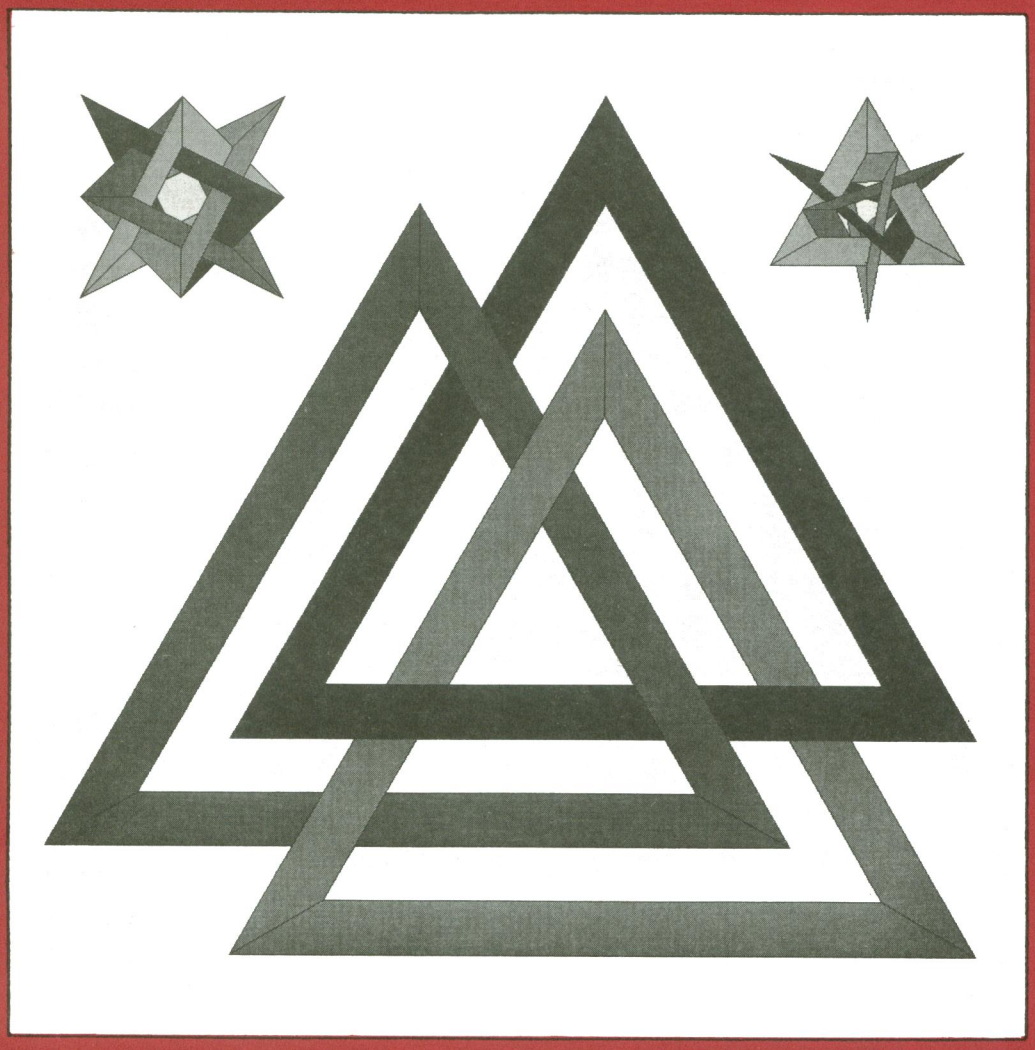
Vol. 69, No. 2 April 1996

MATHEMATICS MAGAZINE



- Inverse Problems for Central Forces
- The Mystery of the Linked Triangles
- Groups, Factoring, and Cryptography

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

Mathematics Magazine aims 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. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 69, pp. 78–79, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Send new manuscripts to Paul Zorn, Editor, Department of Mathematics, St. Olaf College, 1520 St. Olaf Avenue, Northfield, MN 55057-1098. Manuscripts should be laser-printed, with wide spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit three copies and keep one copy. In addition, authors should supply the full five-symbol Mathematics Subject Classification number, as described in *Mathematical Reviews*, 1980 and later. Copies of illustrations should be supplied on separate sheets, both with and without lettering added.

Cover illustration: Linked Triangles, by **Heidi Burgiel. Copyright by The Geometry Center (http://www.geom.umn.edu) for the Regents of the University of Minnesota.**

AUTHORS

Heidi Burgiel is currently a Post-Doctoral Fellow at the Geometry Center in Minneapolis. She received her Ph.D. in mathematics at the University of Washington in Seattle and a B.S. in mathematics at MIT. Her research is in the fields of geometry and combinatorics. When not engaged in research, Heidi works to develop and implement uses for computers in mathematics teaching. She is particularly interested in using the World Wide Web for teaching pre-college geometry and to promote undergraduate research.

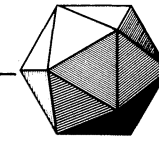
Deborah Franzblau is currently Research Assistant Professor and Education Coordinator at DIMACS (Center for Discrete Mathematics and Theoretical Computer Science), and was previously a member of the mathematics department at Vassar College. Her research interests are in combinatorial optimization; she is currently working on framework rigidity and applications of combinatorial algorithms in materials science. She completed a B.S. in mathematics at U.C. Berkeley and a Ph.D. in mathematics at MIT. She has given several workshops on the topic of this paper to teachers of grades 7–12, and led a 3-week project on linked sculpture for high school students in the DIMACS Young Scholars Program in Discrete Mathematics.

Robert Gutschera received his A.B. in mathematics from Harvard College and his Ph.D. from the University of Chicago. He is now an Assistant Professor of Mathematics at Wellesley College in Wellesley, Massachusetts. His research interests are in differential geometry and dynamical systems. At the same conference where he encountered the mystery of the linked triangles, he was also introduced to origami (by one of his co-authors!), so when he is not teaching or doing research he can often be found making large numbers of origami modules and assembling them into elaborate geometric objects.

Alko Meijer received his B.Sc. degree from the University of Pretoria and his M.Sc. and Ph.D., the latter on the theory of extensions of abelian groups, from the University of South Africa. Since 1971 he has been at the University of Natal. Over the years his interests have shifted from the purest of pure Mathematics to the applications of discrete Mathematics, such as coding theory and, especially, cryptology. He has increasingly come to believe that a philosophy of "art for art's sake" is no longer tenable in mathematics teaching or even research, and, in any case, is unlikely to appeal to students.

After earning a Ph.D. in topology at Columbia University, **Sherman Stein** joined the faculty of the University of California at Davis, when it was expanding. He retired in 1993, when it was shrinking. The present article grew out of reading Newton's *Principia* on the occasion of its tercentenary. In the fall of 1996, John Wiley will be publishing his book, *Strength in Numbers*, designed to make everyone love mathematics.

Vol. 69, No. 2 April 1996



MATHEMATICS MAGAZINE

The *MATHEMATICS MAGAZINE* (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.

The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $16 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $64. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 40% dues discount for the first two years of membership.) The nonmember/library subscription price is $68 per year.

Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Advertising correspondence should be addressed to Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

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

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

PRINTED IN THE UNITED STATES OF AMERICA

ARTICLES

Inverse Problems for Central Forces

S. K. STEIN
University of California at Davis
Davis, CA 95616-8633

Several propositions in Newton's *Principia* tempt one to ask, "What about the converse?" For instance, immediately after Newton showed that a central force implies that area is swept out at a constant rate he obtained its converse (or "inverse"): Area swept out at a constant rate implies a central force. As another example, Newton showed that if the orbits under a central force are conics, then the magnitude of that force varies inversely as the square of the distance. He also established the converse: If the force is inverse-square, then the orbits are conics [**15**]. In 1873 Joseph Bertrand [**3**] went much further, showing that if under a central force all the bounded orbits are periodic, then the force is either linear or inverse-square, that is, its magnitude at a distance r has the form Ar or Br^{-2}, where A and B are constants.

However, two other discoveries in the *Principia* have inspired far more extensive explorations of their inverses, with research beginning with Laplace and continuing for over two centuries to this day. The first discovery is Proposition 60 [**16**; 193]:

> *"If to every point of a spherical surface there tend equal centripetal forces decreasing as the squares of the distances from those points, I say, that a corpuscle placed within the sphere will not be attracted by those forces in any way."*

The second is Proposition 61 [**16**; 193]:

> *"The same things as above, I say, that a corpuscle placed without the spherical surface is attracted towards the centre of the sphere with a force inversely proportional to the square of its distance from that centre."*

Since a ball can be thought of as a union of thin concentric spheres, a similar result holds for it as well.

Each of these two propositions raises three inverse problems. The first proposition, for example, suggests these three questions: "If under a central force a sphere exerts no force in its interior, must that force be inverse-square?" Prescribing the force instead, we might ask, "If under an inverse-square force a surface exerts no force in its interior, must that surface be a sphere?" Most generally, prescribing neither the force nor the surface, we might ask, "If under a central force a surface exerts no force in its interior, must the force be inverse-square and the surface a sphere?" It should be pointed out that only the first of these three questions has been answered.

The second proposition, which concerns the action of a force on external particles, whether for surfaces or solids, also leads to three similar inverse problems. Again, most are not answered.

1. Preliminaries

Throughout we will assume that a unit mass at a point x exerts a force on a unit mass at y that is parallel to the vector $y-x$ (we view points also as vectors) and whose magnitude depends only on the distance between the two points. Letting $r=|y-x|$ and $\hat{r}$ be the unit vector $(y-x)/|y-x|$, we are assuming that there is a scalar function $f(r)$ such that the force is $f(r)\hat{r}$. The function $f(r)$ may assume negative values.

More generally, we will assume that the force exerted by the mass m at x on a unit mass at y is $mf(r)\hat{r}$, and that the total force exerted by several masses at different points is the sum of the individual forces.

Consider a continuous distribution of mass in a region R, with density $\delta(x)$ at x. For instance, assume that R is a surface. A small part of this surface around the point x of area dA has a mass roughly $\delta(x)dA$ and exerts a force approximately $f(r)\hat{r}\,\delta(x)\,dA$ on a unit mass at y. Hence we will use $\int_R f(r)\hat{r}\,\delta(x)\,dA$ to represent the total force exerted by the mass in R on a unit mass at y. In particular, if the surface is a sphere of radius a occupied by a mass of unit density, and described by the usual spherical coordinates, then $dA=a^2\sin\phi\,d\phi\,d\theta$, and the total force is $\int_0^{2\pi}\int_0^{\pi} f(r)\hat{r}a^2\sin\phi\,d\phi\,d\theta$.

The force exerted by a unit mass at the fixed point x, being central, is conservative. Therefore it is the gradient of a scalar function, called the *potential*, which is determined up to a constant. The potential at a point depends only on the distance r from that point to x, and we will denote it $V(r)$. Straightforward differentiation shows that the gradient of $V(r)$ is $V'(r)\hat{r}$. Hence $V'(r)=f(r)$.

If the mass at x is m instead of 1, the associated potential is $mV(r)$. If there are masses m_i at x_i, $1\le i\le n$, the linearity of the gradient shows that $\sum_{i=1}^n m_i V(r_i)$, where $r_i=|y-x_i|$, is a potential for the force exerted by those masses. In the case of a continuous distribution of mass on a surface R, the associated potential would be $\int_R V(r)\,\delta(x)\,dA$, where $r=|y-x|$. Similarly in the case of a solid region it would be $\int_R V(r)\,\delta(x)\,dV$. Newton treated the cases $f(r)$ proportional to r, r^{-2}, r^{-3}, and r^{-5}.

Note that if $f(r)=r^c$, where c is a constant, then

$$f(r)\hat{r}=|y-x|^c\frac{y-x}{|y-x|}=|y-x|^{c-1}(y-x).$$

The case when $f(r)=kr$, where k is a constant, is special, for then any distribution of mass exerts a force on any particle as if all the mass were concentrated at its center of gravity. To see this, note that $f(r)\hat{r}=k(y-x)$. Therefore the force exerted by the mass occupying, say, the solid region R, with density $\delta(x)$, is $\int_R k(y-x)\,\delta(x)\,dV$. Placing the origin of the coordinate system at the center of gravity, we have $\int_R x\,\delta(x)\,dV=0$. Thus the total force reduces to $\int_R ky\,\delta(x)\,dV=ky\int_R\delta(x)\,dV=kMy$, where M is the mass in R. This is the same force as if all the mass were at the center of gravity.

When Newton showed that a sphere, under an inverse-square force, exerts no force at any point P in its interior he used a local-cancellation argument. A narrow double cone with vertex at P intercepts two small patches S_1 and S_2 on the sphere, as shown in FIGURE 1. The forces at P due to the masses in the two patches essentially cancel. (An informal 'infinitesimal' argument goes like this: Let S_i, $i=1,2$ have area dS_i and distance r_i from P. By using the fact that a chord of a circle makes equal angles with the tangents at its endpoints, and similar triangles in planes through P, one shows that S_1 and S_2 are similar, with linear ratio being r_1/r_2. Thus $dS_1/dS_2\approx r_1^2/r_2^2$ or $dS_1/r_1^2\approx dS_2/r_2^2$. This implies that the forces due to the two patches cancel.) This

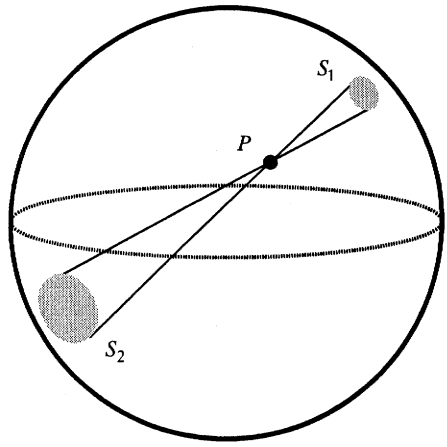


FIGURE 1

result suggests another inverse problem, which we will discuss later.

Each of Sections 2 through 4 is devoted to one type of inverse problem. Section 2 is motivated by the fact that under an inverse-square force a sphere exerts no force in its interior. Section 3 is suggested by the local-cancellation argument. Finally, Section 4 is concerned with an inverse problem associated with the way a ball, under an inverse-square force, attracts external particles. (Throughout, a ball is a solid and a sphere is its surface.)

2. Does no Force Inside a Sphere Imply Inverse Square?

The inverse of Newton's theorem on the force exerted inside a sphere by an inverse-square force played an important role in the early investigations of electricity. Joseph Priestley, in *The History and Present State of Electricity* [**17**], reported in 1767:

> *"I took a small coated phial... and observed that when I held it by the wire, within the electrified cup, it acquired no charge.... May we not infer from this experiment, that the attraction of electricity is subject to the same laws with that of gravitation, and is therefore according to the squares of the distances."*

Henry Cavendish conducted a similar experiment around 1772, but his reasoning was more sophisticated, for he claimed that if the force varied

> *"inversely as some higher power of the distance than the square, the inner globe ought to be in some degree overcharged.... But if the electrical attraction and repulsion is inversely as some lower power of the distance than the square... the inner globe must be undercharged."* [**14**; 8–9, 110–113], [**2**; 98–99]

The assumption that the force varies as some power of the distance is reasonable. As Maxwell put it, "*... if the force were any function of the distance except a power of the distance, the ratio of the forces at two different distances would not be a function of the ratio of the distances....*" [**14**; 86] In other words, if the force is proportional to $f(r)$, physicists expect there to be a function $g(r)$ such that $f(x)/f(y) = g(x/y)$.

Solving this functional equation shows that there are constants A and k such that $f(r) = Ar^k$. Laplace, in the first volume of his *Mécanique Céleste*, published in 1799 [**13**; 298], without assuming that the force varies as a power of the distance, proved that it must be an inverse-square. His proof is expressed in terms of a potential function. The following proof is a variant of Laplace's due to M. Schiffer.

If it can be shown that $V(r) = A + Br^{-1}$ for some constants A and B, it would follow that $f(r) = -Br^{-2}$, and therefore that the force is inverse-square. So assume that *all* spheres exert no force in their interiors. Consider a sphere of radius a whose center is at the origin of the usual spherical and rectangular coordinates with angle ϕ measured from the positive z-axis, having a uniform distribution of mass of density 1. Let P be the point on the z-axis at the height z, $0 < z \leq a$, as in FIGURE 2. (There is no loss of generality since the force exerted by the mass in the sphere is radially symmetric.)

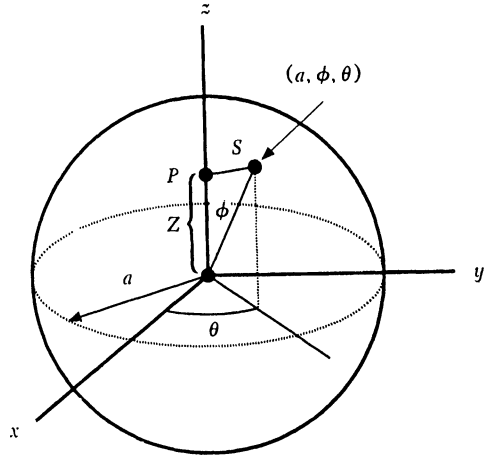


FIGURE 2

Let s be the distance from a point (a, ϕ, θ) on the sphere to the point P. Let $U(z, a)$ be the potential at P due to the mass on the sphere. By the symmetry of the sphere, the potential inside depends only on the distance to the center. Let $U(z, a)$ be the potential at P due to the mass on the sphere. By the assumption that there is no force within the sphere, $U(z, a)$ does not depend on z. We then have

$$U(z, a) = \int_0^{2\pi} \int_0^{\pi} V(s) a^2 \sin \phi \, d\phi \, d\theta, \tag{1}$$

hence

$$U(z, a) = 2\pi a^2 \int_0^{\pi} V(s) \sin \phi \, d\phi. \tag{2}$$

Since $s^2 = a^2 + z^2 - 2za \cos \phi$, we have $s \, ds = za \sin \phi \, d\phi$ and (2) becomes, for $0 < z \leq a$,

$$U(z, a) = \frac{2\pi a}{z} \int_{a-z}^{a+z} V(s) s \, ds,$$

hence

$$zU(z,a) = 2\pi a \int_{a-z}^{a+z} V(s)s\,ds. \tag{3}$$

Since we are assuming that $U(z,a)$ is independent of z, we have

$$U(z,a) = U(0,a) = 4\pi a^2 V(a).$$

Thus (3) becomes

$$4\pi z a^2 V(a) = 2\pi a \int_{a-z}^{a+z} V(s)s\,ds.$$

Letting $F(s) = V(s)s$, we then have

$$4\pi z a F(a) = 2\pi a \int_{a-z}^{a+z} F(s)\,ds. \tag{4}$$

Differentiation of (4) with respect to z gives

$$4\pi a F(a) = 2\pi a[F(a+z) + F(a-z)],$$

hence

$$2F(a) = F(a+z) + F(a-z). \tag{5}$$

Then differentiate (5) with respect to z, obtaining

$$0 = F'(a+z) - F'(a-z). \tag{6}$$

This shows that F' is constant, since for any $0 < x < y$ we have $F'(y) = F'((x+y)/2 + (y-x)/2) = F'((x+y)/2 - (y-x)/2) = F'(x)$. (Here we use the fact that a is arbitrarily large.) Thus there are constants A and B such that

$$F(r) = Ar + B,$$

hence

$$V(r) = A + Br^{-1},$$

and $f(r)$ is therefore $-Br^{-2}$, hence the force is inverse-square.

The preceding proof exploits the assumption that *all* spheres exert zero force in their interiors. If we assume it only for one sphere, of radius a, could we conclude that the force is inverse-square at distances less than $2a$? It is easy to show that the answer is no. To begin, let $V(s)s$ be any nonconstant, smooth function whose graph is symmetric with respect to the point $(a, V(a)a)$. Then

$$\int_{a-z}^{a+z} V(r)r\,dr = 2zV(a)a.$$

Consequently, by the formula preceding equation (3), $U(z,a) = 4\pi a^2 V(a)$. Since the potential function is independent of z (hence independent of r), the force inside the sphere of radius a is 0.

Klamkin and Newman [9], assuming that there is a zero force in the interiors of three spheres, concluded that the force is inverse-square. However, the proof seems to have an unfillable gap.

3. What does Local Cancellation Imply?

Two possible inverses of the local cancellation property of an inverse-square force on a sphere come to mind. In one, the surface is a sphere but the force is to be determined. In the other, the force is inverse-square, but the surface is to be determined. For the first, a sketch and short argument shows that if the ratio of the forces at opposite ends of the cones due to patches on the sphere, S_1 and S_2, approaches 1 as the cone narrows, then the force is inverse-square. (See FIGURE 1.) The case when the force is prescribed as the inverse-square and the surface is to be determined is more involved, as we now show.

Let S be a smooth convex surface and the force be inverse-square. Assume that at each point P in the region surrounded by the surface the opposing forces due to infinitesimal corresponding patches S_1 and S_2 at the ends of any chord through P cancel, in the sense already mentioned. This is equivalent to the assertion that the angles between the chord and the normals to S at the ends of the chord are equal. To show this, let the angle between the chord and the normal to S_i be γ_i, $i = 1, 2$. Let the area of S_i be dA_i and its distance to P be r_i. By elementary geometry,

$$\frac{dS_1 \cos \gamma_1}{dS_2 \cos \gamma_2} \approx \frac{r_1^2}{r_2^2},$$

or

$$\frac{dS_1}{r_1^2} \approx \frac{dS_2}{r_2^2} \frac{\cos \gamma_2}{\cos \gamma_1}.$$

Thus, if the forces due to the two patches cancel, $\cos \gamma_2 / \cos \gamma_1 = 1$, and $\gamma_1 = \gamma_2$. Using the property that the angles between a chord and the normals to S at the ends of the chord are equal, we will show, using an argument due to G. D. Chakerian, that S is a sphere.

Consider the projection of S on a plane Π, which we will take to be the xy-plane. The projection is a region bounded by a curve C, which is the projection of a curve C^* on S. The normals to S at the points on C^* are parallel to Π. (See FIGURE 3.)

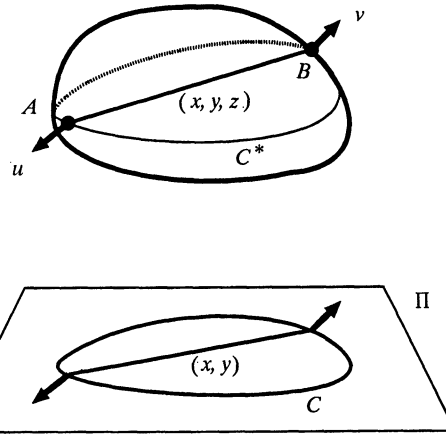


FIGURE 3

The chord AB is represented by the vector (x, y, z). The unit normals at A and B have the form $u = (u_1, u_2, 0)$ and $v = (v_1, v_2, 0)$. Since $(x, y, z)\cdot(u_1, u_2, 0) = (x, y, z)\cdot(v_1, v_2, 0)$, it follows that $(x, y)\cdot(u_1, u_2) = (x, y)\cdot(v_1, v_2)$, which means that the projection of the chord AB makes equal angles with the normals to C at its ends. It is known that C is therefore a circle. (This is a short exercise in the elementary calculus of a curve described in polar coordinates. Place the pole on C and the polar axis on the tangent to C at that point. Consider only chords one of whose ends is at the pole.) Since every projection of S is a disk, S is a sphere, as shown in **[8]** and **[12]**.

Newton also used the cancellation principle when dealing with the force in the region bounded by the inner surface of a spheroidal shell (Corollary 3 of Proposition 91 **[16]**; 221–222]). (A *spheroidal shell* is the region bounded by two similarly situated similar ellipsoids. The ellipsoids have the same centers and one is obtained from the other by a uniform magnification.) Implicit in his reasoning is the conclusion that under an inverse-square force a spheroid exerts no force within the region bounded by the inner ellipsoid. His proof rests on the fact that if P is a point in the interior, any line through P intersects the shell in two line segments of equal lengths. (In FIGURE 4, $DH = IE$.)

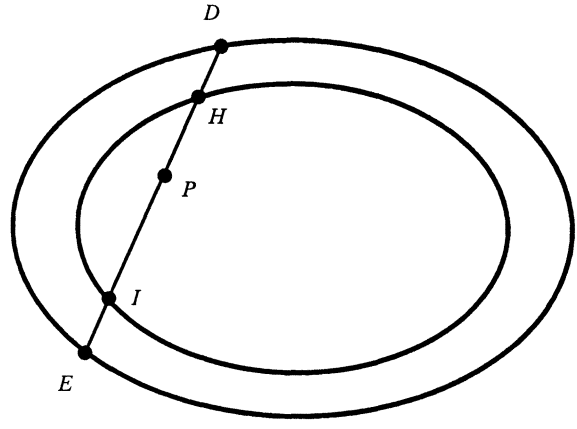


FIGURE 4

(To establish this property, first check it for concentric spheres, then apply an affine map.) By considering narrow opposing cones with vertex P, he then obtains the local cancellation of forces. (Break the parts of the shell within the cones into thin spherical slabs centered at the common midpoint of segments DE and HI.)

One inverse of this result is the following. A shell is bounded by two strictly convex smooth surfaces. If, under an inverse-square force there is a local cancellation of forces at each point in the interior of the inner surface, then the shell must be a spheroid. As we will show, the argument rests on known properties of convex bodies ("known" at least when their boundaries are very smooth).

Using FIGURE 4 now to represent this arbitrary shell, we first note that the local cancellation implies that $DH = IE$. In other words, the two chords, DE and HI, have the same midpoints. It follows (with assumptions of smoothness) that each plane that meets the interior of the inner shell intersects the two surfaces in similarly situated ellipses ([**4**; 40–42, 64, 249]). By **[5]** or **[6**; 91] the two surfaces, which we call S_1 and S_2, are ellipsoids. Now apply an affine mapping that sends the inner surface S_1 to a

sphere S_1'. Let S_2' be the image of the outer surface by that mapping. Then since the cross sections of S_2' by planes through the center of S_1' are circles, S_2' is a sphere, and can easily be seen to have the same center as S_1'. Applying the inverse of the given affine mapping shows that S_1 and S_2 bound a spheroid.

4. Does a Sphere Attracting Like a Point Imply Inverse-Square?

As we mentioned, if the force is linear then any distribution of mass, on a surface or in a solid, homogenous or not, attracts a particle, in its interior or in its exterior, as if all its mass were at its center of gravity. So the assumption that a sphere attracts exterior particles as though the sphere were a point mass does not imply that the force is inverse-square. However, there is an inverse to Newton's theorem, which appears in *Mécanique Céleste* [**13**; 291–297]. Assume that each sphere attracts external particles as if its mass were at a fixed point. By symmetry, that point must be the center of the sphere. Reasoning as in Section 2, we obtain the equation

$$4\pi a^2 V(z) + C(a) = \frac{2\pi a}{z}\int_{z-a}^{z+a} V(s)s\,ds. \tag{7}$$

The function $C(a)$ is introduced since a potential is determined only up to a constant. Again letting $F(r) = V(r)r$, we have

$$4\pi a^2 F(z) + zC(a) = 2\pi a\int_{z-a}^{z+a} F(s)\,ds.$$

(This equation shows that F is infinitely differentiable.) Differentiation with respect to z gives

$$4\pi a^2 F'(z) + C(a) = 2\pi a[F(z+a) - F(z-a)].$$

Differentiating again with respect to z, and simplifying shows that

$$2aF''(z) = F'(z+a) - F'(z-a).$$

Differentiation with respect to a then yields

$$2F''(z) = F''(z+a) + F''(z-a). \tag{8}$$

Differentiating (8) with respect to a shows that $F'''(z+a) = F'''(z-a)$ for all positive a and all z, $0 < a < z$. This implies that F''' is constant, and therefore there are constants A, B, C, D such that $F(r) = Ar^3 + Br^2 + Cr + D$. However, to satisfy (7), B must be 0. Thus $V(r) = Ar^2 + C + Dr^{-1}$, and the magnitude of the force is $2Ar - Dr^{-2}$, hence "linear plus inverse-square." Now, if all balls attract external particles as if their mass were at their centers, then so would spheres, which can be thought of as the difference of two balls of almost the same size. So, if all balls attract external particles as if their masses were at their centers, the force would be the sum of linear and inverse-square forces.

In 1968 Klamkin and Newman [**10**] constructed a nonzero force such that the ball of radius 1 exerts no force anywhere. To do this they considered a potential of the form $\sin(\lambda r)/r$, where λ is any constant other than 0 that satisfies the equation $\tan(\lambda) = \lambda$.

Let B be the ball of radius 1, centered at the origin of rectangular and spherical coordinate systems in their standard relation. The potential at a point a distance $z > 1$

along the z-axis is

$$\int_0^{2\pi}\int_0^1\int_0^{\pi} \frac{\sin\left(\lambda\sqrt{\rho^2+z^2-2\rho z\cos\phi}\right)}{\sqrt{\rho^2+z^2-2\rho z\cos\phi}}\rho^2\sin\phi\,d\phi\,d\rho\,d\theta.$$

The inner integral equals

$$\frac{\rho}{\lambda z}(\cos(\lambda(z-\rho))-\cos(\lambda(z+\rho))).$$

A straightforward integration by parts shows that the triple integral equals

$$\frac{4\pi}{\lambda^3 z}(\sin\lambda-\lambda\cos\lambda)\sin(\lambda z).$$

By the choice of λ, this is 0. (The same result holds for $z<1$.)

In [9] Klamkin and Newman showed that the assumption that two spheres of different radii attract external particles as if the mass of the spheres were concentrated at their centers implies that the force is a sum of linear and inverse-square forces. In 1976 Kondraskov [11] explored another inverse of "a ball attracts exterior particles under an inverse-square force as if it were a point." Keeping the force as inverse-square, he raised the question, "Must the attracting body be a ball, if it attracts external particles as though its mass were at a point?" He showed that if it is also assumed that the boundary of the body is smooth, then the answer is yes. In 1987 Don Zagier obtained the following elementary proof that removes the smoothness condition [19].

Assume that a homogeneous mass of density 1 occupies a region R homeomorphic to a ball and, under an inverse-square force, attracts external particles as if all its mass were at the point O. Place the origin of a rectangular coordinate system at O. Consider the attraction on a unit mass outside R located at the point y. By assumption,

$$\int_R \frac{x-y}{|x-y|^3}dV=\frac{-cy}{|y|^3}, \tag{9}$$

where c is the volume of R. Or, translating this into potentials, we have

$$\int_R \frac{dV}{|x-y|}=\frac{c}{|y|}+d, \tag{10}$$

where d is a constant. (In the integrals y is fixed and x varies over R.) Letting $|y|\to\infty$ in (10) shows that $d=0$, and we have

$$\int_R \frac{dV}{|x-y|}=\frac{c}{|y|}. \tag{11}$$

If O were outside of R the right side of (11) would be unbounded as $y\to O$, but the left side would remain bounded. Thus O is in R. A similar argument shows that O is actually in the interior of R.

Now, if y is not in R, taking the inner product of y with each side of (9) gives

$$\int_R \frac{x\cdot y-|y|^2}{|x-y|^3}dV=\frac{-c}{|y|}.$$

Adding this to (11) gives

$$\int_R \left(\frac{x \cdot y - |y|^2}{|x-y|^3} + \frac{1}{|x-y|} \right) dV = 0.$$

Since $|x-y|^2 = |x|^2 - 2x \cdot y + |y|^2$, this simplifies to

$$\int_R \frac{|x|^2 - x \cdot y}{|x-y|^3} dV = 0. \tag{12}$$

Since the integrand is a continuous function of y, it can be shown that (12) holds even when y is on the boundary of R. Now let B be the largest ball with center O contained in R. Let y be a point on the boundary of R that lies in B, hence on the boundary of B as shown in FIGURE 5. Since a ball attracts an external particle as if its

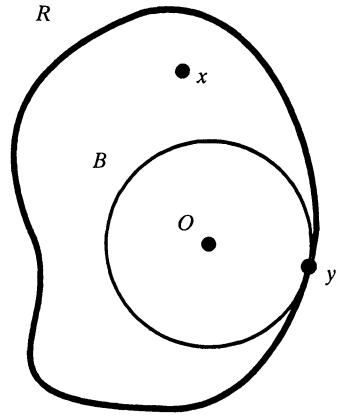


FIGURE 5

mass were at its center, we also have

$$\int_B \frac{|x|^2 - x \cdot y}{|x-y|^3} dV = 0. \tag{13}$$

Subtracting (13) from (12) gives

$$\int_{R-B} \frac{|x|^2 - x \cdot y}{|x-y|^3} dV = 0. \tag{14}$$

For x in $R - B$, $|x|^2 - x \cdot y \geq 0$. Since the integrand in (14) is continuous and nonnegative, it follows that $R - B$ is empty. Thus $R = B$, so in the language of **[1]**, "R must be an orange."

Several generalizations of this result and other inverse problems are discussed in that reference and in **[18**; 214]. But we need not look far for an unsolved inverse problem. After all, the preceding theorem concerns solids, not surfaces. So we should raise the following problem: A homogeneous mass distributed on a surface S homeomorphic to a sphere attracts external particles by the inverse-square law as if all its mass were at a point. Must S be a sphere?

Clearly, the *Principia* still raises inverse problems that are easy to state but challenge the array of analytical tools developed in the three centuries since its appearance.

Acknowledgments. The author wishes to thank G. Donald Chakerian and Anthony Barcellos, who suggested several improvements in the exposition. The illustrations were prepared by Barcellos, using CoHort Graphics.

REFERENCES

1. D. Aharanov, M. M. Schiffer, and L. Zalcman, Potato kugel, *Israel J. Math.* **40** (1981), 331–339.
2. A. J. Berry, *Henry Cavendish*, Hutchison, Alan, Ltd. London, 1960.
3. J. Bertrand, Mécanique analytique, *Comptes Rendus de l'Acad. des Sciences* **77** (1873), 849–853.
4. W. Blaschke, *Vorlesungen über Differential Geometrie*, Vol. 2, Chelsea Publishing Co., New York, 1967.
5. W. Blaschke and G. Hessenberg, Lehrsätze über konvexe Körper, *Deutsche Mathematiker-Vereinigung*, Leipzig Jahresbericht 26 (1917), 215–220.
6. H. Busemann, *The Geometry of Geodesics*, Academic Press, New York, 1955.
7. H. Cavendish, *The Electrical Researches of Henry Cavendish*, Vol. 2, J. C. Maxwell, ed., Cambridge University Press, London, 1879 (Reproduced by Frank Cass and Co., London, 1967).
8. M. Fujiwara, Unendlichviele Systeme Gleichungen mit unendlichvielen Variablen und eine Eigenschaft der Kugel, *Sci. Rep. Tohoku Univ. Ser. 1*, 3 (1914), 199–216.
9. M. S. Klamkin and D. J. Newman, On some inverse problems in potential theory, *Quarterly of Applied Math.* 26 (1968), 277–280. Their argument is marred by an error, one monomial being used to remove *two* constants. However, it is easily repaired, as follows. They wish to solve two equations of the form $g(x+h)-g(x-h)=2hg'(x)+c_h hx$, where h takes on two values and the c_h are corresponding constants. Fortunately, taking the second derivative of these equations removes the linear terms. Their argument applies to g'', and two integrations give g itself.
10. M. S. Klamkin and D. J. Newman, On some inverse problems in dynamics, ibid. 281–283.
11. A. V. Kondraskov, On the uniqueness of a solution of the inverse problem in potential theory, *Dokl. Akad. Nauk SSSR* 245 (1979), 1045–1047; translated in *Soviet Math Dokl.* 20 (1979), 387–390.
12. T. Kubota, Einfach Beweise eines Satzes über die konvexe, geschlossene Fläche, *Sci. Rep. Tohoku Univ., Ser. 1*, 3 (1914), 235–250.
13. P. S. Laplace, *Celestial Mechanics*, Vol. 1, translated by N. Bowditch, Chelsea Publishing Co., New York, 1966.
14. J. C. Maxwell, *A Treatise on Electricity and Magnetism*, Vol. 1, 3rd edition, Oxford, U.K. 1873.
15. M. Nauenberg, Newton's *Principia* and inverse square orbits, *College Math. J.* 25 (1994), 212–221.
16. I. Newton, *Mathematical Principles*, F. Cajori, ed., University California Press, Berkeley, CA, 1934.
17. J. Priestley, *The History and Present State of Electricity*, Vol. 2, 1775, facsimile reprint, Johnson Reprint Corp., New York, 1966.
18. *The Scottish Book*, R. D. Mauldin, ed., Birkhäuser Boston, Boston, MA, 1979.
19. L. Zalcman, Some inverse problems of potential theory, *Contemporary Math.* 63 (1987), 337–350.

The Mystery of the Linked Triangles

H. BURGIEL
Geometry Center
University of Minnesota
Minneapolis, MN 55454

D. S. FRANZBLAU
DIMACS
Rutgers University
P.O. Box 1179
Piscataway, NJ 08855-1179

K. R. GUTSCHERA
Wellesley College
Wellesley, MA 02181-8201

In July of 1993, the authors participated in a workshop[1] for high school teachers led by H. S. M. Coxeter. He began by showing a photograph of John Robinson's sculpture "Intuition," which consists of three hollow triangles linked like the Borromean rings. He then showed a model of the sculpture, stating that in each triangle, the ratio of the lengths of the outer and inner edges was two-to-one. He demonstrated that the model was not rigid, and in fact would "fall" from the position shown in the sculpture to lie flat in the plane if not restrained, and asked us to describe the "lifted" configuration of triangles depicted in the sculpture. Here we explain how to build a similar model, and tell how we discovered that the model was deceptive. Using elementary geometry, trigonometry, and algebra, we show that it is in fact impossible to construct the sculpture unless the ratio of the sides' lengths is strictly smaller than two-to-one. We conclude with some questions about knots and symmetry that arise from this problem.

1. A Linked-Triangle Sculpture and an Intriguing Model

The Australian sculptor John Robinson has built a number of striking outdoor sculptures made of interlocked hollow polygons. He sent a copy of his book of photographs of the sculptures **[1]** to the well-known geometer H. S. M. Coxeter, now at the University of Toronto. One sculpture in particular, called "Intuition" (FIGURE 1), caught Coxeter's attention because it resembled a sculpture (FIGURE 2) by American artist George Odom that he had studied previously **[2]**. Robinson's sculpture is made up of three identical hollow equilateral triangles linked in the manner of Borromean rings (all three triangles are linked, but no pair is linked), and it has an axis of threefold rotational symmetry.

[1]The workshop was organized by Doris Schattschneider and was one of the activities at the NSF-funded (grant number DMS-90 13220) Regional Geometry Institute, organized by Marjorie Senechal and held at Smith College in July, 1993.

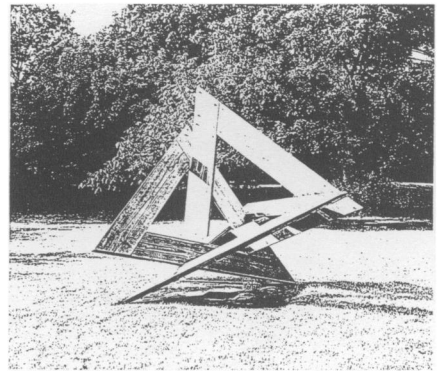


FIGURE 1
Intuition

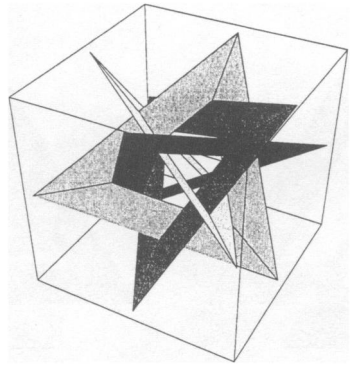


FIGURE 2
Odom's 4-triangle sculpture

To study the sculpture, Coxeter made a model that can be constructed as follows. (This description is adapted from **[2]**, FIGURES 6–9.) Draw three hollow triangles on thin cardboard or heavy paper (copy FIGURE 3), and cut carefully along the solid lines. (Although only one triangle need be cut all the way through, extra lines are added to make cutting easier.) The edges of the outer triangle should be twice as long as those of the inner triangle. (These are the measurements Coxeter assumed, based on his work with Odom's triangle sculpture, although you will see that it helps to make the length of the edge of the outer triangle slightly less than twice the length of the edge of the inner triangle.)

Imagine that the triangles are colored grey, black, and white as in FIGURE 4 (you may wish to actually color them). Tape the grey and black triangles to form closed rings, and lay the grey over the black as in FIGURE 4a. Weave in the white triangle, and tape closed, to get a configuration as in FIGURE 4b. The gray should lie atop the black, the black atop the white, and the white atop the grey. (Be patient; it may take several tries to get the correct linking.)

Now lift the structure, supporting only the outermost corners of the triangles, while preserving threefold symmetry. The opposite corners should slide inward until they stop at a final position, where two outer edges of each triangle touch two inner

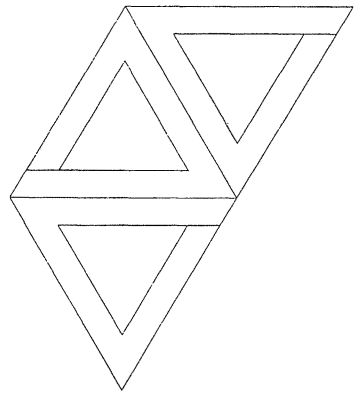


FIGURE 3
Cutting pattern for model

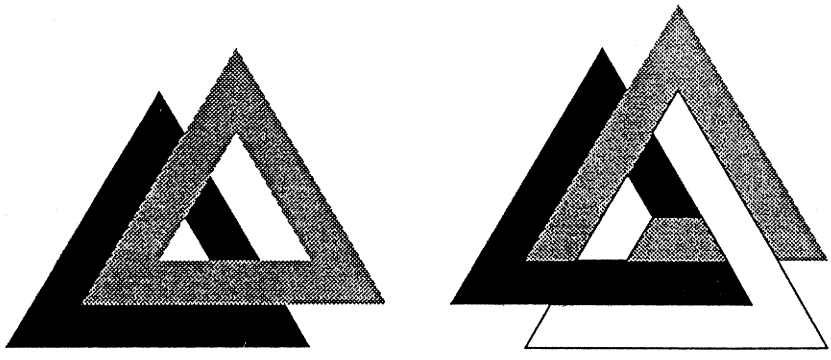


FIGURE 4
Assembling of model (a) 2 triangles, (b) 3 triangles

corners of the next triangle (you may need to shake the model gently). You should now have a three-dimensional model that looks like the original sculpture (FIGURE 1).

If you turn the structure over, and look directly downward, you should see the pattern of FIGURE 5. The point O is the axis of symmetry. Note that points where the boundaries of the triangles meet (points P and Q in FIGURE 6) are *not* the midpoints of edges AB and AC. Coxeter asked why the model stops at that position, and what the lengths AQ and AP must be, given that $AB = AC = BC = 2$ and $DE = DF = EF = 1$. (For simplicity, we use XY to refer either to an edge with endpoints X and Y or its length, depending on the context.)

In the remainder of this paper we describe our approach to Coxeter's problem and our conclusions. We greatly enjoyed working on this problem, and encourage you to work on it before looking at our solution. (Although we chose an algebraic approach to the problem, you could approach this as an experimental problem, by carefully measuring the sculpture, or attempt to solve it using classical geometry; a brief sketch of this last approach is given in Section 4 below.) In the next section, we describe our observations. An outline of our solution is given in Section 3, and in the final section, Section 5, we suggest several related problems.

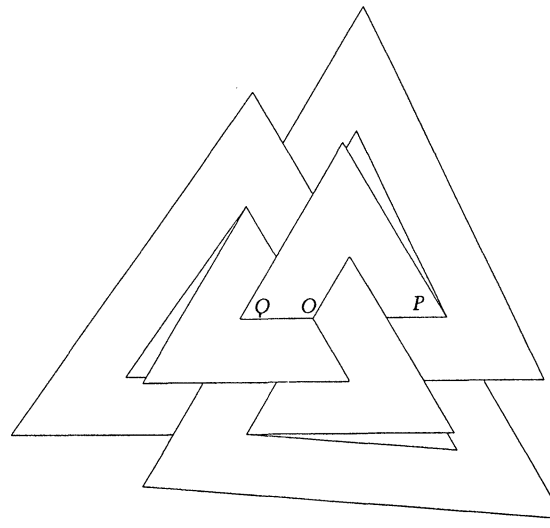


FIGURE 5
View of upside-down model, from above

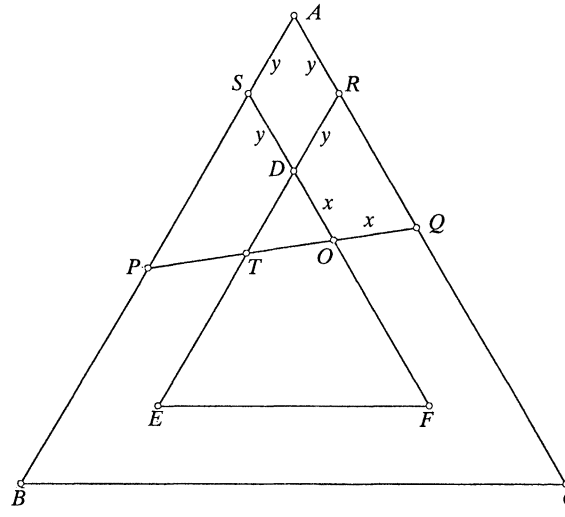


FIGURE 6

2. A Deception Uncovered

When Coxeter presented the problem, we mistakenly believed that he already knew the solution. We were thus surprised when our attempts to solve the problem led to what seemed an impossible conclusion—the sculpture could not exist!

The rotational symmetry of the sculpture allows us to restrict our study to one triangle (FIGURE 6). We assume that at the limit of the sculpture's motion this triangle (triangle ABC) lies against the inner edge PQ of the triangle on top of it. We began our work assuming that $AC = 2$ and $DF = PQ = 1$. During the workshop, we made two crucial observations. First, because of the threefold symmetry about point O, the

lengths OD and OQ must be equal. Second, the triangles APQ, RTQ, and SPO must be similar.

After the workshop we continued to work independently. Eventually we came to the conclusion that the only solution that satisfied all the constraints was the trivial one $AP = AQ = 1$, in which the sculpture lies flat in the plane! We did not believe our result at first, having been thrown off the track by rumors that others in the workshop *had* found a nontrivial solution. Eventually, however, we were convinced by our calculation and realized that we had been tricked by the model.

3. The Mystery Resolved

Despite all mathematical evidence to the contrary, we still had several cardboard sculptures stubbornly doing the impossible. How could this be? Perhaps the flexibility of the cardboard allowed our sculpture to reach a position that it otherwise could not. After some experimentation, we decided that if $DF = 1$ then, in order to obtain a nontrivial solution, AB must be less than 2. Letting $AB = \rho$, where ρ (also equal to AB/DF) is close to two, we were able to show that $\alpha = AQ$ and $\beta = AP$ (these are the numbers that Coxeter wanted) are in fact functions of ρ. From this calculation, one can see that the sculpture can be lifted into a three-dimensional form if, and only if, $1 < \rho < 2$, and also that if $\rho = 2$ the sculpture must lie flat in the plane. The details of our calculation are as follows.

We assume that $PQ = 1$, i.e. that the outer edges of one triangle actually touch the inner corners of the next triangle. (See FIGURE 6.) We also know that $\theta = \angle PAQ = 60°$. Hence, we can apply the law of cosines to triangle APQ to get $\alpha^2 + \beta^2 - 2\alpha\beta\cos\theta = 1$, or

$$\alpha^2 + \beta^2 - \alpha\beta - 1 = 0. \tag{1}$$

Using the threefold rotational symmetry of the model, we see that $OQ = OD$, and by construction $DSAR$ is a rhombus. We describe α and β in terms of the lengths $x = OD$ and $y = DS$, and then use the above relations to simplify (1).

The similarity of triangles APQ and SPO gives us the two relations, namely

$$\frac{AQ}{PQ} = \frac{SO}{PO} \quad \text{and} \quad \frac{AQ}{AP} = \frac{SO}{SP}.$$

When we write all these edge lengths in terms of x and y, we obtain

$$\alpha = \frac{x+y}{1-x} \quad \text{and} \quad \frac{\alpha}{\beta} = \frac{x+y}{\beta-y}, \tag{2}$$

which simplifies to

$$\alpha = \frac{x+y}{1-x} \tag{3}$$

and

$$\beta = \frac{y}{x}. \tag{4}$$

We can now substitute (3) and (4) into (1), to obtain

$$\frac{(x+y)^2}{(1-x)^2} - \frac{y(x+y)}{x(1-x)} + \frac{y^2}{x^2} - 1 = 0.$$

Some algebraic manipulation yields

$$(3y+2)x^2+(-3y-1)x+y=0. \tag{5}$$

The following argument, due to organizer Doris Schattschneider and teacher Elizabeth Whitcomb, allows us to express y in terms of ρ. This will enable us to solve (5) for x in terms of ρ. Drop perpendiculars from points D and E to points U and V on edge AB as shown in FIGURE 7. Since $AB=\rho$ and $UV=DE=1$,

$$y+SU=\frac{1}{2}(AB-UV)=\frac{1}{2}(\rho-1).$$

Since UDS is a 30-60-90 triangle, $SU=y/2$. Substituting this into the equation above and solving gives

$$y=\frac{\rho-1}{3} \quad \text{or} \quad \rho=3y+1. \tag{6}$$

We now write Equation (5) in terms of x and ρ to see

$$(\rho+1)x^2-\rho x+\frac{\rho-1}{3}=0, \tag{7}$$

and thus

$$x=\frac{3\rho\pm\sqrt{12-3\rho^2}}{6(\rho+1)}.$$

Now that we know x and y in terms of ρ, we can use Equations (3) and (4) to see how the sculpture locks into place at a certain point. We can also see that for $\rho=2$, the sculpture cannot exist in three dimensional form.

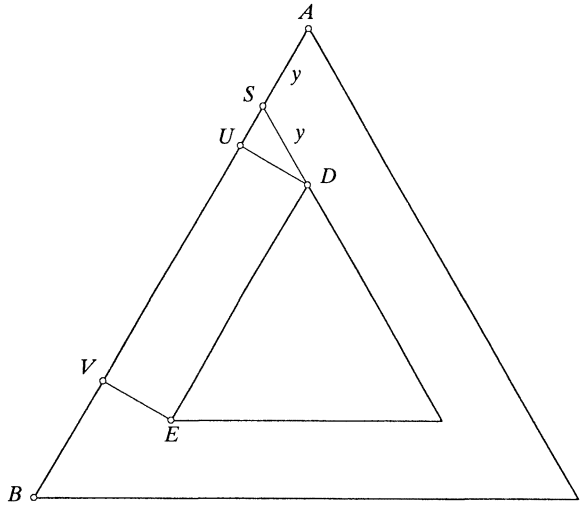


FIGURE 7

For each $\rho=AB$ less than two, there are two possible values of x. (See FIGURE 8.) The smaller root corresponds to the x-value where the sculpture stops moving, and the larger root tells us where the movement would stop if the triangles were arranged grey atop white atop black atop grey instead of grey atop black atop white atop grey,

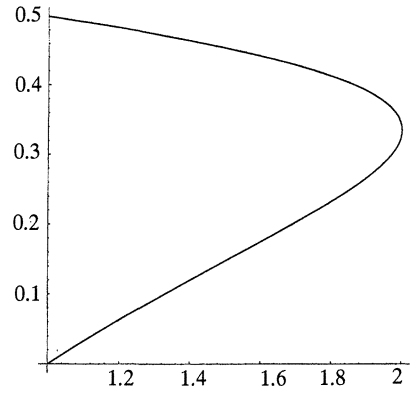


FIGURE 8
Plot of x versus ρ

i.e., if the roles of α and β were reversed. This is plausible intuitively, since neither FIGURE 6 nor the above calculation contains any information regarding which triangles lie atop which others, and yet this certainly affects the final configuration (as one can see by experimenting with the paper model). It can also be verified by solving algebraically for α and β in terms of ρ, yielding

$$\alpha = \frac{1}{6}\left(3\rho \pm \sqrt{12 - 3\rho^2}\right) \quad \text{and} \quad \beta = \frac{1}{6}\left(3\rho \mp \sqrt{12 - 3\rho^2}\right). \tag{8}$$

Note that if $\rho = 2$, then $y = 1/3$ and the only solution to (5) is $x = 1/3$, in which case $\alpha = \beta = 1$. In other words, if $AB = 2$ and $PQ = 1$, the only possible arrangement of hollow triangles with threefold symmetry is the one in which they all lie in the plane.

At the end of [**2**], Coxeter suggests that the ratio should be $(2\sqrt{6} + 1)/3 \approx 1.9663$ to 1 instead of 2 to 1. In our notation, this corresponds to $\rho \approx 1.9663$, $x \approx 0.2959$, $\beta = AP \approx 1.0887$, and $\alpha = AQ \approx 0.8777$. We see from this and from FIGURE 8 that when ρ is close to two, a very small relative change in the value of ρ (from 2.0000 to 1.9663, or 2%) results in a large relative change in x (from 0.3333 to 0.2959, or 11%). This explains why even when one tries to make $\rho = 2$ in the actual sculpture, it often has enough flexibility to "lift up" into Robinson's configuration.

Coxeter also asked why the motion of the model "stops abruptly." Our argument seems to show that during the sliding, the length of PQ is always shorter than the inner length $DF = 1$ (FIGURE 6). If threefold symmetry is maintained, the sliding must stop when $PQ = DF$.

4. A Classical Geometric Approach

Note from (8) that $\alpha + \beta = \rho$ (i.e. $BP = AQ$). Observing this curious fact and thinking about the symmetries it implied, we discovered an alternate approach to the problem, which we sketch briefly.

Consider the two equilateral triangles ABC and DEF (with side lengths ρ and 1 respectively) as before, both centered around the point U (see FIGURE 9). Draw the circle determined by DEF; note that if $\rho = 2$ this circle touches ABC in 3 places, but if $1 < \rho < 2$ then the circle intersects ABC in 6 places (as in FIGURE 9). Symmetrically choose 3 of these intersection points to form a new equilateral triangle VWX. Since

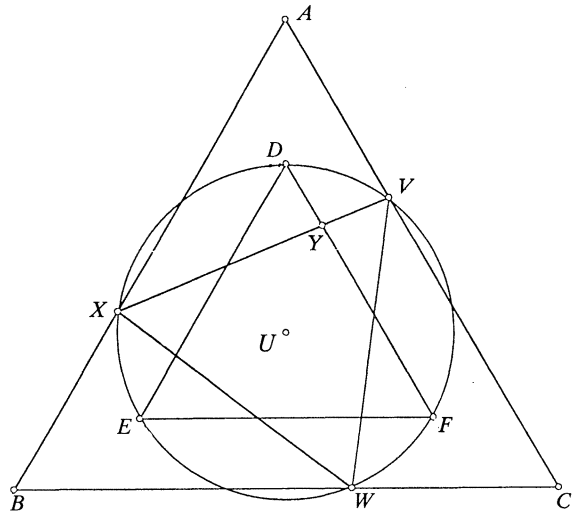


FIGURE 9

VWX is also equilateral, its sides have length 1. Let Y be the intersection of XV and DF as in FIGURE 9. One can show that there is a mirror symmetry along the line through the segment UY and so $DY = VY$. Thus taking PQ to be the segment XV, and letting $O = Y$, we have constructed a "solution to the sculpture," i.e., a segment PQ of length 1, with endpoints on AB and AC, such that $OD = OQ$ (as in FIGURE 6). The fact that $\alpha + \beta = \rho$ follows naturally from the symmetry of this construction (i.e., if $AV = \alpha$ then $VF = XA = \beta$). Now the algebra is greatly simplified: one can use $\alpha + \beta = \rho$ directly with Equation (1) to obtain Equation (8).

To see that any solution is of this form, suppose we have a segment PQ as in FIGURE 6 (so $OD = OQ$). Let U be the common center of the two triangles ABC and DEF. Since $OD = OQ$, we have $UD = UQ$, so Q lies on the circle centered at U that passes through D, E, and F, i.e., Q is the same point as V in FIGURE 9. Since P is the unique point of intersection of the segment AB with the circle of radius 1 about $Q = V$, P must correspond to X, so the segment PQ is indeed the segment XV constructed above.

5. Epilogue

Robinson made another sculpture, "Creation," using three hollow *squares*, still linked as Borromean rings (FIGURE 10). Studying a sheet metal copy[2] of this sculpture suggests that when the ratio of outer to inner edge lengths is four to three, the sculpture is rigid. What other hollow shapes can be assembled with the same linking and threefold symmetry? Perhaps arbitrary regular polygons, circles, or ellipses will work. What ratios between the outer and inner edges of the figure allow construction of these figures? Where do the figures intersect?

[2]Given to University of Washington Professor Lee Stout by the Centre de Recerca, Institut d'Estudis Catalans at Barcelona.

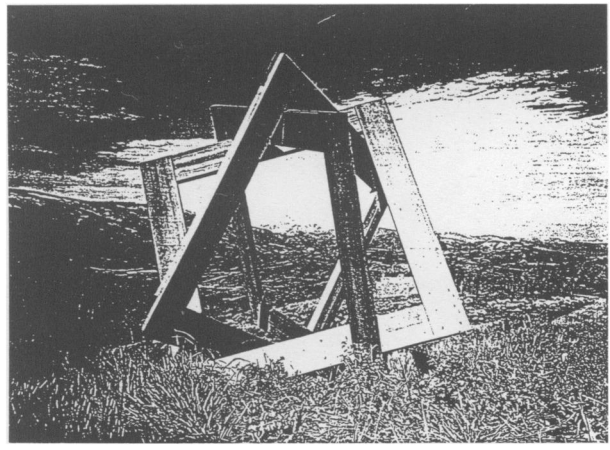


FIGURE 10
Creation

There is no need to limit ourselves to three loops. For example, George Odom's sculpture (FIGURE 2) is constructed from four hollow triangles and displays all the symmetries of a cube [2], including a fourfold rotational symmetry. In [3], Alan Holden discusses the symmetries and rigidity of many different structures made of polygons, which he has built using dowels with circular cross section.

All the sculptures discussed in this paper are made up of three-dimensional objects (not ideal planar triangles and squares); one could also ask how changes in the diameter of the dowel or the thickness of the cardboard affect the properties of the sculptures.

Robinson's triangles and squares are arranged in the configuration of the Borromean rings; no two are linked, but the three hollow polygons are inseparable. In George Odom's sculpture, each pair of triangles is linked. If a set of n closed loops are linked so that all n are linked, but no proper subset of loops are linked, the resulting structure is called a *Brunnian link*. (The Borromean rings are a Brunnian link for $n = 3$.) These are named for Hermann Brunn, who wrote about these links in 1892. (See [4], p. 67.) Are there any other values of n for which there are Brunnian links constructed from n hollow polygons with an axis of n-fold rotational symmetry? In particular, can one build a Brunnian link out of four hollow triangles? What about four squares?

Acknowledgments. We would like to thank Professor H. S. M. Coxeter for posing this problem, and Professor Doris Schattschneider for organizing the program that made this possible. Everyone who attended the workshop contributed to the solution of the problem, and Professor Robert Jamison was responsible for the Geometer's Sketchpad drawing that finally convinced us that our results were correct. We would like to thank the National Science Foundation for sponsoring the Regional Geometry Institute where the workshop took place. We would also like to thank John Robinson for giving us a copy of his book of sculpture photographs [1], several additional photographs, and permission to reproduce them for this article.

REFERENCES

1. J. Robinson, *Symbolic Sculpture*, Edition Limitée, Carouge-Geneva, Switzerland, 1992.
2. H. S. M. Coxeter, Symmetrical Combinations of 3 or 4 Hollow Triangles, *The Mathematical Intelligencer*, Vol. 16, No. 3 (1994), 25–30.
3. A. Holden, *Orderly Tangles: Cloverleafs, Gordian Knots, and Regular Polylinks*, Columbia University Press, New York, 1983, pp. 45–84.
4. D. Rolfsen, *Knots and Links*, Publish or Perish, Inc., Berkeley, CA, 1976, p. 67.

Groups, Factoring, and Cryptography

A. R. MEIJER
University of Natal
Durban 4001, South Africa

If G is a finite group, and $g \in G$, the order of g is, of course, the smallest positive integer n such that $g^n = e$, the identity element of G. Finding the order of an element $g \in G$ is therefore, in theory, not a problem: Just keep multiplying until you get to e. In practice, however, this may not work. In this article, we shall be considering elements where this procedure would involve some 10^{150} multiplications. Even if these multiplications could be carried out at the rate of 1000 billion per second, it would take approximately 3×10^{80} years to arrive at the answer.

The problem of finding the order of an element of a group G becomes much more tractable if the order $|G|$ of G as well as the factorization of $|G|$ are known, since by Lagrange's theorem the order of g is a divisor of $|G|$. Thus, if the factorization of $|G|$ into powers of primes is $|G| = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ then the order of g is $p^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k}$ for some $\beta_1, \dots, \beta_k$ with $\beta_i \le \alpha_i$ for every $i = 1, \dots, k$. Then one only needs to consider $(1 + \alpha_1)(1 + \alpha_2) \dots (1 + \alpha_k)$ possibilities for the order of g. This is a not unreasonable task. In fact, efficient algorithms for determining the order of g under these circumstances exist, such as Algorithm 1.4.3 of [**2**], for example.

This article is concerned with the converse of the above observation: We shall prove that finding the order of a group element is, in general, at least as hard as factoring, in the sense that if one had a fast method of finding orders, one would also have a fast method of factoring. This result may well be part of the cryptological folklore, but this author has not been able to find it explicitly stated.

The observation is of some practical importance, since the security of the RSA public key cryptosystem (which will be described below) depends, as far as anybody knows, on the difficulty of factoring products of large primes. This article arose out of an early, unsuccessful attack on the RSA system and an interesting controversy that is recorded in the papers [**10**], [**6**], [**4**], and [**7**]. (All these papers may conveniently be found together in [**1**].)

The RSA System

Two parties who wish to communicate in secret over a channel to which outsiders may have access must, of necessity, do so by encrypting their messages. Until about 1976, this meant that they first had to agree on the method of encipherment, and, in particular, about the key to be used. Agreement about the key could not be reached while using the insecure channel itself. If large numbers of users want to communicate in secret, this problem of the distribution of keys becomes quite intractable.

In 1976, Diffie and Hellman published their seminal paper [**3**], in which the use of so-called "one way functions" was suggested as a means toward "public key cryptography." In public key cryptography, the keys for encryption and decryption are such that one cannot obtain either, even when given knowledge of the other. Thus it becomes possible for a user to publish the key to be used for encryption, so that anybody who

wishes to send him or her an encrypted message can do so, while keeping the key for decryption secret, so that only he or she can read the messages addressed to him or her.

One of the first of these public key systems, and still "a *de facto* standard in much of the world" [**9**, p. 288], was the RSA system, invented by Rivest, Shamir and Adleman [**8**]. It works as follows:

User R (the receiver) chooses two large primes p and q, finds their product $n = p \times q$ and chooses an integer e that is relatively prime to the Euler totient $\phi(n) = (p-1)(q-1)$. The primes p and q must be chosen large enough to make it "impossible" (that is to say, infeasible from the point of view of computation) to factor n. However, R, who knows p and q, can find (through an application of the Euclidean algorithm) the integer $d \in [1, \phi(n)]$ such that

$$de \equiv 1 \text{ modulo } \phi(n),$$

i.e., $de - 1 = k \times \phi(n)$ for some integer k. He keeps d, as well as $\phi(n)$ and the factorization of n, secret, but R now publishes, for all the world to see, n and e.

Now suppose S (the sender) wishes to send a message m to R. We assume that the message is in fact an integer in the range $[1, n-1]$. (One presumes that a suitable rule for encoding messages into such integers can be agreed upon.) S now encrypts m into $E(m)$ by means of the formula

$$E(m) = m^e \bmod n = x, \text{ say.}$$

Upon receipt of x, R applies the formula

$$D(x) = x^d \bmod n$$

that yields m:

$$\begin{aligned} D(x) &= D(E(m)) \\ &= D(m^e \bmod n) \\ &= m^{ed} \bmod n \\ &= m^{1+k\times\phi(n)} \bmod n \\ &= m \times [m^{\phi(n)}]^k \bmod n \\ &= m \times 1^k \\ &= m. \end{aligned}$$

In the penultimate step, we have used Euler's theorem, which states that for any positive integer n and any integer m (relatively prime to n, but we won't fuss about that too much)

$$m^{\phi(n)} \equiv 1 \text{ modulo } n.$$

This is really a special case of Lagrange's theorem, since, under multiplication modulo n, the integers relatively prime to n form a group of order $\phi(n)$.

Before proceeding, we give an unrealistically small example: Let $n = 23 \times 47 = 1081$, $e = 553$, in which case $\phi(n) = 1012$ and $d = 829$. If $m = 200$, then one may calculate (there are short cuts which we shall not discuss) that

$$E(m) = 200^{553} \bmod 1081 = 623$$

and $D(623) = 623^{829} \bmod 1081 = 200$, as required.

An Attempted Attack on RSA

At the risk of stating the obvious, the system's security depends on the infeasibility of factoring n, since if p and q are known, then so is $\phi(n)$, and an adversary can find the decryption exponent d from the published value of e.

However, shortly after the first publication of the RSA system, Simmons and Norris suggested **[10]** that an attack through repeated encryption, using the published values of e and n, might sometimes succeed. To be specific: an outsider, intercepting x, might calculate

$$\begin{aligned} E(x) &= x^e \bmod n \\ E^2(x) = E(E(x)) &= x^{e^2} \bmod n \\ &\vdots \\ E^k(x) = E(E^{k-1}(x)) &= x^{e^k} \bmod n \\ &\vdots \end{aligned}$$

continuing until $E^l(x) = x$. (This must eventually happen for some l, since, after all, $x, E(x), E^2(x), \ldots$ all belong to the finite set of integers between 1 and $n-1$.) If $E^l(x) = x$, then $E(E^{l-1}(x)) = x$ and since $E(m) = x$ and E is one-to-one, one must have that $m = E^{l-1}(x)$. Thus the outsider would have succeeded in decrypting x.

Trying this on the example above, we find that

$$E^6(x) = 200^{553^6} \bmod 1081 = 623 = x.$$

The adversary, noting the last step, now concludes that the original message must have been $m = E^5(x)$.

How can one avoid an attack of this nature being successful? Before answering this question, we need to look at the group of possible encryption exponents.

The Group of Units of the Ring $\mathbb{Z}/s\mathbb{Z}$

Let s be a natural number and let us denote by $\mathbb{Z}/s\mathbb{Z}$ the ring of integers between 0 and $s-1$, under addition and multiplication modulo s. The units of this ring (its invertible elements) form a group under multiplication, which we shall denote by U. An integer x, $0 \le x \le s-1$ belongs to U if, and only if, x and s are relatively prime, so the group U has $\phi(s)$ elements.

Suppose the factorization of $\phi(s)$ is

$$\phi(s) = 2^{\alpha_0} p_1^{\alpha_1} \ldots p_k^{\alpha_k}$$

where the p_i are distinct primes. It can be shown that the element of U of largest order has as its order the least common multiple $[2^\beta, \phi(p_1^{\alpha_1}), \ldots, \phi(p_k^{\alpha_k})]$ of $2^\beta, \phi(p_1^{\alpha_1}), \ldots, \phi(p_k^{\alpha_k})$. Here

$$\beta = \begin{cases} 0 & \text{if } \alpha_0 \le 1 \\ 1 & \text{if } \alpha_0 = 2 \\ \alpha_0 - 2 & \text{if } \alpha_0 \ge 3 \end{cases}$$

Moreover, every element of U has an order that is a factor of $[2^\beta, \phi(p_1^{\alpha_1}), \ldots, \phi(p_k^{\alpha_k})]$.

A good exposition of these results may be found in Chapter 4 of [5].

Returning to the RSA system, suppose that e has order l in the ring $\mathbb{Z}/\phi(n)\mathbb{Z}$, in other words, suppose that

$$e^l \equiv 1 \bmod \phi(n). \qquad (*)$$

Then $e^l = 1 + t\phi(n)$ for some integer t. For any $x \in [1, n-1]$ one then has that

$$\begin{aligned} E^l(x) &= x^{e^l} \bmod n \\ &= x^{1+t\phi(n)} \bmod n \\ &= x. \end{aligned}$$

As we saw in the previous section, if l is known, this equation enables an outsider to decrypt x.

Clearly, therefore, to ensure that the attack of the previous section is likely to fail, one should choose e in such a way that it has a very large order in the group of units of the ring $\mathbb{Z}/\phi(n)\mathbb{Z}$. This can be ensured by choosing n in such a way that $\phi(\phi(n))$ has very large prime factors, for example, by choosing $n = p \times q$ with

$$\begin{aligned} p &= 2p_1 + 1 \qquad q = 2q_1 + 1 \\ p_1 &= 2p_2 + 1 \qquad q_1 = 2q_2 + 1 \end{aligned}$$

where p, p_1, p_2, q, q_1, and q_2 are all primes. In this case, we have that

$$\begin{aligned} \phi(n) &= \phi(p)\phi(q) \\ &= (p-1)(q-1) \\ &= 4p_1q_1 \end{aligned}$$

so that

$$\begin{aligned} \phi(\phi(n)) &= \phi(4)\phi(p_1)\phi(q_1) \\ &= 2(p_1 - 1)(q_1 - 1) \\ &= 8p_2q_2. \end{aligned}$$

In the next section we shall show explicitly that if n is chosen in this way, "almost" every encryption exponent e will have very large order. Before doing so, two observations may be made.

First, the reader may wonder whether p and q can realistically be prescribed to be of the form given above. The answer to this question is perhaps somewhat unsatisfactory: It is an open question whether an infinite number of primes of this form exist, but in practice there is no difficulty about finding them.

Second, and this is the point we were aiming at in the introduction, if in the group of units of $\mathbb{Z}/n\mathbb{Z}$ one can find an element of known even order, there is a probability of exactly $\frac{1}{2}$ that n can now be factored. For suppose $x^{2k} \equiv 1$ modulo n. Then, putting $y = x^k \bmod n$, we have

$$y^2 - 1 \equiv 0 \bmod n$$

that is

$$(y-1)(y+1) \equiv 0 \bmod pq.$$

Then exactly one of the following four conditions must hold:

$$\begin{array}{ll} (1) & y \equiv 1 \bmod p \\ & y \equiv 1 \bmod q \end{array} \qquad \begin{array}{ll} (2) & y \equiv -1 \bmod p \\ & y \equiv -1 \bmod q \end{array}$$

$$\begin{array}{ll} (3) & y \equiv 1 \bmod p \\ & y \equiv -1 \bmod q \end{array} \qquad \begin{array}{ll} (4) & y \equiv -1 \bmod p \\ & y \equiv 1 \bmod q \end{array}$$

The first two cases merely yield that $y \equiv \pm 1 \bmod n$, which is not helpful. However, in cases (3) and (4), one can easily find the greatest common divisor of $y - 1$ and n, and this is either p or q, so n has been factored.

We remark, parenthetically, that the same argument applies to the ring $\mathbb{Z}/\phi(n)\mathbb{Z}$. Every congruence of the form

$$x^{2k} \equiv 1 \bmod \phi(n) \qquad (**)$$

will, with probability $\frac{1}{2}$, yield a factor of $\phi(n)$. Given enough such congruences we can find $\phi(n)$. This would be serious, since knowledge of $\phi(n)$ will also yield a factorization of n:

Since $\phi(n) = (p-1)(q-1) = n - p - q + 1$ we have

$$p + q = n - \phi(n) + 1.$$

By elementary algebra

$$\begin{aligned} (p-q)^2 &= (p-q)^2 - 4pq \\ &= [n - \phi(n) + 1]^2 - 4q. \end{aligned}$$

Solving these two equations will give one p and q.

Thus, if generating congruences of the form $(**)$ is easy, we carry on doing so until we have found enough factors of $\phi(n)$ to give us a factorization of n in this way.

We summarize the earlier discussion: If a unit of $\mathbb{Z}/\phi(n)\mathbb{Z}$ has "small" order, then the cryptosystem can be broken by repeated encryption. If some units of $\mathbb{Z}/n\mathbb{Z}$ have known "small" even orders, then the system can be broken by factoring n. In the final section, we shall concentrate on the first of these two observations.

The Orders of Units of $\mathbb{Z}/\phi(n)\mathbb{Z}$

We shall consider in detail the case $n = p \times q$, $p = 2p_1 + 1$, $p_1 = 2p_2 + 1$, $q = 2q_1 + 1$, $q_1 = 2q_2 + 1$, where p, p_1, p_2, q, q_1, q_2 are all primes, and show that in this case "almost" every encryption exponent e is safe from attack by repeated encryption. In fact, we shall explicitly calculate how many elements of each order there are. Since $\phi(\phi(n)) = 8p_2q_2$ and every element e has order a divisor of $2p_2, q_2$, we need to consider the possibilities of the order of e being 1, 2, p_2, $2p_2$, q_2, $2q_2$, p_2q_2, and $2p_2q_2$. Recall that $\phi(n) = 4p_1q_1$.

Clearly only $e = 1$ has order 1 modulo $\phi(n)$.

e has order 2 modulo $\phi(n)$ if and only if

$$\begin{aligned} e^2 &\equiv 1 \bmod 4 \\ \text{and } e^2 &\equiv 1 \bmod p_1 \\ \text{and } e^2 &\equiv 1 \bmod q_1, \end{aligned}$$

so that

$$e \equiv \pm 1 \bmod 4$$
$$e \equiv \pm 1 \bmod p_1$$
$$e \equiv \pm 1 \bmod q_1.$$

Each of the eight possibilities gives (by the Chinese Remainder Theorem) a unique solution for e, so there are 7 elements of order 2 in $\mathbb{Z}/\phi(n)\mathbb{Z}$, since we have already counted $e = 1$.

Next consider the congruence

$$e^{p_2} \equiv 1 \bmod \phi(n).$$

This holds if and only if

$$e^{p_2} = 1 \bmod 4$$
$$e^{p_2} = 1 \bmod p_1$$
$$e^{p_2} = 1 \bmod q_1,$$

that is, if and only if

$$e \equiv 1 \bmod 4 \text{ (since } p_1 \text{ is odd)}$$
$$e \equiv \text{a square modulo } p_2$$
$$\text{and } e \equiv 1 \bmod q_1 \text{ (since } p_2 \text{ and } 2q_1 \text{ are relatively prime)}.$$

Since there are p_2 squares modulo p_1, the congruence

$$e^{p_2} \equiv 1 \bmod \phi(n)$$

has p_2 solutions. Again, one of these solutions is $e = 1$, so there are $p_2 - 1$ elements of order p_2 in $\mathbb{Z}/\phi(n)\mathbb{Z}$.

Similarly there are $q_2 - 1$ elements of order q_2.

Next, noting that $e^{2p_2} = e^{p_1 - 1} \equiv 1 \bmod p_1$, for every e, we see that

$$e^{2p_2} \equiv 1 \bmod \phi(n)$$
$$\text{if and only if } e^{2p_2} \equiv 1 \bmod 4$$
$$\text{and } e^{2p_2} \equiv 1 \bmod q_1,$$

that is, if and only if

$$e \equiv \pm 1 \bmod 4$$
$$\text{and } e \equiv \pm 1 \bmod q_1$$

while, modulo p_1, e can take any value between 1 and $p_1 - 1$. Thus there are $4 \times (p_1 - 1) = 8p_2$ different solutions of the congruence

$$e^{2p_2} \equiv 1 \bmod \phi(n),$$

and hence $8p_2 - (p_2 - 1) - 7 - 1 = 7p_2 - 7$ elements of order $2p_2$.

Similarly there are $7q_2 - 7$ elements of order $2q_2$.

Continuing in this way one finds $p_2q_2 - p_2 - q_1 + 1$ elements of order p_2q_2, and consequently (since there are $8p_2q_2$ units in $\mathbb{Z}/\phi(n)\mathbb{Z}$ altogether), $7p_2q_2 - 7p_2 - 7q_2 + 7$ elements which have the maximal possible order $2p_2q_2$.

Putting this differently: If one considers an encryption exponent e "good" if its order is at least p_2q_2, then the proportion of "good" exponents is

$$\frac{8(p_2-1)(q_2-1)}{8p_2q_2}.$$

For implementation of the RSA system, it is suggested that p_2 and q_2 should be of the order of 10^{75}. This will ensure that factoring n will be computationally infeasible. If this is done, and e is chosen at random, the probability of e not being "good" is of the order of

$$\frac{1}{p_2}+\frac{1}{q_2}-\frac{1}{p_2q_2}$$
$$\approx 2\times 10^{-75}.$$

Thus choosing a "bad" encryption exponent e is so unlikely that, should it happen, it may safely be ascribed to willful perversity. In a comment on Rivest's defense [6] following the original proposal by Simmons and Norris [10] of trying repeated encryption, Herlestam [4] claimed that the probability of choosing a "bad" e was in fact 3/4. While it does not seem clear how the figure 3/4 was arrived at, a difference of opinion on whether a value of something is 10^{-75} or 3/4 must constitute some kind of record!

The reader may have noted that we have also proved that most encryption exponents do in fact have even order. Since, however, there does not appear to be any way of determining what that order actually is, this fact does not help in trying to find $\phi(n)$, its factorization, or the factorization of n.

Finally, we want to point out that probabilities of similar orders of magnitude for choosing "bad" values of e are obtained if one merely chooses p_1-1 and q_1-1 to have at least one very large prime factor, that is $p_1=ap_2+1, q_1=bq_2+1$ where p_2, q_2 are large primes and a and b are small (even) integers, not necessarily 2 as in our calculations.

REFERENCES

1. Donald W. Davies, *The Security of Data in Networks*, IEEE Computer Society, Los Angeles, CA 1981.
2. Henri Cohen, *A Course in Computational Algebraic Number Theory*, Springer, New York, Berlin, Heidelberg, 1993.
3. W. Diffie and M. E. Hellman, New directions in cryptography, *IEEE Trans. Info. Theory* 22 (1976), 644–654.
4. T. Herlestam, Critical remarks on some public-key cryptosystems, *BIT* 18 (1978), 493–496.
5. Kenneth Ireland and Michael Rosen, *A Classical Introduction to Modern Number Theory*, Springer, New York, Berlin, Heidelberg, 1990.
6. Ronald L. Rivest, Remarks on a proposed cryptanalytic attack on the M.I.T. public-key cryptosystem, *Cryptologia* 2 (1978), 62–65.
7. Ronald L. Rivest, Critical remarks on "Critical Remarks on some Public-Key Cryptosystems by T. Herlestam," *BIT* 19 (1979), 274–275.
8. R. L. Rivest, A Shamir and L. Adleman, A method for obtaining digital signatures and public key cryptosystems, *Comm. ACM* 21 (1978), 120–126.
9. B. Schneier, *Applied Crypotography*, John Wiley & Sons, Inc., New York, 1994.
10. Gustavus J. Simmons and Michael J. Norris, Preliminary comments on the M.I.T. public-key cryptosystem, *Cryptologia* 1 (1977), 406–414.

NOTES

Using Quadratic Forms to Correct Orientation Errors in Tracking

JACK GOLDFEATHER
Carleton College
Northfield, MN 55057

Introduction The motivation for this note comes from a problem in tracking a moving object in virtual reality systems. For example, in so-called "see-through" virtual reality environments a moving computer-generated "virtual" image is integrated into a real environment, and the placement of the image is based on computing the position and orientation of the observer. There are many techniques for tracking three-dimensional motion, but all of them depend in one way or another on gathering data from the environment which are used as coefficients in some system of equations with the position and orientation parameters as the unknowns. Typically, orientation is expressed as a 3×3 orthogonal matrix, R, interpreted as the change of basis from the fixed world coordinate system to the moving coordinate system.

Errors occur when the coefficients are "noisy," (e.g., when the measuring equipment is capable of returning results accurate to only a few significant digits), and noisy coefficients can produce an R which is not quite orthogonal.

For example, the matrix

$$R = \begin{pmatrix} -0.97451771 & 0.02041436 & 0.03124792 \\ 0.02372552 & 0.97924131 & 0.00034581 \\ -0.03188555 & 0.00102279 & -0.95477235 \end{pmatrix}$$

was actually computed by a tracking system in the graphics lab at the University of North Carolina as the orientation of an ultrasound transducer being used in a trial experiment to create real-time 3D images of a human fetus. Unfortunately, the lengths of the column vectors of R are 0.97532782, 0.97945461, and 0.95528362, and the angles between column vectors are 89.8016998, 90.000003, and 89.999996 degrees, i.e., it is not quite orthogonal. This can create a number of problems for tracking systems. For example, R is used to transform various geometrically-defined data sets to a user's point of view and a non-orthogonal R can skew the resulting image. Even worse, some tracking systems use dynamic models to predict future orientation over short time intervals when there is insufficient time to collect or process data. Integrating over time using a non-orthogonal matrix as an initial condition can produce badly skewed results.

In this note, a method is described for computing an orthogonal matrix that is "nearest" to an "almost orthogonal" matrix. The method has been used successfully in ultrasound transducer tracking to correct skewing errors due to noise.

A Simple Tracking Example Imagine a camera (for simplicity let's assume a pinhole camera) that moves around, and is able to "see" points P in the world whose

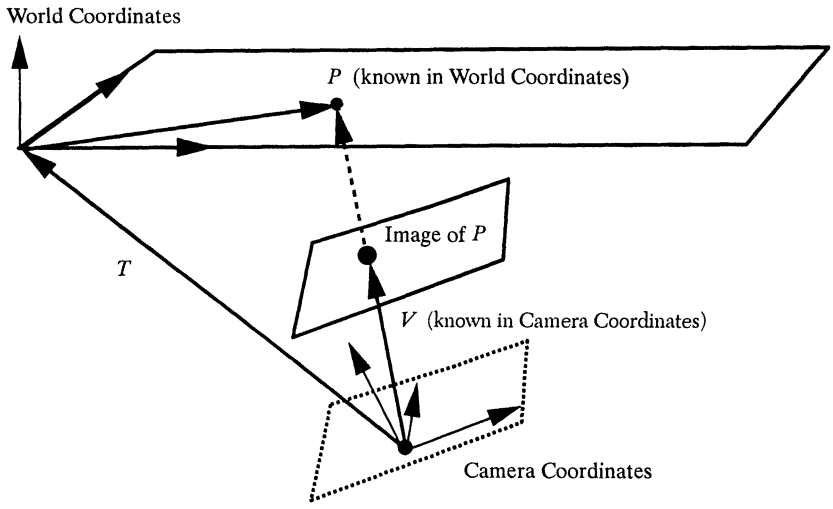


FIGURE 1

fixed world coordinates are known (see FIGURE 1). The camera has its own internal coordinate system, with two orthogonal axes in the image plane; the third orthogonal axis points in the direction the camera is aimed. The camera registers the point P as a vector V drawn from its origin to the image of P on its image plane. We would like to find (i) the camera's position, expressed as a vector T from the camera origin to the world origin; and (ii) the camera's orientation, expressed as a 3×3 orthogonal matrix R that rotates world coordinates into camera coordinates.

The vector equation $T + RP = kV$ relates P to V. It can be understood as follows:

(1) Multiplying P by the unknown rotation matrix R transforms P from the world coordinate system to the camera coordinate system.
(2) The vector sum $T + RP$ is the vector from the camera origin to P.
(3) The same vector can be found by scaling V by an unknown scalar k.

If the camera can see several fixed points, $P_1, P_2, \ldots, P_n$, and n is sufficiently large, the resulting system

$$T + RP_i = k_i V_i,\ i = 1 \ldots n$$

(which we will refer to as the tracking equations) can usually be solved for the unknowns T, R, k_i. How big n has to be is an interesting question and lies at the heart of the next section. For the moment, however, note that both P_i and V_i are not known exactly because both are measured by imprecise instruments. In particular, points look fuzzy through a camera, so their precise position is difficult to determine. It is this "noisy" kind of data that leads to problems.

We will describe two approaches to solving the tracking equations.

The Linear System Approach The simplest way to solve the tracking equations is to treat all 9 entries in

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

as independent parameters. Of course, in an orthogonal matrix the 9 parameters are not independent (e.g. $r_{11}^2 + r_{21}^2 + r_{31}^2 = 1$ and $r_{11}r_{12} + r_{21}r_{22} + r_{31}r_{32} = 0$), but it turns out that there is no need to impose these additional (non-linear) constraints right away. Instead, we impose the constraint that the points P_i all lie in a horizontal plane. In this case, each P_i is of the form $(x_i, y_i, 0)$ which eliminates the third column of R from the tracking equations. (This column can be recovered later by taking the cross product of the first two columns.) The tracking system in the graphics lab at UNC uses tiny lights embedded in a (nearly-planar) ceiling as the tracking points, so the coplanarity constraint is not just a theoretical one. Note, however, that small errors in the planarity of the ceiling lead to more noise in the data.

Once the third column of R is eliminated, and provided that n is sufficiently large, the resulting linear system can be solved (for either a unique solution if the system is consistent or a best fit solution if it is not) by a variety of efficient computer algorithms. It can be shown that $n = 4$ suffices provided that no three of the P_i are collinear, but in practice it is better to overconstrain the system by choosing $n > 4$.

So the good news is that a value for R can be found using a fast computer algorithm. The bad news is that noisy data will produce an R that is not quite orthogonal.

The Quaternion Approach Another approach to solving the tracking equations is to use quaternion parameters (w, x, y, z) and write R in the form

$$(*) \qquad R = \begin{pmatrix} w^2 + x^2 - y^2 - z^2 & 2xy - 2wz & 2xz + 2wy \\ 2xy + 2wz & w^2 + y^2 - x^2 - z^2 & 2yz - 2wx \\ 2xz - 2wy & 2yz + 2wx & w^2 + z^2 - x^2 - y^2 \end{pmatrix}$$

where $w^2 + x^2 + y^2 + z^2 = 1$. The name derives from the fact that if $Q = (w, x, y, z)$ is thought of as a unit quaternion and V is a vector in 3-space, then RV, i.e., "rotating" V by R, is the same as the vector part of the quaternion product QVQ^{-1}. (See [1], [2] for more details.) Quaternion parameters are a good choice for keeping matrices orthogonal, because if the computed $Q_1 = (w_1, x_1, y_1, z_1)$ is not quite a unit vector (remember, coefficients in equations will be noisy), it can be replaced with $Q_2 = Q_1/|Q_1|$. It is not hard to show that among all unit quaternions, Q_2 minimizes

$$f(w, x, y, z) = (w - w_1)^2 + (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2.$$

However, quaternions may be a bad choice because this makes the tracking equations quadratic, not linear, and quadratic systems are much less amenable to a quick computer solution. Indeed, there may not be *any* tractable method for solving certain quadratic systems.

So in many cases, R must be computed directly using the r_{ij} as parameters, and then corrected to an orthogonal matrix.

A Quadratic Forms Solution to Correcting *R* Although the tracking equations may be hard to solve directly using quaternion parameters, we can use them to correct an R found by the linear system method. Suppose

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

has been computed from noisy data and is not orthogonal. We will say an orthogonal matrix

$$S = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{pmatrix}$$

is *nearest* to R if $m = \sum_{i,j}(r_{ij} - s_{ij})^2$ is minimized by S. Since the columns of S have unit length, expanding m gives

$$\begin{aligned} m(S) &= \sum_{i,j} r_{ij}^2 - 2\sum_{i,j} r_{ij}s_{ij} + \sum_{i,j} s_{ij}^2 \\ &= \sum_{i,j} r_{ij}^2 - 2\sum_{i,j} r_{ij}s_{ij} + 3. \end{aligned}$$

Hence, minimizing m is equivalent to maximizing

$$M(S) = \sum_{i,j} r_{ij}s_{ij}.$$

Using the quaternion parameters (w, x, y, z) in $(*)$ for S, multiplying everything out, and rearranging terms, we obtain

$$\begin{aligned} M(w, x, y, z) = M(S) &= \sum_{i,j} r_{ij}s_{ij} \\ &= (r_{11} + r_{22} + r_{33})w^2 + (r_{11} - r_{22} - r_{33})x^2 + (-r_{11} + r_{22} - r_{33})y^2 \\ &\quad + (-r_{11} - r_{22} + r_{33})z^2 + 2(r_{32} - r_{23})wx + 2(r_{13} - r_{31})wy \\ &\quad + 2(r_{21} - r_{12})wz + 2(r_{21} + r_{12})xy + 2(r_{31} + r_{13})xz + 2(r_{32} + r_{23})yz. \end{aligned}$$

This is a quadratic form in (w, x, y, z) and can be written as $M(X) = X^TAX$ where $X = \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix}$, X^T denotes the transpose of X, and

$$A = \begin{pmatrix} r_{11} + r_{22} + r_{33} & r_{32} - r_{23} & r_{13} - r_{31} & r_{21} - r_{12} \\ r_{32} - r_{23} & r_{11} - r_{22} - r_{33} & r_{21} + r_{12} & r_{31} + r_{13} \\ r_{13} - r_{31} & r_{21} + r_{12} & -r_{11} + r_{22} - r_{33} & r_{32} + r_{23} \\ r_{21} - r_{12} & r_{31} + r_{13} & r_{32} + r_{23} & -r_{11} - r_{22} + r_{33} \end{pmatrix}$$

Note that since X is a unit vector, we have converted the problem of finding S to that of maximizing the quadratic form X^TAX on the unit sphere in R^4. A sketch of the solution method follows (see [3] for more details). Since A is a symmetric matrix it can be written in the form

$$A = B^TDB$$

where

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}$$

is the diagonal matrix of (real) eigenvalues of A and $B^T = B^{-1}$. Letting

$$BX = Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

we obtain

$$M(X) = X^TAX = X^TB^TDBX = Y^TDY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 + \lambda_4 y_4^2.$$

Assuming (without loss of generality) that λ_1 is the largest positive eigenvalue, the last expression is maximized when $Y = (1,0,0,0)$ with maximum value λ_1, so the unit eigenvector $X = B^TY$ is the solution we seek. Substituting $X = (w, x, y, z)$ into $(*)$ produces the nearest orthogonal matrix to R.

Note that if R is already orthogonal so that $r_{ij} = s_{ij}$, then

$$M = \sum_{i,j} r_{ij}s_{ij} = \sum_{i,j} r_{ij}^2 = 3$$

so that the largest eigenvalue will be 3. Hence $3 - \lambda_1$ is a measure of how far R is from being orthogonal.

If this method is applied to the matrix in the introduction, the "corrected" orthogonal matrix

$$\begin{pmatrix} -0.99921004 & 0.02256809 & 0.03271062 \\ 0.02259201 & 0.99974470 & 0.00036199 \\ -0.03269410 & 0.00110071 & -0.99946480 \end{pmatrix}$$

is obtained. The associated largest eigenvalue is $\lambda = 2.91006313$.

REFERENCES

1. Patrick Du Val, *Homographies, Quaternions, and Rotations*, Oxford Mathematical Monographs, Oxford University Press, Oxford, UK, 1964, pp. 33–40.
2. Bryant A. Julstrom, Using real quaternions to represent rotations in three dimensions, *UMAP Modules in Undergraduate Mathematics and Its Applications: Module* 652, COMAP, Inc., Lexington, MA, 1992.
3. Ben Noble and James W. Daniel, *Applied Linear Algebra*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1977, pp. 429–430.

The Birth of Period 3, Revisited

JOHN BECHHOEFER
Simon Fraser University
Burnaby, British Columbia V5A 1S6, Canada

Introduction Over the last twenty years, the logistic map,

$$x_{n+1} = rx_n(1 - x_n), \tag{1}$$

has served as an exemplar of nonlinear dynamics [**1**, **2**]. As May stressed some 20 years ago [**3**], the patterns formed by iterates of the logistic map are simple to compute but illustrate the complexities possible in nonlinear dynamics. The bifurcation of the logistic map, which summarizes the long-time dynamics as a function of the control parameter r in Equation 1, is one of the most commonly reproduced images of dynamical systems. (See FIGURE 1.) Also, using this map, Feigenbaum derived his famous renormalization-group theory of scaling exponents. These results were soon shown to apply to real experimental systems, such as fluid undergoing thermal convection [**2**]. Given both the pedagogical value and the scientific importance of the logistic map, exact analytic results concerning its solutions have been collected with great care. One such result is that a period-3 orbit (or 3-cycle), the most prominent of the "periodic windows" in FIGURE 1, is born via a tangent bifurcation at the control-parameter value $r = 1 + \sqrt{8}$.

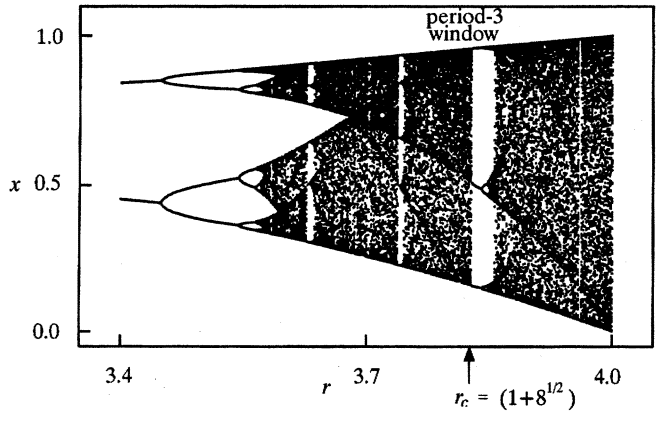


FIGURE 1
Bifurcation diagram of the logistic map. The 3-cycle is indicated on the figure.

A proof of this statement was recently given in these pages by Saha and Strogatz [**1**]. Their method, although straightforward, involves fairly complex algebra and a number of intermediate steps, which led the latter author to wonder whether a simpler proof might not exist. My purpose here is to give just such a proof. An independent (and different!) proof is given in the next article. More precisely, I shall show that given that a tangent bifurcation creates a period-3 cycle somewhere, that 3-cycle must be created at $r = 1 + \sqrt{8}$.

The Proof We first introduce some notation. Let $f(x) = rx(1-x)$ so that Equation 1 becomes $x_{n+1} = f(x_n)$. The second iterate of $f, f(f(x))$, will be denoted as $f^2(x)$, and similarly for higher iterates. Thus, any point of a period-3 orbit p will satisfy $p = f^3(p)$. In FIGURE 2, we plot $f^3(x)$ for a value of r just larger than the onset value for the 3-cycle, r_c. Note that the function $y = f^3(x)$ crosses the diagonal line ($y = x$) eight times. Three crossings, denoted by filled circles, correspond to the stable 3-cycle seen in the long-time dynamics. The nearby set of three crossings denoted by hollow circles correspond to an unstable 3-cycle. If we lower the control parameter r to $r_c = 1 + \sqrt{8}$, the stable and unstable 3-cycles will "collide" and have identical values of x. Finally, the two squares denote values of x that are fixed points of the direct map $f(x)$ (and hence, of course, fixed points of $f^3(x)$).

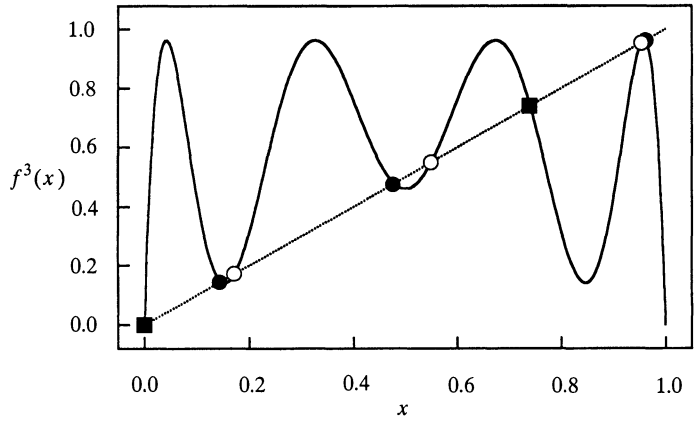


FIGURE 2
Graph of $f^3(x)$ for $r = 3.85$. Filled circles, stable 3-cycles; hollow circles, unstable 3-cycle; filled squares, fixed points of f.

As Saha and Strogatz point out, the condition for a tangent bifurcation may be expressed by the tangency of $y = f^3(x)$ to the line $y = x$. The stability of an orbit is given by the derivative of the map evaluated at any one of the points in the orbit [**2**]. The cycle just goes unstable (via a tangent bifurcation) when that derivative is 1. At a tangent bifurcation, f^3 has slope 1 at each of the 3 points a, b, c of the 3-cycle at $r = r_c$. Thus,

$$\begin{aligned}\frac{d(f^3(x))}{dx} &= \frac{d(f^3(x))}{d(f^2(x))} \cdot \frac{d(f^2(x))}{d(f(x))} \cdot \frac{d(f(x))}{dx} \\ &= \left.\frac{d(f(x))}{dx}\right|_{x=c} \cdot \left.\frac{d(f(x))}{dx}\right|_{x=b} \cdot \left.\frac{d(f(x))}{dx}\right|_{x=a} \\ &= r^3(1-2a)(1-2b)(1-2c) \\ &= 1. \end{aligned} \tag{2}$$

Multiplying this condition out, we have

$$r^3[1 - 2(a+b+c) + 4(ab+bc+ac) - 8(abc)] = 1. \tag{3}$$

Thus, as noted by Saha and Strogatz, we need to find the combinations $\alpha \equiv a+b+c$, $\beta \equiv ab+bc+ac$, and $\gamma \equiv abc$, but *not* a, b, and c individually.

The other condition that r_c must satisfy is that the third iterate of $f(x)$ have a fixed point at $x = a$, b, and c. This motivates us to define an auxiliary function $g(x) = f^3(x) - x$, which has roots at the fixed points of $f(x)$. FIGURE 2 thus shows that for r just larger than r_c, the function $g(x)$ has eight real roots. That is in fact the maximum possible, since $g(x)$ is an eighth-order polynomial ($f^3(x)$ is the third iterate of a quadratic function).

Now the important step: at the value $r = r_c$, the values of x for the stable and unstable 3-cycle collide at a, b, and c. The function $g(x)$ thus has double roots at a, b, and c. This accounts for six of the eight roots of g. Recalling that the two additional fixed points of f are at $x = 0$ and $x = 1 - 1/r$, we know that $g(x)$ must be of the form

$$g(x) \propto x(x - 1 + 1/r)[(x - a)(x - b)(x - c)]^2. \tag{4}$$

Multiplying this out, we find

$$\begin{aligned} g(x) \propto\ & x^8 - [2\alpha + 1 - 1/r]x^7 + [2\beta + \alpha^2 + 2(1 - 1/r)\alpha]x^6 \\ & - [2\gamma - 2\alpha\beta + 2(1 - 1/r)\beta + (1 - 1/r)\alpha^2]x^5 + \ldots \end{aligned} \tag{5}$$

Lower-order terms will not be needed. Notice that the combinations of a, b, and c generated by the expansion of Equation 4 involve only the α, β and γ required for our tangency condition (3).

On the other hand, we may compute $g(x)$ directly by iterating the map f. This gives

$$\begin{aligned} g(x) &= r^3x(1-x)[1 - rx(1-x)][1 - r^2x(1-x)(1 - rx(1-x))] - x \\ &= -r^7[x^8 - 4x^7 + (6 + 2/r)x^6 - (4 + 6/r)x^5 + \ldots]. \end{aligned} \tag{6}$$

Matching coefficients by starting with x^8 and descending in powers, we easily find α, β, and γ:

$$2\alpha = 3 + \frac{1}{r} \tag{7}$$

$$4\beta = \frac{3}{2} + \frac{5}{r} + \frac{3}{2r^2} \tag{8}$$

$$8\gamma = -\frac{1}{2} + \frac{7}{2r} + \frac{5}{2r^2} + \frac{5}{2r^3}. \tag{9}$$

Putting these back in Equation 3 yields

$$r^2 - 2r - 7 = 0 \tag{10}$$

whose sole positive root is $r_c = 1 + \sqrt{8}$, which is our result.

Conclusions Although our proof is simpler than the one given by Saha and Strogatz, it does not readily lend itself to generalization. The result takes advantage of the coincidence of the number of roots of $g(x)$ with the number of points on cycles of immediate interest. Were one to try to use the same method to find the birth of a 5-cycle, for example, one would be faced with a 32nd-order polynomial. But the stable and unstable 5-cycles plus 2 fixed points account for only 12 of those roots. Nonetheless, for the 3-cycle problem, with suitable hints to get students started, the method

outlined here should simplify greatly what had been a difficult homework problem for a course in nonlinear dynamics.

Acknowledgements. This work was supported by an NSERC Research Grant. The author is an Alfred P. Sloan Foundation Fellow. I thank Steven Strogatz for providing me with a pre-publication copy of his paper with Partha Saha and for helpful correspondence.

REFERENCES

1. P. Saha and S. H. Strogatz, The birth of period 3, this MAGAZINE 68 (1995), pp. 42–47.
2. S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA, 1994.
3. R. M. May, Simple mathematical models with very complicated dynamics, *Nature* 261 (1976), pp. 459–467.

Period Three Trajectories of the Logistic Map

WILLIAM B. GORDON
Naval Research Laboratory
Code 5311
Washington, DC 20375

A recent Note in this MAGAZINE **[3]** was concerned with locating the "tangent bifurcation" to the logistic map

$$f(x) = rx(1-x). \tag{1}$$

From graphical considerations, this problem amounts to finding the smallest value of $r \in (0,4)$ for which the map f has a non-trivial 3-periodic orbit. In **[3]** this value is shown to be $r = r_1$, where

$$r_1 = 1 + 2\sqrt{2} \approx 3.8284\,2712\,4746. \tag{2}$$

The purpose of this note is to show how the result (2) can be easily obtained by exploiting the fact that every 3-periodic sequence $\{x(n)\}$ can be written in the form

$$x(n) = \mu + \beta\omega^n + \overline{\beta}\,\overline{\omega}^n, \tag{3}$$

where μ and β are constants, ω is a complex cube root of unity, and the overbars indicate complex conjugation. We shall also give an upper bound for the r-values that support stable 3-periodic orbits, *viz.*, $r = r_2$, where

$$r_2 = 1 + \sqrt{\left[\frac{11}{3} + \left(\frac{1915}{54} + \frac{5\sqrt{201}}{2}\right)^{1/3} + \left(\frac{1915}{54} - \frac{5\sqrt{201}}{2}\right)^{1/3}\right]}$$

$$\approx 3.8414\,9900\,7543. \tag{4}$$

A different proof of (2) is given in this current issue of the MAGAZINE by Bechhoefer **[1]**. We also note that (3) can be viewed as a discrete Fourier transform representation of the 3-periodic orbit $x(n)$; and that discrete Fourier transform techniques have been used in the study of periodic orbits of the Hénon map **[2]**.

outlined here should simplify greatly what had been a difficult homework problem for a course in nonlinear dynamics.

Acknowledgements. This work was supported by an NSERC Research Grant. The author is an Alfred P. Sloan Foundation Fellow. I thank Steven Strogatz for providing me with a pre-publication copy of his paper with Partha Saha and for helpful correspondence.

REFERENCES

1. P. Saha and S. H. Strogatz, The birth of period 3, this MAGAZINE 68 (1995), pp. 42–47.
2. S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA, 1994.
3. R. M. May, Simple mathematical models with very complicated dynamics, *Nature* 261 (1976), pp. 459–467.

Period Three Trajectories of the Logistic Map

WILLIAM B. GORDON
Naval Research Laboratory
Code 5311
Washington, DC 20375

A recent Note in this MAGAZINE **[3]** was concerned with locating the "tangent bifurcation" to the logistic map

$$f(x) = rx(1-x). \tag{1}$$

From graphical considerations, this problem amounts to finding the smallest value of $r \in (0,4)$ for which the map f has a non-trivial 3-periodic orbit. In **[3]** this value is shown to be $r = r_1$, where

$$r_1 = 1 + 2\sqrt{2} \approx 3.8284\,2712\,4746. \tag{2}$$

The purpose of this note is to show how the result (2) can be easily obtained by exploiting the fact that every 3-periodic sequence $\{x(n)\}$ can be written in the form

$$x(n) = \mu + \beta\omega^n + \overline{\beta}\,\overline{\omega}^n, \tag{3}$$

where μ and β are constants, ω is a complex cube root of unity, and the overbars indicate complex conjugation. We shall also give an upper bound for the r-values that support stable 3-periodic orbits, *viz.*, $r = r_2$, where

$$r_2 = 1 + \sqrt{\left[\frac{11}{3} + \left(\frac{1915}{54} + \frac{5\sqrt{201}}{2}\right)^{1/3} + \left(\frac{1915}{54} - \frac{5\sqrt{201}}{2}\right)^{1/3}\right]}$$

$$\approx 3.8414\,9900\,7543. \tag{4}$$

A different proof of (2) is given in this current issue of the MAGAZINE by Bechhoefer **[1]**. We also note that (3) can be viewed as a discrete Fourier transform representation of the 3-periodic orbit $x(n)$; and that discrete Fourier transform techniques have been used in the study of periodic orbits of the Hénon map **[2]**.

In our proof of (2) we substitute (3) into the relation

$$x(n+1) - f(x(n)) = 0, \tag{5}$$

and use the identities $\omega^2 = \overline{\omega}$ and $\overline{\omega}^2 = \omega$, to express the left-hand side of (5) as a linear combination of the three functions $\{1, \omega^n, \overline{\omega}^n\}$. Setting the coefficients of these three functions equal to zero produces the system

$$\begin{aligned} 2\beta\overline{\beta} &= (1 - 1/r)\mu - \mu^2 \\ \overline{\beta}^2 &= (1 - 2\mu - \omega/r)\beta \\ \beta^2 &= (1 - 2\mu - \overline{\omega}/r)\overline{\beta}. \end{aligned} \tag{6}$$

Multiplying the last two equations together and substituting the result into the first equation gives a quadratic equation in μ, *viz.*,

$$(1 - 1/r)\mu - \mu^2 = 2|1 - 2\mu - \omega/r|^2.$$

Solutions to this quadratic are

$$\mu = \frac{3r + 1 \pm \sqrt{r^2 - 2r - 7}}{6r}. \tag{7}$$

The smallest possible value of $r \in (0, 4)$ for which 3-periodic orbits are possible is therefore the positive root of $r^2 - 2r - 7 = 0$; i.e., $r = 1 + 2\sqrt{2}$. This completes our proof of (2).

Let $D = D(r)$ denote the derivative of the third iterate $f^3(x)$ evaluated at a 3-periodic orbit $\{x(n)\}$. In **[1]** and **[3]** the value of r_1 is calculated by solving the equation $D(r) = +1$. We shall now calculate the value of r_2 by solving $D(r) = -1$. First, we express $D(r)$ as an explicit function of r. To this end we have

$$D(r) = r^3(1 - 2x(0))(1 - 2x(1))(1 - 2x(2)),$$

or

$$D(r) = r^3(1 - 2A + 4B - 8C), \tag{8}$$

where for notational ease we set

$$\begin{aligned} A &= x(0) + x(1) + x(2), \\ B &= x(0)x(1) + x(1)x(2) + x(2)x(0), \\ C &= x(0)x(1)x(2). \end{aligned} \tag{9}$$

Using (3) to express A and B in terms of μ and β gives

$$A = 3\mu \qquad \text{and} \qquad B = 3(\mu^2 - |\beta|^2). \tag{10}$$

Now we use a trick. Instead of using (3) to calculate C directly, we use (1) and the 3-periodicity of $\{x(n)\}$ to express C as a function of A and B. From (1),

$$x(n+1)/x(n) = r(1 - x(n)).$$

Multiplying the three equations obtained by setting $n = 0$, 1, and 2, and using the fact that $x(3) = x(0)$, we get

$$1 = r^3(1 - A + B - C).$$

Hence,

$$C = 1 - A + B - 1/r^3. \tag{11}$$

Substituting (11) into (8) we get an expression for $D(r)$ in terms of A and B; using (10) we get an expression for $D(r)$ in terms of μ and $|\beta|^2$. We now use (6) and (7) to express $D(r)$ explicitly as a function of r. The values of r that lead to *stable* orbits can be shown to correspond to the choice of the minus sign in (7), and with this choice we get

$$D(r) = r(2-r)\sqrt{(r^2 - 2r - 7)} - (r^2 - 2r - 8). \tag{12}$$

Now r_2 is a root of the equation $D(r) = -1$. Clearing out the square root in this equation gives the sixth-degree polynomial equation

$$0 = H(r) \equiv r^6 - 6r^5 + 4r^4 + 24r^3 - 14r^2 - 36r - 81. \tag{13}$$

It turns out that $H(r)$ is symmetric about $r = 1$, with

$$H(1+t) = H(1-t) = t^6 - 11t^4 + 37t^2 - 108.$$

Setting $s = t^2$ in the equation $H(1+t) = 0$ gives a cubic equation in s with only one real note, *viz.*,

$$s = \frac{11}{3} + \left(\frac{1915}{54} + \frac{5\sqrt{201}}{2}\right)^{1/3} + \left(\frac{1915}{54} - \frac{5\sqrt{201}}{2}\right)^{1/3}.$$

Equation (4) is then obtained by setting $r_2 = 1 + t = 1 + \sqrt{s}$.

REFERENCES

1. J. Bechhoefer, The birth of period 3, revisited, this MAGAZINE 69 (1996), pp. 115–118.
2. D. L. Hitzl and F. Zele, An exploration of the Hénon quadratic map, *Physica D* 14 (1985), pp. 305–326.
3. P. Saha and S. H. Strogatz, The birth of period three, this MAGAZINE 68 (1995), pp. 42–47.

A Polynomial Taking Integer Values

ROBIN CHAPMAN
University of Exeter
EX4 4QE, United Kingdom

In [2] Sury proves that for integers $a_1 < a_2 < \cdots < a_n$, the expression $\prod_{n \geq i > j \geq 1}(a_i - a_j)\ (i-j)$ is also an integer. (The result follows immediately from the theory of Lie groups; the number turns out to be the dimension of an irreducible representation of $SU(n)$.) Sury gives an elementary but indirect proof, based on the stronger result that $\prod_{n \geq i > j \geq 1}(X^{a_i - a_j} - 1)\ (X^{i-j} - 1) \in \mathbb{Z}[X]$. A direct proof of the original result can be deduced from properties of the Vandermonde determinant and the fact that binomial coefficients are integral.

If we define $\Delta(a_1, a_2, \ldots, a_n) = \prod_{n \geq i > j \geq 1}(a_i - a_j)$, then our task is to show that $\Delta(1, 2, \ldots, n) = \prod_{i=1}^{n}(i-1)!$ divides $\Delta(a_1, a_2, \ldots, a_n)$ whenever $a_1 < a_2 < \cdots < a_n$ are integers. Since $\Delta(a_1, a_2, \ldots, a_n) = \Delta(a_1 + 1, a_2 + 1, \ldots, a_n + 1)$ we may assume that each $a_i \geq 0$. It is well known that $\Delta(a_1, a_2, \ldots, a_n)$ is the value of the *Vandermonde determinant*

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{vmatrix}.$$

(A careful proof of this fact can be found in section 2.9 of [1]; alternatively a good exercise is to obtain this formula by row and column operations and induction.) Applying elementary row operations to this determinant shows that if f_i are any monic polynomials of degree i, $1 \leq i \leq n-1$, then

$$\Delta(a_1, a_2, \ldots, a_n) = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ f_1(a_1) & f_1(a_2) & f_1(a_3) & \cdots & f_1(a_n) \\ f_2(a_1) & f_2(a_2) & f_2(a_3) & \cdots & f_1(a_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{n-1}(a_1) & f_{n-1}(a_2) & f_{n-1}(a_3) & \cdots & f_{n-1}(a_n) \end{vmatrix}.$$

In particular if we choose $f_i(a) = a(a-1)(a+2)\cdots(a-i+1)$, then for non-negative integers a, $f_i(a) = i!\binom{a}{i}$, so each entry in the i-th row of this determinant is divisible by $(i-1)!$. Hence $\Delta(a_1, a_2, \ldots, a_n)$ is divisible by $\prod_{i=1}^{n}(i-1)!$.

REFERENCES

1. David Sharpe, *Rings and Factorization*, Cambridge University Press, Cambridge, UK, 1987.
2. B. Sury, *An integral polynomial*, this MAGAZINE 68, 2 (1995), 134–135.

Simple Proofs for $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$ and $\sin x = x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2\pi^2}\right)$

R. A. KORTRAM
Katholieke Universiteit
6525 ED Nijmegen
The Netherlands

Introduction Almost every textbook on complex analysis or on Fourier series contains a proof of Euler's identity

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Furthermore, there are countlessly many proofs of this result that do not rely explicitly upon complex function theory or Fourier analysis. In the references, we have listed some elementary proofs. Here we shall present a short and simple proof that uses only the definitions of π, sin, cos, and exp, and of course the notion of convergence. Moreover, our proof also gives Euler's identity

$$\sin x = x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2\pi^2}\right).$$

A trigonometric identity Let m and n be positive integers. It follows from the relation

$$\cos(k+1)x + \cos(k-1)x = 2\cos x \cdot \cos kx$$

that there exists a polynomial T_n of degree n such that for all $x \in \mathbb{R}$

$$\cos nx = T_n(\cos x).$$

In particular we have for positive integers k

$$\cos 2kx = T_k(\cos 2x) = T_k(1 - 2\sin^2 x).$$

This, together with the relation

$$\sin(2k+1)x - \sin(2k-1)x = 2\sin x \cdot \cos(2kx),$$

shows that there is a polynomial F_m of degree m such that for all $x \in \mathbb{R}$

$$\sin(2m+1)x = \sin x \cdot F_m(\sin^2 x).$$

Since $\sin(2m+1)\cdot k\pi/(2m+1) = 0$ for $k = 1, 2, \ldots, m$ we conclude that F_m has zeros at $\sin^2 k\pi/(2m+1)$, $k = 1, 2, \ldots, m$. These zeros are distinct, so F_m has no other zeros; thus

$$F_m(y) = F_m(0) \prod_{k=1}^{m} \left(1 - \frac{y}{\sin^2 \frac{k\pi}{2m+1}}\right),$$

and

$$F_m(0) = \lim_{x\to 0} \frac{\sin(2m+1)x}{\sin x} = 2m+1.$$

Therefore we have

$$\sin(2m+1)x = (2m+1)\sin x \prod_{k=1}^{m} \left(1 - \frac{\sin^2 x}{\sin^2 \dfrac{k\pi}{2m+1}}\right);$$

thus

$$\sin x = (2m+1)\sin \frac{x}{2m+1} \prod_{k=1}^{m} \left(1 - \frac{\sin^2 \dfrac{x}{2m+1}}{\sin^2 \dfrac{k\pi}{2m+1}}\right). \tag{1}$$

Comparison of sums and products For all real t we know that $e^t \geq 1+t$. So if $1+t>0$ we see that

$$e^{-t} = \frac{1}{e^t} \leq \frac{1}{1+t}.$$

Let $u<1$. The choice $t = u/(1-u)$ leads to

$$e^{-u/(1-u)} \leq 1-u.$$

For every collection of numbers $u_k \in [0,1)$ we have

$$1 - \sum_k \frac{u_k}{1-u_k} \leq e^{-\sum_k u_k/(1-u_k)} \leq \prod_k (1-u_k) \leq e^{-\sum_k u_k} \leq 1. \tag{2}$$

If we have in addition $\sum_k u_k < 1$, then we even know that

$$e^{-\sum_k u_k} \leq \frac{1}{1+\sum_k u_k},$$

and consequently

$$\frac{\sum_k u_k}{1+\sum_k u_k} \leq 1 - \prod_k (1-u_k) \leq \sum_k \frac{u_k}{1-u_k}. \tag{3}$$

Proof that $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$. Let m and N be positive integers, let $m>N$, and let

$$u_k = \left(\frac{\sin \dfrac{x}{2m+1}}{\sin \dfrac{k\pi}{2m+1}}\right)^2 \qquad k = 1,\ldots,m.$$

Choose x so small that $0 < \sum_k u_k < 1$. It follows from (1) that

$$\prod_{k=1}^{m} (1-u_k) = \frac{\sin x}{(2m+1)\sin \dfrac{x}{2m+1}},$$

and (3) implies that

$$\frac{\sum_k u_k}{1+\sum_k u_k} \le 1 - \frac{\sin x}{(2m+1)\sin\frac{x}{2m+1}} \le \sum_k \frac{u_k}{1-u_k}.$$

Divide by x^2 and let x go to zero. After a short computation we obtain:

$$\frac{1}{6}\left\{1 - \frac{1}{(2m+1)^2}\right\} = \sum_{k=1}^{m} \left(\frac{1}{(2m+1)\sin\frac{k\pi}{2m+1}}\right)^2,$$

and after a rearrangement

$$\left|\frac{1}{6} - \sum_{k=1}^{N}\left(\frac{1}{(2m+1)\sin\frac{k\pi}{2m+1}}\right)^2\right| = \frac{1}{6(2m+1)^2} + \sum_{k=N+1}^{m}\left(\frac{1}{(2m+1)\sin\frac{k\pi}{2m+1}}\right)^2.$$

For $0 \le t \le \frac{\pi}{2}$ we have $\sin t \ge \frac{2}{\pi}t$, hence the right-hand side is less than

$$\frac{1}{6(2m+1)^2} + \sum_{k=N+1}^{m} \frac{1}{(2k)^2} \le \frac{1}{6(2m+1)^2} + \frac{1}{4N}.$$

Let $m \to \infty$; we arrive at

$$\left|\frac{1}{6} - \sum_{k=1}^{N} \frac{1}{k^2\pi^2}\right| \le \frac{1}{4N},$$

and this shows that $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Proof that $\sin x = x\prod_{k=1}^{\infty}(1 - x^2/k^2\pi^2)$. We choose m, N and u_k as in the previous section, but now we take x such that $|x| < \frac{1}{4}N\pi$, and such that $\frac{x}{\pi} \notin \mathbb{Z}$. Again (1) implies that

$$\frac{\sin x}{(2m+1)\sin\frac{x}{2m+1}\prod_{k=1}^{N}(1-u_k)} = \prod_{k=N+1}^{m}(1-u_k),$$

thus we obtain from (2)

$$1 - \sum_{k=N+1}^{m}\frac{u_k}{1-u_k} \le \frac{\sin x}{(2m+1)\sin\frac{x}{2m+1}\prod_{k=1}^{N}(1-u_k)} \le 1. \tag{4}$$

Using again that $\sin t \ge \frac{2}{\pi}t$ for $0 \le t \le \frac{\pi}{2}$ we see that

$$u_k \le \left(\frac{(2m+1)\sin\frac{x}{2m+1}}{2k}\right)^2 \le \left(\frac{x}{2k}\right)^2;$$

thus

$$\frac{u_k}{1-u_k} \le \frac{x^2}{(2k)^2 - x^2} \quad (k > N).$$

Hence

$$\sum_{k=N+1}^{m} \frac{u_k}{1-u_k} \le \frac{x^2}{2N-|x|}.$$

Thus it follows from (4) that

$$1-\frac{x^2}{2N-|x|} \le \frac{\sin x}{(2m+1)\sin\dfrac{x}{2m+1}\displaystyle\prod_{k=1}^{N}(1-u_k)} \le 1.$$

Let $m \to \infty$. Then

$$1-\frac{x^2}{2N-|x|} \le \frac{\sin x}{x\displaystyle\prod_{k=1}^{N}\left(1-\frac{x^2}{k^2\pi^2}\right)} \le 1.$$

Now let $N \to \infty$ and we obtain for $\frac{x}{\pi} \notin \mathbb{Z}$

$$\sin x = x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2\pi^2}\right).$$

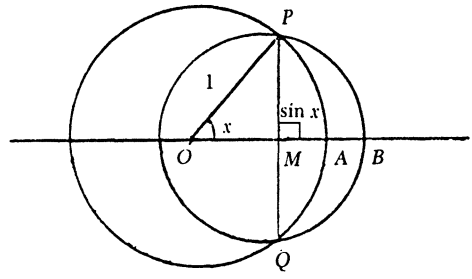
For $\frac{x}{\pi} \in \mathbb{Z}$ this is of course also true.

REFERENCES

1. Tom M. Apostol, A proof that Euler missed: evaluating $\zeta(2)$ the easy way, *The Mathematical Intelligencer* 5:3 (1983), 59–60.
2. Boo Rim Choe, An elementary proof of $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$, *Amer. Math. Monthly* 94 (1987), 662–663.
3. D. P. Giesy, Still another elementary proof that $\sum\frac{1}{k^2}=\frac{\pi^2}{6}$, this MAGAZINE 45 (1972), 148–149.
4. F. Holme, Ein enkel beregning av $\sum_{k=1}^{\infty}\frac{1}{k^2}$, *Nordisk Mat. Tidskr.* 18 (1970), 91–92, 120.
5. G. Kimble, Euler's other proof, this MAGAZINE 60 (1987), 282.
6. K. Knopp, I. Schur, Über die Herleitung der Gleichung $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$, *Archiv der Mathematik and Physik* (3), 27 (1918), 174–176. See also I. Schur, *Gesammelte Abhandelungen* II, Springer-Verlag, New York, 1973, pp. 246–248.
7. Y. Matsuoka, An elementary proof of the formula $\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$, *Amer. Math. Monthly* 68 (1961), 486–487.
8. I. Papadimitriou, A simple proof of the formula $\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$, *Amer. Math. Monthly* 80 (1973), 424–425.
9. E. L. Stark, Another proof of the formula $\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$, *Amer. Math. Monthly* 76 (1969), 552–553.
10. E. L. Stark, $1-\frac{1}{4}+\frac{1}{9}-\frac{1}{16}+\cdots=\frac{\pi^2}{12}$, *Praxis Math.* 12 (1970), 1–3.
11. A. M. Yaglom and I. M. Yaglom, *Challenging Mathematical Problems with Elementary Solutions* II problem 145, San Francisco, CA, 1967.

Proof Without Words:

Jordan's Inequality $\frac{2x}{\pi} \le \sin x \le x,\ 0 \le x \le \frac{\pi}{2}$



$$OB = OM + MP \ge OA \Rightarrow \overset{\frown}{PBQ} \ge \overset{\frown}{PAQ} \ge PQ$$
$$\Rightarrow \pi \sin x \ge 2x \ge 2 \sin x$$
$$\Rightarrow \frac{2x}{\pi} \le \sin x \le x$$

—Feng Yuefeng
Middle School, Hunan Normal University
Changsha, China

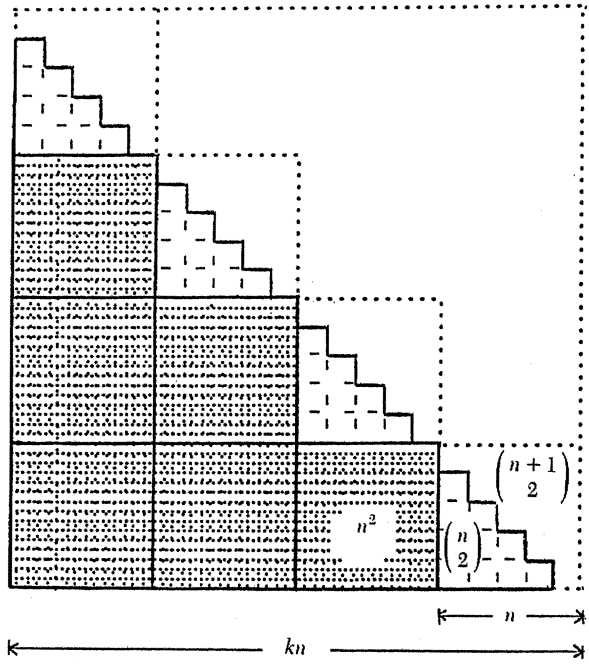
Proof Without Words: Decomposing the Combination $\binom{kn}{2}$

1. $$\binom{n}{2} + \binom{n+1}{2} = n^2$$

2. $$\binom{kn}{2} = k\binom{n}{2} + n^2\binom{k}{2}.$$

COROLLARY.

$$\binom{kn}{2} = \binom{k+1}{2}\binom{n}{2} + \binom{k}{2}\binom{n+1}{2}.$$



—JAMES O. CHILAKA
LONG ISLAND UNIVERSITY
BROOKVILLE, NY 11548

Counting Arrangements of 1's and −1's

D. F. BAILEY
Trinity University
San Antonio, TX 78212

It is well known that the nth Catalan number counts the number of sequences with non-negative partial sums that can be formed from n 1's and n -1's. (See [1].) In this paper we derive a formula for the number of such sequences formed from n 1's and k -1's. In the process we produce a non-standard proof that the nth Catalan number is given by

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Because the numbers we define can be evaluated in a manner similar to the binomial coefficients, we use a symbolism reminiscent of the standard notation for the binomial coefficients.

Definition. Let $\left\{ {n \atop k} \right\}$ denote the number of arrangements of n 1's and k -1's $a_1 a_2 \ldots a_{n+k}$ so that

$$a_1 + a_2 + \cdots + a_i \geq 0 \text{ for all } 1 \leq i \leq n+k.$$

We first establish the following.

LEMMA 1.

(i) $\left\{ {n \atop 0} \right\} = 1$ *for* $n \geq 0$.

(ii) $\left\{ {n \atop 1} \right\} = n$ *for* $n \geq 1$.

(iii) $\left\{ {n+1 \atop k} \right\} = \left\{ {n+1 \atop k-1} \right\} + \left\{ {n \atop k} \right\}$ *for* $1 < k < n+1$.

(iv) $\left\{ {n+1 \atop n+1} \right\} = \left\{ {n+1 \atop n} \right\}$ *for* $n \geq 1$.

Proof. Statement (i) is obvious and (ii) is nearly so. To establish (ii) note that an arrangement of n 1's and a single -1 occupies $n+1$ positions. The -1 may occupy any of these positions save the first. Thus there are n such arrangements.

For (iii) imagine an arrangement of $n+1$ 1's and k -1's with $k \geq 2$. The last element (that is a_{n+1+k}) is either a 1 or a -1. If it is a 1 then the remaining positions are filled with n 1's and k -1's. Hence these positions can be filled in $\left\{ {n \atop k} \right\}$ ways. If, on the other hand, the last element is a -1, the remaining positions are filled with $n+1$ 1's and $k-1$ -1's. This can be done in $\left\{ {n+1 \atop k-1} \right\}$ ways. Thus (iii) is established.

Finally, to establish (iv) consider an arrangement of $n+1$ 1's and $n+1$ -1's. Note that the last element in the arrangement must be a -1. Indeed if this element is a 1

then since

$$a_1 + a_2 + \cdots + a_{2n+1} + a_{2n+2} = 0$$

it must be the case that

$$a_1 + a_2 + \cdots + a_{2n+1} = -1 < 0$$

contrary to our requirement. Hence the first $2n + 1$ positions are filled with $n + 1$ 1's and n -1's. Thus the assertion of (iv) is valid.

Using the results of Lemma 1, we can construct a triangular array containing the numbers $\left\{ {n \atop k} \right\}$ in somewhat the same manner as the Pascal triangle is constructed. In particular, the nth row and kth column will contain $\left\{ {n \atop k} \right\}$. By part (i) of Lemma 1 each row begins with a 1, in the 0th column. Then the next entry (the entry in column 1) in the nth row will be n for $n \geq 1$, according to part (ii). After that (by (iii)) the entry in a position is determined by adding the element directly above the position and immediately to the left of the position until there are no elements directly above. (That is until we reach the nth column of row n.) The final entry of a row is obtained, by (iv), by repeating the element immediately to the left.

Here is the array through row 8.

$k=$	0	1	2	3	4	5	6	7	8
$n=0$	1								
$n=1$	1	1							
$n=2$	1	2	2						
$n=3$	1	3	5	5					
$n=4$	1	4	9	14	14				
$n=5$	1	5	14	28	42	42			
$n=6$	1	6	20	48	90	132	132		
$n=7$	1	7	27	75	165	297	429	429	
$n=8$	1	8	35	110	275	572	1001	1430	1430

We can, however, produce a closed form for $\left\{ {n \atop k} \right\}$. To this end we require three lemmas.

LEMMA 2. *For* $2 \leq k \leq n$ *we have*

$$\left\{ {n \atop k} \right\} = \sum_{i=k}^{n} \left\{ {i \atop k-1} \right\}.$$

Proof. For $n = 2$ our assertion is simply that $\left\{ {2 \atop 2} \right\} = \left\{ {2 \atop 1} \right\}$, which is obviously valid. Therefore we assume the result for some fixed $n \geq 2$ and consider $\left\{ {n+1 \atop k} \right\}$. If $k \leq n$, the assertion follows easily by Lemma 1 and the inductive hypothesis.

If, on the other hand, $k = n + 1$ we have

$$\left\{ {n+1 \atop n+1} \right\} = \left\{ {n+1 \atop n} \right\} = \sum_{i=n+1}^{n+1} \left\{ {i \atop n} \right\}$$

and again we have the desired result.

An easy argument using (ii) and Lemma 2 proves:

LEMMA 3. *If* $n \geq 2$ *then*

$$\left\{ {n \atop 2} \right\} = \frac{(n-1)(n+2)}{2}.$$

LEMMA 4.

$$\sum_{i=k+1}^{n} (i+1-k)(i+2)(i+3)\dots(i+k)$$
$$= \frac{1}{k+1}(n-k)(n+2)(n+3)\dots(n+1+k).$$

Proof. For $n = k+1$ we have the assertion that

$$(k+2-k)(k+1+2)(k+1+3)\dots(k+1+k)$$
$$= \frac{1}{k+1}(k+1-k)(k+1+2)\dots(k+1+k)(k+1+k+1),$$

which reduces to

$$2 = \frac{1}{k+1}(2k+2).$$

The inductive step follows easily to complete the proof.

We are now ready to derive a closed form for $\left\{ {n \atop k} \right\}$.

THEOREM. *For* $n \geq k \geq 2$

$$\left\{ {n \atop k} \right\} = \frac{(n+1-k)(n+2)(n+3)\dots(n+k)}{k!}.$$

Proof. From Lemma 3 we have the result when $k = 2$. So we assume the result for some $k \geq 2$ and consider $\left\{ {n \atop k+1} \right\}$. From Lemma 2 and the inductive hypothesis we have

$$\left\{ {n \atop k+1} \right\} = \sum_{i=k+1}^{n} \left\{ {i \atop k} \right\} = \sum_{i=k+1}^{n} \frac{(i+1-k)(i+2)(i+3)\dots(i+k)}{k!}.$$

By Lemma 4 this last summation is

$$\frac{1}{k!}\frac{1}{k+1}(n-k)(n+2)(n+3)\dots(n+1+k)$$

or equivalently

$$\frac{1}{(k+1)!}(n+1-(k+1))(n+2)(n+3)\dots(n+(k+1)),$$

completing the induction.

As an aside we note that we can now complete what must be the longest possible derivation of the closed form for the nth Catalan number, C_n.

COROLLARY.

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Proof. Set $k = n$ in the formula for $\left\{ {n \atop k} \right\}$.

REFERENCE

1. Richard A. Brualdi, *Introductory Combinatorics*, 2nd edition, Elsevier Science Publishing Co., Inc., New York, NY, 1992.

The Fermat Point of a Triangle

P. G. SPAIN
University of Glasgow
Glasgow G12 8QW, Scotland

The problem Fermat posed the problem of minimizing the sum of the distances from a point to the vertices of a triangle. It was solved by Torricelli, Cavalieri, and others [3]; see also [1], where it is referred to as "Steiner's Problem."

Solution The solution is that there is precisely one point at which the minimum is attained, called the *Fermat point*. It is the unique interior point F of ΔABC for which FA, FB, FC meet at equal angles, if such a point exists; otherwise it is the vertex of the largest angle.

There are several ways to demonstrate this, most of them in terms of classical synthetic geometry. A recent article in this journal [2] provides an interesting discussion and a solution using advanced calculus.

I would like to present a simple proof, self-contained, based on the following elementary lemma concerning a trio of unit vectors summing to zero.

LEMMA. *Suppose that* u, v, w, *are unit vectors such that*

$$u + v + w = 0.$$

Then the angles between the vectors u, v, w *are all equal* (*to* 120°).

Proof. The vectors u, v, w must form the sides of an equilateral triangle.

Construction It is easy to identify the point F for which FA, FB, FC meet at equal angles, when there is one, i.e., when all the angles of the triangle are less than 120°, by the methods of elementary geometry.

Let ABC be a triangle all of whose angles are less than 120°. Then the loci of interior points R such that

$$\angle BRC = 120°, \quad \angle CRA = 120°, \quad \angle ARB = 120° \tag{1}$$

are circular arcs that intersect at the required point F. This construction shows that there can be *only one* such point. It is equally clear that no such point can exist in a triangle when any angle is 120° or more.

COROLLARY.

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Proof. Set $k = n$ in the formula for $\left\{ {n \atop k} \right\}$.

REFERENCE

1. Richard A. Brualdi, *Introductory Combinatorics*, 2nd edition, Elsevier Science Publishing Co., Inc., New York, NY, 1992.

The Fermat Point of a Triangle

P. G. SPAIN
University of Glasgow
Glasgow G12 8QW, Scotland

The problem Fermat posed the problem of minimizing the sum of the distances from a point to the vertices of a triangle. It was solved by Torricelli, Cavalieri, and others [3]; see also [1], where it is referred to as "Steiner's Problem."

Solution The solution is that there is precisely one point at which the minimum is attained, called the *Fermat point*. It is the unique interior point F of ΔABC for which FA, FB, FC meet at equal angles, if such a point exists; otherwise it is the vertex of the largest angle.

There are several ways to demonstrate this, most of them in terms of classical synthetic geometry. A recent article in this journal [2] provides an interesting discussion and a solution using advanced calculus.

I would like to present a simple proof, self-contained, based on the following elementary lemma concerning a trio of unit vectors summing to zero.

LEMMA. *Suppose that* u, v, w, *are unit vectors such that*

$$u + v + w = 0.$$

Then the angles between the vectors u, v, w *are all equal* (*to* 120°).

Proof. The vectors u, v, w must form the sides of an equilateral triangle.

Construction It is easy to identify the point F for which FA, FB, FC meet at equal angles, when there is one, i.e., when all the angles of the triangle are less than 120°, by the methods of elementary geometry.

Let ABC be a triangle all of whose angles are less than 120°. Then the loci of interior points R such that

$$\angle BRC = 120°, \quad \angle CRA = 120°, \quad \angle ARB = 120° \tag{1}$$

are circular arcs that intersect at the required point F. This construction shows that there can be *only one* such point. It is equally clear that no such point can exist in a triangle when any angle is 120° or more.

Notation Let a, b, c be the position vectors of the vertices of the triangle ABC situated in the x, y plane. For each point $R(x, y)$ with position vector r define

$$S(r) = |r-a| + |r-b| + |r-c|.$$

The Fermat problem is, of course, to minimize the function S.

Writing $a = (a_x, a_y)$ we have

$$|r-a|^2 = (x-a_x)^2 + (y-a_y)^2$$

and so

$$2|r-a|\frac{\partial}{\partial x}|r-a| = 2(x-a_x).$$

Similarly, we have analogous expressions for $|r-b|$, $|r-c|$, and for the partial derivatives with respect to y.

Thus the partial derivatives $\partial S/\partial x$, $\partial S/\partial y$ exist for $r \neq a, b, c$, and the directional derivative

$$\begin{aligned} DS(r) &\triangleq (\partial S/\partial x, \partial S/\partial y) \\ &= (r-a)/|r-a| + (r-b)/|r-b| + (r-c)/|r-c|. \end{aligned} \tag{2}$$

Existence of a minimum The function S is positive and continuous everywhere in the plane, and is differentiable at all points except for the vertices. Since the values of S inside the triangle are less than those outside, and since the triangle (edges and interior included) is closed and bounded (i.e., compact), there must be at least one point in the closed triangles at which S attains a global minimum. If R is such a point and not a vertex, then it is a critical point at which $DS(r) = 0$. Note that at each interior point of an edge the derivative DS points away from the opposite vertex, so no interior point of an edge can be a minimum of S. It follows that any point at which S attains its minimum must *either* be an interior point of the triangle *or* a vertex.

Next we need the following result.

PROPOSITION. *If S attains its minimum at any point that is not a vertex, then this point is the unique point that satisfies the relation* (1).

Proof. This follows immediately from the Lemma and the fact that the three terms on the right side of (2) are *unit* vectors.

Two cases We consider the two cases separately: either, **1**: all the angles of the triangle are less than 120°, or, **2**: one angle is 120° or more.

Case 1. Suppose that all the angles of the triangle ABC are less than 120°. Then, by construction above, there is a unique point F (lying inside the triangle) such that FA, FB, FC make angles of 120°.

Now, by the *Law of Cosines* on triangle FAB,

$$\begin{aligned} |AB|^2 = |b-a|^2 &= |f-b|^2 + |f-a|^2 - 2|f-b||f-a|\cos(120°) \\ &= |f-b|^2 + |f-a|^2 + |f-b||f-a| \\ &> (|f-b| + |f-a|/2)^2. \end{aligned}$$

Thus

$$\begin{aligned} S(a) &= |b-a| + |c-a| \\ &> |f-b| + |f-a|/2 + |f-c| + |f-a|/2 \\ &= S(f). \end{aligned}$$

Similarly $S(b) > S(f)$, $S(c) > S(f)$. This shows that S does not attain its minimum at any vertex. It follows that S must attain its minimum at some interior point R; and then, by the Proposition, $DS(r) = 0$. But F is the unique point for which $DS(r) = 0$; so $R = F$.

Case 2. Suppose, alternatively, that one of the angles A, B, C is not less than 120°. Then there can be no point R such that RA, RB, RC make equal angles with each other; so S cannot attain its minimum at an interior point, and must therefore attain its minimum at a vertex. Since the longest side of the triangle is the one opposite the largest angle, we see that S attains its global minimum at the vertex of the largest angle.

We have demonstrated the following.

THEOREM. *If all the angles of the triangle ABC are less than* 120° *then the Fermat point F is the point such that FA, FB, FC meet at* 120°; *otherwise it is the vertex of the largest angle.*

Acknowledgements I am grateful to Messrs. M. Smith and I. Moss for bringing this topic to my attention, to Dr. A. S. Wassermann and Professor Z. Rubinstein for helpful discussions, and to the Institute of Mathematics, Hebrew University of Jerusalem, for affording me shelter and hospitality during a sabbatical year.

REFERENCES

1. R. Courant and H. Robbins, *What is Mathematics*?, Oxford University Press, Oxford, UK, 1941.
2. Mowaffaq Hajja, An advanced calculus approach to finding the Fermat point, this MAGAZINE Vol. 67, 1994, 29–34.
3. J. Pottage, *Geometric Investigations*, Addison-Wesley, Reading, MA, 1983.

Determinants of the Tournaments

CLIFFORD A. McCARTHY
Harvey Mudd College
Claremont, CA 91711

ARTHUR T. BENJAMIN
Harvey Mudd College
Claremont, CA 91711

In a round-robin tournament with n players, each player plays every other player in a game where ties are not possible. The results of the tournament can be summarized by an n by n *tournament matrix* A whose (i, j) entry is 1 if i beat j, -1 if j beat i, and 0 if i equals j. The matrix below represents a tournament where, for example, player 1 beat players 2 and 4, but lost to players 3 and 5. The authors confess that the

Thus

$$\begin{aligned} S(a) &= |b-a| + |c-a| \\ &> |f-b| + |f-a|/2 + |f-c| + |f-a|/2 \\ &= S(f). \end{aligned}$$

Similarly $S(b) > S(f)$, $S(c) > S(f)$. This shows that S does not attain its minimum at any vertex. It follows that S must attain its minimum at some interior point R; and then, by the Proposition, $DS(r) = 0$. But F is the unique point for which $DS(r) = 0$; so $R = F$.

Case 2. Suppose, alternatively, that one of the angles A, B, C is not less than 120°. Then there can be no point R such that RA, RB, RC make equal angles with each other; so S cannot attain its minimum at an interior point, and must therefore attain its minimum at a vertex. Since the longest side of the triangle is the one opposite the largest angle, we see that S attains its global minimum at the vertex of the largest angle.

We have demonstrated the following.

THEOREM. *If all the angles of the triangle ABC are less than* 120° *then the Fermat point F is the point such that FA, FB, FC meet at* 120°; *otherwise it is the vertex of the largest angle.*

Acknowledgements I am grateful to Messrs. M. Smith and I. Moss for bringing this topic to my attention, to Dr. A. S. Wassermann and Professor Z. Rubinstein for helpful discussions, and to the Institute of Mathematics, Hebrew University of Jerusalem, for affording me shelter and hospitality during a sabbatical year.

REFERENCES

1. R. Courant and H. Robbins, *What is Mathematics*?, Oxford University Press, Oxford, UK, 1941.
2. Mowaffaq Hajja, An advanced calculus approach to finding the Fermat point, this MAGAZINE Vol. 67, 1994, 29–34.
3. J. Pottage, *Geometric Investigations*, Addison-Wesley, Reading, MA, 1983.

Determinants of the Tournaments

CLIFFORD A. McCARTHY
Harvey Mudd College
Claremont, CA 91711

ARTHUR T. BENJAMIN
Harvey Mudd College
Claremont, CA 91711

In a round-robin tournament with n players, each player plays every other player in a game where ties are not possible. The results of the tournament can be summarized by an n by n *tournament matrix* A whose (i, j) entry is 1 if i beat j, -1 if j beat i, and 0 if i equals j. The matrix below represents a tournament where, for example, player 1 beat players 2 and 4, but lost to players 3 and 5. The authors confess that the

paper was motivated by *word play* with the hope of determining that determinants and tournaments have more in common than their names suggest. We discovered that in fact the concepts are almost *independent*, but do provide an opportunity to illustrate several powerful types of determinant arguments.

$$\begin{bmatrix} 0 & 1 & -1 & 1 & -1 \\ -1 & 0 & 1 & 1 & -1 \\ 1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ 1 & 1 & -1 & -1 & 0 \end{bmatrix}$$

Proposition 1. *Let A be the matrix representation of a tournament with n players. The determinant of A is zero if, and only if, n is odd.*

Proof. Any tournament matrix A is necessarily skew-symmetric, i.e., $A = -A^T$. Therefore, $\det(A) = \det(-A^T) = (-1)^n \det(A^T) = (-1)^n \det(A)$. When n is odd, $\det(A) = -\det(A)$ and must therefore be zero.

For the case where n is even, recall that to compute the determinant, we can determine it by summing products of the terms in it according to the formula:

$$\det(A) = \sum_{p \in S_n} \operatorname{sign}(p) a_{1,p(1)} a_{2,p(2)} \cdots a_{n,p(n)}$$

where S_n is the set of all permutations on n elements. We shall show that this determinant is odd, and hence nonzero. Since each $a_{i,j}$ is $0, 1$, or -1, so is the product $\operatorname{sign}(p) a_{1,p(1)} a_{2,p(2)} \cdots a_{n,p(n)}$. If $p(i) = i$ for some i, then $a_{i,p(i)}$ is 0, and hence $\operatorname{sign}(p) a_{1,p(1)} a_{2,p(2)} \cdots a_{n,p(n)}$ is 0. So we only need to take the sum over all permutations that do not map any element to itself, since all other permutations contribute zero to the sum. Since for each such permutation p, $\operatorname{sign}(p) a_{1,p(1)} a_{2,p(2)} \cdots a_{n,p(n)}$ is 1 or -1, we can calculate $\det(A)$ modulo 2, simply by counting the number of *derangements*, permutations that do not map any element to itself.

By the principle of inclusion-exclusion, there are

$$\sum_{i=0}^{n} (-1)^i (n-i)! \binom{n}{i}$$

derangements. Since $(n-i)!$ is even for $i \leq n-2$, the previous summation has the same parity as

$$(-1)^{n-1} 1! \binom{n}{n-1} + (-1)^n 0! \binom{n}{n} = -n - 1$$

which is odd. Thus, $\det(A)$ is nonzero.

Here is another simple proof for the case when n is even. Since $\det(A) \bmod 2$ is unaffected by changing (-1)'s into 1's, it suffices to compute the parity of the determinant of the matrix

$$J - I = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix}$$

where I is the n by n identity matrix, and J is the n by n matrix consisting entirely of 1's.

For any matrix C, we have

$$C\vec{x} = \lambda\vec{x} \Rightarrow (C - I)\vec{x} = (\lambda - 1)\vec{x},$$

so the eigenvalues of $J - I$ are all one less than the eigenvalues of J.

The rank of J is 1, so 0 is an eigenvalue with multiplicity $n - 1$. Since n is also an eigenvalue for J (with eigenvector $[1, 1, \ldots, 1]^T$), its multiplicity must be 1. So $J - I$ has the eigenvalue -1 with multiplicity $n - 1$ and the eigenvalue $n - 1$ with multiplicity 1. Hence the determinant of $J - I$ equals the product of its eigenvalues, namely $(-1)^{(n-1)}(n - 1)$, which is odd.

Yet another way to compute $\det(J - I)$ is by performing elementary row and column operations that do not affect the determinant. (This argument can be applied to any square matrix with one number on the main diagonal and another number everywhere else. See for instance, [1].) Adding every row of $J - I$ (except the first) to the first row gives us the matrix

$$\begin{bmatrix} n-1 & n-1 & n-1 & \cdots & n-1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix}.$$

After subtracting the first column of this matrix from all the other columns we obtain the lower triangular matrix below with determinant $(-1)^{(n-1)}(n - 1)$

$$\begin{bmatrix} n-1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}.$$

In fact, a little more can be said about the "zeroness" of $\det(A)$.

Proposition 2. *The nullspace of a tournament matrix A has dimension zero if n is even, and dimension one if n is odd.*

Proof. If n is even, then A is nonsingular and the proposition follows. For odd n, let $\vec{a}_1, \ldots, \vec{a}_n$ be the columns of A. Let C be the $(n - 1) \times (n - 1)$ matrix which results from deleting the last row and column from A. This matrix then corresponds to some tournament on $n - 1$ vertices. Hence, since $n - 1$ is even, the above result implies that C is nonsingular. The columns of this matrix are therefore linearly independent. It follows that the vectors $\vec{a}_1 \cdots \vec{a}_{n-1}$ are linearly independent, since they are the columns of C with an additional component. So the rank of A is at least $n - 1$. If A had rank n, it would be nonsingular, which we know to be false. So the rank of A is $n - 1$, and the dimension of its nullspace is therefore 1.

Acknowledgements. Thanks to Professor R. A. Mena for suggesting the eigenvalue proof of Proposition 1, and the anonymous referees for many valuable suggestions.

REFERENCE

Victor Bryant, *Aspects of Combinatorics*, Cambridge University Press, Cambridge, UK, 1993, p. 182.

Goldbach's Problem in Matrix Rings

GREG BLOY
Pennsylvania State University
University Park, PA 16802

In [2], motivated by the ideas of algebraic K-theory in passing from rings to matrix rings, Vaserstein investigates several questions about the matrix rings $M_n(\mathbb{Z})$ $(n > 1)$ over the integers $\mathbb{Z}$. These questions are generalizations of famous unanswered questions about the underlying ring $M_1(\mathbb{Z}) = \mathbb{Z}$ itself, such as Fermat's last theorem, the Goldbach conjecture, and the number of primes of the form $x^2 + 1$. Nearly as famous as Fermat's last theorem is Goldbach's conjecture that every even integer greater than two is the sum of two primes, which has so far eluded all attempts at proof. Following [2] and the work of Wang [3], we further investigate this mysterious problem in the context of matrix rings.

In the matrix rings $M_n(\mathbb{Z})$, the irreducible elements are precisely those matrices whose determinant is prime in $\mathbb{Z}$ (this can be seen easily by diagonalizing the matrix using elementary row and column operations until it is in "Smith normal form" with each diagonal entry divisible by the previous diagonal entry, see [1]). Thus the natural question to ask in generalizing Goldbach's conjecture to this context is, "can every matrix be written as a sum of two matrices of prime determinant?" In fact, the answer to this question is "yes," but before addressing this, let us first consider a more general result characterizing sums of two matrices with specified determinants. In the subsequent discussion, let n denote an integer greater than one, and for any matrix A let $d(A)$ denote the greatest common divisor of the entries of A. We use the symbol $\langle p \rangle$ to denote the set of all elements of $M_n(\mathbb{Z})$ with determinant p.

THEOREM. *Fix p and $q \in \mathbb{Z}$ and $A \in M_n(\mathbb{Z})$. There exist $M, N \in M_n(\mathbb{Z})$ with $M + N = A$, $\det M = p$, and $\det N = q$ if, and only if, $d(A) | (p + (-1)^{n+1}q)$. That is, $\langle p \rangle + \langle q \rangle = \{A \in M_n(\mathbb{Z}) : d(A) | (p + (-1)^{n+1}q)\}$.*

Proof. The proof follows that given in [2] for even n and in [3] for odd n, with only slight modification. Then we have the following:

COROLLARY 1 (Goldbach's problem in $M_n(\mathbb{Z})$). *Every element of $M_n(\mathbb{Z})$ is the sum of two irreducible elements of $M_n(\mathbb{Z})$.*

Proof. More generally, we have

$$\langle p \rangle + \langle (-1)^n p \rangle = \left\{ A \in M_n(\mathbb{Z}) : d(A) \,|\, \left(p + (-1)^{n+1}(-1)^n p \right) \right\}$$

$$= \{ A \in M_n(\mathbb{Z}) : d(A) \,|\, 0 \} = M_n(\mathbb{Z})$$

for any $p \in \mathbb{Z}$. If we choose p to be prime in $\mathbb{Z}$, this proves the Goldbach conjecture in $M_n(\mathbb{Z})$.

We've seen what happens when we take $p + (-1)^{n+1}q = 0$ in the theorem, since any integer divides zero. It might also be interesting to note what happens when $d(A) = 1$, since one divides every integer, and also since a matrix A of integers chosen at random should usually have $d(A) = 1$.

COROLLARY 2. *If $A \in M_n(\mathbb{Z})$ with $d(A) = 1$, then A can be written as the sum of two matrices with any determinants.*

Proof. For any $p, q \in \mathbb{Z}$, $1 \mid (p + (-1)^{n+1}q)$.

Many open problems still remain from Vaserstein's original paper [2], although some work has been done; the interested reader is urged to try to solve some of them.

REFERENCES

1. Henri Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, New York, 1993.
2. L. N. Vaserstein, Non-commutative number theory, *Contemporary Math.* 83 (1989), 445–449.
3. Jun Wang, Goldbach's problem in the ring $M_n(\mathbb{Z})$, *Amer. Math. Monthly* 99 (1992), 856–857.

A Characterization of Polynomials

K. M. ANDERSEN
Technical University of Denmark
DK-2800 Lyngby, Denmark

In an earlier issue [2] of this MAGAZINE, D. F. Bailey raises an interesting question about the divided difference $f[x_1, x_2, \ldots, x_n]$ on n points $x_1, x_2, \ldots, x_n$ of a function $f(x)$. This quantity is defined recursively by $f[x_1] = f(x_1)$ and

$$f[x_1, x_2, \ldots, x_n] = \frac{f[x_1, \ldots, x_{n-1}] - f[x_2, \ldots, x_n]}{x_1 - x_n}, \tag{1}$$

if $n \geq 2$ (clearly, the points $x_1, x_2, \ldots, x_n$ must be distinct). The mentioned question is whether the property

$$f[x_1, x_2, \ldots, x_n] = h(x_1 + x_2 + \cdots + x_n), \tag{2}$$

where $n \geq 2$ and $h(x)$ is some given function, guarantees that $f(x)$ is a polynomial of degree at most n. The answer is affirmative if $n = 2$, as shown by J. Aczél [1]. It is shown by Bailey [2] to be affirmative also if $n = 3$—provided $f(x)$ is a differentiable function. In this note it is shown to be true for all $n \geq 3$ without any further conditions on $f(x)$ and $h(x)$. As mentioned in the last lines of [2], conversely any polynomial $f(x)$ of degree at most $n \geq 2$ satisfies a relation (2), namely

$$f[x_1, x_2, \ldots, x_n] = \frac{1}{(n-1)!} f^{(n-1)}\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right). \tag{3}$$

In other words, relation (2) *characterizes* polynomials of degree at most $n \geq 2$. First, we prove the identity

$$\sum_{j=1}^{n} x_j^{n-1} \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{1}{x_j - x_k} = 1, \tag{4}$$

where $x_1, \ldots, x_n$ denote distinct points. The proof is by induction. The identity is trivial if $n = 1$ and true if $n = 2$. Suppose it is true for some $n \geq 2$. Using the partial

COROLLARY 2. *If $A \in M_n(\mathbb{Z})$ with $d(A) = 1$, then A can be written as the sum of two matrices with any determinants.*

Proof. For any $p, q \in \mathbb{Z}$, $1 \mid (p + (-1)^{n+1}q)$.

Many open problems still remain from Vaserstein's original paper [2], although some work has been done; the interested reader is urged to try to solve some of them.

REFERENCES

1. Henri Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, New York, 1993.
2. L. N. Vaserstein, Non-commutative number theory, *Contemporary Math.* 83 (1989), 445–449.
3. Jun Wang, Goldbach's problem in the ring $M_n(\mathbb{Z})$, *Amer. Math. Monthly* 99 (1992), 856–857.

A Characterization of Polynomials

K. M. ANDERSEN
Technical University of Denmark
DK-2800 Lyngby, Denmark

In an earlier issue [2] of this MAGAZINE, D. F. Bailey raises an interesting question about the divided difference $f[x_1, x_2, \ldots, x_n]$ on n points $x_1, x_2, \ldots, x_n$ of a function $f(x)$. This quantity is defined recursively by $f[x_1] = f(x_1)$ and

$$f[x_1, x_2, \ldots, x_n] = \frac{f[x_1, \ldots, x_{n-1}] - f[x_2, \ldots, x_n]}{x_1 - x_n}, \tag{1}$$

if $n \geq 2$ (clearly, the points $x_1, x_2, \ldots, x_n$ must be distinct). The mentioned question is whether the property

$$f[x_1, x_2, \ldots, x_n] = h(x_1 + x_2 + \cdots + x_n), \tag{2}$$

where $n \geq 2$ and $h(x)$ is some given function, guarantees that $f(x)$ is a polynomial of degree at most n. The answer is affirmative if $n = 2$, as shown by J. Aczél [1]. It is shown by Bailey [2] to be affirmative also if $n = 3$—provided $f(x)$ is a differentiable function. In this note it is shown to be true for all $n \geq 3$ without any further conditions on $f(x)$ and $h(x)$. As mentioned in the last lines of [2], conversely any polynomial $f(x)$ of degree at most $n \geq 2$ satisfies a relation (2), namely

$$f[x_1, x_2, \ldots, x_n] = \frac{1}{(n-1)!} f^{(n-1)}\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right). \tag{3}$$

In other words, relation (2) *characterizes* polynomials of degree at most $n \geq 2$. First, we prove the identity

$$\sum_{j=1}^{n} x_j^{n-1} \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{1}{x_j - x_k} = 1, \tag{4}$$

where $x_1, \ldots, x_n$ denote distinct points. The proof is by induction. The identity is trivial if $n = 1$ and true if $n = 2$. Suppose it is true for some $n \geq 2$. Using the partial

fraction decomposition of $x^{n-1}/((x-x_1)\dots(x-x_n))$,

$$\frac{x^n}{(x-x_1)\dots(x-x_n)} = x\sum_{j=1}^{n}\frac{x_j^{n-1}}{x-x_j}\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k}$$

with $x = x_{n+1}$ and the induction assumption we get

$$\sum_{j=1}^{n+1} x_j^n \prod_{\substack{k=1\\k\neq j}}^{n+1}\frac{1}{x_j-x_k} = \sum_{j=1}^{n} x_j^{n-1}\left[1+\frac{x_{n+1}}{x_j-x_{n+1}}\right]\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k} + x_{n+1}^n\prod_{k=1}^{n}\frac{1}{x_{n+1}-x_k}$$

$$= 1 - x_{n+1}\sum_{j=1}^{n}\frac{x_j^{n-1}}{x_{n+1}-x_j}\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k}$$

$$+ x_{n+1}\sum_{j=1}^{n}\frac{x_j^{n-1}}{x_{n+1}-x_j}\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k} = 1.$$

This proves the identity (4).

Next, we express the divided difference $f[x_1, x_2, \dots, x_n]$ in terms of the values $f(x_1), f(x_2), \dots, f(x_n)$. Indeed,

$$f[x_1, x_2, \dots, x_n] = \sum_{j=1}^{n} f(x_j)\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k}. \tag{5}$$

Again, the proof is by induction. Clearly, the expression (5) is valid if $n = 1$ or $n = 2$. Suppose it is valid for some $n \geq 2$. Using this on the points $x_1, x_2, \dots, x_n$ and the points $x_2, x_3, \dots, x_{n+1}$ as well, we get from (1)

$$f[x_1, x_2, \dots, x_{n+1}]$$

$$= \frac{1}{x_1-x_{n+1}}\left[\sum_{j=1}^{n} f(x_j)\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k} - \sum_{j=2}^{n+1} f(x_j)\prod_{\substack{k=2\\k\neq j}}^{n+1}\frac{1}{x_j-x_k}\right]$$

$$= \frac{1}{x_1-x_{n+1}}f(x_1)\prod_{k=2}^{n}\frac{1}{x_1-x_k} - \frac{1}{x_1-x_{n+1}}f(x_{n+1})\prod_{k=2}^{n}\frac{1}{x_{n+1}-x_k}$$

$$+ \sum_{j=2}^{n} f(x_j)\frac{1}{x_1-x_{n+1}}\left[\prod_{\substack{k=1\\k\neq j}}^{n}\frac{1}{x_j-x_k} - \prod_{\substack{k=2\\k\neq j}}^{n+1}\frac{1}{x_j-x_k}\right]$$

$$= f(x_1)\prod_{k=2}^{n+1}\frac{1}{x_1-x_k} + f(x_{n+1})\prod_{k=1}^{n}\frac{1}{x_{n+1}-x_k}$$

$$+ \sum_{j=2}^{n} f(x_j)\frac{1}{x_1-x_{n+1}}\left[\frac{1}{x_j-x_1} - \frac{1}{x_j-x_{n+1}}\right]\prod_{\substack{k=2\\k\neq j}}^{n}\frac{1}{x_j-x_k}$$

$$= \sum_{j=1}^{n+1} f(x_j)\prod_{\substack{k=1\\k\neq j}}^{n+1}\frac{1}{x_j-x_k}.$$

This proves the expression (5).

From (4) and (5) we derive two propositions. If $f(x)$ is identically 1, expression (5) becomes

$$\sum_{j=1}^{n} \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{1}{x_j - x_k} = 0, \quad n \geq 2. \tag{6}$$

If $f(x) = x^{n-1}$, $n \geqslant 2$, then $f[x_1, x_2, \ldots, x_n]$ is identically 1; this also follows from (3).

THEOREM. *If for some functions $f(x)$ and $h(x)$, $x \in \mathbb{R}$, the relation (2) holds for any set of n distinct points $x_1, x_2, \ldots, x_n$, $n \geq 2$, then $f(x)$ is a polynomial of degree at most n. Moreover, $h(x) = (1/(n-1)!)f^{(n-1)}(x/n)$, $x \in \mathbb{R}$.*

Proof. By (5) the assumption can be written

$$\sum_{j=1}^{n} f(x_j) \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{1}{x_j - x_k} = h(x_1 + x_2 + \cdots + x_n). \tag{7}$$

From (4), (6), and (7) we get

$$\sum_{j=1}^{n} \left[f(x_j) - f(0) - x_j^{n-1} h(0)\right] \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{1}{x_j - x_k} = h(x_1 + x_2 + \cdots + x_n) - h(0).$$

Hence we introduce the two auxiliary functions

$$g(x) = f(x) - f(0) - x^{n-1} h(0) \qquad \text{and} \qquad h_0(x) = h(x) - h(0). \tag{8}$$

Then $g(0) = h_0(0) = 0$ and

$$\sum_{j=1}^{n} g(x_j) \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{1}{x_j - x_k} = h_0(x_1 + x_2 + \cdots + x_n). \tag{9}$$

Substituting $-x_j$ for x_j, $j = 1, 2, \ldots, n$ in (9) gives

$$\begin{aligned} &\sum_{j=1}^{n} \frac{g(x_j) \pm g(-x_j)}{2} \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{1}{x_j - x_k} \\ &\quad = \frac{1}{2}\left[h_0(x_1 + x_2 + \cdots + x_n) \mp (-1)^n h_0(-x_1 - x_2 - \cdots - x_n)\right] \\ &\quad = h_1(x_1 + x_2 + \cdots + x_n), \end{aligned}$$

where $h_1(x) = \frac{1}{2}[h_0(x) \mp (-1)^n h_0(-x)]$. Notice that $h_1(0) = 0$. This means that the even part and the odd part of $g(x)$ (both being zero at $x = 0$) satisfy a relation of type (9). Clearly, $g(x)$ is a polynomial of degree at most n if the even part and the odd part of $g(x)$ are both polynomials of degree at most n. Hence we may assume that $g(x)$ is either even or odd.

Suppose first that n is odd, $n = 2p + 1$, $p \geq 1$. Choose fixed points $\alpha_1, \alpha_2, \ldots, \alpha_p$ such that $\pm\alpha_1, \pm\alpha_2, \ldots, \pm\alpha_{p-1}, \alpha_p$ are distinct and $\alpha_1, \alpha_2, \ldots, \alpha_{p-1} \neq 0$. Let x

denote a variable point. Substituting $x_1 = x$, $x_2 = -x$, $x_3 = \alpha_1$, $x_4 = -\alpha_1, \ldots, x_{2p-1} = \alpha_{p-1}$, $x_{2p} = -\alpha_{p-1}$ and $x_{2p+1} = \alpha_p$ in (9) we get

$$\frac{g(x)}{2x(x-\alpha_1)(x+\alpha_1)\ldots(x-\alpha_{p-1})(x+\alpha_{p-1})(x-\alpha_p)}$$

$$+\frac{g(-x)}{-2x(-x-\alpha_1)(-x+\alpha_1)\ldots(-x-\alpha_{p-1})(-x+\alpha_{p-1})(-x-\alpha_p)}$$

$$+\sum_{k=1}^{p-1}\left[\frac{g(\alpha_k)}{(\alpha_k-x)(\alpha_k+x)(\alpha_k-\alpha_1)(\alpha_k+\alpha_1)\ldots(\alpha_k+\alpha_k)\ldots(\alpha_k-\alpha_{p-1})(\alpha_k+\alpha_{p-1})(\alpha_k-\alpha_p)}\right.$$

$$\left.+\frac{g(-\alpha_k)}{(-\alpha_k-x)(-\alpha_k+x)(-\alpha_k-\alpha_1)(-\alpha_k+\alpha_1)\ldots(-\alpha_k-\alpha_k)\ldots(-\alpha_k-\alpha_{p-1})(-\alpha_k+\alpha_{p-1})(-\alpha_k-\alpha_p)}\right]$$

$$+\frac{g(\alpha_p)}{(\alpha_p-x)(\alpha_p+x)(\alpha_p-\alpha_1)(\alpha_p+\alpha_1)\ldots(\alpha_p-\alpha_{p-1})(\alpha_p+\alpha_{p-1})}=h_0(\alpha_p),$$

where $x \neq 0, \pm\alpha_1, \ldots, \pm\alpha_{p-1}, \pm\alpha_p$. This can be written

$$\frac{g(x)[x+\alpha_p]+g(-x)[x-\alpha_p]}{2x(x^2-\alpha_1^2)\ldots(x^2-\alpha_p^2)}$$

$$=h_0(\alpha_p)+\frac{g(\alpha_p)}{x^2-\alpha_p^2}\prod_{k=1}^{p-1}\frac{1}{\alpha_p^2-\alpha_k^2}$$

$$+\sum_{k=1}^{p-1}\frac{g(\alpha_k)[\alpha_k+\alpha_p]+g(-\alpha_k)[\alpha_k-\alpha_p]}{2\alpha_k(x^2-\alpha_k^2)}\prod_{\substack{j=1\\j\neq k}}^{p}\frac{1}{\alpha_k^2-\alpha_j^2},$$

by which

$$\frac{g(x)+g(-x)}{2}+\frac{\alpha_p}{x}\frac{g(x)-g(-x)}{2}$$

$$=h_0(\alpha_p)\prod_{k=1}^{p}(x^2-\alpha_k^2)+g(\alpha_p)\prod_{k=1}^{p-1}\frac{x^2-\alpha_k^2}{\alpha_p^2-\alpha_k^2}$$

$$+\sum_{k=1}^{p-1}\frac{g(\alpha_k)[\alpha_k+\alpha_p]+g(-\alpha_k)[\alpha_k-\alpha_p]}{2\alpha_k}\prod_{\substack{j=1\\j\neq k}}^{p}\frac{x^2-\alpha_j^2}{\alpha_k^2-\alpha_j^2}. \qquad (10)$$

This identity is valid for $x \neq 0, \pm\alpha_1, \ldots, \pm\alpha_p$. By inspection, however, it is valid for $x \neq 0, \pm\alpha_p$.

Consider first the case, where $g(x)$ is *even*. We chose $\alpha_p = 0$, by which $g(\alpha_p) = h_0(\alpha_p) = 0$, and get from (10)

$$g(x)=x^2\sum_{k=1}^{p-1}\frac{g(\alpha_k)}{\alpha_k^2}\prod_{\substack{j=1\\j\neq k}}^{p-1}\frac{x^2-\alpha_j^2}{\alpha_k^2-\alpha_j^2}. \qquad (11)$$

This is valid for $x \neq 0$, but inspection shows that (11) also holds with $x = 0$. We conclude that $g(x)$ is a polynomial of degree at most $2p - 2 = n - 3$.

Consider next the case, where $g(x)$ is *odd*. Then we choose $\alpha_p \neq 0$ and get from (10)

$$g(x) = x\left\{\frac{h_0(\alpha_p)}{\alpha_p}\prod_{k=1}^{p}(x^2-\alpha_k^2) + \frac{g(\alpha_p)}{\alpha_p}\prod_{k=1}^{p-1}\frac{x^2-\alpha_k^2}{\alpha_p^2-\alpha_k^2} + \sum_{k=1}^{p-1}\frac{g(\alpha_k)}{\alpha_k}\prod_{\substack{j=1\\ j\neq k}}^{p}\frac{x^2-\alpha_j^2}{\alpha_k^2-\alpha_j^2}\right\}. \tag{12}$$

This is valid for $x \neq 0, \pm\alpha_p$. Since $g(0)=0$ and $g(\pm\alpha_p) = \pm g(\alpha_p)$, inspection shows that (12) holds for all x. We conclude that $g(x)$ is a polynomial of degree at most $2p+1=n$.

From the preceding remark on the even and odd part we conclude that $g(x)$ is a polynomial of degree at most n—provided that n is odd.

Suppose next that n is even, $n = 2p+2$, $p \geq 1$ (recall that the case $n=2$ is already treated in [1]). Let $x_n = x_{2p+2} = 0$ in (9). Since $g(0)=0$ we get

$$\sum_{j=1}^{2p+1}\frac{g(x_j)}{x_j}\prod_{\substack{k=1\\ k\neq j}}^{2p+1}\frac{1}{x_j-x_k} = h_0(x_1+x_2+\cdots+x_{2p+1}). \tag{13}$$

Proceeding as in the previous case with (13) instead of (9), and with the extra restriction that also $\alpha_p \neq 0$, we arrive at the identity (10) with $g(x)$ substituted by $g(x)/x$. If $g(x)$ is even, then $g(x)/x$ is odd, and exactly as in the previous case we infer that $g(x)/x$ for $x \neq 0$ is a polynomial of degree at most $2p+1 = n-1$. Hence $g(x)$ is a polynomial of degree at most n, since $g(0)=0$. If $g(x)$ is odd, then $g(x)/x$ is even, and we get from (10) with $g(x)$ substituted by $g(x)/x$

$$g(x) = x\left\{h_0(\alpha_p)\prod_{k=1}^{p}(x^2-\alpha_k^2) + \frac{g(\alpha_p)}{\alpha_p}\prod_{k=1}^{p-1}\frac{x^2-\alpha_k^2}{\alpha_p^2-\alpha_k^2} + \sum_{k=1}^{p-1}\frac{g(\alpha_k)}{\alpha_k}\prod_{\substack{j=1\\ j\neq k}}^{p}\frac{x^2-\alpha_j^2}{\alpha_k^2-\alpha_j^2}\right\}.$$

This is valid for $x \neq 0, \pm\alpha_p$. Since $g(0)=0$ and $g(\pm\alpha_p) = \pm g(\alpha_p)$, also $x = 0, \pm\alpha_p$ can be allowed. We conclude that $g(x)$ is a polynomial of degree at most $2p+1 = n-1$. In either case $g(x)$ is a polynomial of degree at most n. Hence the same is true for $g(x)$ in the general case—provided that n is even.

It follows from (8) that $f(x)$ is a polynomial of degree at most n, for any $n \geq 2$. This proves the first part of the theorem.

It follows from (9) that $h_0(x)$, $x \in \mathbb{R}$—and by (8) also $h(x)$, $x \in \mathbb{R}$ are continuous functions. This is shown in [1] if $n=2$. Let $n \geq 3$ and let x_0 be arbitrary. Choose distinct points $x_1, x_2, \ldots, x_{n-1}$ such that $x_1+x_2+\cdots+x_{n-1} = 0$ and all different from x_0. Substituting $x_n = x$ in (9) we get

$$h_0(x) = \sum_{j=1}^{n-1}\frac{g(x_j)}{x_j-x}\prod_{\substack{k=1\\ k\neq j}}^{n-1}\frac{1}{x_j-x_k} + g(x)\prod_{k=1}^{n-1}\frac{1}{x-x_k}$$

for x in some interval around x_0. This proves that $h_0(x)$ is continuous at $x = x_0$, giving the desired property. Since $f(x)$ is a polynomial of degree at most n, relation (3) holds. Hence

$$h(x_1+x_2+\cdots+x_n) = \frac{1}{(n-1)!}f^{(n-1)}\left(\frac{x_1+x_2+\cdots+x_n}{n}\right),$$

which is true for all $x_1, x_2, \ldots, x_n$ by the continuity. From this we conclude that

$$h(x) = \frac{1}{(n-1)!} f^{(n-1)}\left(\frac{x}{n}\right), \quad x \in \mathbb{R},$$

i.e., $h(x)$ is a polynomial of degree at most 1. This proves the last part of the theorem.

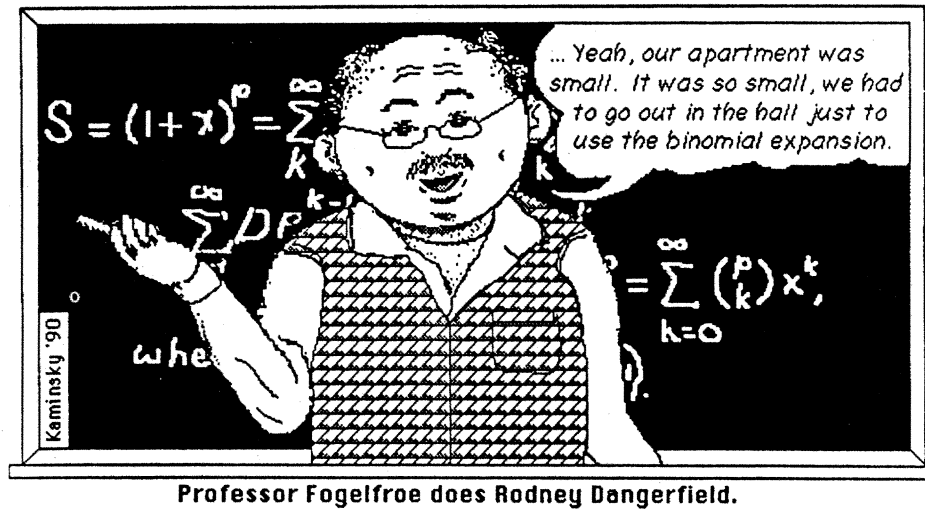
Remark. J. Schwaiger has also solved the problem of Bailey, within a more general frame and by another method (see [3]).

REFERENCES

1. J. Aczél. A mean value property of the derivative of quadratic polynomials—without mean values and derivatives, this MAGAZINE (58) 1985, 42–45.
2. D. F. Bailey. A mean-value property of cubic polynomials—without mean values, this MAGAZINE (65) 1992, 123–124.
3. J. Schwaiger. On a characterization of polynomials by divided differences, *Aequationes Math.* (48) 1994, 317–323.
4. R. L. Graham, *Mathematics Today: Twelve Informal Essays*, L. A. Steen, ed., Springer-Verlag, New York, 1978, 183–211.
5. S. K. Sahni, Algorithms for scheduling independent tasks, *J. Assoc. for Computing Machinery* **23** (1976), 116–127.
6. L. Schrage, *LINDO*, 4th edition, The Scientific Press, South San Francisco, CA, 1991.

Professor Fogelfroe

Professor F. Fogelfroe is Professor of Mathematics at ValuPak™ University, in Margo's Forehead, Minnesota.



Professor Fogelfroe does Rodney Dangerfield.

KEN KAMINSKY
Augsburg College
Minneapolis, MN 55454

which is true for all $x_1, x_2, \ldots, x_n$ by the continuity. From this we conclude that

$$h(x) = \frac{1}{(n-1)!} f^{(n-1)}\left(\frac{x}{n}\right), \quad x \in \mathbb{R},$$

i.e., $h(x)$ is a polynomial of degree at most 1. This proves the last part of the theorem.

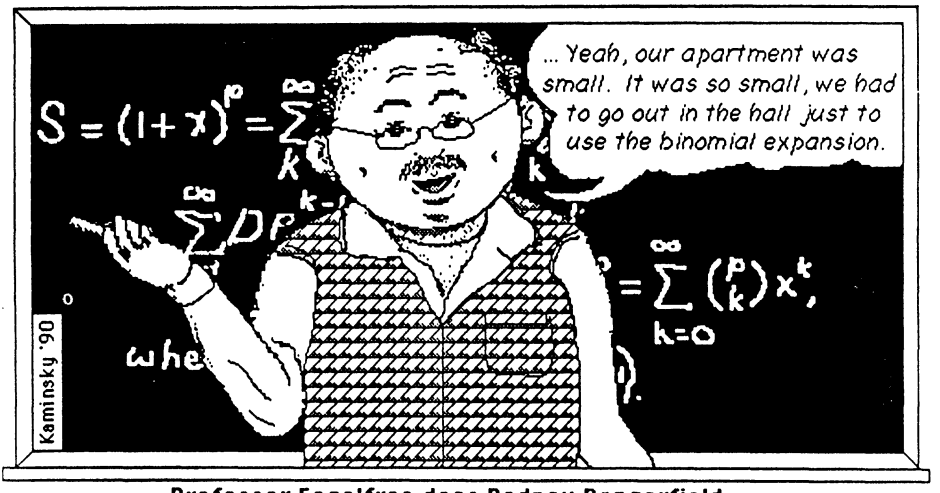
Remark. J. Schwaiger has also solved the problem of Bailey, within a more general frame and by another method (see [3]).

REFERENCES

1. J. Aczél. A mean value property of the derivative of quadratic polynomials—without mean values and derivatives, this MAGAZINE (58) 1985, 42–45.
2. D. F. Bailey. A mean-value property of cubic polynomials—without mean values, this MAGAZINE (65) 1992, 123–124.
3. J. Schwaiger. On a characterization of polynomials by divided differences, *Aequationes Math.* (48) 1994, 317–323.
4. R. L. Graham, *Mathematics Today: Twelve Informal Essays*, L. A. Steen, ed., Springer-Verlag, New York, 1978, 183–211.
5. S. K. Sahni, Algorithms for scheduling independent tasks, *J. Assoc. for Computing Machinery* **23** (1976), 116–127.
6. L. Schrage, *LINDO*, 4th edition, The Scientific Press, South San Francisco, CA, 1991.

Professor Fogelfroe

Professor F. Fogelfroe is Professor of Mathematics at ValuPak™ University, in Margo's Forehead, Minnesota.



Professor Fogelfroe does Rodney Dangerfield.

KEN KAMINSKY
Augsburg College
Minneapolis, MN 55454

PROBLEMS

GEORGE T. GILBERT, *Editor*

ZE-LI DOU, KEN RICHARDSON, and SUSAN G. STAPLES, *Assistant Editors*
Texas Christian University

Proposals

To be considered for publication, solutions should be received by September 1, 1996.

1494. *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, New York, and Ira M. Gessel, Brandeis University, Waltham, Massachusetts.*

Let $n \geq 2$ be a positive integer. Prove that n is prime if and only if $\binom{n-1}{k} \equiv (-1)^k \pmod{n}$ for $k = 0, 1, \ldots, n-1$.

1495. *Proposed by Achilleas Sinefakopoulos, student, University of Athens, Greece.*

Let angles B and C of ΔABC be acute, and let K be a point on arc BC of its circumcircle. Let L be the intersection of chords AK and BC. The feet of the perpendiculars from L to AB and to AC are M and N, respectively. Prove that if the area of ΔABC equals that of quadrilateral $AMKN$, then AK bisects angle A.

1496. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

Find a solution to the differential equation $d^2y/dx^2 = -kx/y^4$, $k > 0$, other than one of the form $y = ax^{3/5}$.

1497. *Proposed by Mihály Bencze, Braşov, Romania.*

Given positive real numbers $\alpha_1, \ldots, \alpha_m$, let $A_1, \ldots, A_m$ be sets of nonnegative integers such that $0 \in A_k$ and $|A_k \cap \{1, 2, \ldots, n\}| \geq \alpha_k \cdot n$ for $k = 1, \ldots, m$ and $n = 1, 2, \ldots$. Prove that

$$\left| \sum_{k=1}^{m} A_k \cap \{1, 2, \ldots, n\} \right| \geq \left(1 - \prod_{k=1}^{m} (1 - \alpha_k)\right) n,$$

where $\sum_{k=1}^{m} A_k = \{a_1 + \cdots + a_m : a_k \in A_k\}$.

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

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

Solutions and new proposals should be mailed to George T. Gilbert, Problems Editor, Department of Mathematics, Box 298900, Texas Christian University, Fort Worth, TX 76129, or mailed electronically (ideally as a LaTeX file) to `g.gilbert@tcu.edu`. *Readers who use e-mail should also provide an e-mail address.*

1498. *Proposed by J. C. Binz, University of Bern, Bern, Switzerland.*

For n a positive integer, express

$$\sum_{j\geq 0}\binom{n-j}{j}r^j(r-1)^{2n-2j} \quad \text{and} \quad \sum_{j\geq 0} j\binom{n-j}{j}r^j(r-1)^{2n-2j}$$

in closed form.

Quickies

Answers to the Quickies are on page 151.

Q847. *Proposed by Ismor Fischer, University of Wisconsin Medical School, Madison, Wisconsin.*

Given a planar angle with vertex O and point P in its interior, construct points A and B on the two rays of the angle so that P is on the line segment AB and ΔAOB has minimal area.

Q848. *Proposed by Homer White, Pikeville College, Pikeville, Kentucky.*

Evaluate $\int_0^{\pi/2} \ln(\sin x)\,dx$ by "low-tech" means.

Q849. *Proposed by Daniel B. Shapiro, The Ohio State University, Columbus, Ohio.*

Is there an $\mathbb{R}$-linear map $\lambda\colon \mathbb{R}[x] \to \mathbb{R}$ which is nonsingular in the sense that if $f \in \mathbb{R}[x]$ and $\lambda(f\cdot g) = 0$ for every $g \in \mathbb{R}[x]$, then $f = 0$?

Solutions

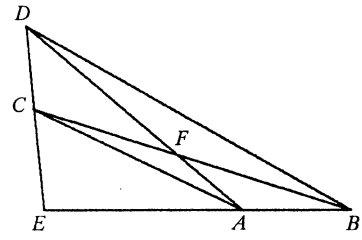
Right Triangle **April 1995**

1469. *Proposed by Roger Izard, Dallas, Texas.*

In triangle EDB, shown below, A and C lie on EB and ED respectively; CB and DA intersect at F. Also,

$$\frac{\text{Area}(\Delta EDB)}{\text{Area}(\Delta ECA)} = 6, \quad DC\cdot AB = 4, \quad \text{and} \quad \text{Area}(\Delta CFA) + \text{Area}(\Delta DFB) = 14/5.$$

Prove that DEB is a right triangle.



I. *Solution by Geoffrey A. Kandall, Hamden, Connecticut.*

Let $[DFB]$ denote the area of ΔDFB, and so forth. Applying Menelaus' theorem to ΔADE and ΔBCE,

$$\frac{FD}{AF} = \frac{DC}{CE} \cdot \frac{EB}{BA} \quad \text{and} \quad \frac{FB}{CF} = \frac{BA}{AE} \cdot \frac{ED}{DC}.$$

Consequently,

$$\frac{[DFB]}{[CFA]} = \frac{FD \cdot FB}{AF \cdot CF} = \frac{EB \cdot ED}{CE \cdot AE} = \frac{[EDB]}{[ECA]}. \tag{1}$$

We also have

$$[DFB] + [FAB] = [DAB] = \frac{1}{2} AB \cdot DE \sin \angle E,$$

$$[CFA] + [FAB] = [CAB] = \frac{1}{2} AB \cdot CE \sin \angle E.$$

Therefore,

$$[DFB] - [CFA] = \frac{1}{2} AB \cdot DC \sin \angle E. \tag{2}$$

In our particular case, we see from (1) that $[DFB] = 6[CFA]$. Since $[CFA] + [DFB] = 14/5$, it follows that $[CFA] = 2/5, [DFB] = 12/5$. It now follows from (2) that $\sin \angle E = 1$, i.e. $\angle E$ is a right angle.

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

For the sake of brevity let us put Area$(\Delta ECA) = [ECA]$, and so forth. Now we are given that

$$[DFB] + [FAB] + [ECA] + [FCD] + [CFA] = 6[ECA] \tag{1}$$

$$AB \cdot CD = 4 \tag{2}$$

$$[DFB] + [CFA] = 14/5. \tag{3}$$

From the figure we see that

$$[FCD] + [CFA] = \frac{1}{2} EA \cdot CD \sin \angle E, [DFB] + [FCD] = \frac{1}{2} EB \cdot CD \sin \angle E. \tag{4}$$

From (2) and (4) it follows that

$$[DFB] - [CFA] = \frac{1}{2} AB \cdot CD \sin \angle E = 2 \sin \angle E. \tag{5}$$

Solving (3) and (5) for $[DFB]$ and $[CFA]$, we obtain

$$[DFB] = 7/5 + \sin\angle E, [CFA] = 7/5 - \sin\angle E. \tag{6}$$

Next we observe that $[FAB]/[CFA] = FB/FC = [DFB]/[FCD]$, i.e.,

$$[FAB][FCD] = [DFB][CFA], \tag{7}$$

and that

$$\frac{[FAB] + [CFA]}{[ECA]} = \frac{AB}{EA} = \frac{[FAB] + [DFB]}{[ECA] + [FCD] + [CFA]},$$

from which it follows that

$$[FAB][FCD] + ([FAB] + [ECA] + [FCD] + [CFA])[CFA] = [ECA][DFB].$$

In view of (7) and (1) this comes down to

$$6[CFA] = [DFB]. \tag{8}$$

Combining (6) and (8) finally yields $\sin\angle E = 1$.

Also solved by Robin Chapman (U.K.), Miguel Amengual Covas (Spain), Kee-Wai Lau (Hong Kong), Achilleas Sinefakopoulos (student, Greece), Harry Weingarten, Po Kee Wong and Adam Wong, Paul J. Zwier, and the proposer.

Submatrices of 0–1 Matrices **April 1995**

1470. *Proposed by John A. Hendrickson, Jr., Academy of Natural Sciences, Philadelphia, Pennsylvania.*

Consider the set of $n \times n$ matrices A, $n \geq 3$, with all entries equal to 0 or 1 and the sum of the entries of each row and of each column equal to 2. Find the maximum and minimum number of quadruples (i, i', j, j'), $i < i'$, for which $A_{ij} = A_{i'j'} = 1$ and $A_{ij'} = A_{i'j} = 0$ as a function of n. For example, there are 12 such quadruples in the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Solution by Harris Kwong, SUNY College at Fredonia, Fredonia, New York.

Define a quadruple (i, i', j, j') to be "good" if it satisfies the given conditions. Let $g(A)$ be the number of good quadruples in A. Since interchanging rows or columns does not alter the value of $g(A)$, we may assume A is in the form

$$\begin{pmatrix} J & & & & & \\ & \ddots & & & 0 & \\ & & J & & & \\ & & & S_{k_1} & & \\ & 0 & & & \ddots & \\ & & & & & S_{k_t} \end{pmatrix},$$

where there are m copies of $J = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and S_k is the $k \times k$ matrix ($k \geq 3$)

$$S_k = \begin{pmatrix} 1 & 1 & 0 & & & & & \\ 1 & 0 & 1 & & & 0 & & \\ 0 & 1 & 0 & \ddots & & & & \\ & & \ddots & \ddots & & & & \\ & & & & 0 & 1 & 0 \\ & 0 & & & 1 & 0 & 1 \\ & & & & 0 & 1 & 1 \end{pmatrix}.$$

Any pair of 1's selected from distinct blocks of A gives a good quadruple. Thus

$$g(A) = 16\binom{m}{2} + 4m\sum_{i=1}^{t} 2k_i + \sum_{1 \leq i < j \leq t} 2k_i \cdot 2k_j + \sum_{i=1}^{t} g(S_{k_i}).$$

To compute $g(S_k)$, observe that for each pair (i, j) with $A_{ij} = 1$, there exist $2k - 5$ pairs (i', j') with $A_{i'j'} = 1$ and $A_{ij'} = A_{i'j} = 0$. Consequently, $g(S_k) = 2k(2k - 5)/2 = 2k^2 - 5k$. Furthermore,

$$\sum_{1 \leq i < j \leq t} 2k_i \cdot 2k_j + \sum_{i=1}^{t} g(S_{k_i}) = g(S_{k_1 + \cdots + k_t}).$$

Therefore $g(A) = g(B)$ for

$$B = \begin{pmatrix} J & & & 0 \\ & \ddots & & \\ & & J & \\ 0 & & & S_{n-2m} \end{pmatrix}.$$

If $n - 2m \geq 5$, compare B with

$$B' = \begin{pmatrix} J & & & 0 \\ & \ddots & & \\ & & J & \\ 0 & & & S_{n-2m-2} \end{pmatrix},$$

which contains $m + 1$ copies of J. We have

$$g(B) = 16\binom{m}{2} + 8m(n - 2m) + 2(n - 2m)^2 - 5(n - 2m) = 2n^2 - 5n + 2m$$

and

$$g(B') = 2n^2 - 5n + 2m + 2.$$

We see that the minimum number of quadruples $q(n)$ is obtained when $A = S_n$. Thus $q(n) = 2n^2 - 5n$. On the other hand, the maximum number of quadruples $Q(n)$

is obtained when A contains as many blocks of J as possible. It follows that

$$Q(n) = \begin{cases} 2n^2 - 4n & \text{if } n \text{ is even,} \\ 2n^2 - 4n - 3 & \text{if } n \text{ is odd.} \end{cases}$$

Also solved by David Callan, Robin Chapman (U.K.), Marty Getz and Dixon Jones, Jerrold W. Grossman, O.P. Lossers (The Netherlands), Kevin McDougal, Robert Patenaude, and the proposer.

Sequence of Floors **April 1995**

1471. *Proposed by Bill Correll, Jr., student, Denison University, Granville, Ohio.*

For positive integers n, define $f(n)$ to be the smallest positive integer j such that

$$\left\lfloor \frac{n^2}{j} \right\rfloor = \left\lfloor \frac{n^2}{j+1} \right\rfloor,$$

where $\lfloor x \rfloor$ denotes the floor function. Let

$$c(n) = \left\lfloor \frac{n^2}{f(n)} \right\rfloor.$$

Prove that

(i) $\{f(n)\}_{n=1}^{\infty}$ consists of all integers except the perfect squares, and
(ii) $f(n) + c(n) = 2n$.

Solution by Anne L. Young, Loyola College, Baltimore, Maryland.

Let k be the integer that satisfies

$$k^2 - k < n \le (k+1)^2 - (k+1) = k^2 + k.$$

We will show that $f(n) = n + k$.

First note that $\lfloor n^2/j \rfloor = q$ if and only if $n^2 = j \cdot q + r$, where $0 \le r < j$. Thus if $\lfloor n^2/j \rfloor = \lfloor n^2/(j+1) \rfloor$,

$$n^2 = j \cdot q + r, \quad \text{where } 0 \le r < j,$$

and

$$n^2 = (j+1) \cdot q + s, \quad \text{where } 0 \le s < j+1.$$

Subtracting gives $q + s = r < j$; hence $q < j$. This in turn implies that $n < j$ and hence $n < f(n)$.

Now for $i \le k$,

$$n^2 = (n+i) \cdot (n-i) + i^2,$$

with $i^2 < n + i$. Thus, $\lfloor n^2/(n+i) \rfloor = (n-i)$. Therefore $f(n) \ge n + k$. Now

$$n^2 = (n+k+1) \cdot (n-k) + (k^2 + k - n).$$

Further,

$$0 \le k^2 + k - n = (k^2 - k) + k + k - n < n + (n+1) + k - n = n + k + 1.$$

Thus $f(n) = n + k$. Moreover, by the definition of k,

$$k^2 < f(n) \le (k+1)^2 - 1.$$

Thus $f(n)$ is never a perfect square. To see that $f(n)$ takes on all other values, let s be an integer such that

$$k^2 < s \le (k+1)^2 - 1.$$

Then our previous work shows that $f(s-k) = s$.

Finally $c(n) = \lfloor n^2/f(n) \rfloor = n - k$. Therefore $f(n) + c(n) = 2n$.

Comment. David Callan refers us to the solution to problem E3340 in *The American Mathematical Monthly* 98 (1991), 62–63, where it is shown that the sequence $\left(n + \left\lfloor (n + n^{1/k})^{1/k} \right\rfloor\right)$ consists of all positive integers that are not kth powers.

Also solved by J. C. Binz (Switzerland), David Callan, Robin Chapman (U.K.), John Christopher, Graham Lord, O.P. Lossers (The Netherlands), Gordon Rice, Heinz-Jürgen Seiffert, Achilleas Sinefakopoulos (student, Greece), and the proposer.

Maximal Area of Quadrilaterals **April 1995**

1472. *Proposed by Erhan Gürel, Middle East Technical University, Ankara, Turkey.*

Let Q denote an arbitrary convex quadrilateral inscribed in a fixed circle, and let $\mathcal{F}(Q)$ be the set of inscribed convex quadrilaterals whose sides are parallel to those of Q. Prove that the quadrilateral in $\mathcal{F}(Q)$ of maximum area is the one whose diagonals are perpendicular to one another.

Solution by Robin Chapman, University of Exeter, Exeter, United Kingdom.

Let Q have vertices A, B, C, D in clockwise order. Let $Q' \in \mathcal{F}(Q)$. Assume that the vertices of Q' are labeled in clockwise order as A', B', C', and D' with $A'B'$ parallel to AB, and so on. Then the angles of Q and Q' are the same, say α, β, γ and δ at the vertices A, B, C and D (or A', B', C' and D'), respectively. By familiar theorems on circles, the chords AC and $A'C'$ have equal length, and so do BD and $B'D'$. Let P be the point of intersection of AC and BD, and let θ be the angle between these two lines. Then the area of Q is

$$\frac{1}{2}(|AP|\cdot|BP| + |BP|\cdot|CP| + |CP|\cdot|DP| + |DP|\cdot|AP|)\sin\theta = \frac{1}{2}|AC|\cdot|BD|\sin\theta.$$

Consequently the area of a quadrilateral in $\mathcal{F}(Q)$ is maximized when the diagonals are perpendicular.

However there may not be a $Q' \in \mathcal{F}(Q)$ whose diagonals are perpendicular. We can only have perpendicular diagonals if the sum of their squared lengths exceeds the square of the diameter of the circle. If this is not true, then the maximal area is assumed by one of the triangles which is a degenerate form of the quadrilaterals in $\mathcal{F}(Q)$. For example, assume that the angles $\alpha, \beta \ge \pi/2$. Then this triangle occurs when $A' = B'$, and has area $\frac{1}{2}|AC|\cdot|BD|\sin(\alpha + \beta - \pi)$.

Also solved by Frank Eccles, Hans Kappus (Switzerland), Nick Lord (England), O. P. Lossers (The Netherlands), James S. Robertson, Noah Rosenberg (student), Achilleas Sinefakopoulos (student, Greece), Ralph I. Vanderslice, Jr., Wai Ling Yee (student, Canada), and the proposer.

Roots of Determinant **April 1995**

1473. *Proposed by Gerald A. Heuer, Concordia College, Moorhead, Minnesota.*

Let $B_n(z)$ denote the determinant of the $(2n+1) \times (2n+1)$ matrix whose entries are given by $b(1, j) = 1$ for all j, $b(j, j) = 2$ for $j = 2, 3, \ldots, 2n+1$,

$$b(i+1, n+i+1) = b(n+i+1, i) = -z \text{ for } i = 1, 2, \ldots n,$$

and all other $b(i,j)=0$. For example,

$$B_1(z)=\det\begin{pmatrix}1&1&1\\0&2&-z\\-z&0&2\end{pmatrix}\qquad B_2(z)=\det\begin{pmatrix}1&1&1&1&1\\0&2&0&-z&0\\0&0&2&0&-z\\-z&0&0&2&0\\0&-z&0&0&2\end{pmatrix}.$$

Find all (complex) roots of $B_n(z)$.

I. *Solution by Yongzhi Yang, University of St. Thomas, St. Paul, Minnesota.*

Let $A_n(z)$ denote the determinant of the $(2n+1)\times(2n+1)$ matrix whose entries are given by $a(1,j)=1$ for all j, $a(j,j)=2$ for $j=2,3,\ldots,2n$, $a(2n+1,2n+1)=0$, $a(i+1,n+i+1)=a(n+i+1,i)=-z$ for $i=1,2,\ldots,n$, and all other $a(i,j)=0$. Beginning with the cofactor expansions by the last rows, it is easy to obtain the following difference equations for $A_n(z)$ and $B_n(z)$:

$$A_n(z)=z^2A_{n-1}(z),\qquad n=2,3,4,\ldots. \tag{1}$$

and

$$B_n(z)=z^2A_{n-1}(z)+4B_{n-1}(z),\qquad n=2,3,4,\ldots, \tag{2}$$

with the initial conditions $A_1(z)=z^2+2z$ and $B_1(z)=z^2+2z+4$.

Evaluating the solution to equation (1) with the corresponding initial condition we obtain

$$A_n(z)=z^{2(n-1)}(z^2+2z)=z^{2n}+2z^{2n-1},\qquad n=1,2,3,\ldots. \tag{3}$$

Substituting (3) into equation (2) and solving with the corresponding initial condition, we obtain $B_n(z)=\sum_{i=0}^{2n}2^iz^{2n-i}$. Solving $B_n(z)=0$ yields

$$z=2\left[\cos\left(\frac{2k\pi}{2n+1}\right)+i\sin\left(\frac{2k\pi}{2n+1}\right)\right],\qquad k=1,2,3,\ldots,2n.$$

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

The roots of $B_n(z)$ are the numbers of the form 2ζ, where $\zeta^{2n+1}=1$ and $\zeta\neq 1$. Since there are $2n$ of these and since the degree of $B_n(z)$ in z equals $2n$, it suffices to find a vector in the kernel of the matrix B corresponding to $B_n(2\zeta)$, for all such ζ. It is easily checked that the vector

$$\mathbf{v}=(1,\zeta^2,\zeta^4,\ldots,\zeta^{4n})^t$$

works. Indeed, $(B\mathbf{v})_1=1+\zeta^2+\cdots+\zeta^{4n}=(1-\zeta^{4n+2})/(1-\zeta)=0$ and $(B\mathbf{v})_i=2\zeta^{2i}-2\zeta\zeta^{2i+2n}=0$.

Comment. Both of the above solutions generalize to the case where the 2's in the diagonal are replaced by y's.

Also solved by J. C. Binz (Switzerland), Stan Byrd and Ronald L. Smith, David Callan, Robin Chapman (U.K.), Charles K. Cook, Heinz-Jürgen Seiffert, Achilleas Sinefakopoulos (student, Greece), Robert K. Stump, and the proposer.

Answers

Solutions to the Quickies on page 144.

A847. The points A and B form the unique line segment having P as its midpoint. One may construct A and B by first constructing the point C for which P is the midpoint of OC, and then constructing lines through C parallel to the two rays of angle O. To see that ΔAOB gives minimal area, consider points A' on ray OA and B' on ray OB such that P is on $A'B'$. We may assume $OA' < OA$. The uniqueness of the segment AB implies $PA' < PB'$, for otherwise there exist a point A'' on segment OA' and a point B'' on ray OB' with P on $A''B''$ and $PA'' = PB''$. Letting $\theta = \angle APA' = \angle BPB'$, we have $\frac{1}{2}PA' \cdot PA \sin\theta < \frac{1}{2}PB' \cdot PB \sin\theta$, hence $\text{Area}(\Delta APA') < \text{Area}(\Delta BPB')$. Adding the area of the common quadrilateral $OA'PB$ gives $\text{Area}(\Delta AOB) < \text{Area}(\Delta A'OB')$. Thus ΔAOB has minimal area.

A848. For $0 < x \le \pi/2$, $\ln(\sin x) = \ln(\sin x/x) + \ln x$. Because $\int_0^{\pi/2} \ln(\sin x/x)\,dx$ is the integral of a continuous function (defining the integrand to be 0 at 0) and the improper integral $\int_0^{\pi/2} \ln x\,dx$ converges, so does $\int_0^{\pi/2} \ln(\sin x)\,dx$. We have

$$\begin{aligned}\int_0^{\pi/2} \ln(\sin x)\,dx &= \int_0^{\pi/4} \ln(\sin x)\,dx + \int_{\pi/4}^{\pi/2} \ln(\sin x)\,dx \\ &= \int_0^{\pi/4} \ln(\sin x)\,dx + \int_0^{\pi/4} \ln(\cos x)\,dx = \int_0^{\pi/4} \ln\left(\frac{\sin 2x}{2}\right) dx \\ &= \int_0^{\pi/4} \ln(\sin 2x)\,dx - \frac{\pi \ln 2}{4} = \frac{1}{2}\int_0^{\pi/2} \ln(\sin x)\,dx - \frac{\pi \ln 2}{4}.\end{aligned}$$

Solving for the integral yield $\int_0^{\pi/2} \ln x\,dx = -\pi \ln 2/2$.

A849. Yes, such a λ exists even when $\mathbb{R}$ is replaced by an arbitrary field K. For example, define λ by $\lambda(x^{2^n}) = 1$ and $\lambda(x^m) = 0$ for m not of the form 2^n. If $f(x) = a_0x^d + a_1x^{d-1} + \cdots + a_d$, $a_0 \neq 0$, and $2^{n-1} > d$, then $\lambda(f(x)\cdot x^{2^n-d}) = a_0 \neq 0$. (In fact, there exists a polynomial f of degree d with $\lambda(f \cdot K[x]) = 0$ if and only if the sequence $(\lambda(x^n))$ satisfies a linear recurrence of degree d.)

REVIEWS

PAUL J. CAMPBELL, *editor*
Beloit College

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

Goldfeld, Dorian, Beyond the last theorem, *The Sciences* (March/April 1996) 34–40.

How complicated must a diophantine equation be to be unsolvable? No algorithm can determine whether an arbitrary diophantine equation in nine variables has solutions (this is a strengthening, by J.P. Jones in 1982, of the negative solution to Hilbert's Tenth Problem by Y. Matiyasevich in 1970). But could there be an algorithm for diophantine equations in, say, three variables? In resolving Fermat's Last Theorem (FLT), Wiles and Taylor proved the Shimura-Taniyama-Weil (STW) conjecture, which "brought the heavy machinery of rigid motions to the theory of elliptic curves"—and whose truth suggests there should be such an algorithm. The key link is the ABC conjecture of J. Oesterlé (University of Paris VI) and David W. Masser (University of Basel), "the most important unsolved problem in Diophantine analysis": *Let sqp denote the square-free part of an integer. For any $n > 1$, the quantity $[sqp(ABC)]^n/C$ attains a minimum value over all A and B relatively prime with $C = A + B$.* The quantity $sqp[(ABC)^2] = sqp(ABC)$ occurs in the conductor of the Frey elliptic curve $y^2 = x(x - A)(x + B)$, which arises in examination of the STW conjecture. The ABC conjecture leads to the solution of many diophantine equations; author Goldfeld, who won the Cole Prize for number theory in 1987, shows how FLT follows easily from it. He concludes by suggesting that diophantine unsolvability could form a basis for practical cryptography.

Cipra, Barry, Calculus reform sparks a backlash, *Science* 271 (16 February 1996) 901–902.

Are calculus students in so-called "calculus reform" classes "learning more and liking it better"? Or do the new courses "water down the subject and coddle students with computers" and foster dependence on calculators for division by 10? Such is a distillation of the issues discussed a panel discussion at the AMS-MAA Meeting in January. Even supporters of "calculus reform" disagree on criteria for judging its success with students: higher grades? spending more time on the course? better understanding (but how do you measure that?)? more favorable attitudes toward mathematics? more students becoming mathematics majors? A more basic scientific question is whether any of the results that are regarded as favorable are anything more than could be expected from the halo effect (experimenter bias in favor of the reforms) and/or the Hawthorne effect (people aware of being in an experiment change their behavior). Even the enhanced interest of university mathematics faculty in calculus instruction could wane, when large NSF grants for calculus experimentation end and the big money and media attention go to to other causes.

Roberts, A. Wayne (ed.), *Calculus: The Dynamics of Change*, MAA, 1996; x + 166 pp, $34.95 (P). ISBN 0–88385–098–2.

It has been eleven years since the conference at Tulane that started the movement to reform instruction in calculus. This book is a brief summary volume of perspectives on what has been tried since then. It includes visions of what a calculus course should be like, essays on how to plan to bring about change, copies of two dozen final examinations, and what the ripple effects may be on high schools, on courses past calculus, and on client disciplines.

van der Poorten, Alf, *Notes on Fermat's Last Theorem*, Wiley, 1996; xvi + 222 pp, $44.95. ISBN 0–471–06261–8.

This book features expanded and revised versions of marvelously informative, illuminating, and and entertaining lectures on FLT that were originally available over the Internet. "It would have been dissatisfying to have had to finish this book with Wiles' proof still incomplete. Fortunately, I suffered writer's block for a sufficiently long time for Wiles to settle the matter." Included are three appendices: a short version of the lectures, "for those who only want to pretend that they have read the rest of this book"; the short story "The Devil and Simon Flagg," by Arthur Porges, in which the devil is thwarted by FLT; and Eric Zorn's humorous reporting, "Math Riots Prove Fun Incalculable."

Devlin, Keith, University of Rochester eliminates Ph.D. program, *Focus* 16 (1) (February 1996) 1, 5. Jackson, Allyn, Downsizing at Rochester: Mathematics Ph.D. program cut, *Notices of the American Mathematical Society* 43 (3) (March 1996) 300–306; Jaffe, Arthur, et al., Demotion of mathematics meets groundswell of protest, 307–313. Holden, Constance, Does Rochester without mathematics add up?, *Science* 271 (19 January 1996) 284. Rochester to eliminate graduate program in mathematics, *SIAM News* 29 (1) (January/February 1996) 2, 4. Rochester news, `http://www.ams.org/committee/profession/rochester/rochester.html`

Can you have a good research university without a graduate program in mathematics? Are mathematicians just knee-jerk reacting to downsizing in academia? The University of Rochester plans to balance its budget by reducing the student body by 20%, the faculty by 10%, and the mathematics faculty by over 50% (through attrition and retirement incentives)—because of elimination of the graduate program in mathematics (linguistics, chemical engineering, and comparative literature go too). Why mathematics? "Modest distinction... dwindling numbers of students [five Ph.D.'s per year]... less than optimal undergraduate instruction (especially in calculus)... minimal linkages with other departments and programs." The plan is to improve calculus instruction by hiring non-tenure-track faculty and using faculty from other disciplines (which already teach graduate courses in applied mathematics for their own students). For its part, the Mathematics Dept. points to no student complaints about teaching, good student evaluations, and little interest from other departments in linkages (the mathematics faculty are predominantly algebraic topologists). The AMS Web site contains a link to the University's Mathematics Dept., the report of an AMS factfinding committee sent to Rochester, a letter from the AMS president to the administration, an AMS Council resolution, and more. The recent domestic "overproduction" of Ph.D.'s, relative to market demand and supplies from abroad, suggests that some Ph.D. programs should go: the "vanity" programs, ones low in quality (to everyone's eyes but their own) and/or below critical mass in numbers of faculty and students. To answer "No!" to the question, "Can you have a good research university without a graduate program in mathematics?," suggests that a major role of such a program is research "linkages" with other departments or even just serving other graduate programs. These are not even options for a department of pure mathematicians in a university where other departments offer the applied mathematics that they feel their students need.

Swetz, Frank, et al., eds., *Learn from the Masters!*, MAA, 1995; x + 303 pp, $23 (P). ISBN 0–88385–703–0.

It was Abel who said, "[I]f one wants to make progress in mathematics one should study the masters." This book is the proceedings of a conference on the history of mathematics that took place in 1988 near where Abel lived briefly and was buried. The book focuses on the use of history as a pedagogical tool in teaching mathematics, with 23 essays on history in school mathematics and in higher mathematics. Everyone who reads this book will learn something that they can take to their classroom and use.

Tucker, Alan C. (ed.), *Models that Work: Case Studies in Effective Undergraduate Mathematics Programs*, MAA, 1996; x + 78 pp, $24 (P). ISBN 0–88385–096–6.

This report summarizes effective practices at some mathematics programs that either attract and train large numbers of mathematics majors, prepare students for advanced study in mathematics, prepare students for teaching school mathematics, or attract and train students from underrepresented groups. The committee that prepared this report visited the institutions, which include two research universities, three other state colleges and universities, two private liberal arts colleges, two historically black institutions, and a community college (for their names, you'll have to get the report). Common features of these effective programs were particular "states of mind" held by most or all faculty (respecting students: teaching the students one has, not the ones one would wish for; caring about students; and enjoying one's career as a collegiate educator), plus a curriculum geared toward student needs instead of faculty values, and an interest in using a variety of pedagogical approaches.

Peterson, Ivars, The song in the stone: Developing the art of telecarving a minimal surface, *Science News* 149 (17 February 1996) cover, 99, 110–111.

Mathematician-sculptor Helaman Ferguson has begun using a "virtual image projector" to translate a surface determined by mathematical equations into instructions on how much material to cut away from an uncarved stone to produce a sculpture of the surface. He has used the projector to make a number of sculptures of the Costa surface, a completely embedded minimal surface of minimal topology.

Gurkewitz, Rona, and Bennett Arnstein, *3-D Geometric Origami: Modular Polyhedra*, Dover, 1995; iv + 73 pp, $6.95 (P). ISBN 0–486–28863–3. Hanson, Robert M., *Molecular Origami: Precision Scale Models from Paper*, University Science Books, 1995; xiii + 223 pp, $22 (P). ISBN 0–935702–30–X.

The first of these books contains diagrams for constructing more than 50 three-dimensional polyhedral models by means of paper-folding (without cutting). The instructions are simplified by the use of a limited syntax of symbols. The second book will interest mainly chemists, as it gives traceable scale-model nets for cut-and-fold models of more than 70 molecules, most at a scale of 3 cm to 1 Å= 100 pm (picometers).

ERRATUM: In Vol. 69, No. 1 (February 1996), p. 77, the year of the article about the golf ball aerodynamics of P.G. Tait should be 1993:

Denley, Chris, and Chris Pritchard, The golf ball aerodynamics of Peter Guthrie Tait, *Mathematical Gazette* (1993) 298–313.

56th ANNUAL WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

A-1. Let S be a set of real numbers which is closed under multiplication (that is, if a and b are in S, then so is ab). Let T and U be disjoint subsets of S whose union is S. Given that the product of any *three* (not necessarily distinct) elements of T is in T and that the product of any three elements of U is in U, show that at least one of the two subsets T, U is closed under multiplication.

Solution. Suppose T is not closed under multiplication. Then there are elements $t_1, t_2 \in T$ with $t_1 t_2 \notin T$, and since S is closed under multiplication, $t_1 t_2 \in U$. Now consider any two elements $u_1, u_2 \in U$; we'll show that $u_1 u_2 \in U$ (and thus that U is closed under multiplication). Suppose $u_1 u_2 \notin U$. Then $u_1 u_2 \in T$, so $t_1 \cdot t_2 \cdot u_1 u_2 \in T$ (as a product of three elements of T), but also $t_1 \cdot t_2 \cdot u_1 u_2 = (t_1 t_2) \cdot u_1 \cdot u_2 \in U$ (as a product of three elements of U), which is a contradiction since T and U are disjoint. So $u_1 u_2 \in U$, and we are done.

A-2. For what pairs (a, b) of positive real numbers does the improper integral

$$\int_b^\infty \left(\sqrt{\sqrt{x+a} - \sqrt{x}} - \sqrt{\sqrt{x} - \sqrt{x-b}} \right) dx$$

converge?

Solution. The integral converges if and only if $a = b$.

Note that the integrand is defined and continuous for all $x \geq b$, so the only issue is convergence at ∞.

Since $(\sqrt{x+a} - \sqrt{x})(\sqrt{x+a} + \sqrt{x}) = a$ and $(\sqrt{x} - \sqrt{x-b})(\sqrt{x} + \sqrt{x-b}) = b$, we can rewrite the integrand as

$$\frac{\sqrt{a}}{\sqrt{\sqrt{x+a} + \sqrt{x}}} - \frac{\sqrt{b}}{\sqrt{\sqrt{x} + \sqrt{x-b}}} = \frac{1}{\sqrt[4]{x}} \left(\frac{\sqrt{a}}{A} - \frac{\sqrt{b}}{B} \right)$$

where

$$A = \sqrt{\sqrt{1 + \tfrac{a}{x}} + 1}, \text{ and } B = \sqrt{1 + \sqrt{1 - \tfrac{b}{x}}}.$$

If $a \neq b$, the quantity in the brackets approaches the nonzero number $\sqrt{\frac{a}{2}} - \sqrt{\frac{b}{2}}$ as $x \to \infty$, so for large enough x the absolute value of the integrand is at least $\frac{c}{2\sqrt[4]{x}}$, where $c = \left| \sqrt{\frac{a}{2}} - \sqrt{\frac{b}{2}} \right|$. The integrand then diverges by comparison with the divergent integral $\int_1^\infty \frac{dx}{\sqrt[4]{x}}$.

On the other hand, if $a = b$, the integrand equals

$$\begin{aligned} \frac{\sqrt{a}}{\sqrt[4]{x}} \left(\frac{1}{A} - \frac{1}{B} \right) &= \frac{\sqrt{a}}{\sqrt[4]{x}} \cdot \frac{1}{AB} \ (A - B) \\ &= \frac{\sqrt{a}}{\sqrt[4]{x}} \cdot \left(\frac{A^2 - B^2}{AB(A+B)} \right) \\ &= \frac{\sqrt{a}}{\sqrt[4]{x}} \cdot \left(\frac{A^4 - B^4}{AB(A+B)(A^2+B^2)} \right) \\ &= \frac{-2a\sqrt{a}}{x\sqrt[4]{x}(AB)(A+B)(A^2+B^2)}. \end{aligned}$$

Since $AB(A+B) \longrightarrow \sqrt{2} \cdot \sqrt{2} \cdot (\sqrt{2} + \sqrt{2}) = 4\sqrt{2}$ as $x \to \infty$, the absolute value of the integrand is then less than $\frac{a\sqrt{a}}{x^{5/4}}$, and the integral converges by comparison with $\int_1^\infty \frac{dx}{x^{5/4}}$.

A-3. The number $d_1 d_2 \ldots d_9$ has nine (not necessarily distinct) decimal digits. The number $e_1 e_2 \ldots e_9$ is such that each of the nine 9-digit numbers formed by replacing just one of the digits d_i in $d_1 d_2 \ldots d_9$ by the corresponding digit e_i $(1 \leq i \leq 9)$ is divisible

by 7. The number $f_1 f_2 \ldots f_9$ is related to $e_1 e_2 \ldots e_9$ in the same way: that is, each of the nine numbers formed by replacing one of the e_i by the corresponding f_i is divisible by 7. Show that, for each i, $d_i - f_i$ is divisible by 7.

[For example, if $d_1 d_2 \ldots d_9 = 199501996$, then e_6 may be 2 or 9, since 199502996 and 199509996 are multiples of 7.]

Solution. Suppose $d_1 \ldots d_9 \equiv a \pmod 7$. Then for each i, $1 \le i \le 9$,

$$\begin{aligned} a \equiv a - 0 &\equiv (d_1 \ldots d_i \ldots d_9) - \\ &\qquad (d_1 \ldots e_i \ldots d_9) \pmod 7 \\ &= 10^{9-i} d_i - 10^{9-i} e_i \pmod 7. \end{aligned}$$

On summing these congruences, we find that $9a \equiv (d_1 d_2 \ldots d_9) - (e_1 e_2 \ldots e_9) \pmod 7$, and therefore $e_1 e_2 \ldots e_9 \equiv -a \pmod 7$.

In a similar manner, starting with $e_1 e_2 \ldots e_9 \equiv -a \pmod 7$, we have

$$-a \equiv 10^{9-i} e_i - 10^{9-i} f_i \pmod 7,$$

and therefore

$$-a \equiv \left(10^{9-i} d_i - a\right) - 10^{9-i} f_i \pmod 7,$$

or equivalently,

$$10^{9-i} d_i \equiv 10^{9-i} f_i \pmod 7.$$

Since 7 and 10 are relatively prime, $d_i \equiv f_i \pmod 7$.

A-4. Suppose we have a necklace of n beads. Each bead is labeled with an integer and the sum of all these labels is $n-1$. Prove that we can cut the necklace to form a string whose consecutive labels $x_1, x_2, \ldots, x_n$ satisfy

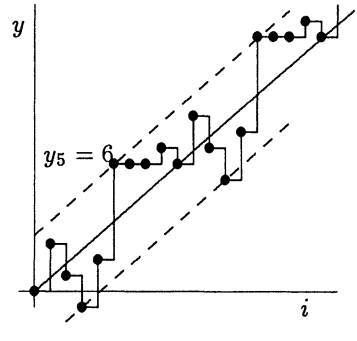
$$\sum_{i=1}^{k} x_i \le k - 1 \qquad \text{for } k = 1, 2, \ldots, n.$$

Solution. We will show that there are just two places where we may cut the necklace. Each is associated with the sense in which we go around the necklace.

Choose an arbitrary starting position, and a sense of rotation, and let the labels be the integers $y_1, y_2, \ldots, y_n$, where $\sum_{i=1}^{n} y_i = n - 1$.

Consider the path in the coordinate plane which starts from the origin, $(0,0)$, moves one space to the right and then vertically to the point $(1, y_1)$, then one space to the right and vertically to the point $(2, y_1 + y_2)$, and so on to the points $(3, y_1 + y_2 + y_3), \ldots, (k, \sum_{i=1}^{k} y_i), \ldots, (n, \sum_{i=1}^{n} y_i = n-1)$. Continue on round the necklace, repeating the pattern: $(n+1, n-1+y_1), (n+2, n-1+y_1+y_2), \ldots$. Choose the point(s) $(K, \sum_{i=1}^{K} y_i)$ which maximize the height above the line $y = \frac{n-1}{n} x$; that is, maximize $\sum_{i=1}^{k} y_i - \frac{n-1}{n} k$. Since $n-1$ and n are relatively prime integers, K is unique modulo n. Relabel the integers, decreasing the subscripts by K, that is, move the origin to the chosen point. Since the slope of the line from this point to any other is at most $\frac{n-1}{n}$ (with equality only at the end of each period), we have achieved our aim: if $\sum_{i=1}^{k} x_i \ge k$, the slope would be at least 1.

Example. Suppose $\{y_1, y_2, \ldots, y_n\} = \{3, -2, -2, 3, 6, 0, 0, 1, -1\}$. The corresponding path is shown below.



The sequence $0,0,1,-1,3,-2,-2,3,6$ has partial sums $0,0,1,0,3,1,-1,2,8 \leq k-1 = 0,1,2,3,4,5,6,7,8$.

To find the cutting place associated with the opposite sense, we don't need to redraw the graph: simply select the *lowest* point below the line to get the shifted sequence $-2,-2,3,-1,1,0,0,6,3$ with partial sums $-2,-4,-1,-2,-1,-1,-1,5,8 \leq 0,1,2,3,4,\ 5,6,7,8$.

A-5. Let $x_1, x_2, \ldots, x_n$ be differentiable (real-valued) functions of a single variable t which satisfy

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n
\end{aligned}$$

for some constants $a_{ij} \geq 0$. Suppose that for all i, $x_i(t) \to 0$ as $t \to \infty$. Are the functions $x_1, x_2, \ldots, x_n$ necessarily linearly dependent?

Solution. We will show that the functions $x_1, x_2, \ldots, x_n$ are necessarily linearly dependent. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{A} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$. Then we are given that $\dfrac{d\mathbf{x}}{dt} = \mathbf{Ax}$. Consider a linear combination $y = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$ of the given functions, where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are constants, possibly complex, to be chosen later. If we set $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, we have $y = \mathbf{v}\cdot\mathbf{x} = \mathbf{v}^T\,\mathbf{x}$ (T = transpose) and thus

$$\frac{dy}{dt} = \mathbf{v}^T \frac{d\mathbf{x}}{dt} = \mathbf{v}^T\mathbf{Ax} = (\mathbf{A}^T\,\mathbf{v})^T\,\mathbf{x}.$$

In particular, if $\mathbf{v}$ is an eigenvector of A^T for the eigenvalue λ, we get

$$\frac{dy}{dt} = (\mathbf{A}^T\mathbf{v})^T\mathbf{x} = (\lambda\mathbf{v})^T\mathbf{x} = \lambda\ \mathbf{v}\cdot\mathbf{x} = \lambda y,$$

so in that case y has the form $y = Ce^{\lambda t}$ for some constant C.

Since we are given that $a_{ij} \geq 0$, in particular, $\mathrm{Trace}(\mathbf{A}^T) = a_{11} + a_{22} + \cdots + a_{nn} \geq 0$, so $\mathbf{A}^T$ has at least one eigenvalue whose real part is nonnegative. Let λ be such an eigenvalue, and let $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be a corresponding eigenvector of $\mathbf{A}^T$. Then, by the above, we have $y = Ce^{\lambda t}$ with $\mathrm{Re}(\lambda) \geq 0$. On the other hand, since $x_i(t) \to 0$ as $t \to \infty$ and y is a linear combination of the x_i, we have $y(t) \to 0$ as $t \to \infty$. But $|e^{\lambda t}| = e^{\mathrm{Re}(\lambda)t} \geq 1$ for $t \geq 0$, so $Ce^{\lambda t} \to 0$ implies $C = 0$. Therefore, a nontrivial linear combination of the x_i is identically zero (note that $\mathbf{v} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is nonzero because it is an eigenvector), so the x_i are linearly dependent, and we are done.

A-6. Suppose that each of n people writes down the numbers $1, 2, 3$ in random order in one column of a $3 \times n$ matrix, with all orders equally likely and with the orders for different columns independent of each other. Let the row sums a, b, c of the resulting matrix be rearranged (if necessary) so that $a \leq b \leq c$. Show that for some $n \geq 1995$, it is at least four times as likely that both $b = a + 1$ and $c = a + 2$ as that $a = b = c$.

Solution. For a positive integer n, and using the other notation of the problem, let $S(n)$ be the statement "it is at least four times as likely that both $b = a + 1$ and $c = a + 2$ as that $a = b = c$." We show that if $S(n)$ is false, then $S(n + 1)$ is true. In particular, $S(n)$ is true for at least one of $n = 1995$, $n = 1996$.

For any positive integer n, let X_n be the number of ways the matrix can be formed so that $b = a + 1$ and $c = a + 2$ (where a, b, c, with $a \leq b \leq c$, are the row sums after rearrangement; incidentally, $a + b + c = 6n$). Let Y_n be the number of ways the matrix can

be formed so that $a = b = c$, and let Z_n be the number of ways with $a = b$ and $c = a+3$.

Our assumption that $S(n)$ is false means that $4Y_n > X_n$. Now note that if a matrix *with* $n+1$ *columns* is formed such that its row sums are all equal, then the first n columns of that matrix form one of the matrices that is counted by X_n. Conversely, for each of the matrices counted by X_n, there is exactly one way to "complete" it to a matrix counted by Y_{n+1}, so we have $Y_{n+1} = X_n$. Similar arguments show that $Z_{n+1} \geq X_n$ (since to row sums $a, a+1, a+2$ one can add $2, 1, 3$ respectively to get $a+2, a+2, a+5$), and $X_{n+1} \geq 6Y_n + 2X_n + 2Z_n$ (since $a+2, a+3, a+4$ can be obtained by adding $1, 2, 3$ *in any order* to $a+1, a+1, a+1$ [and rearranging], or by adding $2, 3, 1$ or $3, 1, 2$ (in that order) to $a, a+1, a+2$ [and rearranging], or by adding 3,2,1 or 2,3,1 to $a, a, a+3$ [and rearranging]).

Therefore, we have

$$\begin{aligned}\frac{X_{n+1}}{Y_{n+1}} = \frac{X_{n+1}}{X_n} &\geq 6\frac{Y_n}{X_n} + 2 + 2\frac{Z_n}{X_n} \\ &\geq 6\frac{Y_n}{X_n} + 2 + 2\frac{X_{n-1}}{X_n} \\ &= 6\frac{Y_n}{X_n} + 2 + 2\frac{Y_n}{X_n} \\ &= 8\frac{Y_n}{X_n} + 2.\end{aligned}$$

But by our assumption, we have $\frac{Y_n}{X_n} > \frac{1}{4}$, so $\frac{X_{n+1}}{Y_{n+1}} \geq \frac{8}{4} + 2 = 4$, so $X_{n+1} \geq 4Y_{n+1}$, and we are done.

B-1. For a partition π of $\{1, 2, \ldots, 9\}$, let $\pi(x)$ be the number of elements in the part containing x. Prove that for any two partitions π and π', there are two distinct numbers x and y in $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ such that $\pi(x) = \pi(y)$ and $\pi'(x) = \pi'(y)$.

[A *partition* of a set S is a collection of disjoint subsets (parts) whose union is S.]

Solution. Suppose there are no two such numbers in $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Then any two numbers in the same part of π must be contained in parts of different size in π', and vice versa. This implies that the largest parts of π and π' have at most three numbers in them (because $1 + 2 + 3 + 4 > 9$). In fact, any two numbers in parts of the same size in π must be contained in parts of different sizes in π'. Therefore, π can have at most one part of size 3, one part of size 2, and at most three parts of size 1. This is impossible for a partition of a set of 9 numbers.

B-2. An ellipse, whose semi-axes have lengths a and b, rolls without slipping on the curve $y = c \sin\left(\frac{x}{a}\right)$. How are a, b, c related, given that the ellipse completes one revolution when it traverses one period of the curve?

Solution. We shall show that $b^2 = a^2 + c^2$.

The ellipse is given parametrically by the equations $x = a\cos\theta$, $y = b\sin\theta$. Using the familiar formula for arc length, the perimeter of the ellipse is

$$\int_0^{2\pi} \sqrt{a^2\sin^2\theta + b^2\cos^2\theta}\, d\theta.$$

The length of one period of the sine curve is

$$\int_0^{2\pi a} \sqrt{1 + \frac{c^2}{a^2}\cos^2\left(\frac{x}{a}\right)}\, dx = \int_0^{2\pi} \sqrt{a^2 + c^2\cos^2\theta}\, d\theta\,.$$

But $a^2 + c^2\cos^2\theta = a^2\sin^2\theta + (a^2+c^2)\cos^2\theta$, so we see that the arc lengths will be equal if and only if $b^2 = a^2 + c^2$, as claimed.

B-3. To each positive integer with n^2 decimal digits we associate the determinant of the matrix obtained by writing the digits in order across the rows. For example, for $n = 2$, to the integer 8617 we associate

$\det\begin{pmatrix} 8 & 6 \\ 1 & 7 \end{pmatrix} = 50$. Find, as a function of n, the sum of all the determinants associated with n^2-digit integers. (Leading digits are assumed to be nonzero; for example, for $n = 2$, there are 9000 determinants.)

Solution. The sum is 45 for $n = 1$, 20250 for $n = 2$, and 0 for $n \geq 3$.

The case $n = 1$ is trivial: $1 + 2 + \cdots + 9 = 45$. Now let $n \geq 2$. Then for each $n \times n$ matrix with entries in $\{0, 1, 2, \ldots, 9\}$ there is another such matrix obtained by interchanging the last two columns. (If this matrix is equal to the original one, its determinant is zero.) Since interchanging two columns in a matrix changes its determinant to the opposite determinant, the sum of *all* determinants of matrices with entries in $\{0, 1, 2, \ldots, 9\}$ is zero. However, we are not supposed to take all such matrices, but only the ones that don't have a 0 in the upper left corner. If $n \geq 3$, interchanging the last two columns doesn't affect that corner, and so the required sum is 0 by the same argument. On the other hand, if $n = 2$, all determinants in the sum cancel *except* those of the form $\det\begin{pmatrix} * & 0 \\ * & * \end{pmatrix}$. These determinants only depend on the diagonal entries, and there are ten of them for each pair of diagonal entries; thus their sum is

$$10 \sum_{i,j=1}^{9} ij = 10 \left(\sum_{i=1}^{9} i \right) \left(\sum_{j=1}^{9} j \right) = 10 \cdot 45 \cdot 45 = 20250,$$

as claimed.

B-4. Evaluate

$$\sqrt[8]{2207 - \cfrac{1}{2207 - \cfrac{1}{2207 - \ldots}}}.$$

Express your answer in the form $\dfrac{a + b\sqrt{c}}{d}$, where a, b, c, d are integers.

Solution. The answer is $\dfrac{3 + \sqrt{5}}{2}$.

Let $F(n) = n - \cfrac{1}{n - \cfrac{1}{n - \cdots}}$, so the problem asks for $\sqrt[8]{F(2207)}$. Note that $F(n) = n - \dfrac{1}{F(n)}$, and solving this for $F(n)$ yields $F(n) = \dfrac{n \pm \sqrt{n^2 - 4}}{2}$. For $n > 2$ we have

$$\frac{n - \sqrt{n^2 - 4}}{2} = \frac{2}{n + \sqrt{n^2 - 4}} < 1 < F(n)$$

so we must have the plus sign: $F(n) = \dfrac{n + \sqrt{n^2 - 4}}{2}$.

Now note that

$$\begin{aligned} (F(n))^2 &= \frac{n^2 + 2n\sqrt{n^2 - 4} + n^2 - 4}{4} \\ &= \frac{n^2 - 2 + n\sqrt{n^2 - 4}}{2} \\ &= \frac{n^2 - 2 + \sqrt{(n^2 - 2)^2 - 4}}{2} \\ &= F(n^2 - 2) \end{aligned}$$

for $n > 2$, since $n > 2$ implies $n^2 - 2 > 2$.

Conversely, if $k > 2$, then we have $k = n^2 - 2$ with $n = \sqrt{k + 2} > 2$, and therefore $F(k) = (F(n))^2$, $\sqrt{F(k)} = F(n) = F(\sqrt{k + 2})$. In particular,

$$\sqrt{F(2207)} = F(\sqrt{2209}) = F(47),$$

$$\sqrt[4]{F(2207)} = \sqrt{F(47)} = F(\sqrt{49}) = F(7),$$

and

$$\begin{aligned} \sqrt[8]{F(2207)} &= \sqrt{F(7)} = F(\sqrt{9}) \\ &= F(3) = \frac{3 + \sqrt{5}}{2}. \end{aligned}$$

B-5. A game starts with four heaps of beans, containing 3, 4, 5 and 6 beans. The two players move alternately. A move consists of taking **either**

a. one bean from a heap, provided at least two beans are left behind in that heap, **or**

b. a complete heap of two or three beans.

The player who takes the last heap wins. To win the game, do you want to move first or second? Give a winning strategy.

Solution. Heaps of 0 or 1 cannot affect the game. In fact, heaps of 1 cannot arise. Heaps of 2 behave as though they were a single bean which may be removed. Heaps of 3 are special. Otherwise a move just removes a bean, and the result depends only on the parity of the total number of beans (counting a heap of 2 as a single bean).

The first player wins by taking one bean from the 3-heap, leaving heaps of 2, 4, 5 and 6 beans, whose "sum" is $1(=2)+4+5+6$ which is even. Now the win is automatic, since the opponent must make the "sum" odd. It doesn't matter what moves are made, *except* that the first player mustn't move in a 4-heap (there is no need to since the sum will always be odd, and all the heaps can't be 4-heaps), and whenever the second player moves in a 4-heap, the first player removes all the remaining beans at the next move.

B-6. For a positive real number α, define

$$S(\alpha) = \{\lfloor n\alpha \rfloor \ : \ n = 1, 2, 3, \dots\}.$$

Prove that $\{1, 2, 3, \dots\}$ cannot be expressed as the disjoint union of three sets $S(\alpha), S(\beta)$, and $S(\gamma)$. [As usual, $\lfloor x \rfloor$ is the greatest integer $\leq x$.]

Solution. Suppose $\alpha < \beta < \gamma$ and $S(\alpha), S(\beta)$, and $S(\gamma)$ disjointly cover $\{1, 2, 3, \dots\}$. Since $\lfloor \alpha \rfloor = 1$, we have $\alpha = 1 + \epsilon$, for some ϵ satisfying $0 \leq \epsilon < 1$.

Let $r > 1$ be the first value not in $S(\alpha)$. We have

$$\lfloor (r-1)\alpha \rfloor = r - 1, \quad \lfloor r\alpha \rfloor = r + 1.$$

Therefore,

$$(r-1)\alpha < r, \quad r\alpha \geq r + 1$$

and

$$1 + \frac{1}{r} \leq \alpha < 1 + \frac{1}{r-1};$$

that is,

$$\frac{1}{r} \leq \epsilon < \frac{1}{r-1}.$$

Fact 1. If $u \notin S(\alpha)$, then the next element missing from $S(\alpha)$ is either $u + r$ or $u + r + 1$ (and the other of $u + r, u + r + 1$ is in $S(\alpha)$). Proof: Suppose $\lfloor t\alpha \rfloor = u - 1$, $\lfloor (t+1)\alpha \rfloor = u + 1$. Let $\delta = (t+1)\alpha - (u+1)$. The next missing element occurs at $u + m$ where m is the smallest integer such that $\delta + (m-1)\epsilon \geq 1$. If $m \leq r - 1$, we have

$$\delta + (m-1)\epsilon < m\epsilon \leq (r-1)\epsilon < 1$$

since $\delta < \epsilon$. Also, for $m = r + 1$,

$$\delta + (m-1)\epsilon = \delta + r\epsilon \geq 1.$$

Therefore $m = r$ or $r + 1$.

Note that $\lfloor \beta \rfloor = r$, so $r \leq \beta < r + 1$.

Fact 2. If $v \in S(\beta)$, then the next element in $S(\beta)$ is $v + r$ or $v + r + 1$.

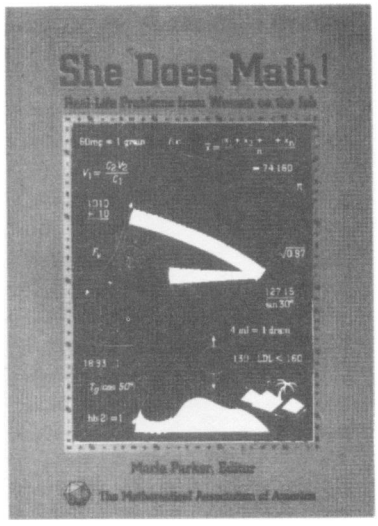
Fact 2 can be proved in the same manner as Fact 1.

Combining Facts 1 and 2 with the fact that $S(\alpha)$ and $S(\beta)$ are disjoint, we conclude that the union of $S(\alpha)$ and $S(\beta)$ is all of $\{1, 2, 3, \dots\}$. Therefore, $\{1, 2, 3, \dots\}$ cannot be expressed as the disjoint union of *three* sets $S(\alpha), S(\beta)$ and $S(\gamma)$.

She Does Math!

Real-Life Problems From Women On The Job

Marla Parker, Editor



She Does Math! presents the career histories of 38 professional women and math problems related to their work. Each history describes how much math the author took in high school and college; how she chose her field of study; and how she ended up in her current job. Each of the women presents problems that are typical of those she has faced in her job. The problems require only high school mathematics for their solution.

Who should have this book?

Your daughter, your granddaughter, your sister, your former math teacher, your students—and young men too. They want to know how the math they study is applied, and this book will show them.

There are many reasons to buy this book:

➤ By reading the career histories of the women profiled in this book, young people will learn that if they take mathematics courses in high school and college they will be qualified to enter interesting technical fields and earn good salaries.
➤ The problems have special appeal to students who are beginning to think about career choices.
➤ The book provides practical information about the job market in an interesting, innovative way.
➤ Strong female role models who work as successful technical professionals are presented.
➤ The problems are interesting and challenging, yet require only high school mathematics. They demonstrate how good math skills are applied to real-life problems.

Read what others have said about She Does Math!

Finally — a practical, innovative, well-written book that will also inspire its readers. The wonder is...it's a mathematics text and a biography! The idea of women telling their own career stories, emphasizing the mathematics they use in their jobs is extremely creative. This book makes me wish that I could go through school all over again!

Anne L. Bryant, Executive Director
American Association of University Women

She Does Math! *will undoubtedly appeal both to students who already enjoy math and want to get a view of potential career paths, and also to students who want to better understand the relevance of their math classes to their future careers. It is an absorbing look into the lives of some very inspiring and talented women!*

Susanne Hupfer and Elisabeth Freeman
Yale University

This collection is a wonderful confirmation that real women do math. They do math in a surprising variety of careers, fully enjoying the challenge and rewards of solving complex problems. This is a book for young women and men, a book for their teachers and parents, a book that informs about the possibilities that mathematics affords to all. It is also a book that will engage you in real-life mathematics! – Doris Schattschneider
Moravian College

272 pp., Paperbound, 1995 ISBN 0-88385-702-2
List: $27.50 MAA Member: $20.95
Catalog Code: SDM/JR

ORDER FROM:
THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW Washington, DC 20036
1-800-331-1622 (301) 617-7800 FAX (301) 206-9789

Membership Code:

____ ____ ____ ____ ____ ____

Name _____

Address _____

City _____

State _____ Zip _____

Qty.	Catalog Code	Price	Amount
_____	SDM/JR	_____	_____
_____	_____	_____	_____
		TOTAL	_____

Payment ☐ Check ☐ VISA ☐ MasterCard

Credit Card No. _____ Expires ___/___

Signature _____

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

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