geq 0}\binom{n-k+1}{i}\binom{i}{k-i}.$$

(Here, $\binom{n}{k}=0$ when $k<0$ or $k>n$.)

I. *Solution by Michael Vowe, Therwil, Switzerland.*

Let

$$f_m(x)=(1-x^3)^m(1-x)^{-m}=(1+x+x^2)^m.$$

We have

$$(1-x^3)^m(1-x)^{-m}=\sum_{j=0}^{m}(-1)^j\binom{m}{j}x^{3j}\sum_{i=0}^{\infty}\binom{m+i-1}{m-1}x^i$$

$$=\sum_{j\geq 0}x^j\sum_{i\geq 0}(-1)^i\binom{m}{i}\binom{m+j-1-3i}{m-1}.$$

If we set $m=n-k+1$ then the coefficient of x^k, $(j=k)$, is the left side of the given identity.

On the other side, we have

$$(1+x+x^2)^m=\sum_{r=0}^{m}\binom{m}{r}x^r(1+x)^r=\sum_{r=0}^{m}\binom{m}{r}\sum_{s=0}^{r}\binom{r}{s}x^{r+s}$$

$$=\sum_{j\geq 0}x^j\sum_{i\geq 0}\binom{m}{i}\binom{i}{j-i}.$$

If we set $m=n-k+1$ then the coefficient of x^k, $(j=k)$, is the right side of the given identity.

II. *Solution by David Callan, University of Bridgeport, Bridgeport, Connecticut.*

With r in place of $n-k+1$ we wish to prove that

$$\sum_{i\geq 0}\binom{r}{i}\binom{r+k-1-3i}{k-3i}=\sum_{i}\binom{r}{i}\binom{i}{k-i}.$$

Recall that $\binom{r+k-1}{k}$ is the number of combinations of k items selected from r with repetitions allowed. Thus $\binom{r-3i+k-1}{k-3i}$ is the number of such combinations in which a specific set of i items appears at least three times. By the inclusion-exclusion principle, the number of ways of choosing k items from r, with repetitions allowed, in which no item occurs more than twice, is just the left side of the identity, namely,

$$\sum_{I}(-1)^{|I|}\binom{\text{no. of combinations of } k \text{ items from } r \text{ in}}{\text{which each item of } I \text{ occurs at least 3 times}}$$

$$=\sum_{i\geq 0}(-1)^{i}\binom{r}{i}\binom{r+k-1-3i}{k-3i},$$

where I in the first sum ranges over all 2^r subsets of the r items (and $|I|$ is the number of items in I).

The right side counts these combinations directly: $\binom{r}{i}\binom{i}{k-i}$ is the number of combinations in which precisely $k-i$ of the items occur twice.

Also solved by Duane M. Broline, David Callan (second solution), Y. H. Harris Kwong, MAPLE (submitted by Shalosh B. Ekhad, Princeton, New Jersey), The Oxford Running Club, and the proposer.

Callan's second solution shows that the identity holds more generally: for arbitrary real numbers n and integral k,

$$\sum_{i}(-1)^{i}\binom{n-k+1}{i}\binom{n-3i}{k-3i}=\sum_{i}\binom{n-k+1}{i}\binom{i}{k-i}.$$

Moreover, this identity holds even without the author's nonstandard assumption that $\binom{r}{k}=0$ when $k>r$ (which is in conflict with the standard definition of binomial coefficients—it does not hold for negative r).

The MAPLE program, written by *Doran Zeilberger, Mathematics Department, Drexel University,* finds a recurrence satisfied by any given sum of products of binomial coefficients. The proof consists of showing that each side of the identity satisfies the same recurrence, as well as the same initial conditions. Zeilberger describes his method in two papers, "A Fast Algorithm for Proving Terminating Hypergeometric Identities," and "The Method of Creative Telescoping" (available from the author).

Inradii Identity **October 1988**

1305. *Proposed by H. Demir and C. Tezer, Middle East Technical University, Ankara, Turkey.*

Let $P_0=B, P_1, P_2,\ldots, P_n=C$ be points, taken in that order, on the side BC of the triangle ABC. If r, r_i, and h denote, respectively, the inradii of the triangles ABC, $AP_{i-1}P_i$, and the common altitude, prove that

$$\prod_{i=1}^{n}\left(1-\frac{2r_i}{h}\right)=1-\frac{2r}{h}.$$

Solution by Jim Francis, University of Washington, Seattle, Washington.

It suffices to prove the case where $n=2$, since the formula then follows by induction.

From Euclidean geometry, we know that the inradius of any triangle is the quotient of its area by its semiperimeter. Hence, if we let $x_1 = BP_1$, $x_2 = P_1C$, $a_1 = AB$, $a_2 = AP_1$, and $a_3 = AC$, then

$$r_1 = \frac{\frac{1}{2}x_1 h}{\frac{1}{2}(x_1 + a_1 + a_2)} = \frac{x_1 h}{x_1 + a_1 + a_2}$$

$$r_2 = \frac{x_2 h}{x_2 + a_2 + a_3},$$

and

$$r = \frac{(x_1 + x_2)h}{x_1 + x_2 + a_1 + a_3}.$$

This implies that

$$\left(1 - \frac{2r_1}{h}\right)\left(1 - \frac{2r_2}{h}\right) = \left(1 - \frac{2x_1 h}{h(x_1 + a_1 + a_2)}\right)\left(1 - \frac{2x_2 h}{h(x_2 + a_2 + a_3)}\right)$$

$$= \frac{(a_1 + a_2 - x_1)(a_2 + a_3 - x_2)}{(a_1 + a_2 + x_1)(a_2 + a_3 + x_2)}.$$

Similarly we have

$$\left(1 - \frac{2r}{h}\right) = \frac{a_1 + a_3 - x_1 - x_2}{a_1 + a_3 + x_1 + x_2}.$$

It remains to show that the right-hand sides of the above two equations are equal, or equivalently, to show that

$$(a_1 + a_2 + x_1)(a_2 + a_3 + x_2)(a_1 + a_3 - x_1 - x_2)$$
$$= (a_1 + a_2 - x_1)(a_2 + a_3 - x_2)(a_1 + a_3 + x_1 + x_2).$$

Expanding and eliminating that which is common to each side, the right side reduces to

$$(-a_1^2 + a_2^2 + x_1^2)x_2 + (-a_3^2 + a_2^2 + x_2^2)x_1,$$

while the left side reduces to the additive inverse of this expression. Thus it remains to show that the above expression is zero. This follows from the law of cosines as follows.

Let $\alpha = \angle AP_1B$. Then

$$-a_1^2 + a_2^2 + x_1^2 = 2a_2x_1 \cos\alpha$$

while

$$-a_3^2 + a_2^2 + x_2^2 = 2a_2x_2\cos(\pi - \alpha) = -2a_2x_2\cos\alpha,$$

and the proof is complete.

Also solved by S. Belbas, Francisco Bellot-Rosado (Spain), Anna Boettcher and Václav Konečný, Duane M. Broline, Michael V. Finn, John F. Goehl, Jr., Francis M. Henderson, J. Heuver (Canada), Hans Kappus (Switzerland), L. Kuipers (Switzerland), Lamar University Problem Solving Group, J. C. Linders (The Netherlands), Vania Mascioni (Switzerland), The Oxford Running Club, Werner Raffke (West Germany), John P. Robertson, Hyman Rosen, Volkhard Schindler (East Germany), Michael Vowe (Switzerland), A. Zulauf (New Zealand), and the proposer.

Coefficients of Formal Power Series **October 1988**

1306. *Proposed by Carl G. Wagner, The University of Tennessee, Knoxville, Tennessee.*

Given the formal power series

$$C(x)=\sum_{n=1}^{\infty} c(n)x^{n},$$

with $c(n)\geq 0$ for all $n\geq 1$, define the sequence $(d(n))_{n=0}^{\infty}$ by

$$(1-C(x))^{-1}=\sum_{n=0}^{\infty} d(n)x^{n}\equiv D(x).$$

Prove that $d(n)>0$ for all n sufficiently large if and only if the ideal generated by $S=\{n:c(n)>0\}$ is equal to the set of integers.

Solution by Duane M. Broline, Eastern Illinois University, Charleston, Illinois.

We first prove the following well-known result.

LEMMA. *If I is the ideal of $\mathbf{Z}$ generated by the positive integers $n_1, n_2,\dots, n_t$, there is an integer N such that whenever $m\in I$ with $m\geq N$ it is possible to find nonnegative integers $a_1, a_2,\dots, a_t$ with $m=\Sigma a_i n_i$.*

Proof of Lemma. Let M be the maximum of $n_1, n_2,\dots, n_t$ and set $N=M\Sigma n_i$. Suppose $m\in I$ and $m\geq N$. Since I is an ideal, m can be written as a linear combination of $n_1, n_2,\dots, n_t$. We must show it is possible to find such a combination with nonnegative numbers as coefficients. To do this, from the set of all representations of $m=\Sigma a_i n_i$ as a linear combination of $n_1, n_2,\dots, n_t$, choose one in which $\Sigma_{a_i<0}|a_i|$ is minimal. (As usual, the empty sum is defined to be 0.) Suppose $a_j<0$ for some j. Since $\Sigma a_i n_i=m\geq N=\Sigma M n_i$, there exists k such that $a_k\geq M\geq n_j$. Hence

$$m=\sum_{\substack{i=1\\ i\neq j,k}} a_i n_i+(a_k-n_j)n_k+(a_j+n_k)n_j.$$

Since $a_k-n_j\geq 0$ and either $a_j+n_k\geq 0$ or $|a_j+n_k|<|a_j|$, the sum of the absolute values of the negative coefficients in this last linear combination is less than that of $\Sigma a_i n_i$ which contradicts our assumption concerning minimality. The result follows.

Continuing with the solution, we have

$$(1-C(x))D(x)=\sum_{n=0}^{\infty}\left(d(n)-\sum_{k=1}^{n} c(k)d(n-k)\right)x^{n}.$$

Since $(1-C(x))D(x)=1$, we have, by comparing coefficients, that $d(0)=1$ and

$$d(n)=\sum_{k=1}^{n-1} c(k)d(n-k)+c(n),\quad \text{for } n>0. \qquad (*)$$

Let I be the ideal generated by $S=\{n:c(n)>0\}$. Assume first that $I\neq\mathbf{Z}$. Then $I=\{pz:z\in\mathbf{Z}\}$, where p is an integer larger than 1. Observe that if $n\notin I$, then $n\notin S$ and $c(n)=0$. We show by induction that $d(pz+1)=0$ for $z\in\mathbf{Z}$. Evidently, $d(1)=c(1)=0$. Suppose, inductively, that $d(pz+1)=0$ for $z=1,2,\dots,r-1$. Thus

$$d(pr+1) = \sum_{k=1}^{pr-1} c(k)d(pr+1-k) + c(pr+1)$$
$$= \sum_{j=1}^{r-1} c(pj)d(p(r-j)+1) = 0.$$

Hence, it is false that $d(n) > 0$ for all n sufficiently large.

Next, suppose that $I = \mathbf{Z}$. It is easy to show that I can be generated by a finite number, $n_1, n_2, \ldots, n_t$ of elements of S. Let $a_1, a_2, \ldots, a_t$ be nonnegative integers. We show, by induction on Σa_i, that $d(\Sigma a_i n_i) > 0$. If $\Sigma a_i = 0$, the result is clear. Assume the statement is true when $\Sigma a_i < r$, for some $r > 0$. Suppose that $\Sigma a_i = r$ and assume, by renumbering if necessary, that $a_1 > 0$. Then

$$d\left(\sum_{i=1}^{t} a_i n_i\right) = d\left(n_1 + \left((a_1-1)n_1 + \sum_{i=2}^{t} a_i n_i\right)\right)$$
$$\geq c(n_1)d\left((a_1-1) + \sum_{i=2}^{t} a_i n_i\right) > 0$$

by the inductive hypothesis, equation $(*)$, and the fact that $c(n) > 0$.

Since $I = \mathbf{Z}$, the lemma shows that the existence of an integer N such that every integer larger than N can be written as $\Sigma a_i n_i$ for nonnegative integers $a_1, a_2, \ldots, a_t$. By the previous paragraph, $d(n) > 0$ for $n > N$, and the proof is complete.

Also solved by David Callan, Jesse Deutsch, Michael V. Finn, H. K. Krishnapriyan, The Oxford Running Club, and the proposer.

Answers

Solutions to the Quickies on p. 275.

A752. From the binomial theorem, it is easy to show that $(x+1)^p \equiv x^p + 1 \pmod p$. Hence, letting $g(t) = t^p - t$, we have $g(x+1) \equiv g(x) \pmod p$. Replacing x by $x+1$ and inductively iterating this replacement, we get $g(x+a) \equiv g(x) \pmod p$ for any positive integer a. This implies that for any positive integer a, $g(a) \equiv g(0+a) \equiv g(0)$ $\equiv 0 \pmod p$, and the proof is complete.

A753. We have

$$\sum_{k=1}^{n} \frac{1}{k+1}\ln(k^2+k)\ln e < \sum_{k=1}^{n} \frac{1}{k+1}\ln(k^2+k)\ln\left(1+\frac{1}{k}\right)^{k+1}$$
$$= \sum_{k=1}^{n} [\ln(k+1) + \ln k][\ln(k+1) - \ln k]$$
$$= \sum_{k=1}^{n} \left[\ln^2(k+1) - \ln^2 k\right] = \ln^2(n+1)$$
$$= \sum_{k=1}^{n} \frac{1}{k}\ln(k^2+k)\ln\left(1+\frac{1}{k}\right)^{k} < \sum_{k=1}^{n} \frac{1}{k}\ln(k^2+k)\ln e.$$

REVIEWS

PAUL J. CAMPBELL, *editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of the mathematics literature. Readers are invited to suggest items for review to the editors.

Flam, Faye, Frothy physics: scrutinizing the laws of suds, *Science News* 136 (29 July 1989), 72-73, 76.

John von Neumann once became interested in the patterns of bubbles in froth and formulated a law of growth for cross-sections of froth bubbles in two dimensions: Bubbles with more than six sides grow larger, while those with fewer sides shrink. His law has been confirmed by recent empirical research, which has also burst other theories about bubbles. Bubbles undergo first a transient phase of evolution, when disorder grows toward maximum entropy, followed by a "scaling state," during which change takes place but the total number of bubbles with a given number of sides stays the same. But bubbles do not necessarily join all their sides at 120-degree angles (thus contradicting the classical result of the "Steiner problem"), and they do not take on fractal patterns (sorry, fractal lovers!).

Mathematics: The straight side of sliced circles, *Science News* 136 (8 July 1989), 31.

Miklós Laczkovich (Eötvös Loránd University) has solved a generalized "squaring the circle" problem formulated by Alfred Tarski in 1925. A circle can be cut up into pieces that can be rearranged (by translations alone!) into a square of the same area, provided one is not limited to ruler-and-compass operations only. Unlike the famous Banach-Tarski paradox, in which the three-dimensional analog can be done with only 5 pieces, Laczkovich's destruction-reconstruction requires about 10^{50} pieces. The proof generalizes to "almost any plane figure with a mathematically well-behaved boundary."

Bentley, Jon, *More Programming Pearls: Confessions of a Coder*, Addison-Wesley, 1988; viii + 207 pp (P). ISBN 0-201-11889-0

Another collection of superb essays on the practice of computing, substantially revised from their first appearance in the "Programming Pearls" column of the *Communications of the ACM.* Topics run the gamut from back-of-the-envelope calculation and bumper-sticker computer science to the value of profilers, I/O fit for humans, and the selection of algorithms. Every major in computer science should read this book (and every math major, too, since many of them will wind up with jobs in computing).

Fienberg, Stephen E. (ed.), *The Evolving Role of Statistical Assessments as Evidence in the Courts*, Springer-Verlag, 1989; xvii + 357 pp, $34. ISBN 0-387-96914-4

Readable committee report on the problems presented by increasing use of evidence of a statistical nature, with recommendations, case studies, and reviews of selected areas of litigation.

Euler, Leonhard, *Introduction to Analysis of the Infinite*, Book I, transl. John D. Blanton, Springer-Verlag, 1988; xv + 327 pp, $49.95. ISBN 0-387-96824-5

This first English translation was occasioned by a talk by André Weil in which he remarked that students of mathematics would profit more from a study of Euler than from modern textbooks. The work treats infinite series, infinite products, and continued fractions—topics Euler regarded as suitable for learning before calculus (but after trigonometry, logarithms, and exponentials). Today's students would benefit by seeing this book after calculus, as they would then be able to appreciate better its spirit of undaunted calculation.

Peterson, Ivars, Inside moves: a new look at the mathematical problem of turning a sphere inside out, *Science News* 135 (13 May 1989), 299, 301.

Bernard Morin (Université Louis-Pasteur) has found a simplified solution to the problem of turning a sphere inside out smoothly (i.e., without creases). His algorithm reduces the problem from the continuous to the discrete, since it suffices to prescribe the transformations of 12 points on the sphere.

Cipra, Barry A., From real numbers to strings of zeros, *Science* 243 (3 March 1989), 1142-1143.

Report on some of the results presented at the January joint mathematics meetings in Phoenix, including computational complexity results for arbitrary ordered rings, the optimality of linear algorithms for linear problems, progress on arithmetic progressions, and statistical analysis of some of the zeros of the Riemann zeta function (we now know from theoretical considerations that at least two-fifths of the zeros lie on the critical line.).

Schwartz, Richard H., *Mathematics and Global Survival*, Ginn, 1989; xxi + 283 pp (P). ISBN 0-536-57540-1

Revision of previous edition reviewed here Sept. 1983, with up-to-date data. Emphasizes critical issues facing the world today—scarcity, hunger, population growth, pollution, waste—as a way to motivate students with practical problems from everyday life. Designed for students who need review of basic mathematics concepts and skills (arithmetic, calculator use, graphical methods), the book also includes an introduction to probability and descriptive statistics. The book is suitable for either college or high school. Students who learn from this book will never have any doubts about the importance and relevance of mathematics.

Friedman, Avner, *Mathematics in Industrial Problems*, Springer-Verlag, 1988; x + 74 pp, $19.80. ISBN 0-387-96860-1

This book is the fruit of the author spending 10 months visiting numerous British industries and national laboratories and meeting with several hundred scientists to discuss mathematical questions that arise in their work. Each of the two chapters is devoted to one problem or set of problems, with accompanying mathematical problems (some solved, some partially solved) and relevant references. As might be expected, understanding the problems requires some background in physics, plus a mathematical background that is comfortable with vector calculus, boundary-value problems, integral equations, and rigorous "engineering mathematics." Since few US math major undergraduates pusue a classical applied math curriculum, the book is more likely to find its way into a senior seminar in mathematics for engineering students, or into a graduate-level applied mathematics seminar.

Knuth, Donald E., et al., *Mathematical Writing*, MAA Notes No. 14, MAA, 1989; 115 pp, $12.50 (P). ISBN 0-88385-062-1

Excerpts from lectures by the authors and the guest lecturers in a course at Stanford on Mathematical Writing (offered in the Computer Science Dept.!). The blurb for the book calls it "an all-out attack on the problem of teaching people the art of mathematical writing." So it is, mixing advice on mathematical expression with general issues of writing, editing, and evaluating writing. Every mathematician who writes on the blackboard, publishes, edits, or referees should read this book; and there should be a copy in every mathematics department's common room.

Li, Yan, and Shiran Du, *Chinese Mathematics: A Concise History*, trans. John N. Crossley and Anthony W.-C. Lun, Oxford U Pr, 1987; xiii + 290 pp, $49.95. ISBN 0-19-858181-5

Short but comprehensive history of mathematics in China—a rare find.

Cannell, D. M., *George Green, Miller and Mathematician* 1793-1841, City of Nottingham Arts Department, 1988; 80 pp, £5.00 (includes postage and handling) (P). (Requests should be sent to: Green's Mill Science Centre, Belvoir Hill, Sneinton, Nottingham NG2 4LF, U.K.) ISBN 0-905634-17-9

A short account of the life of George Green, the Green of Green's theorem and Green's functions, and the restoration of his mill. Green was a miller to the age of 37 and then devoted his life (10 more years) to mathematics, including starting college at age 40. His life is interesting, and this commemorative volume is full of pictures. (See related article, this MAGAZINE 62 (1989), 219-232.)

Peterson, Ivars, A different dimension: The fourth dimension's extraordinary mathematical properties perplex mathematicians, *Science News* 135 (27 May 1989), 328-330.

In dimensions 1, 2, and 3, topological and differentiable manifolds coincide; in dimensions 5 and higher, topological manifolds "come in both the smooth and crinkly varieties, and mathematicians understand when and how the different types occur." Dimension 4—"where Einstein's theories must work and where modern physics resides"—contains manifolds topologically but not differentially equivalent. In 1981 Michael Freedman (UC-San Diego) showed that certain 4-manifolds "can be constructed from simple building blocks and classified entirely on the basis of their quadratic forms." In 1982 Simon Donaldson (Oxford) showed that not all 4-manifolds can be constructed in a smooth way; he also showed that quadratic forms are not sufficient to distinguish between manifolds that are smooth and those that are not, and he developed subtler invariants. Most immediately relevant, he showed that "ordinary 4-dimensional space can be given innumerable smooth descriptions. In other words, there exist exotic 4-manifolds that are topologically but not smoothly equivalent to standard, 4-dimensional Euclidean space." These exotic 4-dimensional spaces "appear to become extremely complicated at great distances and infinitely complex at infinity"; i.e., complexity increases with scale; and recent astronomical observations may be more compatible with an exotic 4-space than with the standard one.

Boyer, Carl B., and Uta C. Merzbach, *A History of Mathematics*, 2nd ed., Wiley, 1989; xviii + 762 pp. ISBN 0-471-09763-2

After 20 years, a revision of Carl Boyer's famous history of mathematics! The first 22 chapters have been left virtually unchanged, but those on the 19th and 20th centuries have (understandably) been revised and expanded, as have the chapter references and general bibliography (all foreign-language references at the end of chapters have been replaced by recent works in English).

NEWS AND LETTERS

EIGHTEENTH U.S.A. MATHEMATICAL OLYMPIAD (Problems)

1. For each positive integer n, let

$$S_n = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n},$$

$$T_n = S_1 + S_2 + S_3 + \cdots + S_n,$$

$$U_n = \frac{T_1}{2} + \frac{T_2}{3} + \frac{T_3}{4} + \cdots + \frac{T_n}{n+1}.$$

Find, with proof, integers $0 < a,b,c,d < 1000000$ such that $T_{1988} = aS_{1989} - b$ and $U_{1988} = cS_{1989} - d$.

2. The 20 members of a local tennis club have scheduled exactly 14 two-person games among themselves, with each member playing in at least one game. Prove that within this schedule there must be a set of 6 games with 12 distinct players.

3. Let $P(z) = z^n + c_1 z^{n-1} + c_2 z^{n-2} + \cdots + c_n$ be a polynomial in the complex variable z, with real coefficients c_k. Suppose that $|P(i)| < 1$. Prove that there exist real numbers a and b such that $P(a+bi) = 0$ and $(a^2 + b^2 + 1)^2 < 4b^2 + 1$.

4. Let ABC be an acute-angled triangle whose side lengths satisfy the inequalities $AB < AC < BC$. If point I is the center of the inscribed circle of triangle ABC and point O is the center of the circumscribed circle, prove that line IO intersects line segments AB and BC.

5. Let u and v be real numbers such that

$$(u + u^2 + u^3 + \cdots + u^8) + 10u^9$$
$$= (v + v^2 + v^3 + \cdots + v^{10}) + 10v^{11} = 8.$$

Determine, with proof, which of the two numbers, u or v, is larger.

LETTERS TO THE EDITOR

Dear Editor:

In a note entitled A nonconstructible isomorphism, this MAGAZINE 61(1988), 187-188, Padmanabhan observed that

(1) The exponential map $x \to e^x$ is an isomorphism from the additive group of real numbers to the multiplicative group of positive real numbers.

(2) There is no isomorphism from the additive group of rational numbers to the multiplicative group of positive rational numbers.

and attempted to show that

(3) There is an isomorphism from the additive group **A** of real algebraic numbers to the multiplicative group **A*** of positive real algebraic numbers, but it is nonconstructible.

Concerning (3), Padmanabhan correctly pointed out that "it is difficult actually to exhibit such an isomorphism." While he did not claim in the body of the note that the isomorphism in (3) is nonconstructible, the title of the note constitutes such a claim, and this is reinforced by his assertion that "To obtain such an isomorphism, we need ... a Hamel basis ... which ... depends upon transfinite induction." The apparent grounds for this latter remark are that we can construct an isomorphism from **A** to **A*** if we can construct bases for **A** and **A*** as vector spaces over the rational numbers; and to prove that an *arbitrary* vector space has a basis, we need transfinite techniques.

In fact, there is an algorithm for computing an isomorphism from **A** to **A***, although there may be no attractive way to write it down. We do not need transfinite induction, nor is there any apparent way transfinite induction could be used because **A** and **A*** are countable. Transfinite techniques would be needed to *extend* the isomorphism to an isomorphism from the additive group of all real numbers to the multiplicative group of all positive real numbers—in this case we might indeed use a Hamel basis of the reals—but that's another story.

What we do need is a procedure for deciding whether or not a given finite number of elements in **A**, or in **A***, are linearly dependent. Once we have such a procedure we can construct a basis, by induction, from an enumeration of the space.

To get such a procedure requires a lot of algebraic number theory. For **A** we can use the general theory of factorial fields developed, for example, in *A Course in Constructive Algebra*, by Mines et al. (Springer Universitext, 1988), where it is shown that if F is a field of characteristic 0, and we can factor polynomials over F, then we can decide whether any finite set S of elements in an algebraic extension of F are linearly dependent over F or not. For **A*** the constructive theory of Dedekind domains developed in the same text, together with a constructive proof of the Dirichlet unit theorem, enable us to construct a basis for the subgroup of **A*** generated by S; if we write the elements of S in terms of that basis, we can decide whether or not they are dependent.

There is a gap in Padmanabhan's argument that **A** and **A*** are infinite-dimensional. He claims that this is obvious because "the various roots of prime numbers are algebraically independent." By 'algebraically independent' he must mean 'linearly independent' as any set of algebraic numbers are algebraically dependent. Moreover, any two roots of the same prime number are dependent in A*. One interpretation of his claim is that the numbers $\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \sqrt{11}, \ldots$ are linearly independent over the rational numbers in both A and A*. For A* this follows easily from the fundamental theorem of arithmetic, but it is not obvious for A.

Fred Richman
New Mexico State University
Las Cruces, NM 88003

Dear Editor:

We reconsidered the Diophantine problem in the note "A Diophantine Equation from Calculus", this MAGAZINE 62 (1989) 97-101, motivated by the idea of avoiding the calculus. This leads to the following simpler treatment:

Maximizing $x(a - 2x)(b - 2x)$ with $a < b$ is the same as maximizing

$$(2k + 2)x \cdot k(a - 2x) \cdot (b - 2x)$$

for any constant $k > 1$. Since the three factors have a constant sum, it follows from the Arithmetic-Geometric Mean Inequality that the maximum product occurs when $(2k + 2)x = k(a - 2x) = (b - 2x)$ (for a treatment and extensions of this result see [1]). Solving for x, we have

$$x = ak/(4k + 2) = b/(2k + 4),$$

so that $b/a = k(k + 2)/(2k + 1)$.

It is clear that x is rational *iff* k is. We now let $k = r/s$ with $r > s$ and $\gcd(r,s) = 1$. Then $a = s(2r + s)$ and $b = r(r + 2s)$, except that if $r \equiv s \pmod 3$, we take $a = s(2r + s)/3$ and $b = r(r + 2s)/3$. This gives all the primitive solutions.

Reference
1. R.P. Boas and M.S. Klamkin, Extrema of polynomials, this MAGAZINE 50 (1977), 75-78.

M.S. Klamkin and Andy Liu
University of Alberta
Edmonton, Alberta
Canada T6G 2G1

Dear Editor:

We have discovered that there is some overlap in the results contained in two papers authored separately by us, both of which have appeared in *Mathematics Magazine*. They are: J. Kiltinen, Linearity of exponentiation, this MAGAZINE 52 (1979), 3-9, and D. Dobbs, Fields with simple binomial theorem, this MAGAZINE 62 (1989), 52-57. These papers deal with rings, domains and fields in which the identity $(a+b)^n = a^n + b^n$ holds for all a and b. Several of our results are essentially the same, although the approaches are somewhat different, as the interested reader may check.

This topic is one which is accessible to undergraduates, and provides interesting questions for exploration by them. It also leads one to explore some more advanced areas of ring theory, providing motivation as well. Because of this, we have both done further investigations in the subject. Through recent correspondence, we have discovered that our independent work has again run along parallel courses and produced additional unpublished, overlapping results.

We plan to draw our results together in a joint paper, which will include work which Kiltinen did in 1979 and which Dobbs and his student, Bobby Orndorff, did in a Research Experiences for Undergraduates project in 1987. Readers interested in learning of the later work prior to its publication may write to either or both of us.

John O. Kiltinen
Northern Michigan University
Marquette, MI 49855-5340

David E. Dobbs
University of Tennessee
Knoxville, TN 37996-1300

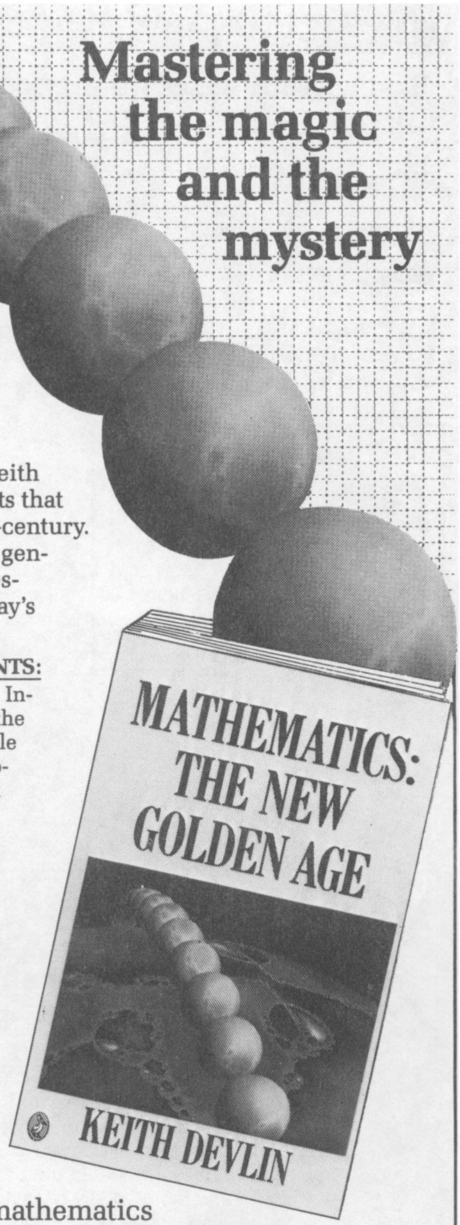
Dear Editor:

In problem #4, 17th U.S.A. Mathematical Olympiad [this MAGAZINE 62 (1989), 210] note that the 6 points $(ABC, A'B'C')$ are *concyclic.* Refer to Altshiller-Court's *College Geometry*, 2nd edition, page 76 and Fig. 52 for proof, with change of notation $A'B'C' \rightarrow KLM$.

Leonard D. Goldstone
808 Sixth Street
Watervliet, NY 12189

1. The stronger result (that the six points are concyclic) was *not* known to us at the time the problem was proposed.
2. This stronger result was discovered by the problem proposer shortly before the 17th Olympiad was printed. We decided to let the problem stand as written. (In our judgement, asking for the stronger result would make the problem harder than we had intended.)
3. Several of the Olympiad contestants found the stronger result.
4. Our published solutions to the 17th U.S.A. Olympiad include the stronger result.

J. Ian Richards
University of Minnesota
Minneapolis, MN 55455



MATHEMATICS

The New Golden Age

Keith Devlin

With a vivid blend of scholarship and originality, Keith Devlin chronicles the ideas, personalities, and events that have revolutionized mathematics in the last quarter-century. Making the most complex concepts accessible to the general reader, Devlin captures in a single volume the essential power and excitement of mathematics in today's "new golden age."

OUTLINE OF CONTENTS:

Prime Numbers, Factoring, and Secret Codes. Sets, Infinity, and the Undecidable. Number Systems and the Class Number Problem. Beauty from Chaos. Simple Groups. Hilbert's Tenth Problem. The Four-Color Problem. Fermat's Last Theorem. Hard Problems About Complex Numbers. Knots and Other Topological Matters. The Efficiency of Algorithms.

0-14-022728-8 304 pp. $8.95

"Devlin has done what most mathematicians consider impossible— convey the excitement and vitality of modern mathematics to the non-specialist."
—Jon Barwise (Stanford University), *Notices of the American Mathematical Society.*

"A remarkable job of describing the profound changes that have taken place in mathematics in the past couple of decades.... The beauty, and the magic, come through clearly."—Herb Wilf (University of Pennsylvania), editor of *American Mathematical Monthly.*

"Devlin's choice of material is excellent, and he is to be praised for the clarity and accuracy with which he presents it."
—Martin Gardner, *The New York Review of Books.*

PENGUIN USA **Academic/Library Marketing**
40 West 23rd Street New York, N.Y. 10010

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, D.C. 20036