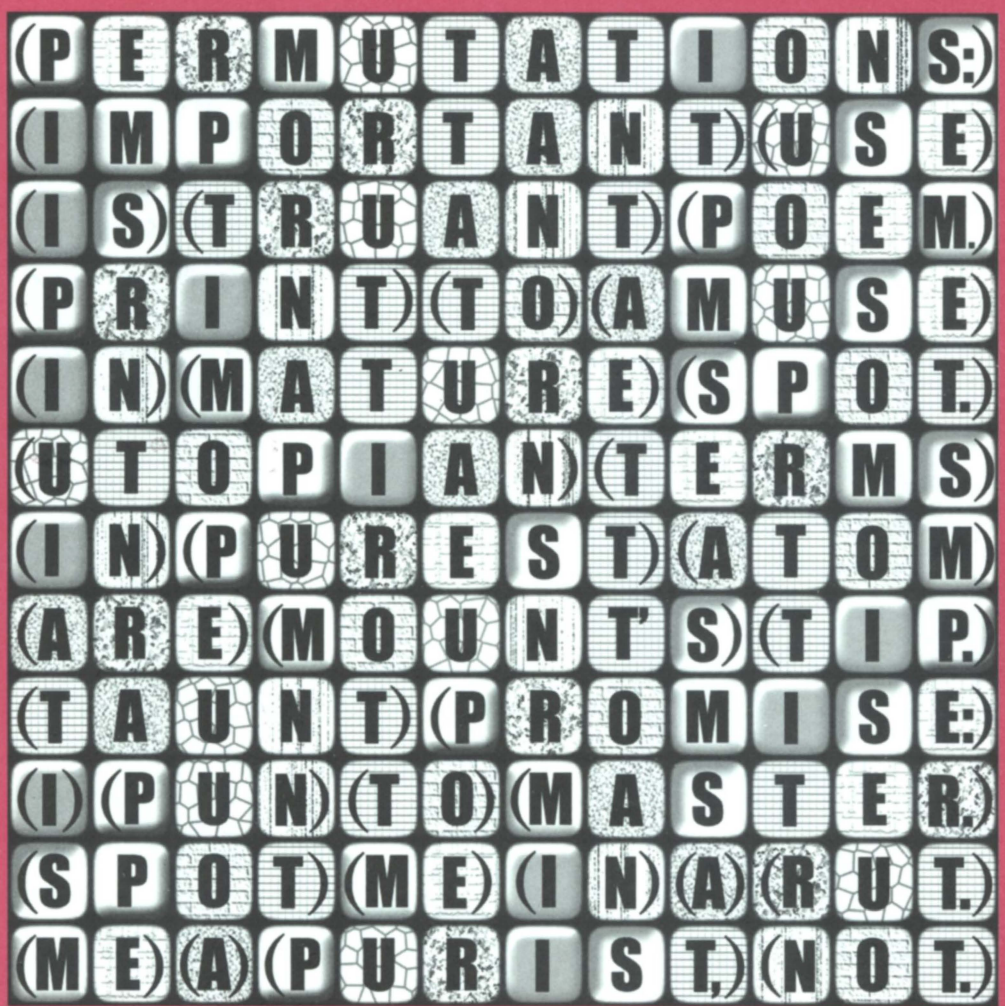
Vol. 77, No. 2, April 2004

MATHEMATICS MAGAZINE



- The Fabulous (11, 5, 2) Biplane
- Perfect Shuffles through Dynamical Systems
- The Geometry of Generalized Complex Numbers

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

EDITORIAL POLICY

Mathematics Magazine aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate 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 among various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 74, pp. 75–76, and is available from the Editor or at www.maa.org/pubs/mathmag.html. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Submit new manuscripts to Frank A. Farris, Editor, Mathematics Magazine, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053-0373. Manuscripts should be laser printed, with wide line spacing, and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should mail three copies and keep one copy. In addition, authors should supply the full five-symbol 2000 Mathematics Subject Classification number, as described in *Mathematical Reviews*.

Cover image, *"PERMUTATIONS" permuted*, by Mary Le, with assistance from Jason Challas and Frank Farris. To illustrate this issue's theme of permutations, the letters of the word *permutations* are rearranged in a dozen different ways.

Mary Le is an undergraduate student at Santa Clara University, with a major in Operations Management Information Systems and double minors in Computer Engineering and Studio Art. Jason Challas lectures on permuted images and computer art at Santa Clara University.

AUTHORS

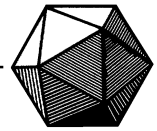
Ezra (Bud) Brown grew up in New Orleans and has degrees from Rice University and Louisiana State University. He has been at Virginia Tech since 1969, with time out for sabbatical visits to Washington, DC and Munich. His research interests include graph theory, the combinatorics of finite sets, and number theory. He has received the MAA MD-DC-VA Section Award for Outstanding Teaching and three MAA awards for expository excellence. As a graduate student, he first met (11, 5, 2), but it was only many years later that he learned of its many intriguing combinatorial connections. He occasionally bakes biscuits for his students, and recently won a karaoke contest.

Daniel Scully received a B.A. in mathematics from Saint John's University in Collegeville, Minnesota in 1975, an M.A. in mathematics from the University of Minnesota in Minneapolis in 1983, and a Ph.D. in matrix theory from Utah State University in Logan, Utah. He has taught at the College of Saint Benedict in Saint Joseph, Minnesota and at Saint Cloud State University where he is a professor and department chair. His main mathematical interests are linear algebra, discrete mathematics, and perfect shuffles. He blames Dr. Danrun Huang for hooking him on symbolic dynamics. His outside interests include convincing his son that even though his cost per ski trip may decrease as n approaches infinity, the total cost does not.

Anthony A. Harkin is a postdoctoral fellow in the Division of Engineering and Applied Sciences at Harvard University. He received his Ph.D. from Boston University in 2001. His current research interests include nonlinear dynamical systems, fluid mechanics, and computational science.

Joseph B. Harkin traces his mathematical heritage to mathematicians at Illinois Institute of Technology during the 1950s and 1960s. During those decades, P. Porcelli empowered his students in the Texas (Moore-Wall) research strategies. Karl Menger promoted a sense of "wonder and awe" in his courses in complex function theory that led to the development of the insights in this paper. Abe Sklar, by his example, always taught us "the next right question to ask."

Vol. 77, No. 2, April 2004



MATHEMATICS MAGAZINE

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 *MATHEMATICS MAGAZINE* to an individual member of the Association is $131. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 20% dues discount for the first two years of membership.)

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 Dave Riska (*driska@maa.org*), Advertising Manager, the Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

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

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

Printed in the United States of America

The Fabulous (11, 5, 2) Biplane

EZRA BROWN
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0123
brown@math.vt.edu

—To Annette L. Brown: Pianist, Mother, Grandmother, and Great-Grandmother *extraordinaire*.

After a workshop for new teaching assistants on innovations in teaching, a new sociology graduate student wandered into my office and asked the question, "Tell me ... how do you make math exciting for students?" By chance, I just happened to have on my computer screen a picture that exhibits some of the symmetries of one of the most intriguing objects in mathematics: the (11, 5, 2) biplane.

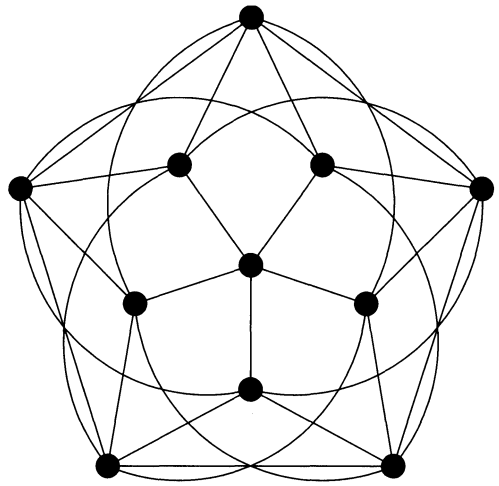


Figure 1 A fascinating picture

I told him of my chagrin on seeing a picture similar to FIGURE 1 (but much prettier, and in color) on the cover of a book [**6**] on combinatorial designs. The picture was lovely, and the reason for my strong feelings was purely selfish: I was trying to construct such a picture, and somebody else thought of it first.

But it wasn't labeled.

It was fun finding a labeling compatible with the symmetries of the biplane. To find generators for the symmetry group of the biplane—which turns out to have a name, $PSL(2, 11)$—was more fun. The best part, however, was learning about the exact connection between the biplane and six pairs of mathematical objects.

We find these six mathematical pairs just outside the boundaries of many traditional courses, where a bit of exploration can lead the curious to all manner of interesting mathematics. A good course in coding theory will mention two pairs of perfect error-correcting codes, namely the Golay codes $\{G_{11}, G_{12}\}$ and $\{G_{23}, G_{24}\}$, but sometimes only in passing. Look past the usual topics in combinatorics into the world of combinatorial designs and you will meet two pairs of Steiner systems, namely $\{S(4, 5, 11), S(5, 6, 12)\}$ and $\{S(4, 7, 23), S(5, 8, 24)\}$. Beyond the first course

in group theory lie two pairs of finite simple groups, namely the Mathieu groups $\{M_{11}, M_{12}\}$ and $\{M_{23}, M_{24}\}$. It was quite a revelation to learn just how these codes, designs, and groups connect with the biplane and with each other.

I told the student all about this, including the reason that the biplane is called a biplane, and he loved it; maybe you will, too.

Difference sets, block designs, and biplanes

The (11, 5, 2) biplane is a collection of the following eleven 5-element subsets of $\{1, 2, 3, 4, 5, 6, 7, 8, 9, X, 0\}$ (we think of X as 10, and we have written $abcde$ for the set $\{a, b, c, d, e\}$:

$$\begin{array}{llll} B_1 = 13459 & B_2 = 2456X & B_3 = 35670 & B_4 = 46781 \\ B_5 = 57892 & B_6 = 689X3 & B_7 = 79X04 & B_8 = 8X015 \\ B_9 = 90126 & B_X = X1237 & B_0 = 02348 & \end{array}$$

The (11, 5, 2) biplane

This is an example of a *block design*, which is an arrangement of v objects called *varieties* into b sets called *blocks*. Each variety appears in exactly r blocks, each block contains exactly k varieties, and each pair of varieties appears together in exactly λ blocks. From the above, we see that $b = v = 11$ and $k = 5$. It is a bit less obvious that $r = 5$ and still less obvious that $\lambda = 2$: for example, 1 appears in blocks B_1, B_4, B_8, B_9, and B_X, and 7 and 0 appear together in blocks B_3 and B_7.

Block designs first appeared in the 1930s in connection with the design of certain agricultural experiments, although they are implicit in the work of Woolhouse [**13**] and Kirkman [**7**] as early as 1844 and 1847, respectively. (These papers are hard to find; a more recent reference is Richard Guy's excellent survey article [**4**].) The parameters b, v, r, k, and λ are not independent: it happens that $bk = vr$ and $r(k-1) = \lambda(v-1)$. Thus, if $b = v$, then $r = k$ and we speak of a (v, k, λ) *symmetric design*. Hence, the (11, 5, 2) biplane is an (11, 5, 2) symmetric design, which explains the numerical part of its name.

Symmetric designs also have the feature that two distinct blocks intersect in exactly λ varieties; for a proof, see Hall [**5**, Section 10.2].

A closer look reveals that we may construct the entire (11, 5, 2) biplane from B_1 by adding a particular integer mod 11 to each element; for example, if we add 5 to each element of B_1 and reduce the results mod 11, we find that

$$\{1+5, 3+5, 4+5, 5+5, 9+5\} \equiv \{6, 8, 9, X, 3\} \equiv B_6 \bmod 11.$$

Now, B_1 is an example of a *difference set*; that is, every nonzero integer mod 11 appears exactly twice among the 20 differences $i - j$ mod 11 for i and j distinct elements of B_1 (in the following, $a \equiv b$ is short for $a \equiv b$ mod 11):

$$\begin{array}{lll} \mathbf{1} \equiv 4-3 \equiv 5-4 & \mathbf{2} \equiv 3-1 \equiv 5-3 & \mathbf{3} \equiv 4-1 \equiv 1-9 \\ \mathbf{4} \equiv 5-1 \equiv 9-5 & \mathbf{5} \equiv 9-4 \equiv 3-9 & \mathbf{6} \equiv 9-3 \equiv 4-9 \\ \mathbf{7} \equiv 1-5 \equiv 5-9 & \mathbf{8} \equiv 9-1 \equiv 1-4 & \mathbf{9} \equiv 1-3 \equiv 3-5 \\ & \mathbf{10} \equiv 3-4 \equiv 4-5. & \end{array}$$

More generally, a (v, k, λ) *difference set* is a k-element subset S of $V = \{0, 1, \ldots, v-1\}$ such that every nonzero integer mod v can be written in exactly λ ways as a difference of elements of S. So, the set $\{1, 3, 4, 5, 9\}$ of nonzero perfect squares mod 11 is an (11, 5, 2) difference set.

In fact, for every prime $p \equiv 3 \bmod 4$, the set Q_p of nonzero perfect squares mod p is a $(p, (p-1)/2, (p-3)/4)$ difference set (a proof appears in [**2**]). For example, you can check that $Q_{23} = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$ is a (23, 11, 5) difference set. (Exercise: Find the five different ways to write 7 as a difference of elements of Q_{23}.)

What is interesting here is that every difference set gives rise to a symmetric design in the following way:

THEOREM 1. *Let* $D = \{x_1, x_2, \ldots, x_k\}$ *be a* (v, k, λ) *difference set. Let* $D_i := \{x_1 + i, \ldots, x_k + i\}$ *where addition is* mod v. *Then the* v *sets* $D_0, \ldots, D_{v-1}$ *are the blocks of a* (v, k, λ) *symmetric design.*

(For a proof, see Hall [**5**, Theorem 11.1.1].) Thus, the (11, 5, 2) difference set gives rise to the (11, 5, 2) symmetric design.

Symmetric designs with $\lambda = 1$ have the property that every pair of varieties determines a unique block and every pair of blocks intersects in a unique variety. Reading *line* for *block* and *point* for *variety* gives us the first two axioms of projective geometry; for this reason, $(v, k, 1)$ designs are called *finite projective planes*, or planes for short. Now for a $(v, k, 2)$ design, every pair of varieties determines exactly two blocks and every pair of blocks intersects in exactly two varieties. For this reason, the blocks and varieties of a $(v, k, 2)$ design are called lines and points, respectively, and the designs themselves are called *biplanes*—and that explains the second part of the (11, 5, 2) biplane's name.

As stated earlier, part of my fascination with the (11, 5, 2) biplane lies both in its symmetries and in the challenge of drawing a picture that will reveal some of its symmetries. By a symmetry of a design, we mean a permutation of the varieties that simultaneously permutes the blocks. For any design, the set of all such permutations is a group called the *automorphism group* of the design. So, first we'll talk about permutations and automorphism groups, and then we'll draw another picture.

The automorphism group of the biplane

A *permutation* on a set Y is a mapping of the set to itself that is one-to-one and onto. An *n-cycle* is an expression of the form $(a_1\ a_2 \ldots a_n)$, where the a_i are distinct. The cycle notation is a standard way to describe permutations on finite sets; here is an example to show how it works. If we write $f = (1\ 3\ 6)(4\ 5)$, it means that $f(1) = 3$, $f(3) = 6$, $f(6) = 1$, $f(4) = 5$, $f(5) = 4$, and $f(x) = x$ for all $x \notin \{1, 3, 4, 5, 6\}$; in this notation, 1-cycles are frequently omitted. In this example, we say that f is a product of two disjoint cycles. Similarly, $g = (1\ 2)$ means that g switches 1 and 2 and leaves everything else fixed. Since permutations are functions, they compose from right to left. If we denote composition by $\circ$, then $f \circ g = (1\ 3\ 6)(4\ 5)(1\ 2)$. This maps 1 to 2, 2 to 3 (since $g(2) = 1$ and $f(1) = 3$), 3 to 6, 4 to 5, 5 to 4, and 6 to 1. We see that $f \circ g = (1\ 2\ 3\ 6)(4\ 5)$ as a product of disjoint cycles.

Let $\mathcal{D}$ be a block design. An *automorphism* of $\mathcal{D}$ is a permutation f of the set V of varieties that is simultaneously a permutation of the set B of blocks. (We say that f *induces* a permutation on B.) For example, the permutation $\tau = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ X\ 0)$ of the set $V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, X, 0\}$ of varieties induces the permutation $\tau' = (B_1\ B_2\ B_3\ B_4\ B_5\ B_6\ B_7\ B_8\ B_9\ B_X\ B_0)$ of the corresponding set of blocks. The set of all such automorphisms is a group under composition, called the *automorphism group* $Aut(\mathcal{D})$ of the design $\mathcal{D}$.

It turns out that there are 660 automorphisms of the (11, 5, 2) biplane. How do we find them all?

In some sense, the automorphism τ is an obvious choice, for the blocks of the biplane were created by repeatedly adding 1 (mod 11) to each member of the difference set $B_1 = \{1, 3, 4, 5, 9\}$. It is not so obvious that the permutation $\mu = (1\ 3\ 9\ 5\ 4)(2\ 6\ 7\ X\ 8)(0)$ also induces a permutation of the blocks—but it does, namely $\mu' = (B_2\ B_4\ B_X\ B_6\ B_5)(B_0\ B_9\ B_3\ B_7\ B_8)(B_1)$.

Clearly, we need a systematic way to find the rest of the automorphisms. We make a four-fold application of that useful and elegant result, the Orbit-Stabilizer Theorem. But first, we need a couple of definitions. Suppose that G is a group of permutations on the set S, let $g \in G$, and let $T \subseteq S$. Then $g(T)$ is the set of images $g(t)$ for all $t \in T$; an element $g \in G$ leaves T *setwise fixed* if $g(T) = T$. The *stabilizer* of T in G, $Stab_G(T)$, is the set of all permutations g in G that leave T setwise fixed. The *orbit* of T, $Orb_G(T)$, is the set of all $Y \subseteq S$ for which $Y = g(T)$ for some permutation $g \in G$. (If $T = \{t\}$, we customarily write $Stab_G(t)$ and $Orb_G(t)$, ignoring the braces.) Let $|A|$ be the number of elements in the set A. Here is the theorem, which follows from the definition of a permutation and from Lagrange's Theorem:

THEOREM 2. (THE ORBIT-STABILIZER THEOREM) *Let G be a finite group of permutations of a set S and let $T \subseteq S$. Then* (a) *$Stab_G(T)$ is a subgroup of G, and* (b) $|G| = |Stab_G(T)| \cdot |Orb_G(T)|$.

We now define the groups G, H, K, and L as follows:

$$\begin{aligned} G &= Aut((11, 5, 2)); \\ H &= Stab_G(B_1) = \{\text{automorphisms in } G \text{ that leave } B_1 \text{ setwise fixed}\}; \\ K &= Stab_H(1) = \{\text{automorphisms in } H \text{ that leave 1 fixed}\}; \\ L &= Stab_K(3) = \{\text{automorphisms in } K \text{ that leave 3 fixed}\}. \end{aligned} \tag{1}$$

By the Orbit-Stabilizer Theorem, L, K, and H are subgroups of K, H, and G, respectively, and since $4 \in B_1$, we see that

$$\begin{aligned} |G| &= |H| \cdot |Orb_G(B_1)| = |K| \cdot |Orb_H(1)| \cdot |Orb_G(B_1)| \\ &= |L| \cdot |Orb_K(3)| \cdot |Orb_H(1)| \cdot |Orb_G(B_1)| \\ &= |Stab_L(4)| \cdot |Orb_L(4)| \cdot |Orb_K(3)| \cdot |Orb_H(1)| \cdot |Orb_G(B_1)|. \end{aligned} \tag{2}$$

If we can show that $|Stab_L(4)| = 1$, $|Orb_L(4)| = 3$, $|Orb_K(3)| = 4$, $|Orb_H(1)| = 5$, and $|Orb_G(B_1)| = 11$, it will follow that $|G| = 1 \cdot 3 \cdot 4 \cdot 5 \cdot 11 = 660$. Let's call it a theorem:

THEOREM 3. *Let G, H, K, and L be as defined above.* (a) *If $\sigma \in H$ and σ fixes* 1, 3, *and* 4, *then $\sigma = I$, the identity map, and $|Stab_L(4)| = 1$.* (b) *$|Orb_L(4)| = 3$, $|Orb_K(3)| = 4$, $|Orb_H(1)| = 5$, and $|Orb_G(B_1)| = 11$.* (c) *$|G| = 660$.*

Proof.

(a) Since $\sigma \in H$, σ fixes B_1 setwise. Now, σ might permute some of the other blocks. We can show that this is false by seeing how it permutes the blocks containing the pairs $\{1, 4\}$, $\{1, 3\}$, and $\{3, 4\}$. Since $B_4 = 46781$, $B_X = X1237$, and $B_0 = 02348$ are the only other blocks containing those pairs, it follows that σ fixes the sets B_4, B_X, and B_0. Thus, σ fixes the subsets $\{6, 7, 8\}$, $\{X, 2, 7\}$, and $\{0, 2, 8\}$ of B_4, B_X, and B_0, respectively. The only way this can happen is if σ fixes the elements 2, 7, and 8. As a consequence, σ also fixes 6, X, and 0, and hence σ fixes

$B_3 = 35670$. It follows that σ fixes 5. Finally, since σ fixes B_1, it must also fix 9, and we conclude that $\sigma = I$, and so $|Stab_L(4)| = 1$.

(b) Let $L = Stab_K(3)$ and let $\alpha \in L$. Then α fixes 1 and 3. The method in (a) shows that any permutation that fixes three distinct points must be the identity map. Hence, either $\alpha = I$ or α cyclically permutes 4, 5, and 9. A little work shows that either α or α^{-1} is equal to (4 5 9)(2 7 X)(0 6 8). It follows that $Orb_L(4) = \{4, 5, 9\}$, and so $|Orb_L(4)| = 3$. A similar argument shows that $K = Stab_H(1)$ contains the permutations $I, \beta = (3\ 4)(5\ 9)(2\ 8)(6\ X)$, $\gamma = (3\ 5)(4\ 9)(2\ 8)(7\ 0)$, and $\beta \circ \gamma$; it follows that $Orb_K(3) = \{3, 4, 5, 9\}$, and so $|Orb_K(3)| = 4$. Next, $H = Stab_H(1)$ contains the powers of $\mu = (1\ 3\ 9\ 5\ 4)(2\ 6\ 7\ X\ 8)$. It follows that $Orb_H(1) = \{1, 3, 4, 5, 9\}$, and so $|Orb_H(1)| = 5$. Finally, G contains the powers of $\tau = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ X\ 0)$; the kth powers of the induced permutation τ' send B_1 to B_k for each k. Hence, $Orb_G(B_1)$ contains all eleven blocks, and we conclude that $|Orb_G(B_1)| = 11$.

(c) We now put the pieces together. By the Orbit-Stabilizer Theorem and Equation (2), we see that

$$\begin{aligned} |G| &= |Stab_L(4)| \cdot |Orb_L(4)| \cdot |Orb_K(1)| \cdot |Orb_H(1)| \cdot |Orb_G(B_1)| \\ &= 1 \cdot 3 \cdot 4 \cdot 5 \cdot 11 = 660, \end{aligned}$$

and we are done. ∎

With so much symmetry, there ought to be a picture that tells us something about the (11, 5, 2) biplane, and FIGURE 1 is where this all began. So let's look at FIGURE 1 with more experienced eyes.

Symmetries of the biplane as revealed in pictures

"Draw a figure." So said that master problem-solver and teacher, George Pólya, in his classic "How To Solve It" [**10**]. We learn so much from figures, so we follow Pólya's lead and return to the picture in FIGURE 1. As we mentioned earlier, the context suggested that it was a picture of the (11, 5, 2) biplane. It is clear that FIGURE 1 is a dressed-up regular pentagon. As such, it is setwise fixed by both a 1/5-turn about the center and reflections about lines through the center. The challenge was to label the figure so that these geometric motions corresponded to symmetries of the (11, 5, 2) biplane, and my efforts were eventually rewarded. In FIGURE 2, the clockwise 1/5-turn about the point 0 and the reflection about the line through 0 and 7 correspond to the automorphisms μ and ρ, respectively, where $\mu = (1\ 3\ 9\ 5\ 4)(2\ 6\ 7\ X\ 8)$ and $\rho = (2\ 8)(3\ 4)(5\ 9)(6\ X)$.

Let us now see just how the figure depicts these automorphisms.

First, consider $\mu = (1\ 3\ 9\ 5\ 4)\ (2\ 6\ 7\ X\ 8)$. As mentioned above, μ induces the permutation $\mu' = (B_2\ B_4\ B_X\ B_6\ B_5)\ (B_0\ B_9\ B_3\ B_7\ B_8)$ on the blocks of the biplane.

Now, look at FIGURE 2. The exterior pentagon joins the five points labeled 1, 3, 9, 4, and 5. This is the block B_1, which is mapped into itself by a 1/5-turn about 0. Next, the dotted lines connect the five points labeled 4, 8, 0, 2, and 3. This is just the block $B_0 = 02348$, and if we rotate the figure about 0 by a 1/5-turn, we see that B_0 is mapped into $B_9 = \{1, 2, 0, 6, 9\}$, B_9 into $B_3 = \{3, 6, 0, 7, 5\}$, B_3 into $B_7 = \{9, 7, 0, X, 4\}$, B_7 into $B_8 = \{5, X, 0, 8, 1\}$, and B_8 into B_0. Finally, the bold lines connect the five points labeled 2, 9, 7, 5 and 8. This is the block $B_5 = \{5, 7, 8, 9, 2\}$, and if we rotate the figure about 0 by a 1/5-turn, we see that B_5 is mapped into $B_2 = \{6, 5, X, 4, 2\}$,

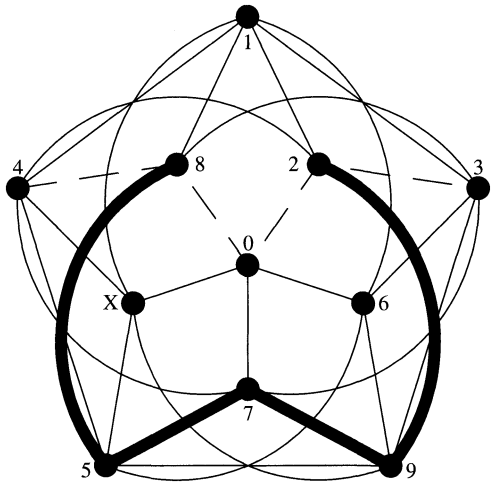


Figure 2 The fabulous (11, 5, 2) biplane

B_2 into $B_4 = \{7, 4, 8, 1, 6\}$, B_4 into $B_X = \{X, 1, 2, 3, 7\}$, B_X into $B_6 = \{8, 3, 6, 9, X\}$, and B_6 into B_5.

Thus, the 1/5-turn about 0 induces the permutation

$$(B_2\ B_4\ B_X\ B_6\ B_5)(B_0\ B_9\ B_3\ B_7\ B_8)(B_1)$$

on the blocks of the biplane. But μ' is exactly this permutation! As for ρ, you can show that the reflection about the line through 0 and 7 induces

$$(B_2\ B_6)(B_4\ B_X)(B_3\ B_7)(B_8\ B_9)(B_0)(B_1)(B_7) = \rho'.$$

Are there ways to draw the (11, 5, 2) biplane that exhibit symmetries other than μ, ρ and others of orders 5 and 2? It is an interesting exercise to find one that exhibits the symmetry of $\alpha = (4\ 5\ 9)(2\ 7\ X)(0\ 6\ 8)$ and has order 3.

We are almost ready to talk about the mathematical pairs connected to the (11, 5, 2) biplane. The most direct path to these pairs leads through a certain matrix associated with the (11, 5, 2) biplane, called the incidence matrix.

One way to describe a block design is by its *incidence matrix*, a $b \times v$ matrix whose (i, j)th entry is 1 or 0 according as the ith block does or does not contain the jth variety. Here is the incidence matrix **M** for the (11, 5, 2) symmetric design. The rows correspond to the blocks in the above order, and the columns correspond to the varieties in the order 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, X:

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

As we shall soon see, the matrix $\mathbf{M}$ is instrumental in constructing the two pairs of Golay codes $\{G_{11}, G_{12}\}$ and $\{G_{23}, G_{24}\}$. So, let's talk about error-correcting codes.

Error-correcting codes

Mathematical schemes to deal with signal errors first appeared in the 1940s in the work of several researchers, including Claude Shannon, Richard Hamming, and Marcel Golay. People at various research labs saw the need for devices that would automatically detect and correct errors in signal transmissions across noisy channels. What they came up with was a new branch of mathematics called *coding theory*—specifically, the study of error-detecting and error-correcting codes. They modeled these signals as sets of n-long strings called *blocks*, to be taken from a fixed alphabet of size q; a particular set of such blocks, or *codewords*, is called a *q-ary code of length n*. If q is a prime number, then a q-ary code of length n is called *linear* if the codewords form a subspace of $\mathbb{Z}_q^n$, the n-dimensional vector space over $\mathbb{Z}_q$, the integers mod q. To *correct* errors means to determine the intended codeword when one has been received incorrectly. Just how this correction happens will vary from code to code.

The fact that d errors in transmission change d characters in a block gives rise to the idea of distance between blocks. If v and w are n-blocks, then the *(Hamming) distance* $D(v, w)$ is the number of positions in which v and w differ. Thus, $D(11001, 10101) = 2$ and $D(1101000, 0011010) = 4$. If I send the block v and you receive the block w, then $D(v, w)$ errors occurred while sending v.

It follows that if the words in a code are all sufficiently far apart in the Hamming distance sense, then we can detect errors. Even better, if we assume that only a few errors are received, then we can sometimes change the received block to the correct codeword. Let us now look at an example of an error-correction scheme.

One way to transmit bits is to send each bit three times, so that our only codewords are 000 and 111. If you receive 010, then it is most likely that I sent 000 and so the intended message was 0; this is the triplication or majority-vote code. Thus, a codeword of length n contains a certain number k of *message bits*, and the other $n - k$ *check bits* are used for error detection and correction. Such a code is called an (n, k) code: the triplication code is a $(3, 1)$ code.

The *minimum distance* of a code is the smallest distance between its codewords; this minimum distance determines the code's error detection and correction features. (Exercise: Show that a code with minimum distance 5 will detect up to 4 errors and correct up to 2. You can then show that a code with minimum distance d will detect up to $d - 1$ errors and correct up to $\lfloor (d-1)/2 \rfloor$ errors.) For an (n, k) code to be efficient, the ratio k/n should be as large as possible, consistent with its error detection and correction capabilities. Maximum efficiency in an (n, k) m-error correcting code occurs when it can correct up to m errors, and no others. Such a code is called *perfect*. Here is a very nice necessary condition—which we can verify—for the existence of a perfect code:

THEOREM 4. *If there exists a q-ary (n, k) perfect m-error-correcting code, then*

$$1 + (q-1)n + (q-1)^2\binom{n}{2} + \cdots + (q-1)^m\binom{n}{m} = q^r$$

for some positive integer r, and $k = n - r$.

Proof. A codeword of length n can have a single error occur in n positions, two errors in $\binom{n}{2}$ positions, and in general m errors in $\binom{n}{m}$ ways. For a q-ary code, there are

$q-1$ ways for a single error to occur at a given position, $(q-1)^2$ ways for two errors to occur at two given positions, and in general $(q-1)^m$ ways for m errors to happen at m given positions. Thus, the total number of ways in which no more than m errors can occur relative to a given codeword is equal to

$$1+(q-1)n+(q-1)^2\binom{n}{2}+\cdots+(q-1)^m\binom{n}{m}. \tag{3}$$

To complete the proof, we need to recognize this as a power of q.

The set of all n-long q-ary strings differing from a given codeword W in at most m positions is called the *sphere of radius m about W*. If a code is perfect, then every n-string lies in a sphere of radius m about some codeword, and the spheres do not overlap. That is, the union of the spheres is equal to the entire space of n-tuples. Since the latter has size q^n, it follows that

$$(\text{number of } m\text{-spheres}) \cdot (\text{size of each } m\text{-sphere}) = q^n.$$

Thus, since q is a prime, the size of an m-sphere must be a power of q, say, q^r, and (3) is satisfied. Finally, every m-sphere is centered about one of the q^k codewords. Since $q^n = q^r \cdot q^k$, it follows that $k = n - r$, and we are done. ■

Now, 11 happens to be the smallest prime number p for which $2^p - 1$ is not a prime. For $p = 2$, 3, 5, and 7, we obtain the primes $2^p - 1 = 3$, 7, 31, and 127, and $2^{11} - 1 = 2047 = 23 \cdot 89$ is composite. But there is ample recompense for the failure of $2^{11} - 1$ to be prime; let's take a closer look:

$$\begin{aligned} 2^{11} &= 1 + 23 \cdot 89 \\ &= 1 + 23(1 + 11 + 11 \cdot 7) \\ &= 1 + 23 + 23 \cdot 11 + 23 \cdot 11 \cdot 7 \\ &= 1 + 23 + \binom{23}{2} + \binom{23}{3}. \end{aligned}$$

In 1949, Golay noted that this is precisely the case $q = 2, n = 23, r = 11$ of Theorem 4. That is, the necessary condition for the existence of a binary $(23, 23 - 11)$ perfect 3-error-correcting code is satisfied.

In the same year, he also noticed that

$$1 + 2 \cdot 11 + 2^2\binom{11}{2} = 1 + 22 + 220 = 243 = 3^5,$$

so that the necessary condition for the existence of a ternary $(11, 11 - 5)$ perfect 2-error-correcting code is satisfied.

Of course, necessary conditions are not always sufficient, but in 1949, Golay constructed two linear codes with the above parameters and two slightly larger linear codes. The binary codes are G_{23} and G_{24}, the (23, 12) Golay code and the (24, 12) extended Golay code; the ternary codes are G_{11} and G_{12}, the (11, 6) Golay code and the (12, 6) extended Golay code.

We can describe an $(n, n - r)$ q-ary linear code as the row space of a matrix of n columns and rank r over $\mathbb{Z}_q$, the so-called *generating matrix* of the code. Let $\mathbf{A}$ be the following 12×24 binary matrix:

$$\mathbf{A} = \left[\begin{array}{cccccccccccc|cccccccccccc}
1&0&0&0&0&0&0&0&0&0&0&0&1&1&1&0&1&1&1&0&0&0&1&\mathbf{0}\\
0&1&0&0&0&0&0&0&0&0&0&0&1&0&1&1&0&1&1&1&0&0&0&\mathbf{1}\\
0&0&1&0&0&0&0&0&0&0&0&0&1&1&0&1&1&0&1&1&1&0&0&\mathbf{0}\\
0&0&0&1&0&0&0&0&0&0&0&0&1&0&1&0&1&1&0&1&1&1&0&\mathbf{0}\\
0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&1&0&1&1&0&1&1&1&\mathbf{0}\\
0&0&0&0&0&1&0&0&0&0&0&0&1&0&0&0&1&0&1&1&0&1&1&\mathbf{1}\\
0&0&0&0&0&0&1&0&0&0&0&0&1&1&0&0&0&1&0&1&1&0&1&\mathbf{1}\\
0&0&0&0&0&0&0&1&0&0&0&0&1&1&1&0&0&0&1&0&1&1&0&\mathbf{1}\\
0&0&0&0&0&0&0&0&1&0&0&0&1&1&1&1&0&0&0&1&0&1&1&\mathbf{0}\\
0&0&0&0&0&0&0&0&0&1&0&0&1&0&1&1&1&0&0&0&1&0&1&\mathbf{1}\\
0&0&0&0&0&0&0&0&0&0&1&0&1&1&0&1&1&1&0&0&0&1&0&\mathbf{1}\\
0&0&0&0&0&0&0&0&0&0&0&1&0&1&1&1&1&1&1&1&1&1&1&\mathbf{1}
\end{array}\right] \tag{4}$$

A is a generating matrix for G_{24}; deleting its last (boldface) column gives a generating matrix for G_{23}.

G_{12} is the row space of the following 12×12 ternary matrix **B**, and G_{11} is the row space of B', obtained from **B** by deleting the last column.

$$\mathbf{B} = \left[\begin{array}{cccccccccccc}
-1&1&1&-1&1&1&1&-1&-1&-1&1&\mathbf{-1}\\
-1&-1&1&1&-1&1&1&1&-1&-1&-1&\mathbf{1}\\
-1&1&-1&1&1&-1&1&1&1&-1&-1&\mathbf{-1}\\
-1&-1&1&-1&1&1&-1&1&1&1&-1&\mathbf{-1}\\
-1&-1&-1&1&-1&1&1&-1&1&1&1&\mathbf{-1}\\
-1&-1&-1&-1&1&-1&1&1&-1&1&1&\mathbf{1}\\
-1&1&-1&-1&-1&1&-1&1&1&-1&1&\mathbf{1}\\
-1&1&1&-1&-1&-1&1&-1&1&1&-1&\mathbf{1}\\
-1&1&1&1&-1&-1&-1&1&-1&1&1&\mathbf{-1}\\
-1&-1&1&1&1&-1&-1&-1&1&-1&1&\mathbf{1}\\
-1&1&-1&1&1&1&-1&-1&-1&1&-1&\mathbf{1}\\
-1&-1&-1&-1&-1&-1&-1&-1&-1&-1&-1&\mathbf{-1}
\end{array}\right]$$

Note that these codes are 6-dimensional subspaces of $\mathbb{Z}_3^{12}$ and $\mathbb{Z}_3^{11}$, respectively, since **B** and **B**′ have rank six. (Arithmetic in $\mathbb{Z}_3$ is just arithmetic mod 3 with the symbols -1, 0 and 1.)

We are now ready to connect the (11, 5, 2) biplane with the Golay code pairs. Let **U** and **V** be the upper rightmost 11×11 submatrices of **A** and **B**, respectively.

Take **U** and change all the 1s on the main diagonal to 0s, and what do you get? You get **M**, the incidence matrix for the (11, 5, 2) biplane.

Take **V** and change all the -1s to 0s. Then, change all the 1s on the main diagonal to 0s, and what do you get? Again, you get **M**.

Thus, from the (11, 5, 2) biplane we are able to construct the pair of binary Golay codes G_{23} and G_{24} and the pair of ternary Golay codes G_{11} and G_{12}.

Nice connections, to be sure, and there are even more connections with Steiner systems, so let's find out about them.

Steiner systems

If n is a positive integer, we use the expressions *n-set* and *n-subset* to mean an n-element set and an n-element subset. A *Steiner system* $S(p, q, r)$ is a collection S of q-subsets of an r-set R, such that every p-set in R is contained in exactly one of the q-sets in S. An $S(1, q, r)$ is just a partition of an r-set into q-sets, so that these exist if and only if r is a multiple of q. It is known that $S(2, 3, r)$s exist if and only if $r \equiv 1$ or 3 mod 6 and $S(3, 4, r)$s exist if and only if $r \equiv 2$ or 4 mod 6. Steiner systems $S(2, 3, r)$ are also block designs known as *Steiner triple systems*; here is $S(2, 3, 7)$, namely the (7, 3, 1) symmetric design with blocks A, B, C, D, E, F, and G:

$$A = 124, \quad B = 235, \quad C = 346, \quad D = 450, \quad E = 561, \quad F = 602, \quad G = 013. \quad (5)$$

(You can construct an $S(3, 4, 8)$ from the $S(2, 3, 7)$: adjoin ∞ to each of the blocks of the $S(2, 3, 7)$, and include the complement in $\{0, 1, 2, 3, 4, 5, 6\}$ of each block in the $S(2, 3, 7)$.)

For $p \geq 4$, the story is different: very few of these are known, and one reason is that there are restrictions on the parameters p, q, and r, namely:

THEOREM 5. *If S is an $S(p, q, r)$ defined on the r-set R, then S contains $\binom{r}{p}/\binom{q}{p}$ q-sets, and for $0 \leq j < p$:*

(a) $\binom{r-j}{p-j}/\binom{q-j}{p-j}$ *is an integer;*

(b) *every j-subset of R belongs to exactly $\binom{r-j}{p-j}/\binom{q-j}{p-j}$ q-sets of S;*

(c) *there exists an $S(p-j, q-j, r-j)$ on an $(r-j)$-subset of R.*

Proof. Each of the $\binom{r}{p}$ p-sets in R belongs to a unique q-set in S and each such q-set contains $\binom{q}{p}$ p-sets; hence, S contains $\binom{r}{p}/\binom{q}{p}$ q-sets. It follows that $\binom{r}{p}/\binom{q}{p}$ is an integer, which establishes (a) for $j = 0$. Now fix $x \in R$. Let $S_x = \{Y | Y \text{ is a } q\text{-set in } R \text{ containing } x\}$. Since each p-set containing x belongs to a unique q-set $Y \in S_x$, it follows that $S'_x = \{Y - \{x\} | Y \in S_x\}$ is a collection of $(q-1)$-subsets of $R - \{x\}$, such that each $(p-1)$-subset of $R - \{x\}$ belongs to a unique $(q-1)$-set in S'_x. In short, S'_x is an $S(p-1, q-1, r-1)$ on the set $R - \{x\}$; by the above, it follows that S_x contains exactly $\binom{r-1}{p-1}/\binom{q-1}{p-1}$ q-sets of S. This establishes (b) and (c) for $j = 1$. Continuing inductively, we see that if an $S(p, q, r)$ exists, then so does an $S(p-j, q-j, r-j)$ for $0 \leq j \leq p-1$; from this, we may deduce (b) and (c) for $0 \leq j < p$. ∎

The $S(p-j, q-j, r-j)$ systems obtained in this way from an $S(p, q, r)$ are said to be *derived* from the $S(p, q, r)$. Every known Steiner system $S(4, q, r)$ is derived from an $S(5, q+1, r+1)$, and very few Steiner systems with $p \geq 4$ are known at all. It turns out that we can use the $(11, 5, 2)$ biplane and the binary Golay codes to construct two pairs of these rare Steiner systems, namely $\{S(4, 5, 11), S(5, 6, 12)\}$ and $\{S(4, 7, 23), S(5, 8, 24)\}$.

We now construct $S(5, 6, 12)$ and the systems derived from it—in particular, $S(4, 5, 11)$—by means of a unified approach, beginning with the $(11, 5, 2)$ biplane. Let $B_1 = \{1, 3, 4, 5, 9\}$, the first block in the $(11, 5, 2)$ biplane, and let $B := B_1 \cup \{\infty\}$. Denote the set $\{0, 1, \ldots, 10\}$ by $[0..10]$. In what follows, addition and subtraction are all mod 11, except that $\infty \pm x = \infty$ for all x. If Y is a set of numbers and m is a number, then we define $Y + m := \{y + m | y \in Y\}$. For example, $B + 6 = \{1+6, 3+6, 4+6, 5+6, 9+6, \infty+6\} = \{7, 9, 10, 0, 4, \infty\}$.

Define the mappings s and σ to be permutations on $[0..10]$ and $[0..10] \cup \{\infty\}$, respectively, by

$$\sigma = (1\ \ 10)(2\ \ 5)(3\ \ 7)(4\ \ 8)(6\ \ 9) \quad \text{and} \quad s = (0\ \ \infty) \circ \sigma.$$

Now, if f is a permutation and Y is a set, then define $f(Y)$ to be $\{f(y) : y \in Y\}$. For example, since $B + 6 = \{7, 9, 10, 0, 4, \infty\}$, we see that

$$\begin{aligned} s(B+6) &= s(\{7, 9, 10, 0, 4, \infty\}) = \{s(7), s(9), s(10), s(0), s(4), s(\infty)\} \\ &= \{3, 6, 1, \infty, 8, 0\}, \quad \text{and so} \\ s(B+6) + 3 &= \{6, 9, 4, \infty, 0, 3\}. \end{aligned}$$

We now construct the Steiner systems as follows:

$$S(5, 6, 12) = \{B + k | k \in [0..10]\} \cup \{s(B + k) + j | j, k \in [0..10]\};$$

$$S(4, 5, 11) = \{B_1 + k | k \in [0..10]\} \cup \{\sigma(B_1 - n) + k : n \in B_1, k \in [0..10]\};$$

$S(3, 4, 10) =$ blocks of $S(4, 5, 11)$ containing 10, with 10 deleted; and

$S(2, 3, 9) =$ blocks of $S(3, 4, 10)$ containing 0, with 0 deleted.

A table on page 100 lists the blocks for $S(4, 5, 11)$ and the list for $S(5, 6, 12)$ is available at the MAGAZINE web site; don't peek until you've tried your hand at constructing them yourself. Notice that the blocks of the (11, 5, 2) biplane appear in $S(4, 5, 11)$ as its first column.

There are many ways to construct $S(5, 8, 24)$ (as, indeed, there are to construct $S(5, 6, 12)$), and one way is to use the Golay code G_{24}. By Theorem 5, if $S(5, 8, 24)$ exists, then it contains $\binom{24}{5}/\binom{8}{5} = 759$ 8-sets. This just happens to be the exact number of codewords of Hamming weight 8 in G_{24}. For example, all rows but the last in the generating matrix **A** (see Equation (4)) are codewords of weight 8. Let us number the columns of **A** with the customary numbering scheme $1, 2, \ldots, 22, 0, \infty$. If $c = c_1 c_2 \cdots c_\infty$ is a weight-8 codeword, then $O_c = \{i | c_i = 1\}$ is an 8-subset of $\{1, 2, \ldots, 22, 0, \infty\}$. The system $S(5, 8, 24)$ consists of these 759 so-called *octads*, and we construct the derived systems as follows:

$S(5, 8, 24) =$ codewords of weight 8 in G_{24};

$S(4, 7, 23) =$ octads of $S(5, 8, 24)$ containing ∞, with ∞ deleted;

$S(3, 6, 22) =$ blocks of $S(4, 7, 23)$ containing 0, with 0 deleted; and

$S(2, 5, 21) =$ blocks of $S(3, 6, 22)$ containing 22, with 22 deleted.

$S(5, 8, 24)$ has many remarkable properties and connections, and we have obviously left out many details. To do justice to this truly amazing object requires quite a journey. Thompson's book [**12**] is an excellent starting point; it certainly was for me.

One of the notable aspects of $S(5, 8, 24)$ is something it shares with the other Steiner systems, namely, a high degree of symmetry. Studying this symmetry leads us to the connection between the (11, 5, 2) biplane and the Mathieu groups.

Automorphisms, transitivity, simplicity, and the Mathieu groups

An *automorphism* of a Steiner system S is a permutation of the underlying r-set that also permutes the q-sets of S among themselves. For example, the permutation $a = (2\ 4)(5\ 6)$ on the set $\{0, 1, 2, 3, 4, 5, 6\}$ is an automorphism of $S(2, 3, 7)$ (see Equation (5)). Using the labeling convention from that equation, you can check that a switches B and C, switches D and F, and leaves A, E, and G fixed. That is, viewed as a permutation on $S(2, 3, 7)$, $a = (B\ C)(D\ F)$.

You may recall that the automorphisms of the (11, 5, 2) biplane form a group under composition, and the same is true for the automorphisms of a Steiner system. As before, we write $Aut(S)$ for the automorphism group of the Steiner system S. In general, Steiner systems have a large number of automorphisms. For example, $S(2, 3, 7)$ consists of seven triples, and yet $Aut(S(2, 3, 7))$ is isomorphic to $PSL(2, 7)$, the group of order 168 generated by the permutations $a = (2\ 4)(5\ 6)$ and $b = (0\ 1\ 2\ 3\ 4\ 5\ 6)$ on the set $\{0, 1, 2, 3, 4, 5, 6\}$. For example, you can show that, as a permutation on $S(2, 3, 7)$, $ab^2 = (A\ B\ F)(C\ E\ G)$.

The automorphism groups of $S(4, 5, 11)$, $S(5, 6, 12)$, $S(4, 7, 23)$, and $S(5, 8, 24)$ are known as the Mathieu groups M_{11}, M_{12}, M_{23}, and M_{24}, respectively. First, we'll learn about their origin and why they are important, and then we'll describe them.

Émile Mathieu (1835–1890) first constructed the groups bearing his name in two papers summarizing work from his doctoral thesis. The Mathieu groups are special in two ways: first, they are *multiply transitive*, and second, they are *simple*—the first of the so-called *sporadic* finite simple groups ever described. Let us see what these terms mean.

A group of permutations G on a set A is called *k-transitive* if for every pair of ordered k-tuples $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ of elements of A, there exists $g \in G$ such that $g(a_i) = b_i$ for $1 \leq i \leq k$. We call G *transitive* (respectively, *multiply transitive*) if it is 1-transitive (respectively, k-transitive for some $k > 1$). A k-transitive group is also $(k-1)$-transitive. For example, the alternating group $A_3 = \{(1), (1\ 2\ 3), (1\ 3\ 2)\}$ is transitive, but not multiply transitive. The group $PSL(2, 7)$ is 2-transitive, but not 3-transitive. The symmetric group S_n, consisting of all permutations on $\{1, 2, \ldots, n\}$, is n-transitive. Now, one special feature of the Mathieu groups is that they are highly transitive. Theorem 6 tells the story; you can find a proof in many texts about finite group theory [**3, 11**].

THEOREM 6.

(a) *If G is 4-transitive, then G is isomorphic to* (i) *a symmetric group S_n for some $n \geq 4$,* (ii) *an alternating group A_n for some $n \geq 6$, or* (iii) *M_n for $n = 11, 12, 23,$ or 24.*

(b) *If G is 5-transitive, then G is isomorphic to* (i) *a symmetric group S_n for some $n \geq 5$,* (ii) *an alternating group A_n for some $n \geq 7$,* (iii) *M_{23}, or* (iv) *M_{24}.*

The Mathieu groups are also *simple*, and to understand what that means, we need to recall an idea from matrix algebra: Two matrices A and B are *similar* if there exists an invertible matrix Q such that $B = Q^{-1}AQ$. We can carry this idea over into groups: two group elements a and b are *conjugate* if there exists a group element g such that $b = g^{-1}ag$. (Remember, all elements of a group are invertible.) For example, if $a = (1\ 2)$, $b = (1\ 3)$, and $g = (1\ 2\ 3)$, then you can show that $b = g^{-1}ag$.

A special property of some subgroups is that of normality: A subgroup H of a group G is *normal* if for all $h \in H$ and for all $g \in G$, H contains $g^{-1}hg$. For example, let $S = \{(1), (1\ 2\ 3), (1\ 3\ 2), (1\ 2), (1\ 3), (2\ 3)\}$, the group of all permutations of $\{1, 2, 3\}$; let $A = \{(1), (1\ 2\ 3), (1\ 3\ 2)\}$ and $C = \{(1), (1\ 2)\}$. You can check that A and C are both subgroups of S, that A is normal, and that C is not normal. If G is an abelian (commutative) group, then all subgroups are normal—for, if $h \in H$ and $g \in G$, then $g^{-1}hg = g^{-1}gh = h \in H$ by commutativity.

A group containing no normal subgroups except itself and the identity subgroup is called *simple*. Just as prime numbers are the (multiplicative) building blocks by which we construct all the integers, so simple groups are the building blocks for constructing all finite groups. A major achievement of twentieth-century mathematics, featuring such luminaries as Chevalley, Feit, Thompson, Conway, Fischer, Gorenstein, and many others, was the complete classification of finite simple groups. The upshot of this effort, spanning some 15,000 journal pages (!), is that all finite simple groups belong to a few well-studied infinite families—except for twenty-six so-called *sporadic* groups. And the Mathieu groups were the very first sporadic groups ever described. Speaking of which:

There are many ways to describe the Mathieu groups; here is one: Let s, t, and u be the permutations defined by

$$s = (0\ \infty)(1\ 10)(2\ 5)(3\ 7)(4\ 8)(6\ 9),$$

$$t = (0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10), \quad \text{and}$$

$$u = (3\ 9\ 4\ 5)(2\ 6\ 10\ 7).$$

Then M_{11} is the group generated by t and u, and M_{12} is the group generated by s, t, and u. And yes, s *is* the same permutation that we used to construct $S(5, 6, 12)$.

Let α, β, and γ be the permutations defined on $\{0, 1, \ldots, 22, \infty\}$ by

$$\alpha = (2\ 16\ 9\ 6\ 8)(4\ 3\ 12\ 13\ 18)(10\ 11\ 22\ 7\ 17)(20\ 15\ 14\ 19\ 21),$$

$$\beta = (0\ 1\ \ldots\ 21\ 22), \quad \text{and}$$

$$\gamma = (0\ \infty)(1\ 22)(2\ 11)(3\ 15)(4\ 17)(5\ 9)(6\ 19)(7\ 13)(8\ 20)(10\ 16)(12\ 21)(14\ 18).$$

Then M_{23} is the group generated by α and β, and M_{24} is the group generated by α, β, and γ. See Thompson [**12**] for more details.

We can now see how M_{11} and M_{12} go together in a pair, and the same is true of M_{23} and M_{24}; to cement their connection with the (11, 5, 2) biplane further, it turns out that the number of elements in each of these groups is divisible by 11.

With that, our whirlwind tour of the (11, 5, 2) biplane, its symmetries, and its connections with six pairs of combinatorial gems is done. Quite a tale!

Questions

- *Where can I go to learn more about these things?* Look in the bibliography. Beth, Jungnickel, and Lenz [**1**] will take you a long way into the world of combinatorial designs, including all the ones mentioned in this paper and many more. Hughes and Piper [**6**] will do the same; they pay special attention to biplanes, and the previously mentioned unlabeled version of FIGURE 2 appears on the cover of their book. Marshall Hall [**5**] has a great deal of information on difference sets. A recent article in this MAGAZINE [**2**] will tell you more about difference sets and squares mod p. MacWilliams and Sloane [**8**] and Pless [**9**] are two standard works on error-correcting codes. Carmichael [**3**] has a whole lot of information about the Mathieu groups, although his presentation is a bit old-fashioned; for a more modern treatment, Rotman [**11**] is one of the best. Finally, Thompson [**12**] has all of these in a wonderfully written book. Happy Reading!
- *You said that $Aut((11, 5, 2))$ is isomorphic to $PSL(2, 11)$. How do you prove that?* It turns out that both $PSL(2, 11)$ and $Aut((11, 5, 2))$ are generated by two elements c and d, for which $c^2 = d^3 = (cd)^{11} = ((cd)^3(cd^2)^3)^2 = 1$, the identity permutation. Two automorphisms of (11, 5, 2) that fill the bill are $c = (1\ 3)(2\ 5)(4\ X)(7\ 9)$ and $d = (3\ 4\ 5)(2\ 6\ 0)(7\ 8\ X)$. Try it and see.
- *Any other tidbits about $G = Aut((11, 5, 2))$?* Here are a few. (1) It so happens that the group $H = Stab_G(B_1)$ has order 60 and is isomorphic to A_5, the alternating group on 5 elements. A picture of the (11, 5, 2) biplane that would show this would indeed be spectacular. (2) Constructing one to depict the automorphism d (of order 3) from the previous bullet is a good warm-up for (1). (3) G contains an automorphism of order 6: find one and draw the associated picture. (4) Recall that $|G| = 660$, which is divisible by 4; does G contain an automorphism of order 4?
- *You told us about four Mathieu groups, but I read somewhere that there are five of them. What is the fifth Mathieu group, and why did you leave it out?* Right you are; it's called M_{22}. This group is a permutation group on a 22-element set that is simple and triply transitive. One way to describe M_{22} is that it is the set of all permutations

in M_{23} that leave 0 fixed. It is a subgroup of index two of the group of automorphisms of $S(3, 6, 22)$. (The additive group of even integers is a subgroup of index two of the integers.) I left it out because it is not part of a pair.

- *Are there any interesting problems associated with the* $(11, 5, 2)$ *biplane, or with biplanes in general?* Several come to mind. Recall that a biplane is a symmetric (v, k, λ) design with $\lambda = 2$. (1) There are biplanes with $v < 11$; find them, find their automorphism groups, and draw some pictures. (2) Two block designs are *isomorphic* if there exists a one-to-one correspondence between the underlying sets of varieties that induces such a correspondence between the sets of blocks. Show that every $(11, 5, 2)$ biplane is isomorphic to the one presented in this paper. (3) Construct a $(16, 6, 2)$ biplane. Then, construct another one not isomorphic to the first one. How do you show that two designs are not isomorphic? Interesting question! (4) One problem is particularly intriguing. Recall that a finite projective plane is a symmetric design with $\lambda = 1$. It turns out that if q is a prime power, then there exists a finite projective plane with parameters $(q^2 + q + 1, q + 1, 1)$; as a consequence, there are infinitely many finite projective planes. So we may ask the question, "Are there infinitely many biplanes?" Nobody knows! Other than the ones alluded to in (1), the only known biplanes are for $k = 5$, 6, 9, 11, and 13. Are there others? Find the answer and become famous.

REFERENCES

1. T. Beth, D. Jungnickel, and H. Lenz, *Design Theory*, 2nd ed., Cambridge University Press, New York, 1999.
2. Ezra Brown, The many names of (7, 3, 1), this MAGAZINE **75** (2002), 83–94.
3. Robert D. Carmichael, *Introduction to the Theory of Groups of Finite Order*, Dover Publications, New York, 1956.
4. Richard K. Guy, The unity of combinatorics, in *Combinatorics Advances*, C. J. Colburn and E. S. Mahmoodian (eds.), Kluwer, 1995, 129–159.
5. Marshall Hall, Jr., *Combinatorial Theory*, 9th ed., Blaisdell Publishing Company, Waltham, MA, 1967.
6. D. R. Hughes and F. C. Piper, *Design Theory*, Cambridge University Press, New York, 1985.
7. T. A. Kirkman, On a problem in combinations, *Camb. Dublin Math. J.* **2** (1847), 191–204.
8. F. Jessie MacWilliams and Neil J. A. Sloane, *The Theory of Error-Correcting Codes*, 2nd reprint, North-Holland Mathematical Library **16**, North-Holland, New York, 1983.
9. Vera Pless, *Introduction to the Theory of Error-Correcting Codes*, 2nd ed., Wiley, New York, 1989.
10. George Pólya, *How To Solve It*, 2nd ed., Doubleday, Garden City, NY, 1957.
11. Joseph J. Rotman, *An Introduction to the Theory of Groups*, 4th ed., Springer-Verlag, New York, 1995.
12. Thomas M. Thompson, *From Error-Correcting Codes Through Sphere Packings to Simple Groups*, Carus Mathematical Monograph No. 21, Mathematical Association of America, Washington, DC, 1983.
13. W. S. B. Woolhouse, Prize question 1733, *Lady's and Gentleman's Diary*, 1844.

13459	07293	03618	0412X	06X59	05784
2456X	183X4	14729	15230	1706X	16895
35670	29405	2583X	26341	28170	279X6
46781	3X516	36940	37452	39281	38X07
57892	40627	47X51	48563	4X392	49018
689X3	51738	58062	59674	504X3	5X129
79X04	62849	69173	6X785	61503	6023X
8X015	7395X	7X284	70896	72615	71340
90126	84X60	80395	819X7	83726	82451
X1237	95071	914X6	92X08	94837	93562
02348	X6182	X2507	X3019	X5948	X4673

The Steiner system S(4, 5, 11) and the (**11**, **5**, **2**) **design** from Brown's article on page 87

Perfect Shuffles through Dynamical Systems

DANIEL SCULLY
Saint Cloud State University
St. Cloud, MN 56301-4498
djscully@stcloudstate.edu

A perfect shuffle on a deck of $2n$ cards is performed by splitting the deck into two packets of n cards each and then interleaving the cards in the two packets together perfectly. This can be done in two ways. If the shuffle moves the top and bottom cards one position into the deck, it is called an *in shuffle*. If, on the other hand, the shuffle leaves the top and bottom cards in their original positions, it is called an *out shuffle*.

There are several ways to model perfect shuffles mathematically, and there is a good deal of mathematical research on perfect shuffles and their applications. We know that they were studied as far back as 1726 [**3**]. Many articles on perfect shuffles are of a recreational-mathematics nature [**1, 10, 15**], and some are directed towards undergraduate mathematics education [**4, 14, 19**]. Authors as renowned as Persi Diaconis, Ron Graham, and Martin Gardner have written on the subject [**7, 8, 9**]. You can find applications to computer science [**2, 5, 16, 20, 21, 22**] as well as magic [**13, 23, 24**]. The best single reference on all of these aspects of perfect shuffles is the book *Magic Tricks, Card Shuffling, and Dynamic Computer Memories* by S. Brent Morris and published by the MAA. It also contains the most complete bibliography the author has found on the subject.

Magicians do perform perfect shuffles. They appeal to magicians because they appear random but are not. When designing a card trick, a practicing magician may be able to use the fact that the 8th card in a deck of 50 returns home after just 3 out shuffles. That magician may want to know what other variations result from changing the deck size. The orbit structures and the orders of perfect-shuffle permutations vary wildly as the deck size changes.

In this article we present a new way to model perfect shuffles that uses the well-known doubling function from dynamical systems. This model shows the intimate relationship between the orbits of the cards under perfect-shuffle permutations and the binary expansions of certain rational numbers from the unit interval. We exploit that relationship to show connections between various perfect shuffles on the various sized decks, and we use it to link deck sizes to orbit lengths. The model generalizes nicely to k-handed perfect shuffles, and we use the model to describe more general (less-than-perfect) riffle shuffles.

Old models One of the first really helpful models for studying perfect shuffles goes as follows: Number the locations of the cards in a deck of size $2n$ from 1 through $2n$ beginning at the top, and let $I_{2n}(x)$ represent the new location after an in shuffle of the card formerly at location x. It is easy to see that

$$I_{2n}(x) = 2x \bmod (2n + 1).$$

A similar result holds for out shuffles. This time the card locations in the deck are numbered from 0 through $2n - 1$ and $O_{2n}(x)$ represents the new location after an out shuffle of the card formerly at location x. Then

$$O_{2n}(x) = 2x \bmod (2n - 1).$$

Since $0 \equiv 2n - 1 \bmod (2n - 1)$, we must view the above formula with the added understanding that $O_{2n}(0) = 0$ and $O_{2n}(2n - 1) = 2n - 1$, as the top and bottom cards are not moved by an out shuffle. This little problem is not that strange; it also occurs in ones-complement computer arithmetic with the confusing dual representation of zero as 0 and -0, and for essentially the same reason.

Through composition of functions, we see that if k in shuffles are performed on a deck, the card that originated at location x is moved to location

$$I_{2n}^k(x) = 2^k x \bmod (2n + 1).$$

The *order* of a permutation equals the smallest positive number k of iterations of the permutation that brings every permuted element back to its original location. In this case, the order of the in shuffle on a deck of size $2n$, denoted by $|I_{2n}|$, equals the smallest positive integer k such that $2^k x \equiv x \bmod (2n + 1)$ for all $x = 1, 2, \ldots, 2n$. That is, $|I_{2n}|$ equals the smallest positive integer k such that $2n + 1$ divides $2^k x - x = x(2^k - 1)$ for all $x = 1, 2, \ldots, 2n$. But note that if $2n + 1$ divides $1(2^k - 1)$, then $2n + 1$ divides $x(2^k - 1)$ for all x. This tells us that after repeated in shuffles, all cards return home whenever the top card returns home. This proves the following lemma:

LEMMA 1. *After repeated in shuffles on an even-sized deck of cards, whenever the top card returns to the top, all other cards return to their original locations.* (The proof above yields a similar result for repeated out shuffles except that the top card (location 0) is not important; rather, we look for when the card just under the top (location 1) returns to its original location.)

These models are particularly helpful when dealing with questions that involve in shuffles exclusively or out shuffles exclusively. Their disadvantages result from the different numbering schemes and the different moduli. This becomes apparent when dealing with questions that involve both in and out shuffles. For example, suppose you want to find the shortest sequence of in and out shuffles that moves the 30th card in a deck of 52 to the top of the deck. It is difficult to use the above models to answer such a question. For the answer to this question, see [**18**], in which the following model is used:

In a deck of size $2n$, number the locations of the cards from 0 through $2n - 1$. If $O_{2n}(x)$ and $I_{2n}(x)$ represent the new locations of the card formerly at location x after an out shuffle and an in shuffle respectively, then the following two formulas describe the movements of the cards in the deck that result from out and in shuffles:

$$O_{2n}(x) = \begin{cases} 2x \bmod 2n & \text{if } 0 \le x \le n - 1 \\ 2x + 1 \bmod 2n & \text{if } n \le x \le 2n - 1 \end{cases}$$

$$I_{2n}(x) = \begin{cases} 2x + 1 \bmod 2n & \text{if } 0 \le x \le n - 1 \\ 2x \bmod 2n & \text{if } n \le x \le 2n - 1. \end{cases}$$

This model has a particularly nice wrap-around property when dealing with a deck of size 2^t [**7**]. In this case, let $x_1 x_2 \ldots x_t$ be the t-digit binary expansion of x, then $O_{2^t}(x_1 x_2 \ldots x_t) = x_2 x_3 \ldots x_t x_1$ and $I_{2^t}(x_1 x_2 \ldots x_t) = x_2 x_3 \ldots x_t \overline{x_1}$ where $\overline{x_i} = 1$ if $x_i = 0$ and $\overline{x_i} = 0$ if $x_i = 1$. With our new model, we obtain a similar result that applies to decks of all sizes.

New model Consider the function S defined by

$$S(x) = \begin{cases} 2x & \text{if } 0 \le x \le \frac{1}{2} \\ 2x - 1 & \text{if } \frac{1}{2} < x \le 1 \end{cases}$$

and pictured in FIGURE 1. This is known in the field of dynamical systems as the *doubling function* or as the *doubling function* mod 1. For our purposes the title *shuffle function* seems more appropriate.

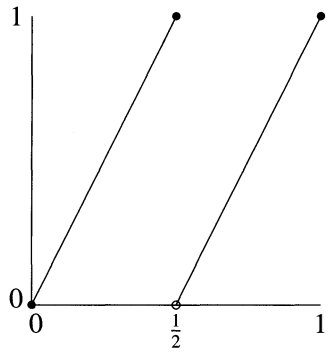


Figure 1 Shuffle function

We use the function S to describe both in and out shuffles on decks of all even sizes. Later, we use S to describe perfect shuffles on decks of odd size as well.

Out Shuffle In a deck of size $2n$, label the positions in the deck with the rational numbers

$$0 = \frac{0}{2n-1}, \frac{1}{2n-1}, \frac{2}{2n-1}, \ldots, \frac{2n-1}{2n-1} = 1$$

beginning at the top. Then $S(x)$ provides the new position, after the completion of an out shuffle, of the card that occupied position x before the shuffle.

In Shuffle The function S is used in the same way to describe in shuffles except that the card positions are labeled

$$\frac{1}{2n+1}, \frac{2}{2n+1}, \ldots, \frac{2n}{2n+1}.$$

We refer to the rational-number label as the *position* of the card and the *card number* as the positive-integer label of the card when we label the cards with the natural numbers starting with 1 at the top of the deck and advancing downward. To further avoid confusion, card number i is referred to as the ith card. Ordinal numbers (1st, 2nd, 3rd, etc.) are used in this article to indicate card number. The card positions are rational numbers from the unit interval.

EXAMPLE 1. *In a deck of* 52 *cards, the* 30*th card occupies position* 29/51 *in this out-shuffle labeling, so an out shuffle sends it to position* $S(29/51) = 7/51$, *which represents the* 8*th card. On the other hand, the* 30*th card occupies position* 30/53 *in the in-shuffle labeling, and* $S(30/53) = 7/53$, *so after an in shuffle it becomes the* 7*th card in the deck.*

Intuitively, it is quite clear why the function S works so well to describe perfect shuffles. The line segments in its graph (FIGURE 1) have a slope of 2, so the function stretches points from the same half interval apart, and the vertical drop in the second piece (the effect of mod 1) interleaves the two halves together perfectly.

Perfect shuffles can also be performed on odd-sized decks. This is done on a deck of size $2n-1$ by dividing it into two packets of size n and $n-1$ and then by interleaving the two packets together perfectly. If the larger packet comes from the top of the deck, the shuffle must leave the top card out on top and so it is called an *out shuffle*. Otherwise, the top card moves down one place, giving an *in shuffle*.

For an out shuffle on $2n-1$ cards, we label the positions of the cards with the rational numbers with denominator $2n-1$, including 0 but excluding 1. For in shuffles, 1 represents the last card and no card is represented by 0. Applying S to these domains correctly models the shuffles.

We can simplify some of the details by observing that the four different perfect shuffles (ins and outs on even- and odd-sized decks) are modeled by applying the same function S to the four different domains defined in TABLE 1. These domains differ only in whether they include either or both of the endpoints of the unit interval.

TABLE 1: Shuffle permutations (r an odd integer)

Domain	Endpoints	Shuffle	Deck Size
$A_r = \left\{\frac{0}{r}, \frac{1}{r}, \ldots, \frac{r}{r}\right\}$	0 & 1	out	$r+1$
$B_r = \left\{\frac{1}{r}, \frac{2}{r}, \ldots, \frac{r}{r}\right\}$	1	in	r
$C_r = \left\{\frac{0}{r}, \frac{1}{r}, \ldots, \frac{r-1}{r}\right\}$	0	out	r
$D_r = \left\{\frac{1}{r}, \frac{2}{r}, \ldots, \frac{r-1}{r}\right\}$	neither	in	$r-1$

We call these sets *domains with denominator* r or, for various values of r, *domains with odd denominators*. We use the generic notation E_r to refer to any one of the four domains with denominator r, and the notation S_r to refer to the shuffle function S when restricted to any of these four domains. Two domains E_r and E_s are said to be *of the same type* if they include the same endpoints 0 and 1. That is, they are both used to describe the same type of perfect shuffle (both out shuffles on even decks or both in shuffles on odd decks, etc.) but on different sized decks. Since S is a bijection from each of these finite domains back to itself, S_r is called a *shuffle permutation*.

The function S has been well studied in the field of dynamical systems [**6**], and many of the ideas in this article are borrowed from dynamical systems and applied to perfect shuffles. Though it is quite uncommon in the study of dynamical systems to consider functions on finite domains, it is quite common to take simple functions like S and vary their domains to obtain interesting and complex results. For this reason, it seemed important to include *dynamical systems* in the title.

Because the same function S is used to describe perfect shuffles on decks of all sizes, this model is particularly nice for finding relationships among the various perfect shuffles on different sized decks. As a first rather trivial example of this, it is immediately clear that within an out shuffle on $r+1$ cards, there is an in shuffle on r cards if we ignore the top card only, an out shuffle on r cards if we ignore the bottom card only, and an in shuffle on $r-1$ cards if we ignore both. Though the behaviors of these shuffles appear different when we look at the effects of the various shuffles on the card numbers, the effects of the shuffles on the card positions (rational-number labels) are identical. The same function applied to the same points produces the same results. We summarize this in the following lemma:

LEMMA 2. *Given an odd positive integer r and a deck of $r-1$, r, or $r+1$ cards, label the card positions with the rational numbers in the sets A_r, B_r, C_r, or D_r, defined in* TABLE 1, *beginning at the top. Then the applications of the shuffle function S*

to these domains are permutations on these domains and describe the movements of the cards upon the applications of a perfect shuffle as indicated on TABLE 1. *The effects of all four perfect shuffles on the card positions of their corresponding decks are essentially the same except for the inclusion or exclusion of one or the other or both of the fixed cards at the top and bottom in the positions of the rational numbers* 0 *and* 1.

Binary representation

A fact that you know and can easily prove about decimal expansions of rational numbers is illustrated as follows:

$$.1234 = \frac{1{,}234}{10{,}000} \quad \text{and} \quad .\overline{1234} = \frac{1{,}234}{9{,}999}.$$

In general, $.\overline{x_1x_2\ldots x_m} = x_1x_2\ldots x_m/(10^m - 1)$. The property carries over to other bases, and if $x_1x_2\ldots x_m$ is a base-b expansion of a positive integer x, then in base b, $.\overline{x_1x_2\ldots x_m} = x/(b^m - 1)$. We are particularly interested in binary expansions ($b = 2$).

Let x be a real number between 0 and 1 with a binary expansion of $.x_1x_2x_3\ldots$. For rational numbers with two binary expansions, we agree to use the one that does not terminate. So, we would express 1/4 in binary as $.00111\ldots = .00\overline{1}$ rather than $.01 = .01\overline{0}$. If $x_1 = 0$, then $0 \le x \le 1/2$ and so $S(x) = 2x = .x_2x_3x_4\ldots$. On the other hand, if $x_1 = 1$, then $1/2 < x \le 1$ so again $S(x) = 2x - 1 = .x_2x_3x_4\ldots$. Either way, the shuffle function S acts on the binary expansion of x by dropping the first digit. In symbolic dynamics [**12**] such functions are called *one-sided shift maps*.

The effect of S on a repeating binary expansion is particularly nice, since

$$S(.\overline{x_1x_2\ldots x_m}) = .\overline{x_2\ldots x_mx_1}.$$

This is the key to finding the number of elements in an *orbit*, the set of images under repeated applications of S.

EXAMPLE 2. *Find the orbit of the* 15*th card of a* 36*-card deck under the out-shuffle permutation.*

The position of the 15th card is $14/35 = 2/5$. Its binary expansion is $.\overline{0110}$, so its orbit has length 4. We see that this equals the number of digits in the repeating block of the binary expansion. Its orbit is

$$.\overline{0110} \to .\overline{1100} \to .\overline{1001} \to .\overline{0011} \to .\overline{0110} \to \ldots.$$

Expressed in fractions with denominators of 35, we get

$$\frac{14}{35} \to \frac{28}{35} \to \frac{21}{35} \to \frac{7}{35} \to \frac{14}{35} \to \ldots.$$

Translated to card numbers, we get

$$15\text{th} \to 29\text{th} \to 22\text{nd} \to 8\text{th} \to 15\text{th} \to \ldots.$$

When modeling card shuffles, we are interested in the effects of S on rational numbers with odd denominators. We know that rational numbers have repeating binary expansions, but the repeating block does not always begin with the first digit.

In dynamical systems, we say a point x has *period* k under S if there is a positive integer k such that $S^k(x) = x$. Such a point is said to be *periodic*. If k is the smallest such positive integer, we say that x has *prime period* k. If a point x is not periodic

but another point in its orbit is, we say that x is *eventually periodic*. Points that are neither periodic nor eventually periodic are said to be *aperiodic*. It is clear that rational points with binary expansions that begin repeating right away are periodic under S, and the period equals the length of the repeating block. Rational points that begin their repeating blocks later in their binary expansions are eventually periodic, and the irrationals are aperiodic because their binary expansions never repeat.

The following number-theoretic lemma shows that all of the domains with odd denominators consist entirely of points periodic under S. Thus this new model has a binary-expansion aspect that is reminiscent of the wrap-around property discussed earlier with one of the old models, but it applies to decks of all sizes, rather than just those of size 2^t. The lemma is easily proved using the shuffle function.

LEMMA 3. *The binary expansion of a rational number between 0 and 1 starts repeating right away if and only if the rational number can be expressed as a ratio of two integers with an odd denominator.*

Proof. In binary, if $x = .\overline{x_1x_2\ldots x_m}$, then as mentioned above,

$$x = \frac{x_1x_2\ldots x_m}{2^m - 1}.$$

Since $2^m - 1$ is odd, binary expansions that start repeating right away represent rational numbers with odd denominators.

Conversely, suppose that x is a rational number between 0 and 1 with an odd denominator and in binary $x = .x_1\overline{x_2x_3\ldots x_m}$ with $x_1 \neq x_m$, that is, the repeating block begins immediately after the first digit of the binary expansion, then

$$S(x) = .\overline{x_2x_3\ldots x_m}, \quad S^2(x) = .\overline{x_3\ldots x_mx_2}, \cdots$$
$$S^{m-1}(x) = .\overline{x_mx_2\ldots x_{m-1}} = .x_m\overline{x_2x_3\ldots x_m}, \quad S^m(x) = .\overline{x_2x_3\ldots x_m}.$$

Since $x_1 \neq x_m$, $S^{m-1}(x) \neq x$, but $S^m(x) = S(x)$. This would imply that a perfect-shuffle permutation is not one to one, a contradiction. Similar arguments show that the repeating block cannot begin after any finite delay. ∎

We summarize with the following theorem that we already proved:

THEOREM 1. *Let x be an element of a domain with denominator r and let $.\overline{x_1x_2\ldots x_m}$ be the binary expansion of x written as the shortest possible repeating block. Then the length of the orbit of x under the perfect shuffle permutation S_r is equal to m, the length of the repeating block. Further, its orbit, expressed in binary, is*

$$.\overline{x_1x_2\ldots x_m} \to .\overline{x_2\ldots x_mx_1} \to \ldots \to .\overline{x_mx_1\ldots x_{m-1}} \to \ldots$$

Also, the length of the orbit of x equals the smallest value of m such that the denominator of x when written in lowest terms divides $2^m - 1$.

Theorem 1 shows us how the card position, specifically the binary expansion of that rational number, carries the information concerning the orbit of the card. All cards in positions with the same lowest-terms denominator have orbits of the same length.

We must keep in mind that the fractions that can be expressed in binary as repeating blocks of length m include those that can be expressed in blocks of lengths that divide m (just as in decimals where we know that $.\overline{123123} = .\overline{123}$). In these cases their lowest-terms denominators divide $2^m - 1$, but they also divide $2^t - 1$ where t is a proper divisor of m.

Let *packet 0* be the top packet of the deck and *packet 1* the bottom packet. The card positions in packet 0 are labeled with rational numbers from the interval [0, 1/2]. These have binary expansions that begin with 0. The positions in packet 1 have binary expansions that begin with 1 and come from the interval (1/2, 1]. Dividing the deck into disjoint packets in this way is an example of a *Markov partition* from symbolic dynamics [**12**]. The *itinerary* of a card under a particular shuffle is the sequence of packet numbers visited by the card as it travels through its orbit in the order the packets are visited. It is clear, by the wrap-around effect of S on the binary expansions of rational numbers with odd denominators, that if x is the card position with a binary expansion of $.\overline{x_1 x_2 \ldots x_m}$, then the itinerary of the card in position x is $x_1, x_2, \ldots, x_m, x_1, x_2, \ldots$. That is, the binary expansion of the card position gives the itinerary, and the position of a card is completely determined by its itinerary. Thus, given two cards, repeating a perfect shuffle some number of times will place them in different packets.

There is another point worthy of mention that is illustrated well in Example 2. It is called the *induced orbit rule*. Since $14/35 = 2/5 = 6/15$ and since the same shuffle function S is used to describe perfect shuffles on all sized decks, the 15th card in a deck of size 36 has the same out-shuffle orbit as the 3rd card in a deck of size 6 and the 7th card in a deck of size 16. These orbits are also the same as the in-shuffle orbits of the 14th, 2nd, and 6th cards in decks of size 34, 4, and 14, respectively. Similar connections can be made with orbits in odd-sized decks. This phenomenon is not a new discovery [**19**], but the new model makes it very easy to understand. Cards in the same rational-number positions in various sized decks have the same orbits. Though the card numbers are very different, the card positions are identical, so the orbits are identical.

We can use the new model, Theorem 1, and some inclusion-exclusion counting arguments to determine the cycle structure of perfect-shuffle permutations. Example 3 illustrates.

EXAMPLE 3. *Determine the cycle structure of the out-shuffle permutation on a deck of size* 36.

The 36 positions in the deck are given in the set

$$A_{35} = \left\{ \frac{0}{35}, \frac{1}{35}, \ldots, \frac{35}{35} \right\}.$$

By Theorem 1, the length of the orbit of an individual card in position x is the length of the repeating block in the binary expansion of the fraction x. This, again by Theorem 1, is the smallest exponent m for which x can be written as a rational number with a denominator of $2^m - 1$. The smallest exponent m for which 35 divides $2^m - 1$ is $m = 12$ (this is probably most easily found by calculating the binary expansion of 1/35 or, equivalently, applying S to 1/35 repeatedly to find its itinerary), so $1/35 = 117/4,095 = 117/(2^{12} - 1)$, and the 2nd card returns home after 12 out shuffles.

By Lemma 1, when the card just under the top returns home after repeated out shuffles, all cards return home, so $|O_{36}| = |S_{35}| = 12$ and every card has period 12. But we are interested in the *prime*, or shortest periods. Since these have lengths that divide 12, the possible prime orbit lengths are 12, 6, 4, 3, 2, and 1.

By Theorem 1, finding out how many cards have each of these orbit lengths is simply a matter of counting how many fractions in A_{35} can be expressed with denominators of $2^t - 1$ for $t = 1, 2, 3, 4, 6, 12$. This can be done by finding the greatest common divisors of 35 and $2^t - 1$.

We begin at the top to check for points of period 6. Since $2^6 - 1 = 63$ and $\gcd(35, 63) = 7$, the 35ths in the unit interval that are also 63rds are the 7ths. There

are eight of them $0/7, 1/7, \ldots, 7/7$. They are the elements of A_{35} that have period 6 (though not necessarily prime period 6). The repeating blocks in their binary expansions have length 6 or a divisor of 6 (that is, 3, 2, or 1).

The only other cards in this deck with periods smaller than 12 and not counted in the eight above (the 7ths) would have to have a period of 4. As above, $2^4 - 1 = 15$ and $\gcd(35, 15) = 5$, so the 35ths in the unit interval that are also 15ths are the 5ths. There are six of them, but two of them, $0/5$ and $5/5$ are counted in the 7ths above. Thus in an inclusion-exclusion manner we count $36 - 8 - 6 + 2 = 24$ cards with periods of 12. That gives us two 12-cycles.

Using the same reasoning for the smaller divisors, we conclude that the disjoint cycle structure of O_{36} consists of two 12-cycles, one 4-cycle, two 3-cycles, and two 1-cycles. It is interesting to note that the permutation has no 6-cycles. This is because $2^3 - 1 = 7$ and $\gcd(35, 7) = 7$ also, so all period-6 cards in this deck have prime period 3 or 1.

The two 1-cycles represent the top and bottom cards that remain fixed by the out shuffle. The in shuffle on 34 cards has the same cycle structure except that the two 1-cycles are eliminated. The in and out shuffles on 35 cards each leave out one of the two 1-cycles.

Guaranteeing a t-cycle For developing magic tricks, it is nice to know how many perfect shuffles of a particular type must be repeated in order to return some, but not necessarily all, cards back to their original positions. In order for a card to return home for the first time after t perfect shuffles, the perfect-shuffle permutation must contain a t-cycle. That is, some of the positions in the deck must be rational numbers that can be expressed with denominators of the form $2^t - 1$. Thus, when those rational numbers are expressed in lowest terms, their denominators must be divisors of $2^t - 1$. This condition is necessary but not sufficient because any such lowest-terms denominator that happens to be a divisor of $2^t - 1$ may also be a divisor of $2^m - 1$ for some m that is a proper divisor of t, thus making the cycle length shorter than t. This is illustrated in Example 3 where O_{36} contains no 6-cycles.

Using the method of exhaustion, this simple idea can be exploited to obtain simple conditions on the deck size that are necessary and sufficient to guarantee a t-cycle for any given t. We illustrate with the example $t = 4$.

EXAMPLE 4. *Find a necessary and sufficient condition on the deck size that assures a perfect shuffle on that deck contains a* 4*-cycle.*

In order to have a 4-cycle in the disjoint cycle decomposition of a perfect-shuffle permutation on a domain E_r, the domain must contain rational numbers that can be written with denominators of $2^4 - 1 = 15$. This is possible only if these rational numbers, when written in lowest terms, have denominators that divide 15. Thus the lowest-terms denominators we seek are 1, 3, 5, and 15. But 1 divides $2^1 - 1$, and 3 divides $2^2 - 1$, so any lowest-terms fractions with these denominators generate 1-cycles (fixed points) and 2-cycles (transpositions) respectively, not 4-cycles. This leaves denominators of 5 and 15. Clearly, the domain E_r contains rational numbers with denominators of 5 or 15 upon reduction to lowest terms if and only if 5 divides r or 15 divides r. Both cases are covered by requiring that 5 divide r, so we have the following result:

The disjoint cycle decomposition of a perfect-shuffle permutation on a domain with denominator r contains a 4*-cycle if and only if* 5 *divides* r.

Using $r = 45$ as an example, since 5 divides 45, the domains with denominators of 45 contain elements with orbits of length 4. Thus, the out shuffle on 46 cards, the in shuffle on 44 cards, and both the in and out shuffles on 45 cards contain 4-cycles.

By applying the reasoning process illustrated in Example 4 for $t = 4$ to other values of t, we obtain Theorem 2.

THEOREM 2. *Let r be an odd natural number and let S_r represent a perfect-shuffle permutation on a domain with denominator r. The following table lists the values $k_1, k_2, \ldots, k_p$ for various values of t in the statement: S_r contains a t-cycle if and only if k_1 divides r or k_2 divides r or ... k_p divides r.*

t	$k_1, k_2, \ldots, k_p$	t	$k_1, k_2, \ldots, k_p$	t	$k_1, k_2, \ldots, k_p$
2	3	9	73	16	257
3	7	10	11, 93	17	131071
4	5	11	23, 89	18	19, 27, 219
5	31	12	13, 35, 45	19	524387
6	9, 21	13	8191	20	25, 41, 55, 155
7	127	14	43, 129	21	49, 337, 889
8	17	15	151, 217	22	69, 267, 683

That is, S_r contains a 2-cycle if and only if 3 divides r, S_r contains a 3-cycle if and only if 7 divides r, and S_r contains a 12-cycle if and only if 13, 35, or 45 divide r.

Order of a perfect shuffle By Lemma 1 and Theorem 1 the order $|S_r|$ of a perfect-shuffle permutation S_r on a domain with denominator r equals the length of the repeating block of integers in the binary expansion of $1/r$. Performing the long division in base two or, equivalently, finding the orbit of $1/r$ under S is not difficult, but it can be time consuming if r is large. The job can be made easier if r can be factored. A simple and natural construction with the new model provides the necessary insight.

THEOREM 3. *Let r and s be odd natural numbers and let S_r and S_s be perfect-shuffle permutations on domains with denominators r and s respectively. Then,*

$$|S_{\mathrm{lcm}(r,s)}| = \mathrm{lcm}(|S_r|, |S_s|).$$

Proof. Let $t = \mathrm{lcm}(r, s)$, $i = |S_r|$, $j = |S_s|$, and $k = \mathrm{lcm}(i, j)$. Also, let E_r, E_s, and E_t be domains with denominators r, s, and t respectively of the same type.

It is clear that E_t is the smallest domain with an odd denominator that contains $E_r \cup E_s$. Since $|S_r|$ and $|S_s|$ equal the periods of $1/r$ and $1/s$ respectively under S, we know that S_t contains both an i and a j-cycle, so $|S_t|$ is a multiple of both i and j. Therefore, $|S_t| \geq \mathrm{lcm}(i, j) = k$.

Choose a such that $k = ai$. Since $|S_r| = i$, we know that r divides $2^i - 1$, and since $2^k - 1 = (2^i)^a - 1 = (2^i - 1)[(2^i)^{a-1} + \cdots + 1]$ we know that r divides $2^k - 1$. Similarly, s divides $2^k - 1$. Therefore t divides $2^k - 1$. But $|S_t| = \min\{m : t \text{ divides } 2^m - 1\}$. Thus, $|S_t| \leq k$. Putting the two inequalities together gives us $|S_t| = k$, and our theorem is proved. ∎

Theorem 3 is a standard number-theoretic improvement over an earlier result [**19**] that states: If r and s are relatively prime, then $|S_{rs}| = \mathrm{lcm}(|S_r|, |S_s|)$.

EXAMPLE 5. *Find the order of the out shuffle on a deck of size* 100.

Factoring $100 - 1$, we get $99 = 3^2 \cdot 11$. The binary expansions for $1/9$ and $1/11$ are $.\overline{000111}$ and $.\overline{0001011101}$ respectively with repeating block lengths of 6 and 10, thus the order of the out shuffle on a deck of 100 cards equals $\mathrm{lcm}(6, 10) = 30$.

Using Theorem 3, we reduce the problem of finding $|S_r|$ for any odd r to that of finding $|S_{p^\alpha}|$, where p^α is a power of an odd prime. As pointed out by Rosenthal [**19**], an examination of several examples suggests that $|S_{p^\alpha}| = p^{\alpha-1}|S_p|$. If this were indeed

the case, our problem could be reduced further to that of finding $|S_p|$. Unfortunately, it is not true. We can, however, prove Theorem 4, which shows that for p an odd prime, the sequence $|S_p|, |S_{p^2}|, \ldots$ is ultimately geometric with a common ratio of p. Before the sequence turns geometric, it is constant.

THEOREM 4. *Let p be an odd prime. Suppose m is a positive integer with the property that $|S_{p^m}| = |S_p|$, but $|S_{p^{m+1}}| \neq |S_p|$, then $|S_{p^{m+t}}| = p^t|S_p|$ for all nonnegative integers t.*

Proof. As mentioned earlier, $|S_r| = \min\{j : r \text{ divides } 2^j - 1\}$. It is important to recall that every j in this set is a multiple of $|S_r|$. This is most easily seen from elementary group theory. If a group element raised to a power equals the identity, then that power is a multiple of the order of the element [**11**]. Here, S_r, a permutation, is the group element, and composition is the operation.

Let $k = |S_p|$. The proof is by contradiction, so we assume that there exists a nonnegative integer t such that $|S_{p^{m+t}}| = kp^t$, but $|S_{p^{m+t+1}}| \neq kp^{t+1}$. Without loss of generality, suppose t is the first such nonnegative integer with this property. Since the domains with odd denominators $A_{p^{m+t}} \subseteq A_{p^{m+t+1}}$, we know that $|S_{p^{m+t}}|$ divides $|S_{p^{m+t+1}}|$, so $|S_{p^{m+t+1}}| = akp^t$ for some positive integer a.

Note that

$$2^{kp^{t+1}} - 1 = (2^{kp^t})^p - 1 = (2^{kp^t} - 1)\left[(2^{kp^t})^{p-1} + (2^{kp^t})^{p-2} + \cdots + 1\right].$$

Since $|S_{p^{m+t}}| = kp^t$, we know that p^{m+t} divides $2^{kp^t} - 1$, the first factor in the above factorization. This will be useful shortly, but it also implies that p divides $2^{kp^t} - 1$, and so $2^{kp^t} \equiv 1 \bmod p$. This tells us that $[(2^{kp^t})^{p-1} + (2^{kp^t})^{p-2} + \cdots + 1] \equiv 0 \bmod p$ since all p terms in the sum are congruent to 1. Thus, p divides $[(2^{kp^t})^{p-1} + (2^{kp^t})^{p-2} + \cdots + 1]$. Since p^{m+t} divides the first factor in the above factorization and p divides the second, p^{m+t+1} must divide their product $2^{kp^{t+1}} - 1$, and thus kp^{t+1} is a multiple of $|S_{p^{m+t+1}}|$. That is, $abkp^t = kp^{t+1}$ for some positive integer b. Canceling, we get $ab = p$, so $a = p$ or $a = 1$. We are assuming that $|S_{p^{m+t+1}}| \neq kp^{t+1}$, so $a \neq p$. Therefore $a = 1$ and $|S_{p^{m+t+1}}| = kp^t = |S_{p^{m+t}}|$. Though we are not yet finished, this does tell us that $t > 0$ since $|S_{p^{m+1}}| \neq |S_{p^m}|$.

Since $|S_{p^{m+t+1}}| = kp^t$, we know that p^{m+t+1} divides $2^{kp^t} - 1$. But, since $t > 0$,

$$2^{kp^t} - 1 = (2^{kp^{t-1}})^p - 1 = (2^{kp^{t-1}} - 1)\left[(2^{kp^{t-1}})^{p-1} + (2^{kp^{t-1}})^{p-2} + \cdots + 1\right].$$

And since $|S_{p^{m+t}}| = kp^t$, we know that $|S_{p^{m+t}}| > kp^{t-1}$, so p^{m+t} does not divide $(2^{kp^{t-1}} - 1)$, the first factor above. Thus, p^2 divides the second factor $[(2^{kp^{t-1}})^{p-1} + (2^{kp^{t-1}})^{p-2} + \cdots + 1]$. This is the fact that we now disprove to complete our proof by contradiction.

By the minimality of t and the fact that $t > 0$, we know that $|S_{p^{m+t-1}}| = kp^{t-1}$, so p^{m+t-1} divides $2^{kp^{t-1}} - 1$, which in turn implies that p divides $2^{kp^{t-1}} - 1$, and so $2^{kp^{t-1}} \equiv 1 \bmod p$. This implies that $2^{kp^{t-1}} = cp + 1$ for some integer c. So,

$$(2^{kp^{t-1}})^{p-1} + (2^{kp^{t-1}})^{p-2} + \cdots + 1 = (cp+1)^{p-1} + (cp+1)^{p-2} + \cdots + 1.$$

And

$$\begin{aligned}(cp+1)^n &= (cp)^n + \binom{n}{n-1}(cp)^{n-1} + \cdots n(cp) + 1 \\ &\equiv n(cp) + 1 \bmod p^2\end{aligned}$$

for all positive integers n, since all the terms but the last two are multiples of p^2. Thus,

$$\begin{aligned}
&(2^{kp^{t-1}})^{p-1} + (2^{kp^{t-1}})^{p-2} + \cdots + 1 \\
&\quad \equiv \left[(p-1)cp + 1\right] + \left[(p-2)cp + 1\right] + \cdots + [cp + 1] + 1 \bmod p^2 \\
&\quad \equiv \left[(p-1)cp + (p-2)cp + \cdots + cp\right] + p \bmod p^2 \\
&\quad \equiv \frac{p(p-1)}{2}cp + p \bmod p^2 \\
&\quad \equiv p \bmod p^2.
\end{aligned}$$

So, p^2 does not divide $(2^{kp^{t-1}})^{p-1} + (2^{kp^{t-1}})^{p-2} + \cdots + 1$. A contradiction. Therefore, $|S_{p^{m+t}}| = p^t|S_p|$ for all nonnegative integers t. ■

The first prime number p for which $|S_{p^\alpha}| \neq p^{\alpha-1}|S_p|$ is $p = 1093$ where $|S_{p^2}| = |S_p|$, but the geometric sequence begins right after that since $|S_{p^3}| = p|S_p|$ [**19**]. To the best of the author's knowledge, the next prime p (if it exists) for which $|S_{p^2}| = |S_p|$ is unknown. It is also unknown to the author whether there exist primes p for which $|S_{p^3}| = |S_p|$. Theorem 4 is stated in [**19**] along with a different proof of the special case $m = 2$.

k-handed perfect shuffles

Perfect shuffles can be generalized as follows. Split a deck of kn cards into k packets ($k \geq 2$) of n cards each. Number the packets from 0 through $k - 1$ with packet number 0 containing the top n cards and then proceed downward. The cards are then interleaved, one card from each packet, according to a fixed and repeated pattern. That is, after the shuffle, the top k cards consist of the top cards from each of the original packets ordered according to some permutation of the packet numbers $0, 1, \ldots, k - 1$. The next k cards of the shuffled deck are the second cards from each packet ordered by the same permutation. This permutation is repeated throughout the shuffle. This shuffle is called a *k-handed perfect shuffle*. Under this generalization, it is easy to see that there are $k!$ different k-handed perfect shuffles on a deck of kn cards (one shuffle for each permutation on the set of packet numbers).

When $k = 2$ we have two perfect shuffles, the out and the in shuffles. The out shuffle corresponds to the identity permutation on the set of packet numbers $\{0, 1\}$ (indicated by $[0, 1]$), since the top card after the shuffle comes from packet 0 and the next comes from packet 1. On the other hand, the in shuffle corresponds to the permutation $[1, 0]$ in which the packet order is reversed.

If k is allowed to get large, we can describe many permutations this way. Clearly, any permutation on n cards can be described as an n-handed perfect shuffle.

We generalize the two standard 2-handed out and in shuffles on $2n$ cards to two k-handed perfect shuffles on kn cards in the following natural way: The *k-handed out shuffle* is the k-handed perfect shuffle that corresponds to the identity permutation $[0, 1, \ldots, k - 1]$ of the packet numbers, and the *k-handed in shuffle* on kn cards is the k-handed perfect shuffle that corresponds to the reverse permutation $[k - 1, k - 2, \ldots, 0]$ of the packet numbers.*

Our new model also generalizes in an obvious way to describe k-handed out and in shuffles on decks of size kn. To describe the k-handed out shuffle, label the kn card

*EDITOR'S NOTE: This permutation is given in line notation, with brackets and commas added for clarity. Information about the various notations for permutations appears on page 129.

positions with the rational numbers

$$\frac{0}{kn-1}, \frac{1}{kn-1}, \ldots, \frac{kn-1}{kn-1},$$

then the card in position x is moved to position $S^{(k)}(x)$ by the k-handed out shuffle where

$$S^{(k)}(x) = kx \bmod 1.$$

The function $S^{(k)}$ is called the *k-handed shuffle function*. The graph of the 4-handed shuffle function is shown in FIGURE 2.

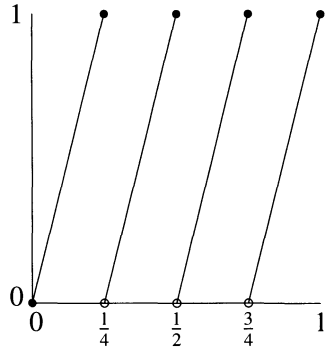


Figure 2 The 4-handed shuffle function

The same function models the k-handed in shuffle, but the positions must be labeled by the rational numbers

$$\frac{1}{kn+1}, \frac{2}{kn+1}, \ldots, \frac{kn}{kn+1}.$$

So, in describing the out and in k-handed shuffles on decks of size kn, the denominators of the rational numbers in the domains need not be odd. Rather, they must be one less or one more than the deck size kn.

The other k-handed perfect shuffles can be described using $S^{(k)}$ as well as k-handed perfect shuffles on decks of sizes that are not exact multiples of k. The k-handed perfect shuffle on $kn-1$ cards that corresponds to the identity permutation of the packet numbers and in which the last packet has one fewer card than the others is described by $S^{(k)}$ on the domain

$$\left\{\frac{0}{kn-1}, \frac{1}{kn-1}, \ldots, \frac{kn-2}{kn-1}\right\},$$

and the shuffle that corresponds to the permutation $[1, 2, \ldots, k-1, 0]$, where the top packet is short one card is described with the domain

$$\left\{\frac{1}{kn-1}, \frac{2}{kn-1}, \ldots, \frac{kn-1}{kn-1}\right\}.$$

Some are more difficult to describe, as Example 6 illustrates.

EXAMPLE 6. *The 4-handed perfect shuffle on a deck of 8 cards that corresponds to the permutation* [1, 3, 2, 0] *of packet numbers is modeled by* $S^{(4)}$ *when it is applied to the domain*

$$\left\{\frac{25}{255}, \frac{59}{255}, \frac{70}{255}, \frac{100}{255}, \frac{145}{255}, \frac{179}{255}, \frac{206}{255}, \frac{236}{255}\right\}.$$

At this point, it is less than clear how one arrives at this domain to describe this shuffle, but we will give directions in the next section. However, you can try applying $S^{(4)}$ to this set and see that it models the shuffle correctly.

When considering both even- and odd-sized decks, a 2-handed perfect shuffle may have 0, 1, or 2 fixed cards. This model makes it very clear how many fixed cards a k-handed perfect shuffle may have.

THEOREM 5. *A k-handed perfect shuffle can have at most k fixed cards.*

Proof. A proof without words is given in FIGURE 3. ■

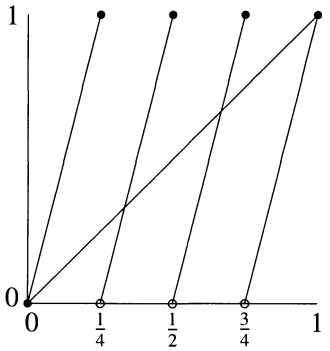


Figure 3 Proof of Theorem 5 without words

Of course, some or all of the points of intersection may not be included in the domain, so some k-handed perfect shuffles may have fewer than k fixed cards, but k is definitely an upper bound. Finding those fixed points is simply a matter of solving for x in the equations $kx - t = x$ for $t = 0, 1, \ldots, k-1$. The solutions are $x = t/(k-1)$ for $t = 0, 1, \ldots, k-1$. How many fixed points actually fall in the domain depends on the denominator r used in the domain. For out shuffles on decks of size $kn = r + 1$, the number of fixed points is $\gcd(r, k-1) + 1$. For in shuffles on decks of size $kn = r - 1$, the number is $\gcd(r, k-1) - 1$.

The properties analogous to those of 2-handed perfect shuffles are illustrated in the following two examples. To simplify some of the confusing details in these examples, we assume that we are working on out shuffles on decks of $kn = r + 1$ cards.

EXAMPLE 7. *Find the orbit and itinerary of the* 13*th card in a deck of size* 85 *under the* 5*-handed out shuffle, and find the cycle structure of the shuffle permutation* $S_{84}^{(5)}$.

Here, instead of using a binary expansion, we work in base 5. The base-5 expansion for the position $12/84 = 1/7$ (occupied by the 13th card) is $.\overline{032412}$, so its itinerary is 0, 3, 2, 4, 1, 2, 0, ... and its orbit in base-5 expansions is

$$.\overline{032412} \to .\overline{324120} \to .\overline{241203} \to .\overline{412032} \to .\overline{120324} \to .\overline{203241} \to \ldots.$$

Translating to fractions we get

$$\frac{1}{7} \to \frac{5}{7} \to \frac{4}{7} \to \frac{6}{7} \to \frac{2}{7} \to \frac{3}{7} \to \frac{1}{7} \ldots.$$

To find their card numbers, these fractions must be written with denominators of 84. This gives us

$$13\text{th} \to 61\text{st} \to 49\text{th} \to 73\text{rd} \to 25\text{th} \to 37\text{th} \to 13\text{th} \to \ldots.$$

The base-5 expansion for $1/84$ (the position of the 2nd card) is $.\overline{001221}$, so $|S_{84}^{(5)}| = 6$ and all 85 elements of the domain

$$\left\{\frac{0}{84}, \frac{1}{84}, \ldots, \frac{84}{84}\right\}$$

have orders 6, 3, 2, or 1. Since $\gcd(84, 5^3 - 1) = 4$, $\gcd(84, 5^2 - 1) = 12$, and $\gcd(84, 5^1 - 1) = 4$, $S_{84}^{(5)}$ has 5 points of period 3 or 1, 13 points of period 2 or 1, and 5 points of period 1. That yields five fixed points (five 1-cycles), $13 - 5 = 8$ points of period 2, (four 2-cycles), zero $5 - 5 = 0$ points of period 3, (no 3-cycles), and $85 - 5 - 8 - 0 = 72$ points of period 6 (twelve 6-cycles).

This gives a new and interesting interpretation of long division. Any sixth grader who uses long division to find the decimal expansion of the fraction $13/59$ is unwittingly finding the itinerary of the 14th card in a deck of size 60 under a 10-handed out shuffle!

EXAMPLE 8. *Find a necessary and sufficient condition on r to guarantee that $S_r^{(5)}$ contains a 6-cycle.*

In order to contain an orbit of length 6, the domain A_r must contain a rational number that has a base-5 expansion that consists of a repeating block of length 6. That is, it can be expressed in lowest terms with a denominator that divides $5^6 - 1 = 2^3 \cdot 3^2 \cdot 7 \cdot 31$. But those that divide $5^3 - 1 = 2^2 \cdot 31$, $5^2 - 1 = 2^3 \cdot 3$, or $5^1 - 1 = 2^2$ fall in 3-cycles, 2-cycles, or 1-cycles, respectively. By checking all possible factors of $5^6 - 1$ and ruling out those that are also factors of $5^3 - 1$, $5^2 - 1$, or $5^1 - 1$, we find that $S_r^{(5)}$ contains a 6-cycle if and only if 7, 9, 93, or 248 divide r. In addition, we can say that $|S_r^{(5)}| = 6$ if and only if at least one of those four conditions is satisfied and r divides $5^6 - 1$.

Riffle shuffles

A 2-*handed riffle shuffle* is the most common shuffle used to mix cards in a deck. It is performed by dividing the deck into two packets of consecutive cards (not necessarily equal in size) and interlacing the cards of the two packets together in some manner (not necessarily perfectly). The important property that makes a shuffle a riffle shuffle is that if card x is above card y and they are in the same packet before the shuffle, then x is also above y after the shuffle. So, in this sense, a riffle shuffle is an order-preserving shuffle. It preserves the order of the cards that were in the same packet before the shuffle. Of course, *k-handed riffles shuffles* can be defined analogously. Clearly, the k-handed perfect shuffles are all examples of k-handed riffle shuffles.

We mentioned earlier that perfect shuffles enjoy the property that no two distinct cards in a deck can have the same itineraries. Some less-than-perfect riffle shuffles enjoy this property also. We say that a riffle shuffle is *good* if every card in the deck has a distinct itinerary under the shuffle. A riffle shuffle that is not good is *poor*.

Suppose x and y are two cards in a deck that have identical itineraries under a poor riffle shuffle. If x and y are not adjacent, then, by the fact that the cards within a packet are consecutive and by the order-preserving property of riffle shuffles, all

cards between x and y must have the same itinerary as x and y. In addition, because of the order preserving property for cards within a packet, these cards remain in the same order as they were originally found in the deck. It is as though the cards in these blocks are glued together and act as a single card in a smaller deck under a good riffle shuffle.

Next we show that for each good k-handed riffle shuffle, there is a domain (subset of the unit interval [0, 1]) on which $S^{(k)}$ describes the shuffle and in which the cards in packet j have positions in $[j/k, (j+1)/k]$.

Our approach is straightforward and has a dynamical-systems flavor. We number the packets from 0 through $k-1$. By tracking the itinerary, we can locate the rational number needed to describe that card's position in the deck, as the following examples illustrate. We use the notation $\langle a_1 a_2 \ldots a_m \rangle \langle b_1 b_2 \ldots b_n \rangle$ to indicate that cards numbered $a_1, a_2, \ldots, a_m$ are in the top packet in that order and cards $b_1, b_2, \ldots, b_n$ are in the bottom packet in order.

EXAMPLE 9. *Find the domain that, together with S describes the following 2-handed riffle shuffle on 8 cards. Number the 8 cards 1 through 8 beginning at the top. Split the deck into two packets of 4 cards each and interlace them as follows:* $\langle 1, 2, 3, 4 \rangle \langle 5, 6, 7, 8 \rangle \to \langle 5, 1, 2, 6 \rangle \langle 3, 7, 4, 8 \rangle$.

Note that this is not a perfect shuffle because cards 1 and 2 are adjacent both before and after the shuffle, but it is a 2-handed riffle shuffle (order preserving). Note also that the cards numbered 1 and 2 form the only adjacent pair that is not separated by the shuffle, but the locations of the two cards are changed by the shuffle, so another application of the shuffle will separate that pair also. Thus, this is a good riffle shuffle.

We can write this shuffle as a product of disjoint cycles $(1, 2, 3, 5)(4, 7, 6)(8)$. The top packet is packet 0, and the bottom packet 1, so we get the 1st card's itinerary to be $0, 0, 0, 1, \ldots$ since the top three cards begin in packet 0 and the 5th card begins in packet 1. After that the itinerary repeats. The binary expansion for the rational-number position of the 1st card must, therefore, be $.\overline{0001}$, so the 1st card is in position $1/15$. It follows that the 2nd card must be in position $.\overline{0010} = 2/15$, the 3rd card must be in position $.\overline{0100} = 4/15$, and the 5th card must be in position $.\overline{1000} = 8/15$. Similarly, the itinerary of the 4th card is $0, 1, 1, \ldots$, so its position must be $.\overline{011} = 3/7$, placing the 7th and 6th cards in positions $6/7$ and $5/7$ respectively. The 8th card is fixed in packet 1, so its position is $.\overline{1} = 1$. Therefore, the domain on which S describes this riffle shuffle is

$$\left\{ \frac{1}{15}, \frac{2}{15}, \frac{4}{15}, \frac{3}{7}, \frac{8}{15}, \frac{5}{7}, \frac{6}{7}, 1 \right\}.$$

Though this is only an example, it is clear that the method of using the itinerary to determine the position can be applied to any good riffle shuffle.

Just by finding a common denominator, we see that any good riffle shuffle is part of a larger perfect shuffle. That is, we can expand the domain (add cards to the deck) in such a way as to make the shuffle on the larger deck a perfect shuffle without interfering with the positions of the cards in the original deck (card numbers will change, but not the rational-number positions). In this example, since $\text{lcm}(15, 7) = 105$, we see that this shuffle is a subshuffle of the perfect out shuffle on 106 cards. This is a Cayley-like result [**11**]. Cayley's theorem states that every finite group is a subgroup of a permutation group.

THEOREM 6. *A riffle shuffle can be embedded into a perfect shuffle if and only if it is good.*

We can clearly see the problem this model has dealing with poor riffle shuffles. Since the itinerary determines the position in the model, identical itineraries imply identical positions. On the other hand, all itineraries are distinct under a good riffle shuffle, so each card has its own unique position. If one allows the cards in these blocks to be identified (glued together) and thought of as a single card in a smaller deck, then the shuffle becomes good and the model can be used to describe the shuffle.

EXAMPLE 10. *Use the methods of this section to find the domain for the* 4*-handed perfect shuffle on* 8 *cards that corresponds to the permutation* [1, 3, 2, 0] *of packet numbers from Example 6.*

This shuffle can be described in more detail by $\langle 1,2\rangle\langle 3,4\rangle\langle 5,6\rangle\langle 7,8\rangle \to \langle 3,7\rangle\langle 5,1\rangle\langle 4,8\rangle\langle 6,2\rangle$. In disjoint cycles it becomes $(1,4,5,3)(2,8,6,7)$. The itinerary of the 1st card is 0, 1, 2, 1, ..., so the base-4 expansion of the position of the 1st card is $.\overline{0121} = (16+8+1)/(4^4-1) = 25/255$. Using the wrap-around effect of $S^{(4)}$ on the base-4 expansion, we get the other three positions in its orbit to be 100/255, 145/255, and 70/255. The itinerary of the 2nd card is 0, 3, 2, 3, ..., so the base-4 expansion of its position is $.\overline{0323} = 59/255$, and the remaining positions in its orbit are 236/255, 179/255, and 206/255. This gives us the domain we presented in Example 6.

The order-preserving (riffle) nature of the shuffle is also crucial for $S^{(k)}$ to be able to describe the shuffle. The graph of $S^{(k)}$ consists of line segments with positive slopes, so the order of the cards within a packet must be preserved by $S^{(k)}$. If the itineraries of the cards are given, the positions are determined, so their order within the packet is fixed. On the other hand, any permutation can be broken up into increasing subsequences (though some subsequences may contain as few as one term), so every permutation on a deck of cards can be thought of as a k-handed riffle shuffle for some k. In fact, as mentioned earlier, every permutation on n cards can be thought of as an n-handed perfect shuffle.

When performing a (k-handed) riffle shuffle, the first thing you do is break the deck up into packets. It seems natural, therefore, to think of the packets (and their packet numbers) as being fixed before the shuffle (permutation) is performed. At times, however, it is possible to accomplish the same permutation by redefining the packets. This is a way getting around the problem of a poor riffle shuffle.

EXAMPLE 11. *The shuffle*

$$\langle 1,2\rangle\langle 3,4\rangle\langle 5,6\rangle \to \langle 5,6\rangle\langle 1,2\rangle\langle 3,4\rangle$$

is a simple example of a poor 3*-handed riffle shuffle. The* 1*st and* 2*nd cards are never separated throughout their entire orbits. Their itineraries are identical. The same is true about the* 3*rd and* 4*th cards, and the* 5*th and* 6*th cards. But, if we redefine the packets*

$$\langle 1,2,3,4\rangle\langle 5\rangle\langle 6\rangle \to \langle 5,6,1,2\rangle\langle 3\rangle\langle 4\rangle$$

we can accomplish the same permutation as a good riffle shuffle without increasing the number of hands (packets) needed to perform it. It is described by $S^{(3)}$ *on the domain*

$$\left\{\frac{1}{26}, \frac{1}{13}, \frac{3}{26}, \frac{3}{13}, \frac{9}{26}, \frac{9}{13}\right\}.$$

More work is needed to determine the minimum number of hands k needed to describe a given permutation of cards as a good riffle shuffle or as a perfect shuffle.

REFERENCES

1. I. Adler, Make up your own card tricks, *J. Recreational Math.* **6** (Spring 1973), 87–91.
2. A. Aho and J.D. Ullman, Dynamic memories with rapid random and sequential access, *IEEE Trans. on Computers* **C-20** (Mar. 1974), 272–276.
3. Anonymous, *The Whole Art and Mystery of Modern Gaming Fully Expos'd and Detected*, J. Roberts, London, 1726.
4. I. C. Bivens, ed., Student research projects, *College Math. J.* **22** (Mar. 1991), 144–147.
5. P. Y. Chen, D. H. Lawrie, P. C. Yew, and D. A. Padua, Interconnection networks using shuffles, *Computer* (Dec. 1981), 55–64.
6. R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd ed., Addison-Wesley, Reading, MA, 1989.
7. P. Diaconis, R. L. Graham, and W. M. Kantor, The mathematics of perfect shuffles, *Advances in Applied Mathematics* **4** (1983), 175–193.
8. M. Gardner, Mathematical games: Can the shuffling of cards (and other apparently random events) be reversed?, *Scientific American* **215** (Oct. 1966), 114–117.
9. M. Gardner and C. A. McMahan, Riffling casino checks, this MAGAZINE **50** (Jan. 1977), 38–41.
10. M. Gardner, *Mathematical Carnival*, MAA, Washington DC, 1989.
11. I. N. Herstein, *Topics in Algebra*, 2nd ed. Wiley, New York, 1975.
12. D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, Cambridge, 1995.
13. E. Marlow, *The Faro Shuffle*, Ireland Magic Co., Chicago, 1958.
14. S. Medvedoff and K. Morrison, Groups of perfect shuffles, this MAGAZINE **60** (Feb. 1987), 3–15.
15. S. B. Morris, Faro shuffling and card placement, *J. Recreational Math.* **8** (1975), 1–7.
16. S. B. Morris, A. Valliere III, and R.A. Wisniewski, Processes for random and sequential accessing in dynamic memories, *IEEE Tran. on Computers* **C-28** (Mar. 1979), 225–237.
17. S. B. Morris, *Magic Tricks, Card Shuffling, and Dynamic Computer Memories*, MAA, Washington DC, 1998.
18. S. Ramnath and D. Scully, Moving card i to position j with perfect shuffles, this MAGAZINE **69** (Dec. 1996), 361–365.
19. J. W. Rosenthal, Card shuffling, this MAGAZINE **54** (Mar. 1981), 64–67.
20. H. S. Stone, Parallel processing with the perfect shuffle, *IEEE Trans. on Computers* **C-20** (Feb. 1971), 153–161.
21. ———, Dynamic memories with enhanced data access, *IEEE Trans. on Computers* **C-21** (Apr. 1972), 359–366.
22. ———, Dynamic memories with fast random and sequential access, *IEEE Trans. on Computers* **C-24** (Dec. 1975), 1167–1174.
23. P. Swinford, *Faro Fantasy*, Haley Press, Connersville, IN, 1968.
24. ———, *The Cyberdeck*, Haines' House of Cards, Cincinnati, 1986.

The editor wishes to thank Curtis Bennett of Loyola Marymount University for his expert advice in coordinating our theme of permutations for this issue. His careful reading of the relevant articles and notes helped greatly to improve our offerings.

Geometry of Generalized Complex Numbers

ANTHONY A. HARKIN
Division of Engineering and Applied Sciences, Harvard University
Cambridge, MA 02138
harkin@deas.harvard.edu

JOSEPH B. HARKIN
SUNY Brockport
Brockport, NY 14420
jharkin@brockport.edu

Alternative definitions of the imaginary unit i other than $i^2 = -1$ can give rise to interesting and useful complex number systems. The 16th-century Italian mathematicians G. Cardan (1501–1576) and R. Bombelli (1526–1572) are thought to be among the first to utilize the complex numbers we know today by calculating with a quantity whose square is -1. Since then, various people have modified the original definition of the product of complex numbers. The English geometer W. Clifford (1845–1879) developed the "double" complex numbers by requiring that $i^2 = 1$. Clifford's application of double numbers to mechanics has been supplemented by applications to noneuclidean geometries. The German geometer E. Study (1862–1930) added still another variant to the collection of complex products. The "dual" numbers arose from the convention that $i^2 = 0$ [**11**]. Well known in kinematics is the use of dual number methods for the analysis of spatial mechanisms, robotic control, and virtual reality [**4, 5, 10**].

The ordinary, dual, and double numbers are particular members of a two-parameter family of complex number systems often called binary numbers or generalized complex numbers, which are two-component numbers of the form

$$z = x + iy \quad (x, y \in \mathbb{R}) \quad \text{where} \quad i^2 = iq + p \quad (q, p \in \mathbb{R}).$$

It can be shown that generalized complex number systems are isomorphic (as rings) to the ordinary, dual, and double complex numbers when $p + q^2/4$ is negative, zero, and positive, respectively (FIGURE 1) [**11**].

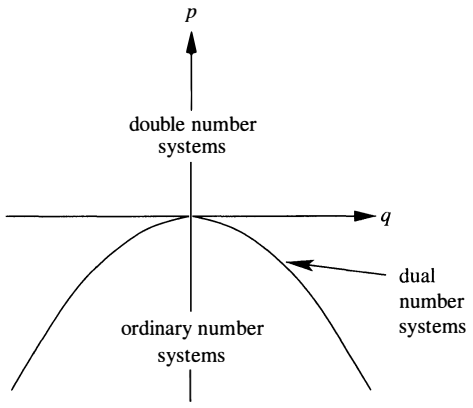


Figure 1 Generalized complex numbers are isomorphic (as rings) to the ordinary, dual, and double numbers.

In this article we study the geometry of a one-parameter family of generalized complex number systems in which $i^2 = p$, so that $q = 0$ and $-\infty < p < \infty$. Those who know the geometries of Laguerre and Minkowski will recognize that they arise naturally from generalized complex planes. Moreover, interrelations among the various complex products become obvious when the story of these planes unfolds.

Generalized complex multiplication

In what follows, we will let i denote a formal quantity, subject to the relation $i^2 = p$. Let $\mathbb{C}_p$ denote the system of numbers

$$\mathbb{C}_p = \{x + iy : x, y \in \mathbb{R}, \quad i^2 = p\}.$$

Addition and subtraction in this *p-complex plane* are defined, as usual, componentwise. Multiplication is also as we would expect, distributing multiplication over addition and using $i^2 = p$. Still, it will be helpful later on to introduce specific notation for this *p-multiplication*. So, for $z_1, z_2 \in \mathbb{C}_p$, we denote the product by

$$M^p(z_1, z_2) = (x_1x_2 + py_1y_2) + i(x_1y_2 + x_2y_1).$$

This definition yields the ordinary, Study, and Clifford products as p is equal to -1, 0, and 1.

Ordinary product: $(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2)$
Study product: $(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2) + i(x_1y_2 + y_1x_2)$
Clifford product: $(x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 + y_1y_2) + i(x_1y_2 + y_1x_2)$

We note that $\mathbb{C}_p$, under addition and p-multiplication, is a field only for $p < 0$. The *p-magnitude* of a generalized complex number $z = x + iy \in \mathbb{C}_p$ is defined to by the nonnegative real number

$$\|z\|_p = \sqrt{|M^p(z, \bar{z})|} = \sqrt{|x^2 - py^2|}$$

where an overbar denotes the usual complex conjugation.

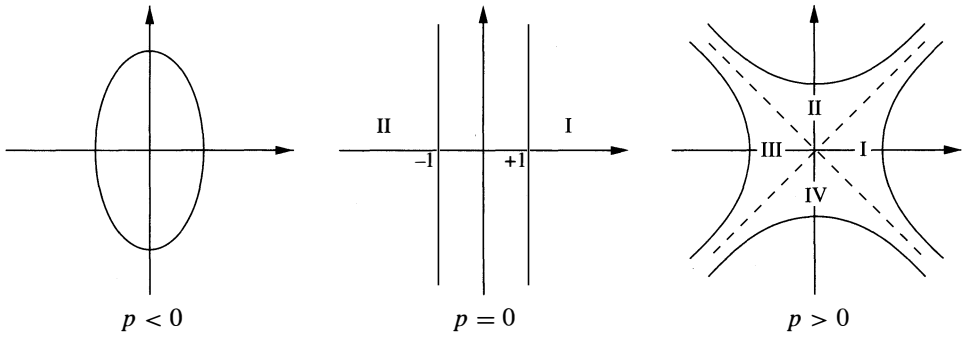


Figure 2 Unit circles in $\mathbb{C}_p$

Unit "circles" are defined by requiring $\|z\|_p = 1$ as in FIGURE 2. When $p < 0$ we obtain unit ellipses of the form $x^2 + |p|y^2 = 1$, and refer to $\mathbb{C}_p$ $(p < 0)$ as an elliptical complex number system. In the special case $p = -1$, the p-complex plane corresponds to the Euclidean plane. For $\mathbb{C}_0$, where $\|z\|_0^2 = x^2$, the unit circle is the

set of z where $x = \pm 1$. The space $\mathbb{C}_0$ is the parabolic complex number system whose p-complex plane corresponds to the Laguerre plane. The parabolic complex plane is naturally divided in half by the imaginary axis. The right-half plane of $\mathbb{C}_0$ will be referred to as branch I and the left half-plane branch II. Unit circles in $\mathbb{C}_p$ $(p > 0)$ are hyperbolas of the form $|x^2 - py^2| = 1$ whose asymptotes are $y = \pm x/\sqrt{p}$ (dashed lines in FIGURE 2). The spaces $\mathbb{C}_p$ $(p > 0)$ are referred to as hyperbolic complex number systems. For the special case $p = 1$, the p-complex plane is the well-known Minkowski plane. The asymptotes of the unit circles naturally separate the hyperbolic complex planes into four regions labeled branches I, II, III, and IV as shown in FIGURE 2.

Generalized trigonometry

Much of the geometrical insight into the ordinary complex plane is facilitated by the trigonometric form of a complex number. The same is true for generalized complex planes. Therefore, we now examine a trigonometry suitable for computations with generalized complex numbers.

Measures of angles The generalized complex number, $z = x + iy$, determines a ray $\overrightarrow{OT}$ as shown in FIGURE 3. Let the point N be the intersection of the ray $\overrightarrow{OT}$ and the unit circle in $\mathbb{C}_p$ (for now, suppose that z lies in the first hyperbolic branch). The p-argument of z, θ_p, is defined to be twice the Euclidean area of the shaded sector OMN determined by the arc MN and the radii $\overline{OM}$ and $\overline{ON}$. (The meanings of the words *sector*, *arc*, and *radius* should be clear from the picture.) Define the ratio $\sigma \equiv y/x$; then the geometric definition of angular measure yields formulae involving familiar inverse tangent functions:

$$\theta_p = \begin{cases} \frac{1}{\sqrt{|p|}} \tan^{-1}(\sigma\sqrt{|p|}), & p < 0 \\ \sigma, & p = 0 \\ \frac{1}{\sqrt{p}} \tanh^{-1}(\sigma\sqrt{p}), & p > 0 \text{ (branch I, III).} \end{cases}$$

The various factors of $\sqrt{p}$ simply account for the scaling of the unit ellipses and hyperbolas. Observe that angular measure can also be expressed succinctly as a power series:

$$\theta_p = \sum_{n=0}^{\infty} \frac{p^n}{2n+1} \sigma^{2n+1}, \qquad |\sigma|\sqrt{|p|} < 1.$$

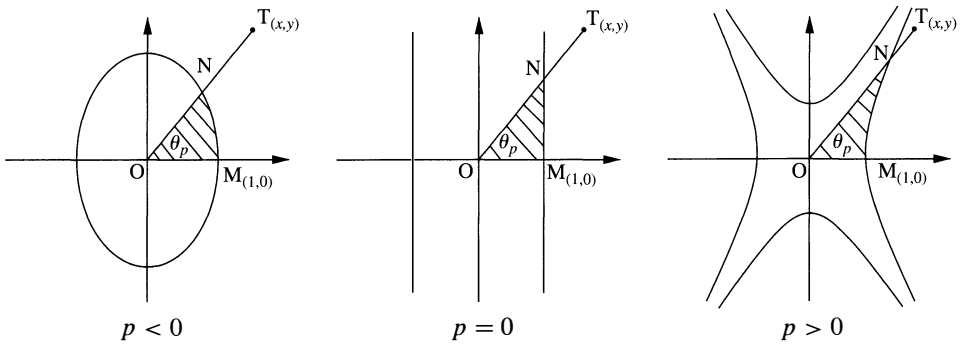
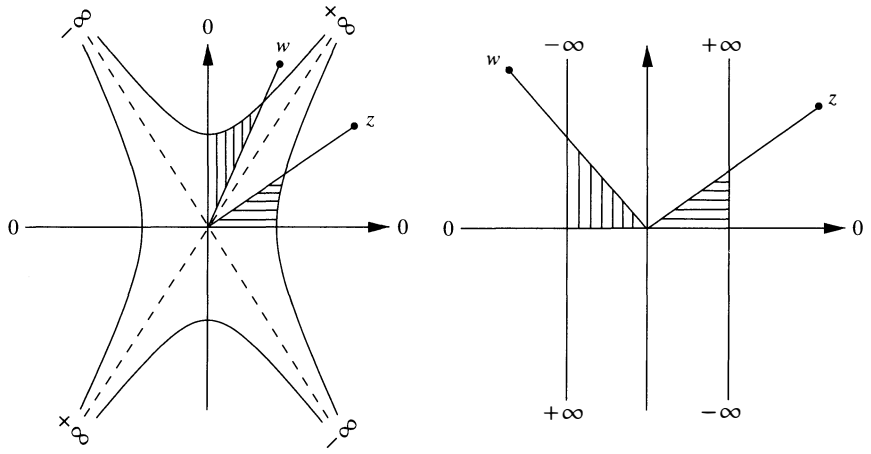


Figure 3 Elliptic, parabolic, and hyperbolic angles



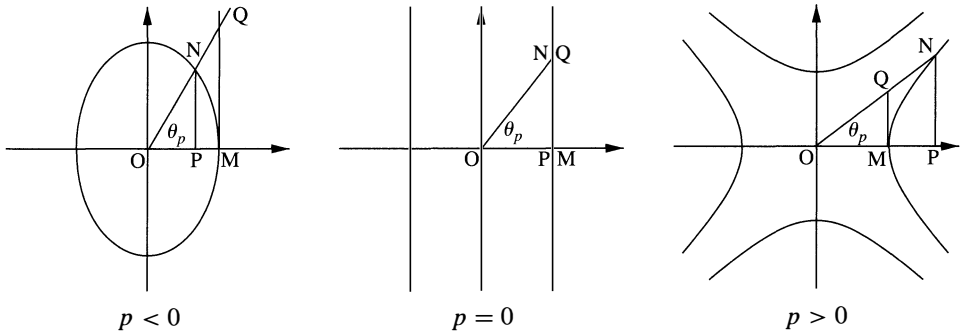
$p > 0$ $p = 0$

Figure 4 Angular measure extended to the whole hyperbolic and parabolic complex planes

The extension of angular measure throughout the entire parabolic and hyperbolic complex planes is suggested by FIGURE 4 and consists of some simple bookkeeping. Within each of the four branches of the hyperbolic complex plane (left FIGURE 4), angular measure is determined with respect to the half-axis that lies within the particular branch, and θ_p varies from $-\infty$ to $+\infty$ in the manner labeled on the asymptotes. Thus, for a hyperbolic complex number in branch II or IV, the angular measure is given by $\theta_p = (1/\sqrt{p}) \tanh^{-1}[1/(\sigma\sqrt{p})] = (1/\sqrt{p}) \coth^{-1}[\sigma\sqrt{p}]$. For example, in $\mathbb{C}_3$ the argument of $w = 2 + 5i$ is $\theta_3 = (1/\sqrt{3}) \tanh^{-1}[2/(5\sqrt{3})] \approx 0.1358$ which is measured from the positive imaginary axis.

In both of the branches of the parabolic complex plane (right FIGURE 4), angular measure is given by $\theta_p = \sigma = y/x$. When the real part of a parabolic complex number is negative, then its angular measure is referenced with respect to the negative part of the real axis and the unit circle in $\mathbb{C}_0$. Hence, the orientation of angles in branch II is opposite that in branch I, as indicated in the figure.

Trigonometric functions From the point N on the unit circle in $\mathbb{C}_p$ drop the perpendicular $\overline{\text{NP}}$ to the radius $\overline{\text{OM}}$ (FIGURE 5). At the point M draw a line tangent to the unit circle. Let Q be the point of intersection of the tangent and the line through $\overline{\text{ON}}$. The



$p < 0$ $p = 0$ $p > 0$

Figure 5 Geometric definitions of cosp, sinp, and tanp

lengths of the segments $\overline{\mathrm{OP}}$, $\overline{\mathrm{NP}}$, and $\overline{\mathrm{QM}}$ are defined to be the p-cosine (cosp), p-sine (sinp), and p-tangent (tanp), respectively. These geometric definitions give familiar expressions for the p-trigonometric functions:

$$\operatorname{cosp}\theta_p = \begin{cases} \cos(\theta_p\sqrt{|p|}), & p < 0 \\ 1, & p = 0 \text{ (branch I)} \\ \cosh(\theta_p\sqrt{p}), & p > 0 \text{ (branch I)} \end{cases}$$

and

$$\operatorname{sinp}\theta_p = \begin{cases} \frac{1}{\sqrt{|p|}}\sin(\theta_p\sqrt{|p|}), & p < 0 \\ \theta_p, & p = 0 \text{ (branch I)} \\ \frac{1}{\sqrt{p}}\sinh(\theta_p\sqrt{p}), & p > 0 \text{ (branch I).} \end{cases}$$

From the proportion QM/OM = NP/OP, we see that

$$\operatorname{tanp}\theta_p = \frac{\operatorname{sinp}\theta_p}{\operatorname{cosp}\theta_p}.$$

When $p = -1$ we find that the definitions reduce to the traditional circular trigonometric functions. Moreover, when $p = 1$ the familiar hyperbolic functions are recovered.

The parabolic and hyperbolic trigonometric functions on the other branches of their respective complex planes can be naturally defined in terms of the trigonometric functions on branch I. In the parabolic complex plane, define $\operatorname{cosp}_{\mathrm{II}}\theta_p = -\operatorname{cosp}_{\mathrm{I}}\theta_p$ and $\operatorname{sinp}_{\mathrm{II}}\theta_p = -\operatorname{sinp}_{\mathrm{I}}\theta_p$, where the subscripts are a convenient way to keep track of branches. In the hyperbolic complex planes, let

$$\operatorname{cosp}_{\mathrm{II}}\theta_p = \frac{i}{\sqrt{p}}\operatorname{cosp}_{\mathrm{I}}\theta_p, \quad \operatorname{cosp}_{\mathrm{III}}\theta_p = -\operatorname{cosp}_{\mathrm{I}}\theta_p, \quad \operatorname{cosp}_{\mathrm{IV}}\theta_p = -\frac{i}{\sqrt{p}}\operatorname{cosp}_{\mathrm{I}}\theta_p$$

and

$$\operatorname{sinp}_{\mathrm{II}}\theta_p = \frac{\sqrt{p}}{i}\operatorname{sinp}_{\mathrm{I}}\theta_p, \quad \operatorname{sinp}_{\mathrm{III}}\theta_p = -\operatorname{sinp}_{\mathrm{I}}\theta_p, \quad \operatorname{sinp}_{\mathrm{IV}}\theta_p = -\frac{\sqrt{p}}{i}\operatorname{sinp}_{\mathrm{I}}\theta_p.$$

The Maclaurin expansions for cosp and sinp (branch I) are given by

$$\operatorname{cosp}\theta_p = \sum_{n=0}^{\infty}\frac{p^n}{(2n)!}\theta_p^{2n}$$

and

$$\operatorname{sinp}\theta_p = \sum_{n=0}^{\infty}\frac{p^n}{(2n+1)!}\theta_p^{2n+1}.$$

A generalized Euler's formula is obtained by comparing these Maclaurin series with the formal power series expansion for $e^{i\theta_p}$, recalling that $i^2 = p$:

$$e^{i\theta_p} = \operatorname{cosp}\theta_p + i\ \operatorname{sinp}\theta_p.$$

Trigonometric identities The identity $|\operatorname{cosp}^2\theta_p - p\ \operatorname{sinp}^2\theta_p| = 1$ is evident, since $|x^2 - py^2| = 1$ is the form of a unit circle in $\mathbb{C}_p$. The next candidates for generalization are the addition laws for cosp and sinp. Let θ_p and ϕ_p be angular measures (in branch I

when $p = 0$ or $p > 0$). Then

$$\operatorname{cosp}(\theta_p \pm \phi_p) = \operatorname{cosp}\theta_p \operatorname{cosp}\phi_p \pm p \operatorname{sinp}\theta_p \operatorname{sinp}\phi_p$$
$$\operatorname{sinp}(\theta_p \pm \phi_p) = \operatorname{sinp}\theta_p \operatorname{cosp}\phi_p \pm \operatorname{cosp}\theta_p \operatorname{sinp}\phi_p.$$

Demonstrating these is straightforward, since the formulas for cosp and sinp reduce, in each case, to situations where addition laws are known.

We also observe that when $p < 0$ the p-trigonometric functions are periodic with period $2\pi/\sqrt{|p|}$. In particular, let θ_p be an angular measure with $p < 0$ and $k = 0, 1, 2, 3, \ldots$, then

$$\operatorname{cosp}(\theta_p + 2k\pi/\sqrt{|p|}) = \operatorname{cosp}\theta_p$$
$$\operatorname{sinp}(\theta_p + 2k\pi/\sqrt{|p|}) = \operatorname{sinp}\theta_p.$$

Interpretation of generalized complex multiplication

The trigonometric forms of the real and imaginary parts of $z = x + iy$ in $\mathbb{C}_p$ are

$$x = r_p \operatorname{cosp}\theta_p$$
$$y = r_p \operatorname{sinp}\theta_p,$$

where $r_p = \|z\|_p$ is the p-magnitude of z, and θ_p is the p-argument of z. Therefore, the trigonometric form of a generalized complex number is

$$z = x + iy = r_p(\operatorname{cosp}\theta_p + i \operatorname{sinp}\theta_p).$$

The geometric significance of p-multiplication now becomes clear. Suppose we have two complex numbers in $\mathbb{C}_p$, for example $z = \|z\|_p(\operatorname{cosp}\theta_p + i \operatorname{sinp}\theta_p)$ and $w = \|w\|_p(\operatorname{cosp}\phi_p + i \operatorname{sinp}\phi_p)$. Using the definition of p-multiplication and then recalling the addition laws for cosp and sinp, we obtain

$$M^p(z, w) = \|z\|_p\|w\|_p\big(\operatorname{cosp}(\theta_p + \phi_p) + i \operatorname{sinp}(\theta_p + \phi_p)\big).$$

Hence the p-length of the product is the product of the p-lengths and the p-argument of the product is the sum of the p-arguments. Therefore, the product of two generalized complex numbers can be obtained via rotation and amplification, and it should be emphasized that the rotation is along a generalized circle in $\mathbb{C}_p$. More specifically, suppose we wish to multiply z with w as in FIGURE 6. The product $M^p(z, w)$ is derived geometrically by rotating z through an angle $\phi_p = \arg w$ along the generalized circle of radius $\|z\|_p$, and then expanding by a factor of $\|w\|_p$.

In FIGURE 7 we present pictorially a few concrete examples of the geometry of generalized complex multiplication. In each plot, the two complex numbers labeled with circles are being multiplied to produce the third complex number marked with an asterisk. The unit circles are shown for reference.

For complex products in the hyperbolic and parabolic complex planes, evaluating cosp and sinp requires keeping track of the branch into which the product falls. To see an example of how this can be done, we'll examine the multiplication of the two dual numbers in the $p = 0$ case of FIGURE 7. In that case, the definition of p-multiplication gives,

$$M^0(2 + 3i, -1 + i) = -2 - i.$$

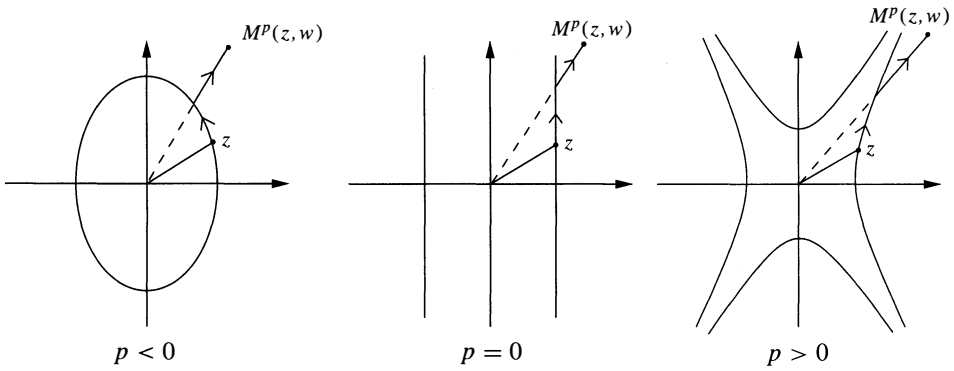


Figure 6 Multiplication is accomplished by rotation and amplification

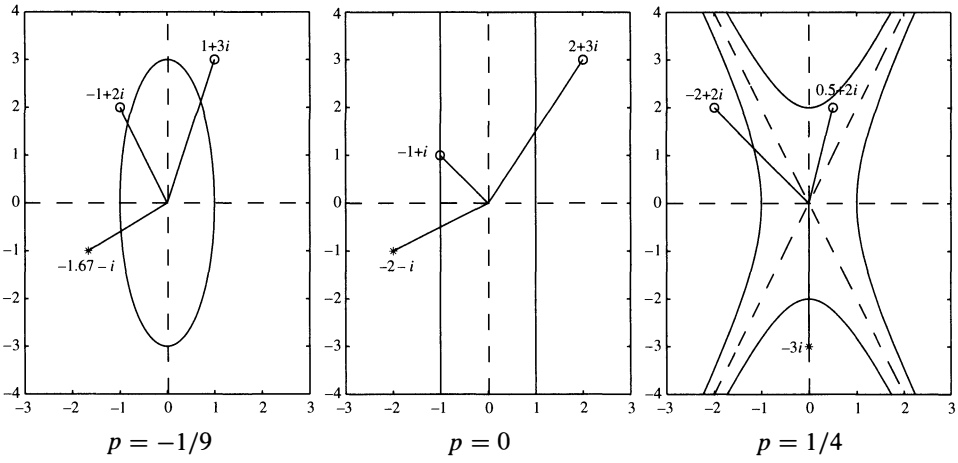


Figure 7 Geometric illustration of generalized complex multiplication

Alternatively, we can multiply the trigonometric forms of the numbers. The modulus of $2 + 3i$ in $\mathbb{C}_0$ is $\|2 + 3i\|_0 = 2$ and the argument is $\theta_p = 3/2$, which by definition of θ_p is twice the area of the triangle bounded by the real axis, the unit circle, and the ray connecting the origin to $2 + 3i$. So, the trigonometric form is $2 + 3i = 2(\cos\text{p}_{\text{I}} \frac{3}{2} + i \sin\text{p}_{\text{I}} \frac{3}{2})$. Similarly, $-1 + i = 1(\cos\text{p}_{\text{II}}(-1) + i \sin\text{p}_{\text{II}}(-1))$, which we rewrite as $(-\cos\text{p}_{\text{I}}(-1) - i \sin\text{p}_{\text{I}}(-1))$. We leave it to the reader to verify that multiplying the numbers in these yields $-2 - i$ as the product.

Generalized rotations and special relativity As seen in the previous section, the generalized complex product typically involves both an expansion (or contraction) and a generalized rotation. In the specific case where $\|w\|_p = 1$, the generalized complex product, $M^p(z, w)$, represents a pure rotation of z in $\mathbb{C}_p$. A pure rotation in $\mathbb{C}_p$ can be thought of as motion of the point z restricted to the generalized circle with radius $\|z\|_p$.

Generalized rotations can be applied to the theory of special relativity. In two-dimensional special relativity, an event that occurs at time t and at a space coordinate x is denoted by the spacetime point (t, x). Consider the generalized complex number $z = t + ix$ to be the spacetime coordinate of an event in $\mathbb{C}_p$, where $p = 1/c^2$ ($c \equiv$ speed of light). Let V represent the velocity of a coordinate frame (t', x') in uniform motion with respect to the inertial coordinate frame (t, x) of the event. If we now let

$w = 1 - iV$, then the pure rotation represented by the product

$$M^p\left(z, \frac{w}{\|w\|_p}\right) = \left[\frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}}\right] + i\left[\frac{x - Vt}{\sqrt{1 - V^2/c^2}}\right] = t' + ix'$$

yields the Lorentz coordinate transformations of two-dimensional special relativity. Hence, the Lorentz transformations of two-dimensional special relativity are simply rotations in the hyperbolic complex plane. In fact, if a velocity parameter, ϕ_p, is defined by $\tanp \phi_p = -V$, then the Lorentz transformation can be succinctly expressed as multiplication of $z = t + ix$ by $e^{i\phi_p}$. An article by Fjelstad [**6**] further explores the connection of hyperbolic complex numbers to special relativity.

Powers and roots of generalized complex numbers

Generalized De Moivre formulas allow us to compute powers and roots of complex numbers in $\mathbb{C}_p$.

THEOREM 1. (POWERS OF GENERALIZED COMPLEX NUMBERS) *For $z \in \mathbb{C}_p$ and n a positive integer,*

$$z^n = \left[r_p(\cosp \theta_p + i\ \sinp \theta_p)\right]^n = r_p^n\big(\cosp(n\theta_p) + i\ \sinp(n\theta_p)\big).$$

The proof is left to the reader, as it follows easily by induction, using the laws for p-multiplication and addition.

We will state two theorems concerning the computation of nth roots of complex numbers. The first theorem applies to complex numbers in $\mathbb{C}_p$ ($p < 0$) and the second theorem covers the cases when $p > 0$ and $p = 0$. For $p < 0$, the trigonometric functions are $2\pi/\sqrt{|p|}$ -periodic, leading to the following theorem on the extraction of nth roots of elliptical complex numbers.

THEOREM 2. (ROOTS OF ELLIPTICAL COMPLEX NUMBERS) *For z in $\mathbb{C}_p (p < 0)$ and n a positive integer,*

$$\begin{aligned} z^{\frac{1}{n}} &= \left[r_p(\cosp \theta_p + i\ \sinp \theta_p)\right]^{\frac{1}{n}} \\ &= r_p^{\frac{1}{n}}\left(\cosp\left(\frac{\theta_p + 2k\pi/\sqrt{|p|}}{n}\right) + i\ \sinp\left(\frac{\theta_p + 2k\pi/\sqrt{|p|}}{n}\right)\right), \end{aligned}$$

where $k = 0, 1, 2, 3, \ldots, (n-1)$.

Proof. An application of the generalized De Moivre formula for powers yields

$$\begin{aligned} &\left[r_p^{\frac{1}{n}}\left(\cosp\left(\frac{\theta_p + 2k\pi/\sqrt{|p|}}{n}\right) + i\ \sinp\left(\frac{\theta_p + 2k\pi/\sqrt{|p|}}{n}\right)\right)\right]^n \\ &\quad = r_p\big(\cosp(\theta_p + 2k\pi/\sqrt{|p|}) + i\ \sinp(\theta_p + 2k\pi/\sqrt{|p|})\big) \\ &\quad = r_p(\cosp \theta_p + i\ \sinp \theta_p). \end{aligned}$$

■

FIGURE 8(a) displays the three elliptical ($p = -1/4$) cube roots of $2 + 5i$. The roots determine three sectors of equal area in the root-ellipse: $x^2 + y^2/4 = \|2 + 5i\|_{-1/4}^{2/3} \approx 2.172$. In general, each set of n complex roots on a root-ellipse in $\mathbb{C}_p$ ($p < 0$) partitions the root-ellipse into n sectors of equal area. A comparison of elliptical ($p = -6$) and

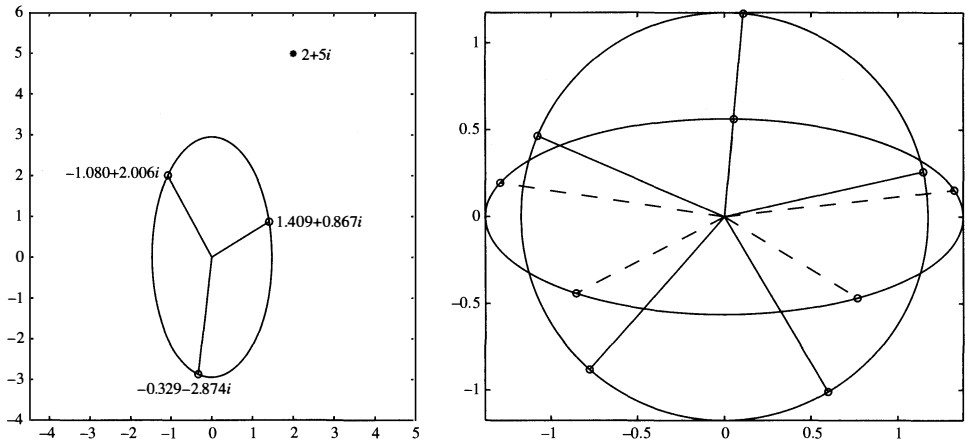


Figure 8 Illustrations of the elliptical De Moivre theorem

circular ($p = -1$) roots is shown in FIGURE 8(b), which shows five elliptical fifth roots of $1 + 2i$ and the five circular fifth roots of $1 + 2i$.

The lack of periodicity in the trigonometric functions when $p > 0$ and when $p = 0$ permits a slightly modified De Moivre theorem for the computation of nth roots of parabolic and hyperbolic complex numbers.

THEOREM 3. (ROOTS OF PARABOLIC AND HYPERBOLIC COMPLEX NUMBERS) *For $z \in \mathbb{C}_p$ ($p > 0$ or $p = 0$) and n a positive integer,*

$$z^{\frac{1}{n}} = \left[r_p(\operatorname{cosp}\theta_p + i\ \operatorname{sinp}\theta_p)\right]^{\frac{1}{n}} = r_p^{\frac{1}{n}}\left(\operatorname{cosp}\left(\frac{\theta_p}{n}\right) + i\ \operatorname{sinp}\left(\frac{\theta_p}{n}\right)\right).$$

In the next four examples we illustrate the disparate outcomes that result from the lack of periodicity in the parabolic ($p = 0$) and hyperbolic ($p > 0$) trigonometric functions. FIGURE 9(a) displays the two square roots of $2 + 3i$ in $\mathbb{C}_0$. The two vertical lines drawn are actually the parabolic circle whose radius is given by $\|2 + 3i\|_0^{1/2}$. In FIGURE 9(b), we find only a single cube root of $3 + 2i$ in $\mathbb{C}_1$. There are no others. This cube root lies in branch I on the root-hyperbola given by $|x^2 - y^2| = \|3 + 2i\|_1^{2/3} \approx 1.71$. Since the square of any hyperbolic or parabolic complex number lands in branch I, then the cube of the number winds up back in its original branch. In light of this observation, we see that each hyperbolic and each parabolic complex number has exactly one cube root. Moreover, when n is an odd positive integer, every hyperbolic and parabolic complex number has exactly one nth root.

In FIGURE 9(c), we illustrate the existence of four fourth roots of $4 - 3i$ when $p = 1$. And finally, in FIGURE 9(d), we display the four distinct square roots of $2 + 2\sqrt{3}\,i$ with $p = 1/4$. Note that $\|2 + 2\sqrt{3}\,i\|_{1/4} = 1$ implies that all of the roots lie on the unit hyperbola in $\mathbb{C}_{1/4}$. When n is an even positive integer then hyperbolic complex numbers in branch I have exactly four nth roots (one in each branch), and hyperbolic complex numbers in the other branches have no nth roots. Similarly, every parabolic complex number in branch I has two nth roots when n is even, and parabolic complex numbers in branch II have no even nth roots. The total number of nth roots of a generalized complex number is summarized in TABLE 1. The situation becomes more complicated when looking for solutions of polynomials that are defined over parabolic or hyperbolic complex number systems [**1**].

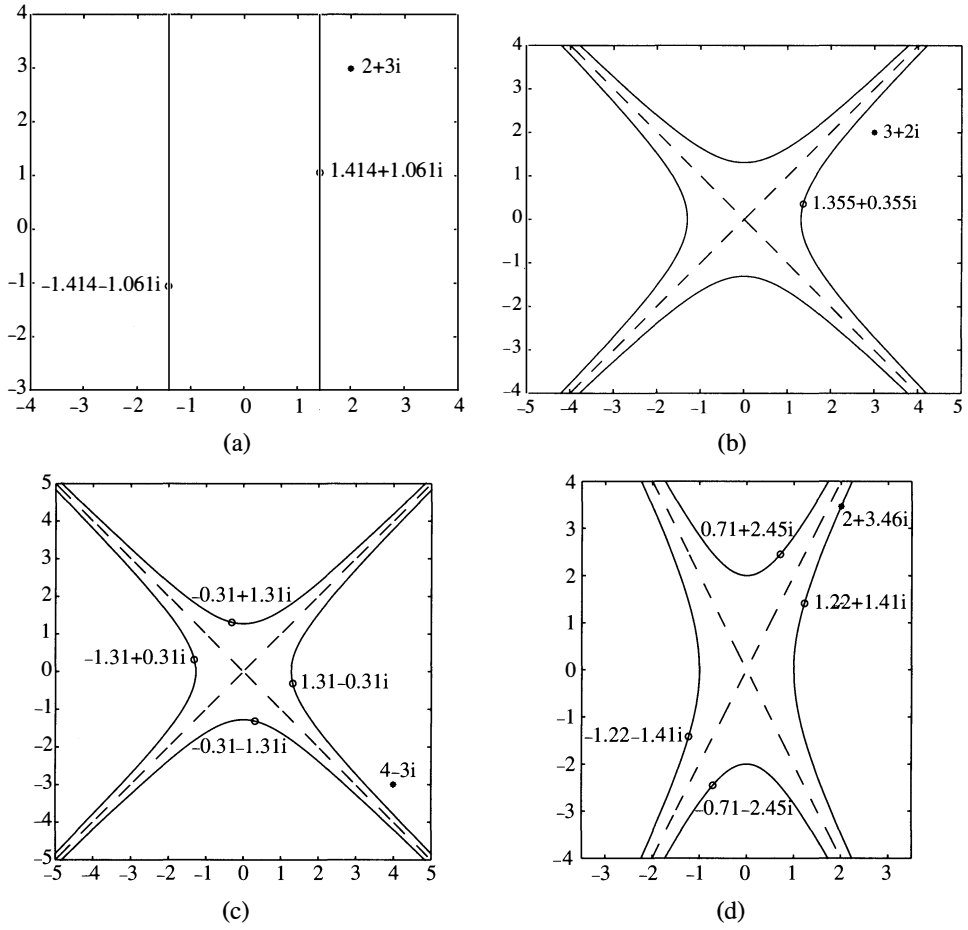


Figure 9 Illustrations of the generalized De Moivre theorem for parabolic and hyperbolic roots

TABLE 1: Number of nth Roots of $z \in \mathbb{C}_p$

$p < 0$		n roots
$p = 0$	$z \in$ branch I	$z \in$ branch II
n even	2 nth roots	0 nth roots
n odd	1 nth root	1 nth root
$p > 0$	$z \in$ branch I	$z \in$ branch II, III, or IV
n even	4 nth roots	0 nth roots
n odd	1 nth root	1 nth root

Functions of a generalized complex variable

At this point, one might wonder about a generalization of the theory of complex analytic functions. We make a few brief observations about analyticity in $\mathbb{C}_p$.

The p-derivative of a function f of a generalized complex variable $z \in \mathbb{C}_p$ is defined, as usual, by

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z},$$

provided this limit exists independent of the manner in which $\Delta z \to 0$, excluding approaches on which the quotient is not defined. Recall that a function $f = u + iv$ of an ordinary complex variable $z = x + iy$ is analytic on a region, D, if and only if it satisfies

$$i\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} \quad \text{on } D.$$

Suppose that f is a function of a generalized complex variable, then we say that $f = u + iv$ is p-analytic when its real and imaginary parts satisfy generalized Cauchy-Riemann equations,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = p\frac{\partial v}{\partial x}.$$

Moreover, if these partial derivatives are continuous, then the real and imaginary parts of f are order-p harmonic:

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{p}\frac{\partial^2 u}{\partial y^2} = 0, \qquad \frac{\partial^2 v}{\partial x^2} - \frac{1}{p}\frac{\partial^2 v}{\partial y^2} = 0.$$

In the special case $p = -1$, the real and imaginary parts of f satisfy Laplace's equation, and when $p = 1$, the real and imaginary parts of f satisfy a wave equation. Study referred to analytic functions of a dual variable ($p = 0$) as *synectic* functions.

As an example, consider the exponential e^z in $\mathbb{C}_p$:

$$e^z = e^{x+iy} = e^x e^{iy} = e^x(\operatorname{cosp} y + i \operatorname{sinp} y) = u + iv.$$

Since the real and imaginary parts are

$$u = e^x \operatorname{cosp} y \qquad v = e^x \operatorname{sinp} y$$

and since the derivatives of cosp and sinp are given by

$$\frac{d}{dy}(\operatorname{cosp} y) = p \operatorname{sinp} y \qquad \frac{d}{dy}(\operatorname{sinp} y) = \operatorname{cosp} y,$$

it can be verified that the Cauchy-Riemann equations hold. Thus e^z is p-analytic.

Integration is defined on rectifiable curves. When $f(z)$ is differentiable and C is a closed curve, it can be shown that

$$\oint_C f(z)\,dz = 0$$

for all spaces $\mathbb{C}_p$. However, Cauchy's integral formula does not hold in the parabolic or hyperbolic complex planes, as discussed by Deakin [**3**].

Acknowledgments. The authors are grateful for the valuable suggestions of the referees.

REFERENCES

1. H. H. Cheng and S. Thompson, *Proceedings of the 1996 ASME Design Engineering Technical Conference and Computers in Engineering Conference*, Irvine, CA, 1996.

2. W. K. Clifford, *Mathematical Papers* (ed. R. Tucker), Chelsea Pub. Co., Bronx, NY, 1968.
3. M. A. B. Deakin, this MAGAZINE **39**:4 (1966), 215–219.
4. F. M. Dimentberg, *The Screw Calculus and its Applications in Mechanics*, Izdat. "Nauka", Moscow, USSR, 1965.
5. I. S. Fischer and A. S. Fischer, *Dual-Number Methods in Kinematics, Statics and Dynamics*, CRC Press, 1998.
6. P. Fjelstad, *Am. J. Phys.* **54**:5 (1986), 416–422.
7. L. Hahn, *Complex Numbers and Geometry*, Math. Assoc. of America, Washington DC, 1994.
8. T. Needham, *Visual Complex Analysis*, Clarendon Press, Oxford, 1997.
9. H. Schwerdtfeger, *Geometry of Complex Numbers*, University of Toronto Press, Toronto, 1962.
10. E. Study, *Geometrie der Dynamen*, Leipzig, 1903.
11. I.M. Yaglom, *Complex Numbers in Geometry*, Academic Press, New York, 1968.

Permutation Notations

Permutations can be thought of as shuffles or rearrangements, but they are most easily described as one-to-one functions from a set onto itself. For example, take your two hands and match them as follows: pinkies to pinkies, fourth fingers to thumbs, index fingers to middle fingers. Numbering the fingers one through five the same way on each hand (and cheating a little by calling the thumb a finger), we get a function:

$$f(1) = 4, \quad f(2) = 3, \quad f(3) = 2, \quad f(4) = 1, \quad f(5) = 5.$$

This is all one needs for certain applications. But authors Deutsch, Johnson, and Thanatipanonda use *line notation*, which is simply a list of the values of the function in order: 43215. This works well for permutations of small sets, but author Scully uses line notation where some elements have names like k and $k - 1$. For clarity, brackets and parentheses can be used: [4, 3, 2, 1, 5].

A longer version of line notation uses two lines in a before-and-after display, like this:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 1 & 5 \end{bmatrix}.$$

Both line notation and function notation obscure some valuable information about the cycles that occur upon repeated applications of a permutation. This is apparent in *cycle* notation. Our finger permutation would be written as (14)(23)(5) or more simply (14)(23). This notation is read as "1 goes to 4, which goes back to 1; 2 goes to 3, which goes back to 2; 5 goes to 5." When an element is omitted, it is understood to stay fixed.

Both the two-line notation and cycle notation were introduced by Cauchy in 1815. You can read a translated excerpt from his paper in *The History of Mathematics: A Reader*, edited by John Fauvel and Jeremy Gray, Macmillan Press in association with The Open University, 1987, pp. 506–507.

Incidentally, the finger permutation described above is the starting point for "compound eensy-weensy spider."

NOTES

Create Your Own Permutation Statistics

EMERIC DEUTSCH
Polytechnic University
Brooklyn, NY 11201
deutsch@duke.poly.edu

WARREN P. JOHNSON
Bates College
Lewiston, ME 04240
wjohnson@bates.edu

A *permutation* of length n is, for us, a list of the numbers $\{1, 2, \ldots, n\}$ in some order, so that (for example) 27163854 is a permutation of length 8.* An *inversion* in a permutation is any pair of numbers, not necessarily consecutive, that are "out of order" in the sense that the larger number occurs before the smaller one; thus 7 and 5 are an inversion in 27163854, and you can check that there are eleven others. We say that the *inversion number* of 27163854 is 12.

The inversion number occurs at least implicitly in the definition of a determinant as a sum over permutations—inversions by which a permutation differs from increasing order are equivalent to row exchanges needed to change a permutation matrix into the identity—and the concept dates back to Cramer's pioneering work on determinants in 1750 [**4**]. (The term *inversions* seems to date from expository work on determinants by Gergonne in 1813 [**9**] and Garnier [**8**] the following year; Cramer called them *dérangements*. See Muir [**15**] for more details.) It is probably the second best known example of what is called a *permutation statistic*, which is just a function (typically nonnegative and integer-valued) defined on permutations. The best known example is the number of cycles, and Thanatipanonda's note in this issue discusses what may be the third best known example.

There are some wonderful results about inversions [**10, 17, 18**], but it is not our purpose to discuss them here. Rather we want to point out an easy way to construct another class of permutation statistics: Given a property $\mathcal{P}$ that a permutation π may or may not have, we can record the length of the longest initial segment of π that has property $\mathcal{P}$. We will prove a simple result about this type of statistic, and look at some nice examples.

We will try to maintain a distinction between permutations and sequences. By a *sequence* of length n we shall mean a list of n distinct positive integers, which may or may not be $\{1, 2, \ldots, n\}$, in some order. To each sequence of length n we associate a permutation of length n in the natural way, by relabeling the smallest number in the sequence as 1, the second smallest number as 2, and so on, relabeling the largest number as n. We call this permutation the *reduction* of the sequence; for example, the reduction of 428396 is 315264. We also define the *truncation* of a sequence σ to be σ with its last element deleted; for example, the truncation of 2371 is 237. We will focus on properties $\mathcal{P}$ satisfying two conditions:

*EDITOR'S NOTE: Information about the various notations for permutations appears on page 129.

(i) σ has property $\mathcal{P}$ if and only if the reduction of σ has property $\mathcal{P}$;

(ii) if σ has property $\mathcal{P}$, so does its truncation (we assume this always holds if σ has length 1).

Sometimes it will be convenient to identify a property with the sequences that possess it. Thus, rather than looking at the property that a sequence increases, we can consider the set of all increasing sequences, which we will call $\mathcal{P}_1$; or, instead of considering the property of beginning with an even number, we can look at the set $\mathcal{Q}_1$ of all sequences that begin with an even number. Note that $\mathcal{Q}_1$ fails condition (i) (if $\sigma = 428396$, for example), while $\mathcal{P}_1$ satisfies (i) and (ii).

Here are some more examples: let $\mathcal{Q}_2$ be the set of all sequences of length at most four, let $\mathcal{Q}_3$ be the set of all sequences of length exactly four, let $\mathcal{Q}_4$ be the set of all sequences whose smallest element is the first, and let $\mathcal{Q}_5$ be the set of all sequences whose smallest element is the last. Then $\mathcal{Q}_3$ and $\mathcal{Q}_5$ both fail condition (ii) (if $\sigma = 2371$, for example), but $\mathcal{Q}_2$ and $\mathcal{Q}_4$ satisfy both (i) and (ii).

The main examples We are primarily interested in six further examples that satisfy (i) and (ii). These come naturally in three pairs, though there are other connections among them. One of our first pair of examples is fairly well known: let $\mathcal{P}_2$ denote the set of *up-down* sequences, which are sequences $a_1a_2 \ldots a_n$ that alternately rise and fall, starting with a rise (that is, $a_1 < a_2, a_2 > a_3, a_3 < a_4$, and so on); an example is 2956374. For a variation on this idea consider 5974286, which is an example of what we will call a *zigzig* sequence. These sequences have the property that the sequence of elements from the smallest to the last (286 in this case) is up-down, and if we cut this piece off then the remainder 5974 still has this property (4 is vacuously up-down), and continues to have it as we keep cutting pieces off at the right (597 is up-down). We denote the set of all zigzig sequences by $\mathcal{P}_3$.

Let $\mathcal{P}_4$ denote the set of *123-avoiding* sequences, which are sequences $a_1a_2 \ldots a_n$ that contain no increasing subsequences of length 3. Note that the elements of the subsequence need not be in consecutive positions: 3746 has no three consecutive elements increasing, but it is not 123-avoiding since it has the subsequence 346. Let $\mathcal{P}_5$ denote the set of *132-* and *231-avoiding* sequences, with the same interpretation as before; these are sequences with no interior local maxima, such as 963145. These two examples, and others that also satisfy (i) and (ii), were considered by Simion and Schmidt [**16**].

Our last pair of examples requires more discussion to explain the term *stack-sortable sequence*, but is pretty enough to be worth a little space. We will call $\mathcal{P}_6$ the set of all stack-sortable sequences, and $\mathcal{P}_7$ the set of all 2-stack-sortable sequences. Stack-sortable permutations were studied by Knuth [**14**], and the generalization to 2 stacks was made by West [**19**].

To see what these are, consider the permutations of {1, 2, 3}, namely 123, 132, 213, 231, 312, and 321. If possible, we want to put all of them in the order 123 with the aid of a "stack"; each element, starting with the first, is moved to the stack and then to the output. 123 is already in order (we could call it a 0-stack-sortable permutation, but this is synonymous with increasing), so we need only move each element in turn to the stack and then immediately to the output. As West pointed out, we can code the sorting by using a left parenthesis when an element is put on the stack, and a right parenthesis when it is taken off, so the sorting of 123 would be coded ()()(). 132 can be sorted by the moves with code ()((): 1 goes on the stack and then off, 3 goes on the stack, then 2 goes on top of 3, then 2 comes off, then 3 comes off. 213 can be sorted by (())(), 312 by (()()), and 321 by ((())). Since a number has to go on the stack before it can come off, these are *well-formed* sets of parentheses—at any point in the sequence

we have had at least as many left parentheses as right parentheses. To reconstruct the permutation from the code, pair each right parenthesis with the closest left parenthesis that precedes it and is not already in a pair, working from left to right. If the ith right parenthesis is paired with the jth left parenthesis then put element i in position j.

The exceptional permutation is 231, for the only way to get 3 at the end is by the sequence of moves ()(()), but this transforms it to 213 instead of 123. (Note also that the permutation we would recover from ()(()) as above is 132, not 231.) It follows that any sequence with a 231-type subsequence is not stack sortable. Conversely, it is not difficult to prove that a 231-avoiding sequence is stack sortable (use induction on the length of the sequence, or see [**14**, p. 533] or [**19**]).

The sorting algorithm can be described succinctly: if the stack is empty, put the next element in the sequence on the stack. Otherwise, compare the next element in the sequence to the element on top of the stack. If the top element on the stack is smaller, take it off; otherwise put the next element in the sequence on the stack. Since this procedure avoids putting an element on top of a smaller element on the stack, it is the unique way to sort a stack-sortable sequence. If we apply the algorithm to a non-stack-sortable sequence, then we won't succeed in sorting it, but we may get a sequence that is itself stack-sortable; for example, 231 is changed into 213, which is stack-sortable. West defined a 2-stack-sortable permutation to be one that can be sorted by two applications of the above sorting algorithm, and he proved that π is 2-stack-sortable if and only if it avoids the subsequence 2341, and also avoids 3241 unless it is part of 35241. Note that it would be possible to sort 2341 in two passes by changing it on the first pass to 1432, which is stack-sortable; still, 2341 is not said to be 2-stack-sortable since the sorting algorithm first changes it into 2314, and then into 2134.

The main result Let $\mathcal{S}_n$ denote the set of all permutations of $\{1, 2, \ldots, n\}$. For each $\pi \in \mathcal{S}_n$ we define $\ell_{\mathcal{P}}(\pi)$ to be the length of the longest initial segment of π belonging to $\mathcal{P}$. The reader can verify that for the permutation $27163854 \in \mathcal{S}_8$, the values of this statistic corresponding to the nine properties $\mathcal{Q}_2$, $\mathcal{Q}_4$, $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, $\mathcal{P}_4$, $\mathcal{P}_5$, $\mathcal{P}_6$, $\mathcal{P}_7$ are 4, 2, 2, 7, 7, 5, 2, 2, and 6, respectively. We will enumerate the $n!$ permutations of $\{1, 2, \ldots, n\}$ according to the values of the statistic $\ell_{\mathcal{P}}$. In other words, we want to determine the numbers

$$\alpha_{n,k} \stackrel{\text{def}}{=} |\{\pi \in \mathcal{S}_n \,:\, \ell_{\mathcal{P}}(\pi) = k\}| \quad (1 \le k \le n).$$

We will use the following terminology: given a property $\mathcal{P}$ satisfying (i) and (ii), call a permutation π *good* if it has $\mathcal{P}$, and let a_n be the number of good permutations of length n, so that $\alpha_{n,n} = a_n$. If π lacks $\mathcal{P}$ but the truncation of π has it, then we say π is *almost good*. We denote the number of almost good permutations of length n by b_n, so that $b_n = \alpha_{n,n-1}$ for $n \ge 2$, and we define $b_1 = 0$.

Suppose we have a good permutation π of $\{1, 2, \ldots, n-1\}$. If we put any j from $\{1, 2, \ldots, n\}$ at the end of π and relabel k as $k+1$ for each k with $j \le k \le n-1$, then we get either a good or an almost good permutation of $\{1, 2, \ldots, n\}$. Considering all the possible values of j, this implies

$$b_n + a_n = na_{n-1} \quad \text{if } n \ge 2. \tag{1}$$

On the other hand, for $k < n$ there is a simple connection between $\alpha_{n,k}$ and b_{k+1}. To construct a permutation of length n with $\ell_{\mathcal{P}} = k$ we can choose any $k+1$ elements from $\{1, 2, \ldots, n\}$ and arrange them in an almost good sequence, and then put the remaining $n-k-1$ elements on the end in any order. Considering all the possible

ways of doing this, we have

$$\alpha_{n,k} = \binom{n}{k+1} b_{k+1}(n-k-1)! = \frac{n!}{(k+1)!} b_{k+1}. \tag{2}$$

We might also think of (2) probabilistically: we must have

$$\frac{\alpha_{n,k}}{n!} = \frac{b_{k+1}}{(k+1)!}$$

since both sides equal the probability that a permutation of length n ceases to be good after exactly k elements. Combining (1) and (2) we get the formula

$$\alpha_{n,k} = \begin{cases} \frac{n!}{(k+1)!}[(k+1)a_k - a_{k+1}] & \text{if } 1 \le k \le n-1; \\ a_n & \text{if } k = n. \end{cases}$$

We could rewrite this as

$$\frac{\alpha_{n,k}}{n!} = \frac{a_k}{k!} - \frac{a_{k+1}}{(k+1)!} \quad \text{for } 1 \le k \le n-1,$$

telling us that the probability that a permutation of some length $n > k$ goes bad at element $k+1$ equals the probability that the first k elements are good minus the probability that the first $k+1$ elements are good. Thus $\alpha_{n,k}$ can be found whenever a_n is known. We conclude by giving a_n for each of $\mathcal{P}_1$–$\mathcal{P}_7$, along with some of the more interesting $\alpha_{n,k}$:

$\mathcal{P}_1$: $a_n = 1$, because there is only one permutation of length n with increasing entries. Therefore

$$\alpha_{n,k} = \begin{cases} \frac{n!k}{(k+1)!}, & \text{if } 1 \le k \le n-1; \\ 1, & \text{if } k = n. \end{cases}$$

This result is sometimes referred to as the enumeration of permutations according to the length of the first run.

$\mathcal{P}_2$: $a_n = E_n$, where $\tan x + \sec x = \sum_{n=0}^{\infty} E_n \frac{x^n}{n!}$. This was first proved by André [**1, 2**], and one can also find it elsewhere [**10**, p. 169; **12**; **11**, p. 377; **17**, p. 149; **18**, p. 74]. The quantities E_n are often called Euler numbers and can be found in various places in Euler's work. For example, the first fourteen of them are given on p. 416 of [**5**]; they are 1, 1, 1, 2, 5, 16, 61, 272, 1385, 7936, 50521, 353792, 2702765, 22368256. Forder [**7**] gives a simple proof that, starting at $n = 2$, the last digits repeat the pattern 1, 2, 5, 6.

$\mathcal{P}_3$: $a_n = E_{n+1}$, where E_n is the nth Euler number [**13**]. We leave the calculation of the $\alpha_{n,k}$ in this pair of examples as an exercise.

$\mathcal{P}_4$: $a_n = \frac{1}{n+1}\binom{2n}{n}$, which is a Catalan number [**16**; **18**, p. 260]. Hence, we have

$$\alpha_{n,k} = \begin{cases} \frac{n!}{(k+1)!}\binom{2k}{k-2}, & \text{if } 1 \le k \le n-1; \\ \frac{1}{n+1}\binom{2n}{n}, & \text{if } k = n. \end{cases}$$

$\mathcal{P}_5$: $a_n = 2^{n-1}$. The elements that precede 1 must be decreasing, and those that follow 1 must be increasing, so there is a one to one correspondence between these permutations and subsets of $\{2, 3, \ldots, n\}$. In this example one has the curious fact that $\alpha_{n,2} = \alpha_{n,3}$ for all $n \geq 4$. This happens exactly when $a_4 = 8a_3 - 12a_2$, so it also occurs with $\mathcal{P}_3$.

$\mathcal{P}_6$: $a_n = \frac{1}{n+1}\binom{2n}{n}$ again. That Catalan numbers count sequences of well-formed parentheses is a well-known fact also due to André [**3**] (or see [**19**] or [**18**, p. 261]). This is the source of the reflection principle in probability—see [**6**, p. 72].

It is easy to see that 231-avoiding permutations of $\{1, 2, \ldots, n\}$ are equinumerous with 132-avoiding permutations (read backwards) or with 213-avoiding permutations (subtract every number in the permutation from $n + 1$) or with 312-avoiding permutations (do both), and it is equally clear that 123-avoiding permutations are equinumerous with 321-avoiding permutations, but it is not so clear that all six classes should be equinumerous. A nice bijective proof of this appears in [**16**], and Stanley gives references for two more proofs [**18**, p. 260].

$\mathcal{P}_7$: $a_n = \frac{2(3n)!}{(2n+1)!(n+1)!}$. This is one of the prettier results in enumeration in the last fifteen years, conjectured by West [**19**] and first proved by Zeilberger [**20**]; Stanley's text also gives an account [**18**, p. 275]. Here we have

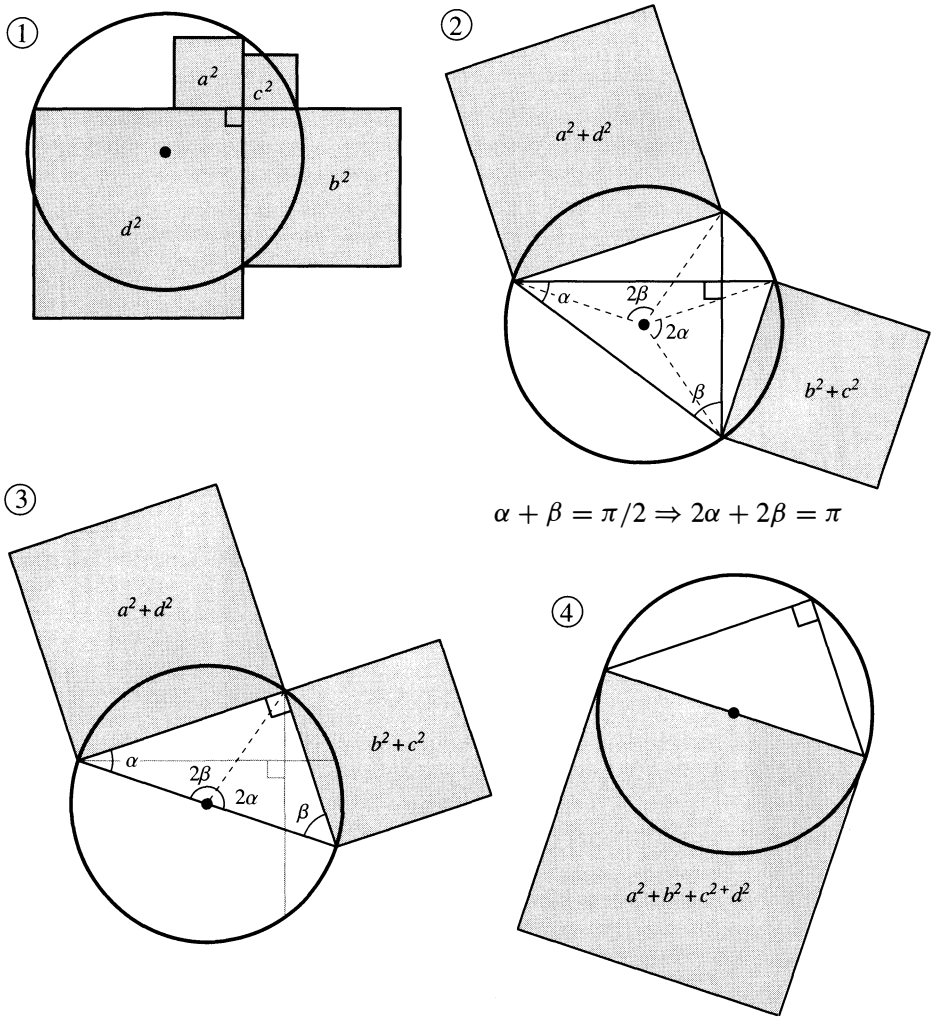
$$\alpha_{n,k} = \frac{6(4k+3)}{2k+3}\,\frac{n!}{(k+2)!}\binom{3k-1}{k-3} \quad \text{if } 1 \leq k \leq n-1.$$

REFERENCES

1. D. André, Développements de sec x et de tang x, *C.R. Acad. Sci. Paris* **88** (1879), 965–967.
2. ———, Sur les permutations alternées, *J. Math.* **7** (1881), 167–184.
3. ———, Solution directe du problème résolu par M. Bertrand, *C.R. Acad. Sci. Paris* **105** (1887), 436–437.
4. Gabriel Cramer, *Introduction a l'analyse des Lignes Courbes algébriques*, Geneva, 1750.
5. Leonhard Euler, *Opera Omnia*, Series Prima, vol. XIV, B.G. Teubner, Berlin, 1925.
6. William Feller, *An Introduction to Probability Theory and its Applications*, vol. 1, Wiley, New York, 1968.
7. H. G. Forder, Mathematical Notes, *Math. Gazette* **14** (1928), 233.
8. J. Garnier, *Analyse Algébrique*, 2nd ed., Paris, 1814.
9. Joseph Diez Gergonne, Développement de la théorie donnée par M. Laplace pour l'élimination au premier degré, *Annales de Mathématiques* **4** (1813), 148–155.
10. Ian P. Goulden and David M. Jackson, *Combinatorial Enumeration*, Wiley, New York, 1983.
11. Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics*, 2nd ed., Addison-Wesley, Reading, MA, 1994.
12. Richard Grassl, Euler numbers and skew-hooks, this MAGAZINE, **66**:3 (1993), 181–188.
13. Warren P. Johnson, Some polynomials associated with up-down permutations, *Discrete Math.* **210** (2000), 117–136.
14. Donald E. Knuth, *The Art of Computer Programming, Volume 1: Fundamental Algorithms*, Addison-Wesley, Reading, MA, 1969.
15. Thomas Muir, *The Theory of Determinants in the Historical Order of Development*, vol. 1, reprinted by Dover, New York, 1960.
16. Rodica Simion and Frank W. Schmidt, Restricted permutations, *European J. Combin.* **6** (1985), 383–406.
17. Richard P. Stanley, *Enumerative Combinatorics*, Vol. 1, Cambridge University Press, Cambridge, 1997.
18. ———, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, Cambridge, 1999.
19. Julian West, Sorting twice through a stack, *Theoret. Comput. Sci.* **117** (1993), 303–313.
20. Doron Zeilberger, A proof of Julian West's conjecture that the number of 2-stack-sortable permutations of length n is $\frac{2(3n)!}{(2n+1)!(n+1)!}$, *Discrete Math.* **102** (1992), 85–93.

Proof Without Words: Four Squares with Constant Area

If two chords of a circle intersect at right angles, then the sum of the squares of the four segments formed is constant (the square of the length of the diameter).



$\alpha + \beta = \pi/2 \Rightarrow 2\alpha + 2\beta = \pi$

——ROGER B. NELSEN
LEWIS & CLARK COLLEGE
PORTLAND, OR 97219

Inversions and Major Index for Permutations

THOTSAPORN THANATIPANONDA
Rutgers University,
Piscataway, NJ 08854
thot@math.rutgers.edu

It is well known that there are $n!$ permutations of the set $\{1, 2, \ldots, n\}$. Stern [**4**] asked the question of how many inversions there are in these $n!$ permutations. An *inversion* is the occurrence of a larger number before a smaller one in the line notation for the permutation.* For example, the permutation 4132 has 4 inversions: three from 4 appearing before 1, 2, and 3, and one from 3 appearing before 2. The count of inversions in a permutation is one example of a permutation statistic.

Stern's question was answered by Terquem [**5**] who showed that the total number of inversions in all the permutations of $\{1, 2, \ldots, n\}$ is

$$I(n) = \frac{n!\, n\,(n-1)}{4}. \tag{1}$$

Shortly after this, Rodrigues [**3**] gave two other solutions to Stern's problems. More importantly, he found the generating function for these inversions: If $I(n, k)$ denotes the number of permutations of $\{1, 2, \ldots, n\}$ with k inversions, then

$$\sum_{k\geq 0} I(n, k)q^k = 1(1+q)\cdots(1+q+\cdots+q^{n-1}). \tag{2}$$

Rodrigues's work was largely ignored and then forgotten. MacMahon [**2**] studied a more general problem. He studied the inversion problem for multisets, that is, he allowed the integers $1, \ldots, n$ to be repeated. In addition, he introduced a second permutation statistic, which is now called the major index.

For a permutation of $\{1, 2, \ldots, n\}$ a *descent* occurs when a larger number appears immediately before a smaller number. The *major index* of a permutation is the sum of the locations of the first number in each pair that is a descent. The example given before, $\underline{4}1\underline{3}2$, has major index $1 + 3 = 4$, since a descent from 4 to 1 occurs in the first space and another from 3 to 2 in the third space. While the number of inversions and the major index are the same for 4132, this is usually not true. The permutation $\underline{43}12$ has 5 inversions and its major index is $1 + 2 = 3$.

Here are the data for the permutations of $\{1, 2, 3\}$:

Permutations	Inversions	Major Index
123	0	0
132	1	2
213	1	1
231	2	2
312	2	1
321	3	3

*EDITOR'S NOTE: Information about the various notations for permutations appears on page 129.

Notice that the number of permutations with k inversions is equal to the number of permutations with major index equal to k. In this case, the two statistics are said to be *equidistributed*. MacMahon proved this in the more general case of multisets by showing that both statistics have the same generating function.

Schutzenberger suggested privately to Foata that there should be a combinatorial way to show that these two statistics are equidistributed. Foata [**1**] found a bijection that does this. However, Foata's bijection is not easily understood. In a combinatorics course at the University of Wisconsin, Madison, Professor Richard Askey asked if anyone could find a combinatorial argument to show the equidistribution of these two statistics for the set $\{1, 2, \ldots, n\}$. My solution is the subject of this article. We will begin with some background information on inversions and permutations.

Permutations and inversions Let $I(n)$ denote the number of inversions of the $n!$ permutations of $\{1, 2, \ldots, n\}$. There are a number of ways to find the value of $I(n)$. Rodrigues [**3**] organized the set of permutations into pairs where one was a reversal of the other, namely, pairs of the form $p_1\, p_2 \ldots p_n$ and $p_n\, p_{n-1} \ldots p_1$. Any two indices p_i and p_j will appear as an inversion in one but not the other of the pair. Thus, half of the total number of $n!n(n-1)/2$ pairs of indices in all permutations are inversions. This means that there are $n!n(n-1)/4$ inversions.

A second way is due to Terquem, who gave a recurrence formula for $I(n+1)$ in terms of $I(n)$. Each permutation $p_1\, p_2 \ldots p_n$ of $\{1, 2, \ldots, n\}$ gives rise to $n+1$ permutations of $\{1, 2, \ldots, n, n+1\}$ by inserting $n+1$ before any of the p_i or after p_n. Since inserting $n+1$ before k adds $n+1-k$ inversions to each permutation, we have

$$I(n+1) = \underbrace{I(n)}_{n+1 \text{ last}} + \underbrace{\big(I(n)+n!\big)}_{n+1 \text{ second to last}} + \underbrace{\big(I(n)+2n!\big)}_{n+1 \text{ third to last}} + \cdots + \underbrace{\big(I(n)+nn!\big)}_{n+1 \text{ first}}. \tag{3}$$

Simplifying gives

$$I(n+1) = (n+1)I(n) + \frac{(n+1)!\,n}{2}. \tag{4}$$

An easy calculation shows that the formula in (1) satisfies the recurrence in (4). If we did not already know this, we could solve the recurrence by imitating a well-known method from ordinary differential equations. First solve the homogeneous equation $I(n+1) = (n+1)I(n)$ to get $I(n) = c\,n!$. Then adapt the technique of variation of parameters setting

$$I(n) = J(n)\,n!.$$

Substituting this into the recurrence, we get $J(n+1) = J(n) + n/2$, which implies

$$J(n) = J(1) + \frac{1+2+\cdots+(n-1)}{2} = J(1) + \frac{1}{2}\binom{n}{2}.$$

Since $J(1) = 0$, this gives the same formula as above.

Rodrigues' generating function is also easy to find. If

$$G_n(q) = \sum_{k\geq 0} I(n,k)q^k,$$

then the same insertion argument that proved (3) gives

$$G_{n+1}(q) = G_n(q)(1 + q + \cdots + q^n). \tag{5}$$

Iteration of (5) and the case $n = 2$ gives Rodrigues' formula

$$\sum_{k\geq 0} I(n,k)q^k = 1(1+q)\cdots(1+q+\cdots+q^{n-1}). \tag{6}$$

Rodrigues first solved Stern's problem by differentiating (6) with respect to q and then setting $q = 1$. We leave this as an exercise for the reader.

Equidistribution of inversions and major index We will use mathematical induction to prove that the number of inversions and the major index are equidistributed. Specifically, we will assume that the two statistics are equidistributed for permutations of length n and see what happens when we form new permutations by inserting the number $n + 1$ at any of the possible positions. The $n = 1$ case is trivial, since there are no inversions and the major index is zero.

When we insert $n + 1$ into a permutation of $\{1, 2, \ldots, n\}$ that has k inversions, the resulting number of inversions follows an easy pattern: If we insert $n + 1$ at the end, there are still k inversions; when we insert this element at the jth position, there will be $k + j$ inversions. Thus the number of inversions when inserting $n + 1$ into each of the $n + 1$ places in this permutation give permutations with $k, k + 1, \ldots, k + n$ inversions.

The pattern is not as transparent with the major index, but it gives rise to the same set of numbers. To see what happens, consider the following example, when 8 is inserted into the permutation 5724631 of $\{1, 2, \ldots, 7\}$, whose major index is $2 + 5 + 6 = 13$.

Permutation	Major Index
5724631$\underline{8}$	$2+5+6=13$
5724638$\underline{1}$	$2+5+7=14$
5724683$\underline{1}$	$2+6+7=15$
5728463$\underline{1}$	$2+5+6+7=20$
5728463$\underline{1}$	$2+4+6+7=19$
5782463$\underline{1}$	$3+6+7=16$
5872463$\underline{1}$	$2+3+6+7=18$
8572463$\underline{1}$	$1+3+6+7=17$

Notice that the set of numbers that appear as major indices is $\{13, 14, 15, \ldots, 20\}$. The number of inversions is not the same as the major index permutation by permutation; after all, there is no reason for the number of inversions and the major index of 5724631 to be the same. However, we will be done if we can explain why, when the major index of the original permutation is k, the set of major indices for all the insertions of $n + 1$ consists of exactly the numbers $k, k + 1, \ldots, k + n$.

We categorize the insertions according to two cases.

CASE (a). *Inserting the number $n + 1$ between the descents in the permutation or at either end.* Here is one example.

Permutation	Major Index
5724631$\underline{8}$	$2+5+6=13$
5724638$\underline{1}$	$2+5+7=14$
5724683$\underline{1}$	$2+6+7=15$
5782463$\underline{1}$	$3+6+7=16$
8572463$\underline{1}$	$1+3+6+7=17$

In this case, we can see that after inserting the number $n+1$ in each of the descents from right to left in the permutation, the major index increases by one each time.

CASE (b). *Inserting the number $n+1$ between the ascents in the permutation.* We continue with the same example.

Permutation	Major Index
5724$\underline{8}$631	$2+5+6+7=20$
572$\underline{8}$4631	$2+4+6+7=19$
5$\underline{8}$724631	$2+3+6+7=18$

In this case, we can see that after inserting the number $n+1$ in each of the ascents from right to left in the permutation, the major index decreases by one each time.

We will now prove that what we noticed in the examples is true in general.

Proof for Case (a). We show that the major index increases by one each time we move the insertion point for the number $n+1$ from between one descent to the next descent to the left.

For the general case, consider the third and fourth rows of the example for Case (a). Two things happen when we move the location of $n+1$:

(i) *At the new position of $n+1$.* In the new permutation, the number $n+1$ adds a value to the major index, while the major index of the descent where $n+1$ was inserted is gone. So the major index increases by one. For example in 57$\underline{8}$24631, we gained 3 from inserting 8 and lost 2 as the major index from 7 is gone. So the major index increases by one.

(ii) *At the original position of $n+1$.* Although the major index from having $n+1$ in the original permutation is gone, we get the same major index after the insertion point for $n+1$ is moved to the next descent on the left. For example, 57246$\underline{8}$31 has a contribution of 6 to the major index from number 8, which is lost when we move the 8. But in 57$\underline{8}$24631 we gain the 6 back again from the 6 in the sixth position.

So from (i) and (ii) we see that the major index increases one each time we move the number $n+1$ from a descent to the next descent on the left. ■

Proof for Case (b). We show that each time we move the insertion point for $n+1$ one more ascent to the left, the major index decreases by one.

The general case here is very much like the second and third rows of the example for Case (b). Two things happen when inserting $n+1$ into one of the permutations from Case (b):

(i) The number $n+1$ moves r position(s) to the left causing the major index to decrease by r. In the example, the number 8 moves 2 positions to the left causing the major index to decrease by 2.

(ii) The number $n+1$ will pass $r-1$ descents. So it will push each of these $r-1$ descents one position to the right and the major index increases by $r-1$. In the example, there is one decreasing sequence "72" that is pushed to the right.

From (i) and (ii) the change in major index is given by $-r+(r-1)=-1$. So the major index decreases by one each time we insert the number $n+1$ between an ascent and the next increasing sequence on the left. ■

Conclusion Consider inserting $n+1$ into a permutation with m descents.

Case (a) The major index increases from the major index of the original permutation by $0, 1, \ldots, m+1$.

Case (b) When we insert $n+1$ in the first ascent on the right, we add n to the major index. This is because we add ℓ from inserting $n+1$ at position ℓ and gain $n-\ell$ for each of the $n-1$ descents that are pushed to the right by the insertion. With each successive insertion the major index decreases by one, so the major index of the original permutation has increased by $n, n-1, \ldots, m$.

We have shown that the major index gains $0, 1, 2, \ldots, n$ upon inserting the number $n+1$ at all possible positions in any permutation of $\{1, 2, \ldots, n\}$. This is exactly what happens with inversion numbers, so we have established the induction step, thereby proving that the major index and the number of inversions are equidistributed.

REFERENCES

1. D. Foata, On the Netto inversion number of a sequence, *Proc. Amer. Math. Soc.* **19** (1968), 236–240.
2. P. A. MacMahon, The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects, *Amer. J. Math.* **35** (1913), 281–322.
3. G. Rodrigues, Note sur les Inversions, ou dérangements produits das les permutations, *J. Math. Pures Appl.* **4** (1839), 236–240.
4. M. Stern, Aufgaben, *Jour. für reine und angew. Math.* **18** (1838), 100.
5. M. O. Terquem, Solution d'un Probléme de combinaison, *J. de Math. Pures et Appl.* **3** (1838), 559–560.

Permutations in the Sand

MARK D. SCHLATTER
Centenary College of Louisiana
Shreveport, LA 71104
mschlat@centenary.edu

In chapter 4 of his book [**2**], Paulus Gerdes introduces the sona drawings of the Chokwe people. These drawings (as in FIGURE 1) are made in the sand and have been associated with Chokwe initiation rites. Gerdes asks a typical mathematical question: Can these figures be drawn in a single motion without lifting your finger and without changing direction at a crossing point?

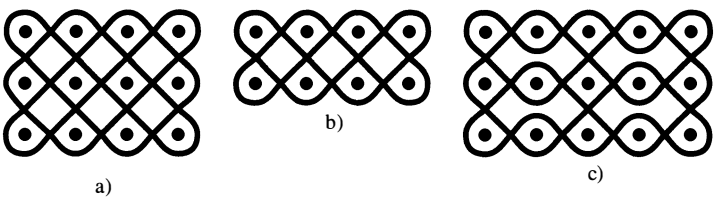


Figure 1 Three sona drawings

In this paper, we will show how permutations can be used to answer this question. In particular, we will use permutations to prove a conjecture of Gerdes' from his book [**2**]

about a class of sona called *lion's stomach designs.* As we do this, we will build up a set of tools that can be used to analyze a wider class of sona drawings.

Mirror curves To start, we explain what a sona drawing is, or more specifically, how it can be constructed. Roughly speaking, we can create each sona drawing in this article by tracing paths that bounce off carefully placed mirrors. For example, if you start at the points marked S in FIGURE 2 and head up and to the right at a 45 degree angle, you recreate the designs in FIGURE 2 by bouncing off the mirrors, that is, the sides of the box or the straight line segments. (The rectangular array of dots in the box is not necessary, but helps in locating mirrors and the path.) Note that both the first and third sona drawings can be completed without lifting your pencil, but you need to start at two places and draw two curves to recreate FIGURE 1b. We will think of sona drawings as unions of *mirror curves* (see Jablan [**4**] for a general definition of mirror curves and many examples). For our purposes, we consider mirror curves on rectangular arrays. To start, a *mirror box* is a rectangle in the plane with corners at $(0, 0)$, $(2n, 0)$, $(0, 2m)$, and $(2n, 2m)$ for $m, n \in \mathbb{N}$. The dots forming a rectangular array are placed at points (h, k) where h and k are odd numbers less than, respectively, $2n$ and $2m$. A mirror may be placed at any location (h, k) within the box as long as either h or k is even (but not both) and $1 \leq h \leq 2n - 1$, $1 \leq k \leq 2m - 1$. (In other words, mirrors are between pairs of dots.) Mirrors have either horizontal or vertical orientation. As an example, the mirror box in FIGURE 2c has an upper right corner at $(10, 6)$, a dot in the upper right corner at $(9, 5)$, and an upper right mirror oriented horizontally at $(7, 4)$.

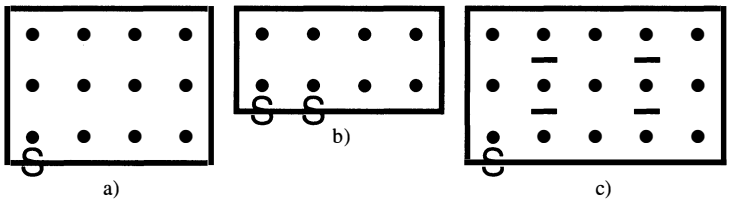


Figure 2 Mirror boxes

Consider tiling the mirror box with unit squares whose vertices have integer coordinates. We can create a graph where the vertices are the midpoints of the sides of these squares and the edges are the line segments connecting the midpoints of adjacent sides. (FIGURE 1a shows a rounded version of such a graph.) We define a mirror curve to be a cycle in this graph with the property that the direction of travel can only be changed at mirrors and the sides of the box. A sona drawing is complete when all possible mirror curves have been drawn. So, for example, FIGURE 1b is the union of two mirror curves. We identify the size of the drawing by referring to the array of dots. Thus, FIGURE 1a is a 3×4 sona drawing.

Gerdes refers to sona drawings like FIGURE 1a and FIGURE 1b (that is, sona drawings with no mirrors on rectangular arrays) as *plaited-mat designs.* FIGURE 1c is called a *lion's stomach design.* A lion's stomach design is built on a rectangular m by n array where n is odd with columns that alternate between no mirrors and horizontal mirrors in all possible locations. We start and end with a column that has no mirrors. In his book [**2**], Gerdes builds a number of lion's stomach designs to make the following conjecture:

CONJECTURE. The number of mirror curves needed in a m by n lion's stomach design is 1 if $4|(n - 1)$ and is m otherwise.

We will introduce permutations in order to prove this conjecture.

Why use permutations? We start with some simplifying assumptions. We assume that all mirrors are horizontal and that no mirrors occur in the first or last column of dots. (Both of these assumptions are true of the plaited-mat designs and lion's stomach designs studied by Gerdes.) Under these assumptions, we can demonstrate that the number of mirror curves needed for a sona drawing can be detected from a permutation that we will define using the interior columns of the drawing. This number is an example of a permutation statistic.

As an example, look at the sona drawings in FIGURE 3. Each drawing consists of two mirror curves. To the right of each drawing, we have split the drawing into the first column, the interior columns, and the last column. We number the strands that make up the drawing to the right of the first column after the crossing point. We then see what position a strand has moved to after it has passed through the interior columns. For example, in both drawings strand 1 has moved to the position first held by strand 2, and strand 2 has moved to the position first held by strand 4. In other words, we have a permutation of the strand numbers. For both drawings, this induced permutation, in cycle notation, is (1243)(56).* Note that since all mirrors are horizontal, a strand starting from the first column goes across to the last column and then returns.

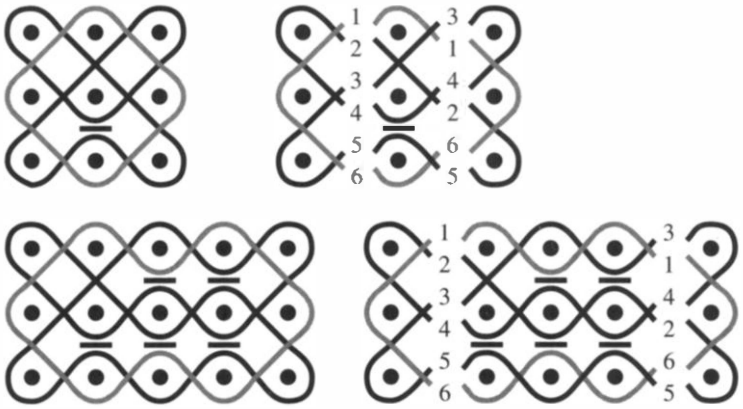


Figure 3 Examining permutations

Note that both drawings have the same induced permutation, even though they have different column arrangements. Because in both drawings, the strands end up in the same positions after passing through the interior columns, the strands also have the same connections to the portions of the curves in the first and last columns. Thus, given our assumptions, we have demonstrated the following theorem:

THEOREM 1. *If the interior columns of two sona drawings induce the same permutation on the strands, the drawings will require the same number of mirror curves.*

Our strategy is now set: we will determine the permutation induced on strands by the interior columns of lion's stomach designs. We will then find sona drawings that induce the same permutation where the number of mirror curves is easy to calculate.

A note before we move on: permutations have been used to analyze mirror curves before. Jaritz [**5**] used permutations with row strands to analyze what we are calling

*EDITOR'S NOTE: Information about the various notations for permutations appears on page 129.

plaited-mat designs and some generalizations. Gerdes [3] has used permutations to look at sona drawings in unpublished work.

The columns of lion's stomach designs To find the permutation induced on the strands by the interior columns of a lion's stomach design, we look at the permutation induced by each interior column. After we have found the permutations for each interior column, we will find the induced permutation for the entire set of interior columns by composing the permutations for adjacent columns. Remember that for lion's stomach designs, we alternate columns with no mirrors with columns with horizontal mirrors in all possible positions. FIGURE 4 shows the left-hand edge of a sona drawing with m rows. Once again we have numbered the strands after they cross and shaded the strands in pairs to help distinguish them. We first show the new positions of these strands after a column with no mirrors. We can write this permutation as $\sigma = (2\ 4\ 6 \ldots 2m-2\ 2m\ 2m-1 \ldots 3\ 1)$. In the same figure, we show the new positions of the strands after a column with horizontal mirrors. We obtain the permutation $\tau = (1\ 2)(3\ 4)\ldots(2m-1\ 2m)$.

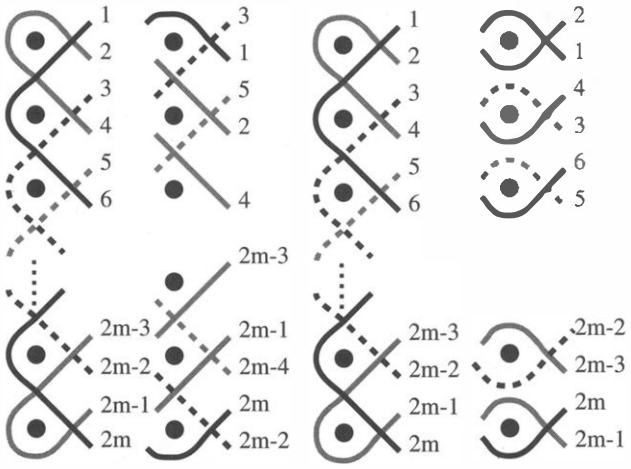


Figure 4 Our two permutations

Now, given any m by n lion's stomach design, we associate a word $x_1x_2 \ldots x_{n-2}$ where each x_i is either σ or τ and corresponds to an interior column. (Of course, in a lion's stomach design we know something about the word—σ and τ will alternate—but we will look at all possible words using these permutations.)

We can reduce the words using multiplication as composition on the right. It is easy to check that σ^{2m} and τ^2 are the identity permutation (you can even see the latter in FIGURE 3) and with a little bit of work, we can also check that $\tau\sigma = \sigma^{2m-1}\tau$. Given the word associated with any drawing constructed as above, we therefore reduce it to the form $\sigma^j\tau^k$ where $k = 0$ or $k = 1$ and $j \geq 0$. Thus, by Theorem 1, if we know the number of mirror curves needed for the sona drawings associated with these words, we know the number of mirror curves needed for all lion's stomach designs. But to find the number of curves needed for these words, we need a different technique.

The "cutting off" algorithm We need the following theorem:

THEOREM 2. *Given any $m \times n$ sona drawing with no mirrors, the number of mirror curves needed is* $\gcd(m, n)$.

A proof of this theorem is not explicitly found in Gerdes' book [2], but he discusses the ingredients for a proof by induction in the "cutting off squares" algorithm on pp. 176–179. Since we will use a similar technique, we review the algorithm here.

The basic principle involved is that we may cut off a portion of a sona drawing and not change the number of mirror curves needed *if* we only cut off loops of a mirror curve. As an example, look at FIGURE 5. This 2×4 sona drawing requires two mirror curves. If we cut off the 2×2 square on the right, we remove a loop from each of the two mirror curves. Thus the remaining drawing still requires two mirror curves.

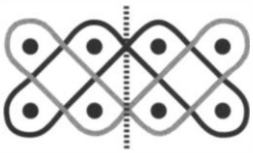
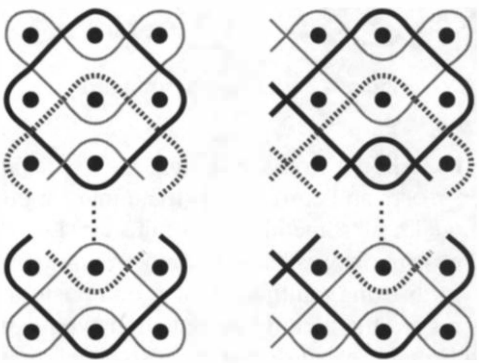


Figure 5 The "cutting off" algorithm

Given any $m \times n$ sona drawing, we can cut off squares until a square drawing remains. It is easy to check that the number of mirror curves needed for a $k \times k$ drawing with no mirrors is k. In other words, if $f(m, n)$ is the number of mirror curves needed to complete an $m \times n$ sona drawing with no mirrors, we know that $f(k, k) = k$ and $f(m, n) = f(m, n - m)$ where $n > m$. Theorem 2 can then be proved by induction by noting that $\gcd(k, k) = k$ and $\gcd(m, n) = \gcd(m, n - m)$.

Our base cases Recall that for our lion's stomach designs, the reduced words had the form $\sigma^j \tau^k$ where $k = 0$ or $k = 1$ and $j \geq 0$. We look at the possible cases using the above tools.

Case 1: The reduced word is τ. A quick look at FIGURE 6a shows that the design will require as many curves as rows.



a) b)

Figure 6 Two cases

Case 2: The reduced word has the form σ^j. In other words, the original lion's stomach design induces the same permutation as a drawing with no mirrors on an $m \times (j + 2)$ array. (We add the 2 to account for the columns on the ends.) By Theorem 2, the number of curves required will be $\gcd(m, j + 2)$.

Case 3: The reduced word has the form $\sigma^j \tau$. FIGURE 6b shows the last three columns of the drawing associated with the reduced word. We can use the "cutting

off" algorithm and remove the last three columns without changing the number of mirror curves. We therefore produce a drawing with $j - 2$ interior columns and no mirrors that requires $\gcd(m, j)$ mirror curves. (Once again, we have added in the two columns on the ends.)

We can now prove the conjecture. Lion's stomach designs have the associated words $\tau\sigma\tau$, $\tau\sigma\tau\sigma\tau, \ldots$ or, more generally, $\tau(\sigma\tau)^k$ for an $m \times (2k + 3)$ drawing. Since $\tau\sigma\tau = \sigma^{2m-1}$, it is easy to check that $\tau(\sigma\tau)^k$ equals τ if k is even and equals σ^{2m-1} if k is odd. Therefore a lion's stomach design with $4j + 3$ columns will have the associated word $\tau(\sigma\tau)^{2j}$, the reduced word τ, and by Case 1 will require as many mirror curves as rows. A lion's stomach design with $4j + 1$ columns will have the associated word $\tau(\sigma\tau)^{2j-1}$, the reduced word σ^{2m-1}, and by Case 2 will require $\gcd(2m + 1, m)$ (or 1) mirror curve. Besides proving the conjecture, we can now determine the number of mirror curves needed for any rectangular sona drawing that has the two types of above columns as interior columns.

Open questions While the machinery of permutations works for lion's stomach designs, whether or not their application can determine the number of mirror curves needed for other sona drawings is an open question. We can associate a permutation with any interior column whose mirrors are horizontal. However, the strategy we have used (finding a small set of reduced words for which the number of mirror curves can be computed) may not be feasible for general rectangular sona drawings.

Another avenue of exploration is vertical mirrors. FIGURE 7 shows a 3 by 6 sona drawing that Gerdes calls a *chased chicken design* [**2**]. While this drawing uses vertical mirrors, we can associate a permutation with each pair of adjacent interior columns (as shown on the right). Gerdes has proven [**1**] that an m by n chased chicken design (where m is odd and n is even) requires $\gcd((m + 1)/2, (n + 2)/2)$ mirror curves. The interested reader can prove this with fixed values of m using permutations—whether or not permutations can be applied to prove the fact for any value of m is an open question.

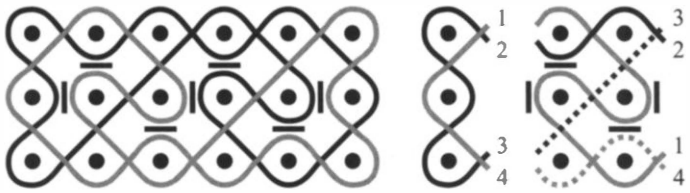


Figure 7 A chased chicken design

REFERENCES

1. P. Gerdes, *Geometria Sona*, Instituto Superior Pedagógico, 1993.
2. ———, *Geometry From Africa: Mathematical and Educational Explorations*, Mathematical Association of America, 1999.
3. ———, personal communication.
4. S. Jablan, *Mirror Curves*, http://members.tripod.com/~modularity/index.html.
5. W. Jaritz, Über Bahnen Auf Billardtischen - Oder: Eine Mathematische Untersuchung Von Ideogrammen Angolanischer Herkunft, *Berichte der Mathematisch-Statistischen Sektion im Forschungzentrum Graz*, **207** 1983, 1–22.

The Minimal Polynomials of $\sin\left(\frac{2\pi}{p}\right)$ and $\cos\left(\frac{2\pi}{p}\right)$

SCOTT BESLIN
Nicholls State University
Thibodaux, LA 70310
scott.beslin@nicholls.edu

VALERIO DE ANGELIS
Xavier University of Louisiana
New Orleans, LA 70125
vdeangel@xula.edu

Every student of trigonometry knows that if $n = 1, 2, 3, 4$, or 6, then $\cos(2\pi/n)$ is a rational number, and so it is the root of a first degree polynomial with integer coefficients. Another way of expressing this is to say that for the values of n listed above, $\cos(2\pi/n)$ is an *algebraic number* of *algebraic degree* one. Similarly well known is the fact that for $n = 5$, 8, or 12, $\cos(2\pi/n)$ is a root of a quadratic, irreducible polynomial with integer coefficients, and we express this by saying that for $n = 5$, 8, or 12, $\cos(2\pi/n)$ is an algebraic number of algebraic degree two. (The case $n = 5$ may not be as popular as the others; for those who wonder, $\cos(2\pi/5)$ is a root of the polynomial $4x^2 + 2x - 1$).

As the reader will no doubt have guessed at this point, if a real (or complex) number is a root of an irreducible polynomial of degree n with integer coefficients, we say that it is an *algebraic number* with *algebraic degree* n. The irreducible polynomial in question (where *irreducible* means that it cannot be factored into lower degree polynomials with integer coefficients) is its *minimal polynomial.* So the previous paragraph can be rephrased by saying that the algebraic degree of $\cos(2\pi/n)$ is widely known for $n = 1$, 2, 3, 4, 5, 6, 8, and 12.

But probably few trigonometry students know that 1, 2, 3, 4, and 6 are the *only* positive integer values of n for which the algebraic degree of $\cos(2\pi/n)$ is one, and 5, 8, and 12 are the only values for which it is two. In fact, there is a general formula to compute the algebraic degree of $\cos(2\pi/n)$ and $\sin(2\pi/n)$ (see the remark at the end).

The purpose of this note is to exhibit explicit expressions for the minimal polynomials of $\cos(2\pi/n)$ and $\sin(2\pi/n)$ in the special case where n is a prime.

We will show that if $p > 2$ is a prime number, then the minimal polynomial of $\sin(2\pi/p)$ is

$$S_p(x) = \sum_{k=0}^{(p-1)/2} (-1)^k \binom{p}{2k+1} (1-x^2)^{\frac{p-1}{2}-k} x^{2k},$$

of degree $p - 1$, and the minimal polynomial of $\cos(2\pi/n)$ is

$$C_p(x) = S_p\left(\sqrt{\frac{1-x}{2}}\right),$$

of degree $(p - 1)/2$. All arguments are elementary, and make no use of field theory. The main tool is Eisenstein's criterion for irreducibility of polynomials [**1**, p. 160], which we recall below.

EISENSTEIN'S CRITERION. *Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial with integer coefficients. If there is a prime p that divides each of $a_0, a_1, \ldots, a_{n-1}$, while p does not divide a_n, and p^2 does not divide a_0, then $f(x)$ is irreducible.*

Here is a brief outline of the proof: if the given conditions are satisfied and $f(x)$ factors as $f(x) = g(x)h(x)$, consider all coefficients modulo p (that is, disregard multiples of p). Then $f(x)$ has the form cx^n, and so (because p is prime) the same must be true of $g(x)$ and $h(x)$, hence p divides the constant terms of both $g(x)$ and $h(x)$. But then p^2 divides the constant term of $f(x)$.

We proceed to the derivation of $S_p(x)$. Let $p > 2$ be a prime number. Using Euler's identities $e^{i\theta} = \cos\theta + i\sin\theta$ and $\sin\theta = (e^{i\theta} - e^{-i\theta})/2i$, and setting $x = \sin\theta$, we derive an expression for $\sin(p\theta)/\sin\theta$ in terms of x. To simplify

$$\frac{\sin(p\theta)}{\sin\theta} = \frac{1}{2ix}\left[\left(\sqrt{1-x^2} + ix\right)^p - \left(\sqrt{1-x^2} - ix\right)^p\right],$$

we use the binomial theorem to write

$$\left(\sqrt{1-x^2} \pm ix\right)^p = \sum_{k=0}^{p} (\pm i)^k \binom{p}{k} \left(1-x^2\right)^{(p-k)/2} x^k.$$

Then, in the bracketed expression, we find that the even terms of the sum will cancel. Reindexing, we obtain

$$\frac{\sin(p\theta)}{\sin\theta} = \sum_{k=0}^{(p-1)/2} (-1)^k \binom{p}{2k+1} \left(1-x^2\right)^{(p-1)/2-k} x^{2k} = S_p(x). \qquad (*)$$

While the first equality in $(*)$ has been derived only in the case that $|x| = |\sin\theta| \le 1$, we can clearly define $S_p(x)$ by the second equality for any real value of x. Note that $S_p(x)$ is a polynomial in x^2 with integer coefficients, and of degree $p-1$. This polynomial is closely related to the classical Chebyshev polynomials of the second kind $U_n(x)$, defined by $U_n(\cos\theta) = \sin((n+1)\theta)/\sin\theta$, because $S_p(x) = U_{p-1}(\sqrt{1-x^2})$.

From $(*)$, we have $S_p(\sin\theta) = \sin(p\theta)/\sin\theta$ for all θ, and so $S_p(\sin(2\pi/p)) = \sin(2\pi)/\sin(2\pi/p) = 0$. Hence to show that $S_p(x)$ is the minimal polynomial of $\sin(2\pi/p)$ we only need to prove that $S_p(x)$ is irreducible. Clearly $S_p(0) = p$, so that the condition on the constant term required by Eisenstein's criterion is satisfied. If $k < (p-1)/2$, then the binomial coefficient $\binom{p}{2k+1} = p(p-1)\cdots(p-2k)/(2k+1)!$ is a multiple of p, because all prime factors of $(2k+1)!$ are less than p, so $(p-1)\cdots(p-2k)/(2k+1)!$ must be an integer. Since the term corresponding to $k = (p-1)/2$ in $(*)$ is $(-1)^{(p-1)/2}x^{p-1}$, it does not contribute to any term of degree less than $p-1$. We conclude that all coefficients of $S_p(x)$ other than the leading one are divisible by p.

Since the degree of $S_p(x)$ is $p-1$, $x^{p-1}S_p(1/x)$ is a polynomial in x whose constant term is the leading coefficient of $S_p(x)$. We find from $(*)$ that

$$x^{p-1}S_p\left(\frac{1}{x}\right) = \sum_{k=0}^{(p-1)/2} (-1)^k \binom{p}{2k+1} \left(x^2-1\right)^{(p-1)/2-k},$$

and so the leading coefficient of $S_p(x)$ is

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

To evaluate this sum, note that

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

We see that the leading coefficient of $S_p(x)$ is $(-1)^{(p-1)/2}2^{p-1}$, and from Eisenstein's criterion we conclude that $S_p(x)$ is irreducible, and therefore it is the minimal polynomial of $\sin(2\pi/p)$.

Now consider $C_p(x) = S_p(\sqrt{(1-x)/2})$, of degree $(p-1)/2$ (recall that $S_p(x)$ is a polynomial in x^2, so $S_p(\sqrt{x})$ is a polynomial in x). We then have $C_p(\cos\theta) = S_p(\sin(\theta/2)) = \sin(p\theta/2)/\sin(\theta/2)$, and so $C_p(\cos(2\pi/p)) = \sin(\pi)/\sin(\pi/p) = 0$. It is a simple exercise in the use of trigonometric identities to show (by induction, for example) that for each positive integer m, we have $\sin[(m+1/2)\theta]/\sin(\theta/2) = 1 + 2\sum_{k=1}^{m} \cos(k\theta)$. Using this with $m = (p-1)/2$, we find that

$$C_p(\cos\theta) = 1 + 2 \sum_{k=1}^{(p-1)/2} \cos(k\theta).$$

The terms of this sum define the Chebyshev polynomials of the first kind $T_k(x)$, given by $T_k(\cos\theta) = \cos(k\theta)$, which are easily seen to have integer coefficients. So, we conclude that $C_p(x)$ has integer coefficients. In order to prove that $C_p(x)$ is the minimal polynomial of $\cos(2\pi/p)$ it only remains to show that it is irreducible. If $f(x)$ were a factor of $C_p(x)$, then $f(1-2x)$ would be a factor of $S_p(\sqrt{x})$. But the coefficient of x^k in $S_p(\sqrt{x})$ is the coefficient of x^{2k} in $S_p(x)$, so the same application of Eisenstein's criterion as before shows that $S_p(\sqrt{x})$ is irreducible. This concludes our derivation of the minimal polynomials of $\sin(2\pi/p)$ and $\cos(2\pi/p)$.

We remark that the polynomials $S_p(x)$ and $C_p(x)$ are also primitive, in the sense that the greatest common factor of their coefficients is one. This is not guaranteed by Eisenstein's criterion (as can be seen, for example, if we multiply the polynomials by 2), but it is easily checked for $S_p(x)$ because the constant term is p while the leading coefficient is a power of 2, and for $C_p(x)$ because $C_p(0) = C_p(\cos(\pi/2)) = \sin(p\pi/4)/\sin(\pi/4) = \pm 1$.

We list below $S_p(x)$ and $C_p(x)$ for the first few values of p:

$$S_3(x) = -4x^2 + 3$$
$$S_5(x) = 16x^4 - 20x^2 + 5$$
$$S_7(x) = -64x^6 + 112x^4 - 56x^2 + 7$$
$$S_{11}(x) = -1024x^{10} + 2816x^8 - 2816x^6 + 1232x^4 - 220x^2 + 11$$
$$S_{13}(x) = 4096x^{12} - 13312x^{10} + 16640x^8 - 9984x^6 + 2912x^4 - 364x^2 + 13$$
$$C_3(x) = 2x + 1$$
$$C_5(x) = 4x^2 + 2x - 1$$
$$C_7(x) = 8x^3 + 4x^2 - 4x - 1$$
$$C_{11}(x) = 32x^5 + 16x^4 - 32x^3 - 12x^2 + 6x + 1$$
$$C_{13}(x) = 64x^6 + 32x^5 - 80x^4 - 32x^3 + 24x^2 + 6x - 1$$

Remark: The general formula to compute the algebraic degree of $\sin(2\pi/n)$ and $\cos(2\pi/n)$ in terms of Euler's totient function ϕ for all $n > 2$ is as follows [**2**, p. 289]:

1. The degree of the cosine in question is $\deg(\cos(2\pi/n)) = \frac{1}{2}\phi(n)$.
2. If $n \neq 4$, and we write $n = 2^r m$, where m is odd, then

$$\deg(\sin(2\pi/n)) = \begin{cases} \phi(n) & \text{if } r = 0 \text{ or } 1 \\ \frac{1}{4}\phi(n) & \text{if } r = 2 \\ \frac{1}{2}\phi(n) & \text{if } r \geq 3. \end{cases}$$

We leave it to the reader to check that this formula gives the right answer for the well-known cases mentioned at the beginning of this Note.

Acknowledgment. We thank the referees for a thorough and very helpful review, including the suggestion to use $x = \sin\theta$ instead of $x = \cos\theta$ in order to simplify the derivation of $S_p(x)$, and for pointing out that Eisenstein's criterion can be checked without fully expanding all the terms of the sum.

REFERENCES

1. I. N. Herstein, *Topics in Algebra*, 2nd ed., Xerox College Publishing, Lexington, MA, 1975.
2. Paulo Ribenboim, *Algebraic Numbers*, Wiley-Interscience, 1972.

Final Digit Strings of Cubes

DANIEL P. BIEBIGHAUSER
Vanderbilt University
Nashville, TN 37240
Dan.Biebighauser@vanderbilt.edu

JOHN BULLOCK
University of Nebraska
Lincoln, NE 68588-0323
jbullock@math.unl.edu

GERALD A. HEUER
Concordia College
Moorhead, MN 56562
heuer@cord.edu

Let s be the string of decimal digits 31415926535 . . . 5275045519 formed by the first billion digits of π. You have probably wondered idly whether there is an integer n such that n^3 ends in s. Or, to take another example, let s be formed by concatenating the decimal digits of $999^3, 998^3, \ldots, 2^3, 1^3$, and ask the same question. The fourth annual North Central Section/MAA Team Contest [**1**] of November 2000, contained the following problem: Is there an integer n such that n^3, in decimal form, ends in 2000 ones? The answer to all three of the above questions is affirmative. Indeed, if s is any string of decimal digits ending in 1, 3, 7, or 9, there is an integer n such that n^3 ends in s, as we'll see in the next section.

What about other final digit strings? If the last digit of s is anything other than 1, 3, 7, or 9, there may or may not be a cube ending in s. The precise conditions under which there is such a cube are rather interesting, and are given in three subsequent sections. From the material there you can conclude that there are cubes ending

in 2000 8s but none ending in 22, 66, or 444. There are cubes ending in 2 preceded by 2000 1s but none ending in 114, 116, or 118. It is easy to see that cubes ending in 5 end in 125, 375, 625, or 875. From Theorem 3 below and the lemmas preceding it, it follows that there are cubes ending in 25 preceded by any string of 1s and 6s, and there are cubes ending in 75 preceded by any string of 3s and 8s. But there are none ending in 33125 or 44375.

What is special about 1, 3, 7, and 9 is that they are relatively prime to 10. The problem of final digit strings of cubes divides naturally into four cases, depending on whether the last digit is in {1, 3, 7, 9}, {2, 4, 6, 8}, {5}, or {0}. We shall examine them in that order.

The situation with powers other than cubes is similar when the exponents are relatively prime to 10, as 3 is. Squares, fifth powers, and so on, will be somewhat different. For instance, all squares end in 0, 1, 4, 5, 6, or 9, and there are just 22 different two-digit strings with which squares end. Some projects along these lines are suggested in the final section.

A somewhat corresponding result for initial sequences of digits is the following: If $k_1k_2\ldots k_d$ is any finite sequence of decimal digits with $k_1 \neq 0$, and a is any positive integer not a power of 10, then some integral power of a has $k_1k_2\ldots k_d$ as its initial sequence of digits. **[2, 3]**

Final digit 1, 3, 7, or 9 To introduce the idea used in the proof of Theorem 1, let's look at a step in the special case of the NCS Team Contest problem. Suppose we've found that $71^3 = 357911$, which ends in 11, and we want to use this to find an integer whose cube ends in 111. We try a number of the form $c71$; that is, $71 + 100c$:

$$(71 + 100c)^3 = 71^3 + 3(71)^2100c + 3(71)10000c^2 + 1000000c^3.$$

The last two terms will have no effect on the last three digits of this number; the first two terms are $357911 + 1512300c$, and modulo 1000 we have

$$(71 + 100c)^3 \equiv 911 + 300c.$$

The last two digits remain 11, and the digit before that is $9 + 3c$, modulo 10. To make this a 1 we need $3c \equiv 2 (\mathrm{mod}\ 10)$, and $c = 4$ does the job. Indeed, $471^3 = 104487111$ ends in 111. The existence of such a c is assured by the fact that 3 is relatively prime to 10.

Throughout the paper we shall use d for the length of (that is, the number of digits in) the string s. We shall freely switch between s as a string of digits and the integer represented by that string.

THEOREM 1. *Let d be a positive integer and let s be a string of d decimal digits ending in 1, 3, 7, or 9. Then there is an integer n such that n^3 ends in s.*

Proof. The proof is by induction on the length d of the string. For $d = 1$ we exhibit the cubes $1^3 = 1$, $7^3 = 343$, $3^3 = 27$, and $9^3 = 729$. Suppose, then, that the assertion is true for $d = r$, and consider a string $s = k_{r+1}k_rk_{r-1}\ldots k_1$ where k_1 is 1, 3, 7, or 9. By the induction hypothesis there is an integer m with m^3 ending in $k_rk_{r-1}\ldots k_1$. Consider

$$(m + c \cdot 10^r)^3 = m^3 + 3m^2c \cdot 10^r + 3mc^210^{2r} + c^310^{3r}, \tag{1}$$

with c a decimal digit to be determined. The last two terms in (1) end in more than r zeros, and the final r digits of the sum are those of m^3. The digit preceding that is $b_{r+1} + 3m^2c (\mathrm{mod}\ 10)$, where b_{r+1} is the digit preceding k_r in m^3. Because m is

relatively prime to 10, so is $3m^2$. Therefore $3m^2$ has a multiplicative inverse modulo 10, and we may choose c from $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ so that $b_{r+1} + 3m^2c \equiv k_{r+1} \pmod{10}$. For such c and $n = m + c \cdot 10^r$ we have n^3 ending in s. ∎

Note that if m in (1) has exactly r digits, then $m + c \cdot 10^r$ is the $(r+1)$-digit integer obtained by putting the digit c in front of m. Thus, the proof shows that if s is a d-digit string ending in 1, 3, 7, or 9, then there is an integer of exactly d digits (allowing initial zeros) whose cube ends in s.

There is an interesting connection here to elementary group theory. The strings s of length d may be regarded as the elements of $\mathbb{Z}/10^d\mathbb{Z}$, where $\mathbb{Z}$ is the ring of integers, and those strings ending in 1, 3, 7, or 9, being relatively prime to 10^d, constitute precisely the multiplicative group of invertible elements of $\mathbb{Z}/10^d\mathbb{Z}$. The fact that all of them occur as final digit sequences of cubes is another way of saying that the cubing map is bijective on this abelian group (that is, there are no elements of order 3).

Final digit 2, 4, 6, or 8 Whether there exists a cube ending in s when s ends in 2, 4, 6, or 8 depends on the relationship between the length d of s and the number of factors of 2 in s. It may be summarized by saying that if the number of factors of 2 in s is a multiple of 3, there is always a cube ending in s. Otherwise, there is such a cube if and only if there are at least as many factors of 2 in s as there are digits in s. For example, there is no cube ending in $314 = 2 \cdot 157$, but there is an integer n with n^3 ending in $3141592 = 2^3 \cdot 392699$. Indeed, $n = 4881798$ is such an integer. (The proofs of Theorem 1 and Lemma 2 enable one to find this n without having to resort to trial and error.)

We begin with two fundamental lemmas.

LEMMA 1. *Let s be a string of d decimal digits ending in 2, 4, 6, or 8. If $d \geq 3$ and n^3 ends in s, then $8 \mid s$.*

Proof. Let $n = 2m$. We have $n^3 = 8m^3 = s + 10^d k$ for some integer $k \geq 0$, and because $d \geq 3$ we see that 8 is a factor of 10^d, and therefore of s. ∎

LEMMA 2. *Let s be a string of d decimal digits ending in 2, 4, 6, or 8. If $s = 8t$ and there is an integer m such that m^3 ends in t, then there exists n^3 ending in s. (Here if s has initial zeros, t is assumed to have an equal number of them. For instance, if $s = 0024$, then $t = 003$.)*

Proof. We note that t has either d or $d-1$ digits. Suppose first that t has d digits. Then there are integers m and k such that $m^3 = t + k \cdot 10^d$, and $(2m)^3 = 8t + 8k \cdot 10^d = s + 8k \cdot 10^d$, so $(2m)^3$ ends in s.

In the remaining case t has $d-1$ digits and we may write $m^3 = t + k \cdot 10^{d-1}$. Then $d \geq 2$, and if c is any nonnegative integer,

$$(2m + c \cdot 10^{d-1})^3 = 8m^3 + 12m^2c \cdot 10^{d-1} + 6mc^2 10^{2d-2} + c^3 10^{3d-3}. \qquad (2)$$

Because $d \geq 2$ we have $3d - 3 > 2d - 2 \geq d$, and the last two terms in (2) have at least d final zeros. Therefore

$$\begin{aligned}(2m + c \cdot 10^{d-1})^3 &\equiv 8m^3 + 12m^2c \cdot 10^{d-1} \pmod{10^d} \\ &\equiv 8(t + k \cdot 10^{d-1}) + 12m^2c \cdot 10^{d-1} \pmod{10^d} \\ &\equiv s + (8k + 12m^2c)10^{d-1} \pmod{10^d}.\end{aligned}$$

Observe that for $(2m + c \cdot 10^{d-1})^3$ to end in s it suffices that $8k + 12m^2c$ be a multiple of 10. Because the integers mod 5 are a field, and $m^2 \not\equiv 0 \pmod 5$, we may

choose c from $\{0, 1, 2, 3, 4\}$ so that $m^2c \equiv k \pmod 5$. Then

$$8k + 12m^2c \equiv 3k + 2m^2c \equiv 3k + 2k \equiv 5k \pmod 5,$$

and because $8k + 12m^2c$ is even, it is a multiple of 10. Thus, for such a c, $(2m + c \cdot 10^{d-1})^3$ ends in s. ∎

THEOREM 2. *Let s be a string of d digits, $d \geq 1$, with s ending in 2, 4, 6, or 8. Write $s = 8^p t$, with $p \geq 0$ and $t \not\equiv 0 \pmod 8$.*

(A) If t is odd, there is an integer n with n^3 ending in s.

(B) Assume that $t \equiv 2,\ 4$ or $6 \pmod 8$.

(i) If $d \leq 3p + 1$, then there is n^3 ending in s.

(ii) If $d = 3p + 2$, then there is n^3 ending in s iff $t \equiv 4 \pmod 8$.

(iii) If $d \geq 3p + 3$, then there is no integer n with n^3 ending in s.

It seems worth repeating here that Theorem 2 states that when s ends in 2, 4, 6, or 8, if the number of factors of 2 in s is a multiple of 3, then there is always a cube ending in s. Otherwise there is such a cube if and only if there are at least as many factors of 2 in s as there are digits in s.

An interesting exercise at this point is to verify the claims made about cubes ending in 2, 4, 6, or 8 in the second introductory paragraph.

Proof of Theorem 2:

(A) Obviously t cannot end in 5, so this is immediate from Lemma 2 and Theorem 1.

(B) (i) We treat first the subcase $d \leq 3p$. Consider the integer s' obtained by putting the digit 1 in position $3p + 1$ in front of s (with intermediate zeros if $d < 3p$):

$$s' = 10^{3p} + s = 8^p \left(\frac{10^{3p}}{2^{3p}} + t \right) = 8^p (5^{3p} + t) = 8^p t'.$$

Observe that t' is odd. By part (A), there is n^3 ending in s', and therefore ending in s.

Now suppose that $d = 3p + 1$. This time let

$$s' = c \cdot 10^{3p+1} + s = 8^p \left(\frac{c \cdot 10^{3p+1}}{2^{3p}} + t \right) = 8^p (2c \cdot 5^{3p+1} + t),$$

with c to be determined. Note that, modulo 8, 5^r is 1 if r is even and is 5 if r is odd. In both cases, $2 \cdot 5^r \equiv 2 \pmod 8$. Thus $2c \cdot 5^{3p+1} \equiv 2c \pmod 8$, and we may choose c from $\{1, 2, 3\}$ to make $2c + t \equiv 0 \pmod 8$; that is, $2c \cdot 5^{3p+1} + t \equiv 0 \pmod 8$, and thus $s' = 8^q t'$, where $q \geq p + 1$ and the number of digits in s' is $d' = 3p + 2 < 3q$, and $8 \nmid t'$. By the first subcase above (if t' is even), or by part (A) (if t' is odd), there is n^3 ending in s' and therefore ending in s.

(ii) Let $d = 3p + 2$ and $t \equiv 4 \pmod 8$. We again append the digit 1 to the left end of s. This yields the string

$$s' = 10^{3p+2} + s = 8^p \left(\frac{10^{3p+2}}{2^{3p}} + t \right) = 8^p (4 \cdot 5^{3p+2} + t).$$

Now, 5^{3p+2} is either 1 or 5 modulo 8; in either case, $4 \cdot 5^{3p+2} \equiv 4 \pmod 8$, and we have $4 \cdot 5^{3p+2} + t \equiv 0 \pmod 8$. Thus $s' = 8^q t'$, with s' having length $d' = 3p + 3$, $q \geq p + 1$, and $8 \nmid t'$. Then by part (i), or (A), according as t' is even or odd, we know there exists n^3 ending in s', and therefore ending in s.

Conversely, suppose that $d = 3p + 2$ and that there is an integer n with n^3 ending in s:

$$n^3 = k \cdot 10^{3p+2} + s = 8^p k \left(\frac{10^{3p+2}}{2^{3p}} \right) + 8^p t = 2^{3p}(4k \cdot 5^{3p+2} + t). \quad (3)$$

We want to show that $t \equiv 4 (\mathrm{mod}\, 8)$. From (3) we see that $2^{3p} \mid n^3$, so $2^p \mid n$. Write $n = 2^p m$. Then

$$m^3 = 4k \cdot 5^{3p+2} + t. \quad (4)$$

But, as we've seen, $4 \cdot 5^{3p+2} \equiv 4 (\mathrm{mod}\, 8)$, so $4k \cdot 5^{3p+2}$ is 0 or 4 modulo 8, while t is assumed to be 2, 4, or 6 modulo 8. Because m is even we know that $m^3 \equiv 0 (\mathrm{mod}\, 8)$, so (4) shows that t can only be 4 modulo 8.

(iii) The case $p = 0$ is covered by Lemma 1. Thus assume that $d \geq 3p + 3$ and $p \geq 1$. Suppose, for purposes of a contradiction, that there is n^3 ending in s:

$$n^3 = s + k \cdot 10^d = 8^p t + k \cdot 10^d.$$

Because $d \geq 3p + 3$ we know that 2^{3p} divides 10^d, and therefore divides n^3, so $n = 2^p m$ for some integer m. Then

$$8^p m^3 = n^3 = 8^p t + k \cdot 10^d = 8^p t + 8^p k \left(\frac{10^{3p}}{2^{3p}} \right) 10^{d-3p},$$

so $m^3 = t + k \cdot 5^{3p} \cdot 10^{d-3p}$, and $d - 3p \geq 3$. Therefore $10^{d-3p} \equiv 0 (\mathrm{mod}\, 8)$, and $m^3 \equiv t (\mathrm{mod}\, 8)$. But this contradicts Lemma 1, provided that t has at least 3 digits. Well, s has $3p + 3$ digits, and $t = s/8^p > s/10^p$, so t has at least $2p + 3$ digits, and we have a contradiction. Thus no cube exists ending in s. ■

In view of Theorem 1 we may, for example, conclude from Lemma 2 that there is a cube with final four digits $0024 = (8)(003)$, as well as one with final 2004 digits $00\ldots024$. From Theorem 2 we see, for instance, that all multiples of 16 from $0016 = (16)(001)$ to $9984 = (16)(624)$ that do not end in 0 are final digit sequences of cubes because $d \leq 4$ and $16 = 2^4$.

Final digit 5 The results for cubes ending in 5 are quite similar to those for cubes ending in 2, 4, 6, or 8. Roughly speaking, the roles of the primes 2 and 5 are interchanged. Once again the results can be summarized by saying that if the number of factors of 5 in s is a multiple of 3, then there is always a cube ending in s, and otherwise there is a cube ending in s if and only if there are at least as many factors of 5 in s as there are digits in s. We will omit proofs here, since they are similar to the others. However, as an aid to readers who would like to try constructing the proof of Theorem 3 below, we include the statements of the preliminary lemmas that we used. Proving these and Theorem 3 would be an important preparation for students who would like to tackle any of the projects in the final section. We will be happy to supply more details to anyone who requests them.

LEMMA 3. *Let s be a string of $d = 3$ digits ending in 5. If n^3 ends in s, then s is one of* 125, 375, 625, *and* 875.

LEMMA 4. *If $d = 4$ and s ends in* 125, 375, 625, *or* 875, *then there is a cube ending in s.*

LEMMA 5. *If $d \geq 3$ and n^3 ends in s, (and s ends in 5), then* $125 \mid s$.

LEMMA 6. *If $s = 125t$ (t odd) and there is an integer m such that m^3 ends in t, then there is an integer n such that n^3 ends in s.*

THEOREM 3. *Let s be a string of d digits, $d \geq 1$, with s ending in 5. Write $s = 125^p t$, $p \geq 0$, where $125 \nmid t$.*

(A) If $5 \nmid t$, then there is an integer n with n^3 ending in s.

(B) Assume that $5 \mid t$.

(i) If $d \leq 3p + 1$, then there is n^3 ending in s.

(ii) If $d = 3p + 2$, then there is n^3 ending in s if and only if $25 \mid t$.

(iii) If $d \geq 3p + 3$, then there is no cube ending in s.

In other words, when s ends in 5, if the number of factors of 5 in s is a multiple of 3, then there is always a cube ending in s; otherwise, there is a cube ending in s if and only if there are at least as many factors of 5 in s as there are digits in s.

EXERCISE. Verify the statements made in the introduction about cubes ending in 5.

Final digit 0 Now that we know which digit strings s with final digit different from 0 are final digit strings of cubes, the corresponding question for strings ending in 0 is easy. If n^3 ends in 0, then n ends in 0. Write $n = 10^p m$, where m is not a multiple of 10. Then $n^3 = 10^{3p} m^3$, which has exactly $3p$ final zeros. We have, therefore, the following theorem:

THEOREM 4. *Let s be a string of digits ending in 0. Then s is the final digit string of a cube if and only if both of the following conditions hold: (i) the number of final zeros is a multiple of 3, and (ii) there is a cube m^3 ending in the digit string s' obtained by removing the final zeros of s.*

Exercises and projects for further study The work in this paper can be extended in various ways. Some extensions will be quite routine, but others may require some non-routine investigation. For starters, the proofs of the lemmas and theorems dealing with final digit 5 will be good exercises, not entirely routine. We offer several possibilities here for further student exercises or research projects.

THEOREM 5. *Let d be a positive integer and let s be a string of d decimal digits ending in 1, 3, 7, or 9. Let $r > 2$ be an integer, the decimal representation of which ends in 1, 3, 7, or 9. Then there is an integer n such that n^r ends in s.*

Proof. Exercise 1.

THEOREM 6. *Let s be a string of digits in base p, where p is a prime, and s does not end in 0. Let r be a positive integer that is not a multiple of p. Then there is an integer n such that n^r in base p ends in s.*

Proof. Exercise 2.

What if r is a multiple of p?

PROJECT 1. Theorem 5 is the counterpart of Theorem 1 for higher powers than cubes, with exponents ending in 1, 3, 7, or 9. Investigate the counterparts of Theorems 2, 3, and 4 for such powers.

PROJECT 2. Investigate final digit sequences for squares of integers; for fifth powers. For squares, for example, there are the following 22 two-digit endings: 00; 01, 21, 41, 61, 81; 04, 24, 44, 64, 84; 25; 16, 36, 56, 76, 96; 09, 29, 49, 69, 89. What digit sequences can precede these?

PROJECT 3. Investigate final digit sequences for integer powers in base 8; in base 12.

Acknowledgment. We are indebted to two anonymous referees for helpful suggestions.

REFERENCES

1. http://www.usiouxfalls.edu/academic/maa/contest/2000/Probs00.pdf
2. G. A. Heuer, Rational numbers generated by two integers, *Amer. Math. Monthly* **78** (1971), 996–997.
3. Leo Moser and Nathanial Macon, On the distribution of first digits of powers, *Scripta Math.* **16** (1950), 290–292.

Letter to the Editor

In "The Operator $(x\frac{d}{dx})^n$ and Its Applications to Series" (vol. **76**:5, December 2003), Peter Knopf mentions an expansion of $n!$ as an example of the occurrence of Bernoulli numbers (p. 370). That expansion is equivalent to the well-known expansion of $\log(n!)$, commonly known as Stirling's series. But these expansions are merely asymptotic; they do not converge. Although the terms of the series initially decrease in absolute value, they eventually increase without bound. Of course, this behavior then limits the accuracy to which $n!$ can be approximated using partial sums of the series. Therefore, the statement that such an expansion "allows us to compute $n!$ to arbitrarily good precision" is incorrect. Furthermore, the correct symbol to show the relationship between $n!$ and the asymptotic expansion is not $=$, but rather $\sim$, where "$\sim$" is read as "is asymptotic to." More information can be found in de Bruijn's *Asymptotic Methods in Analysis*, Erdélyi's *Asymptotic Expansions*, etc.

David W. Cantrell
221 25th Ave. East
Tuscaloosa, AL 35404-2528

MATHEMATICS MAGAZINE Editor Search

The MAA seeks to identify candidates to succeed Frank Farris as Editor of MATHEMATICS MAGAZINE when his term expires in December 2005. The new editor would be Editor Elect in 2005, handling all new manuscript submissions, and would serve as Editor for the five years 2006–2010.

Questions about the position and its workload can be addressed to Frank Farris (FFarris@scu.edu); questions about MAA support for the editor's work can be addressed to the MAA's Director of Publications, Don Albers (dalbers@maa.org).

Each applicant should submit a resume, names of references, and a statement of interest to the chair of the Search Committee,

Deanna Haunsperger, Department of Mathematics and Computer Science
Carleton College
One North College
Northfield, MN 55057, dhaunspe@carleton.edu.

Nominations are also welcome if you know someone who would be an outstanding editor. Applications and nominations will be accepted until the position is filled, although preference will be given to applications received by early May.

PROBLEMS

ELGIN H. JOHNSTON, *Editor*
Iowa State University

Assistant Editors: RĂZVAN GELCA, Texas Tech University; ROBERT GREGORAC, Iowa State University; GERALD HEUER, Concordia College; VANIA MASCIONI, Ball State University; PAUL ZEITZ, The University of San Francisco

Proposals

To be considered for publication, solutions should be received by September 1, 2004.

1691. *Proposed by Murray S. Klamkin, University of Alberta, Edmonton, Alberta, Canada.*

Let p, r, and n be integers with $1 < r < n$, and let k be a positive constant. Determine the maximum and minimum values of

$$\sum_{j=1}^{n} \frac{t_j^p}{1+kt_j},$$

where $x_i \geq 0$, $1 \leq i \leq n$ with $x_1 + x_2 + \cdots + x_n = 1$, and $t_j = x_j + x_{j+1} + \cdots + x_{j+r-1}$, where $x_{i+n} = x_i$.

1692. *Proposed by Mario Catalani, Department of Economics, University of Torino, Torino, Italy.*

Let $F_n = F_n(x, y)$ and $L_n = L_n(x, y)$ be the bivariate Fibonacci and Lucas polynomials, defined by

$$F_0 = 0, \quad F_1 = 1, \quad F_n = xF_{n-1} + yF_{n-2}, \quad n \geq 2$$
$$L_0 = 2, \quad L_1 = x, \quad L_n = xL_{n-1} + yL_{n-2}, \quad n \geq 2.$$

Assume that $x \neq 0$, $y \neq 0$, and $x^2 + 4y \neq 0$. Prove that

$$F_n\left(L_{2m+1}, y^{2m+1}\right) = \frac{F_{n(2m+1)}(x, y)}{F_{2m+1}(x, y)} \quad \text{and} \quad F_n\left(L_{2m}, -y^{2m}\right) = \frac{F_{2mn}(x, y)}{F_{2m}(x, y)}.$$

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.

Solutions and new proposals should be mailed to Elgin Johnston, Problems Editor, Department of Mathematics, Iowa State University, Ames IA 50011, or mailed electronically (ideally as a LaTeX file) to ehjohnst@iastate.edu. All communications should include the readers name, full address, and an e-mail address and/or FAX number.

1693. *Proposed by Erwin Just (Emeritus) and Norman Schaumberger (Emeritus), Bronx Community College of the City University of New York, Bronx, NY.*

Let $A = (p, q)$, $B = (p^2, q^2)$, and $C = (p^3, q^3)$ be the vertices of a nondegenerate triangle.

a. For how many pairs (p, q) is triangle ABC equilateral?

b. If p or q is rational, can triangle ABC be equilateral?

1694. *Proposed by Óscar Ciaurri, Universidad de La Rioja, La Roija, Spain.*

For $\alpha \geq 1$, a sequence $\{b_n\}_{n\geq 0}$ is defined by

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

A sequence $\{a_n\}_{n\geq 0}$ is then defined by $a_0 = 1$ and, for $n \geq 1$, by $\sum_{k=0}^{n} a_{n-k} b_k = 0$. Find the value of $\limsup_{n\to\infty} \sqrt[2n]{|a_n|}$.

1695. *Proposed by Shalom Feigelstock, Bar-Ilan University, Ramat-Gan, Israel.*

A field F is a real field if -1 cannot be written as a sum of squares. Prove that a field F is a real field if and only if for every $n \times m$ matrix A with entries from F, $\operatorname{rank}(A) = n$ implies that AA^T is invertible.

Quickies

Answers to the Quickies are on page 162.

Q939. *Proposed by Daniel Ashlock, Iowa State University, Ames, IA.*

Let n, k be positive integers with $n \geq 3$ and $k < n/2$. The *generalized Petersen graph* $P_{n,k}$ is a graph on $2n$ vertices and is defined as follows. Let the vertices be two copies of $\mathbb{Z}_n$, denoted respectively by $\{0, 1, 2, \ldots, (n-1)\}$ and $\{0', 1', 2', \ldots, (n-1)'\}$. Edges have, as endpoints, any of the following pairs of vertices: (i) (a, b) with $a - b \equiv 1 (\mathrm{mod}\, n)$, (ii) (a', b') with $a' - b' \equiv k (\mathrm{mod}\, n)$, (iii) (a, a'). The girth of a graph is defined to be the length of the shortest cycle in the graph. Prove that the girth of the generalized Petersen graph $P_{n,k}$ is less than or equal to 8.

Q940. *Proposed by Jim Delany, California Polytechnic State University, San Luis Obispo, CA.*

For positive integers m and n, define

$$\phi_m(n) = \begin{cases} \phi(n) & n \mid m \\ 0 & \text{otherwise,} \end{cases}$$

where ϕ is Euler's totient function. Prove that $\sum_{d|n} \phi_m(d) = \gcd(m, n)$.

Solutions

Extremes on a Sphere **April 2003**

1667. *Proposed by Murray S. Klamkin, University of Alberta, Alberta, Canada.*

Let a, b, and c be nonnegative constants. Determine the maximum and minimum values of

$$\sqrt{a^2x^2 + b^2y^2 + c^2z^2} + \sqrt{a^2y^2 + b^2z^2 + c^2x^2} + \sqrt{a^2z^2 + b^2x^2 + c^2y^2},$$

subject to $x^2 + y^2 + z^2 = 1$.

I. *Solution by Michael Andreoli, Miami-Dade Community College, North Campus, Miami, FL.*

Let $S = \sqrt{a^2x^2 + b^2y^2 + c^2z^2} + \sqrt{a^2y^2 + b^2z^2 + c^2x^2} + \sqrt{a^2z^2 + b^2x^2 + c^2y^2}$. We show that if $x^2 + y^2 + z^2 = 1$, then the maximum and minimum values of S are $\sqrt{3(a^2 + b^2 + c^2)}$ and $a + b + c$, respectively.

Because the square root function is concave, Jensen's inequality implies

$$\frac{1}{3}S \le \sqrt{\frac{1}{3}\left((a^2x^2 + b^2y^2 + c^2z^2) + (a^2y^2 + b^2z^2 + c^2x^2) + (a^2z^2 + b^2x^2 + c^2y^2)\right)}$$

$$= \sqrt{\frac{1}{3}(a^2 + b^2 + c^2)(x^2 + y^2 + z^2)} = \sqrt{\frac{1}{3}(a^2 + b^2 + c^2)}.$$

Thus $S \le \sqrt{3(a^2 + b^2 + c^2)}$. Because equality holds if $x = y = z = 1/\sqrt{3}$, this value of S is the maximum value.

Because $x^2 + y^2 + z^2 = 1$, it also follows from Jensen's inequality that

$$\sqrt{a^2x^2 + b^2y^2 + c^2z^2} \ge \sqrt{a^2}x^2 + \sqrt{b^2}y^2 + \sqrt{c^2}z^2 = ax^2 + by^2 + cz^2,$$

and similar inequalities hold for $\sqrt{a^2y^2 + b^2z^2 + c^2x^2}$ and $\sqrt{a^2z^2 + b^2x^2 + c^2y^2}$. It follows that

$$S \ge (a + b + c)(x^2 + y^2 + z^2) = a + b + c.$$

Because equality occurs when $x = 1$ and $y = z = 0$, this value is the minimum value for S.

II. *Solution by Michel Bataille, Rouen, France.*

Let

$$u = \sqrt{a^2x^2 + b^2y^2 + c^2z^2}, \quad v = \sqrt{a^2y^2 + b^2z^2 + c^2x^2}, \quad w = \sqrt{a^2z^2 + b^2x^2 + c^2y^2},$$

and $S = u + v + w$, and observe that $u^2 + v^2 + w^2 = a^2 + b^2 + c^2$ when $x^2 + y^2 + z^2 = 1$. Hence, by the Cauchy-Schwarz inequality,

$$S = u + v + w \le \sqrt{1^2 + 1^2 + 1^2}\sqrt{u^2 + v^2 + w^2} = \sqrt{3(a^2 + b^2 + c^2)},$$

with equality when $x = y = z = 1/\sqrt{3}$.

Again by the Cauchy-Schwarz inequality,

$$uv = \sqrt{(ax)^2 + (by)^2 + (cz)^2}\sqrt{(cx)^2 + (ay)^2 + (bz)^2} \ge cax^2 + aby^2 + bcz^2,$$

with analogous inequalities for vw and wu. Summing we find

$$uv + vw + wu \ge (ab + bc + ca)(x^2 + y^2 + z^2) = ab + bc + ca.$$

Thus

$$S^2 = a^2 + b^2 + c^2 + 2(uv + vw + wu) \ge a^2 + b^2 + c^2 + 2(ab + bc + ca)$$
$$= (a + b + c)^2,$$

so $S \ge a + b + c$. Equality holds when $x = 1$ and $y = z = 0$.

Also solved by Robert A. Agnew, Roy Barbara (Lebanon), Jean Bogaert (Belgium), Mario Catalani (Italy), Con Amore Problem Group (Denmark), Chip Curtis, Knut Dale (Norway), Daniele Donini (Italy), Robert L. Doucette, Charles Flemming, Ovdiu Furdui, Julien Grivaux (France), Chris Hill, Danrun Huang, Steve Kaczkowski, Elias Lampakis (Greece), Kee-Wai Lau (China), Magli Pierluigi (Italy), Heinz-Jürgen Seiffert (Germany), Achilleas Sinefakopoulos, Michael Vowe (Switzerland), Paul Weisenhorn (Germany), Li Zhou, and the proposer. There were six incorrect submissions.

The Signs Are A' Changin **April 2003**

1668. *Proposed by Steve Butler, Provo, UT.*

Let f be a real valued function defined on an open interval I containing $[a, b]$. Assume that f has a continuous second derivative on I and that there is a single line tangent to the graph of $y = f(x)$ at $(a, f(a))$ and $(b, f(b))$. Prove that if $f''(x)$ is not identically zero on (a, b), then $f''(x)$ must change sign at least twice on (a, b).

Solution by Robert Doucette, McNeese State University, Lake Charles, LA.

Define the function g on I by

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) - f(a).$$

Then $f''(x) = g''(x)$ for all $x \in I$. Because g'' is not identically 0 on (a, b), g is not identically 0 on (a, b). Because $g(a) = g(b) = 0$, g is not monotone on (a, b). Thus g' must change sign on (a, b), so we can find real numbers $x_1, x_2 \in (a, b)$ with $g'(x_1) < 0$ and $g'(x_2) > 0$. We may assume that $x_1 < x_2$. If not, replace g by $-g$. By the Mean Value Theorem there exist real numbers x_3, x_4, x_5 with $a < x_3 < x_1 < x_4 < x_2 < x_5 < b$ and such that

$$g''(x_3) = \frac{g'(x_1) - g'(a)}{x_1 - a}, \quad g''(x_4) = \frac{g'(x_2) - g'(x_1)}{x_2 - x_1},$$

and

$$g''(x_5) = \frac{g'(b) - g'(x_2)}{b - x_2}.$$

Because $g'(a) = g'(b) = 0$, we have $g''(x_3) < 0$, $g''(x_4) > 0$, and $g''(x_5) < 0$. Therefore, f'' changes sign at least twice in (a, b).

Note. Some readers attempted to prove that there must be two points at which f'' changes sign. However, this need not be true. As an example, consider the function f on $[0, 6]$ with $f(0) = f'(0) = 0$ and such that the graph of $y = f''(x)$ is the polygonal path joining the points $(0, 1)$, $(1, 0)$, $(2, 0)$, $(3, -1)$, $(4, 0)$, $(5, 0)$, and $(6, 1)$.

Also solved by Tsehaye Andebrhan, Michel Bataille (France), Marco Carone (Canada), Con Amore Problem Group (Denmark), Chip Curtis, Knut Dale (Norway), Robert L. Doucette, Ovidiu Furdui, Julien Grivaux (France), James Henderson, Ming-Lun Ho, Danrun Huang, Stephen Kaczkowski, Patricia L. Kiihne, Elias Lampakis (Greece), Allen Mauney, James M. Meehan, Northwestern University Math Problem Solving Group, Rolf Richberg (Germany), Jawad Sadek, Cornelius Stallman, Gerald Thompson, Xiaoshen Wang, Paul Weisenhorn (Germany), Paul Wilcock, Michael Woltermann, Wang Yi (China), Li Zhou, and the proposer. There was one solution with no name, and one unreadable submission.

A Similar Pair **April 2003**

1669. *Proposed by Ali Nabi Duman (student), Bilkent University, Turkey.*

Let ABC be a triangle and let E be the midpoint of $\overline{BC}$. A circle passing through A and C intersects $\overline{BA}$ and $\overline{BC}$ in points G and E respectively. Let D be the midpoint of $\overline{EC}$. A line through D and perpendicular to $\overline{BC}$ intersects $\overline{AC}$ at F, with $3AF = FC$. Prove that triangle FDG is similar to triangle ABC.

Solution by Marty Getz and Dixon Jones, University of Alaska Fairbanks, Fairbanks, AK.

Let H be the point on $\overline{AB}$ such that $\overline{HF}$ is parallel to $\overline{BC}$. Because $HF = BC/4 = DC = ED$, quadrilateral $HDCF$ is a parallelogram and quadrilateral $HEDF$ is a rectangle. Next observe that in quadrilaterals $HGED$ and $AGEC$,

$$\angle HGE = \angle AGE, \quad \angle GED = \angle GEC, \quad \angle EDH = \angle ECA, \quad \text{and} \quad \angle DHG = \angle CAG.$$

Thus, because $AGEC$ is cyclic, the same is true of $HGED$, so G is on the circumcircle of $HEDF$. Because quadrilateral $HGDF$ is cyclic, we have

$$\angle DFG = \angle DHG = \angle CAB \quad \text{and} \quad \angle FGD = \angle FHD = \angle ACB.$$

It follows that $\triangle FDG$ is similar to $\triangle ABC$.

Also solved by Herb Bailey, Michel Bataille (France), Gerald D. Brown, Alper Cay (Turkey), Knut Dale (Norway), Daniele Donini (Italy), Robert L. Doucette, Barbara Falkowski, Julien Grivaux (France), John G. Heuver (Canada), Geoffrey A. Kandall, Victor Y. Kutsenok, Elias Lampakis (Greece), Junaid N. Mansuri, Julio Castiñeira Merino (Spain), Rolf Richberg (Germany), Francisco Bellot Rosado and María Ascensión López Chamorro (Spain), Helen Skala, Paul Weisenhorn (Germany), Michael Woltermann, Wang Yi (China), Li Zhou, and the proposer.

Cosines and ϕ **April 2003**

1670. *Proposed by Erwin Just (Emeritus) and Norman Schaumberger (Emeritus), Bronx Community College of the City University of New York, Bronx, NY.*

Let $n \geq 3$ be an odd integer and let $\{a_1, a_2, \ldots, a_{\phi(n)}\}$ be the set of positive integers less than n and relatively prime to n. Prove that

$$\left| \prod_{k=1}^{\phi(n)} \cos\left(\frac{a_k \pi}{n}\right) \right| = \frac{1}{2^{\phi(n)}}.$$

Solution by Mikhail Goldenberg and Mark Kaplan, The Ingenuity Project at Baltimore Polytechnic Institute, Baltimore, MD.

Because n is odd and $\{a_k : 1 \leq k \leq \phi(n)\}$ is reduced residue system modulo n, the same is true of the set $\{2a_k : 1 \leq k \leq \phi(n)\}$. Thus, for each k, $2a_k = a_{l(k)} + \epsilon n$, where $\epsilon \in \{0, 1\}$ and $\{l(k) : 1 \leq k \leq \phi(n)\}$ is a permutation of $\{1, 2, \ldots, \phi(n)\}$. Now let

$$A = \left| \prod_{k=1}^{\phi(n)} \cos\left(\frac{a_k \pi}{n}\right) \right| \quad \text{and} \quad B = \left| \prod_{k=1}^{\phi(n)} \sin\left(\frac{a_k \pi}{n}\right) \right|,$$

and note that $B \neq 0$. Then

$$\begin{aligned} 2^{\phi(n)} AB &= \left| \prod_{k=1}^{\phi(n)} 2 \sin\left(\frac{a_k \pi}{n}\right) \cos\left(\frac{a_k \pi}{n}\right) \right| = \left| \prod_{k=1}^{\phi(n)} \sin\left(\frac{2a_k \pi}{n}\right) \right| \\ &= \left| \prod_{k=1}^{\phi(n)} \sin\left(\frac{a_{l(k)} \pi}{n} + \epsilon \pi\right) \right| = \left| \prod_{k=1}^{\phi(n)} \sin\left(\frac{a_{l(k)} \pi}{n}\right) \right| = B. \end{aligned}$$

Hence, $A = 1/2^{\phi(n)}$.

Also solved by Tsehaye Andebrhan, Roy Barbara (Lebanon), Michel Bataille (France), Jany C. Binz (Switzerland), Jean Bogart (Belgium), Minh Can, John Christopher, Chip Curtis, Con Amore Problem Group (Denmark), Jim Delany, Daniele Donini (Italy), Robert L. Doucette, Ovidiu Furdui, Julien Grivaux (France), Chris Hill, Elias Lampakis (Greece), Northwestern University Math Problem Solving Group, Magli Pierluigi (Italy), Rolf Richberg (Germany), Achilleas Sinefakopoulos, Nicholas C. Singer, SMSU Problem Solving Group, Albert Stadler (Switzer-

land), H. T. Tang, Andrew Uzzell, Paul Weisenhorn (Germany), Chu Wenchang and Di Claudio Leontina Veliana (Italy), Michael Woltermann, Li Zhou, Paul Zwier, and the proposer.

Integer Triangles and the Golden Mean **April 2003**

1671. *Proposed by M. N. Deshpande, Institute of Science, Nagpur, India.*

Let $\mathcal{T}$ be the set of triangles ABC for which there is a point D on BC such that segments AB, BD, AD, DC, and AC have integral length and $\angle ACD = \frac{1}{2}\angle ABC = \frac{1}{3}\angle ADB$.

(a) Characterize the sets $\{a, b, c\}$ that are sets of side lengths of triangles in $\mathcal{T}$.

(b) Find the triangle of minimum area in $\mathcal{T}$.

Solution by Jim Delany, California Polytechnic State University, San Luis Obispo, CA.

A triangle is in $\mathcal{T}$ if and only if its side lengths satisfy

$$(a, b, c) = \left((p^2 - q^2)^2 k,\ pq(p^2 - q^2)k,\ q^2(p^2 - q^2)k\right),$$

were p, q, k are positive integers, p and q are relatively prime, and $\phi q < p < 2q$, where $\phi = 2\cos(\pi/5) = (\sqrt{5} + 1)/2$. The triangle of minimal area is produced when $(p, q, r) = (5, 3, 1)$ and has side lengths $(a, b, c) = (256, 240, 144)$.

First consider $\triangle ABC$. Let a, b, c be the lengths of the sides opposite $\angle A$, $\angle B$, $\angle C$, respectively, let $\angle C = \theta$ and let $t = 2\cos\theta$. Then $\angle B = 2\theta$ and $\angle A = \pi - 3\theta$, so $0 < \theta < \pi/3$ and $1 < t < 2$. By the Law of Cosines,

$$t = 2\cos\theta = \frac{a^2 + b^2 - c^2}{ab}.$$

This is a rational number, so let $t = p/q$ where $(p, q) = 1$. Then $q < p < 2q$. By the Law of Sines,

$$\frac{c}{\sin\theta} = \frac{b}{\sin 2\theta} = \frac{b}{2\sin\theta\cos\theta} = \frac{a}{\sin 3\theta} = \frac{a}{3\sin\theta - 4\sin^3\theta}.$$

Hence

$$a = \frac{c(3\sin\theta - 4\sin^3\theta)}{\sin\theta} = c(4\cos^2\theta - 1) = (t^2 - 1)c$$

and

$$b = \frac{2c\sin\theta\cos\theta}{\sin\theta} = tc.$$

Because $a = (p^2/q^2 - 1)c = (p^2 - q^2)c/q^2$ and q and $p^2 - q^2$ are relatively prime, we must have $q^2 \mid c$. Let $c = mq^2$. Then

$$a = m(p^2 - q^2), \qquad b = tc = mpq, \quad \text{and} \quad c = mq^2,$$

where p and q are relatively prime positive integers and $q < p < 2q$.

Now consider the point D. Because $\angle ADB = 3\theta$, we have $\angle BAD = \pi - 5\theta$, so $t > 2\cos\pi/5 = \phi$. In addition, $\angle CAD = 2\theta$, so $\triangle DAC$ is similar to $\triangle ABC$. Thus $CD/b = b/a$, so $CD = b^2/a = mp^2q^2/(p^2 - q^2)$. Because CD is an integer and $p^2 - q^2$ is relatively prime to p and q, it must be the case that $(p^2 - q^2) \mid m$. Let $m = (p^2 - q^2)k$. Then $CD = p^2q^2k$. Using similar triangles we also find $AD = bc/a = mpq^3/(p^2 - q^2) = pq^3k$. Replacing m by $(p^2 - q^2)k$ in the above expres-

sions for a, b, and c leads to the characterization

$$a = (p^2 - q^2)^2 k, \qquad b = pq(p^2 - q^2)k, \qquad c = q^2(p^2 - q^2)k,$$

where k is an integer and $\phi q < p < 2q$.

To minimize the area $\mathcal{A}$ of the triangle we minimize

$$\begin{aligned}\mathcal{A}^2 &= \frac{1}{4}a^2b^2\sin^2\theta = \frac{1}{4}a^2b^2\left(1 - \left(\frac{t}{2}\right)^2\right)\\ &= \frac{1}{4}\left((p^2 - q^2)^4k^2\right)\left(p^2q^2(p^2 - q^2)^2k^2\right)\left(\frac{4q^2 - p^2}{4q^2}\right)\\ &= \frac{1}{16}p^2(p^2 - q^2)^6(4q^2 - p^2)k^4.\end{aligned}$$

For minimal area it is clear that $k = 1$. The only acceptable pair p, q with $q < 4$ is $p = 5$, $q = 3$. For these values $a = 256$, $b = 240$, $c = 144$ and $\mathcal{A}^2 = (5120\sqrt{11})^2 = 288{,}358{,}400$. To see that this is a minimum, note that since $p > \phi q > 3q/2$ we have $p^2 > 9q^2/4$ and $p^2 - q^2 > 5q^2/4$. Also, because $p \le 2q - 1$ we have $4q^2 - p^2 \ge 4q^2 - (2q - 1)^2 = 4q - 1$. Thus, with $k = 1$,

$$\mathcal{A}^2 > \frac{1}{16}\left(\frac{9}{4}q^2\right)\left(\frac{5}{4}q^2\right)^6(4q - 1) = \frac{9 \cdot 5^6}{4^9}q^{14}(4q - 1).$$

When $q \ge 4$ we have $A^2 \ge 9 \cdot 5^6 \cdot 4^5 \cdot 15 = 2{,}160{,}000{,}000$. Thus the minimum area occurs when $p = 5$, $q = 3$, and $k = 1$.

Also solved by Jany C. Binz (Switzerland), Chip Curtis, Daniele Donini (Italy), Robert L. Doucette, Ming-Lun Ho, Elias Lampakis (Greece), Paul Weisenhorn (Germany), Li Zhou, and the proposer. Michel Bataille (France) solved part 1. There was one submission with no name and one incorrect submission.

Answers

Solutions to the Quickies from page 157.

A939. Consider the list of vertices

$$0, 1, 1', (k + 1)', k + 1, k, k', 0', 0.$$

Each adjacent pair in the list gives the endpoints of an edge in $P_{n,k}$. This list exhibits a cycle of length 8 unless $k = 1$. If $k = 1$, then the list of vertices $0, 0', 1', 1, 0$ exhibits a cycle of length 4. Thus, the girth of the generalized Petersen graph is less than or equal to 8.

A940. It is well known that for positive integer N, $\sum_{d|N}\phi(d) = N$. Thus,

$$\sum_{d|n}\phi_m(d) = \sum_{\substack{d|n\\ d|m}}\phi(d) = \sum_{d|\gcd(m,n)}\phi(d) = \gcd(m, n).$$

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.

Borwein, Jonathan, and David Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A K Peters, 2003; x + 288 pp, $45. ISBN 1–56881–211–6.

This book presents "the rationale and historical context of experimental mathematics," together with "*accessible* examples of modern mathematics where intelligent computing plays a significant role." (A companion volume, *Experimentation in Mathematics: Computational Paths to Discovery*, offers additional examples.) Topics include prime numbers and the zeta function, normality of numbers, lots about pi "and its friends," high-precision arithmetic, and constructive approaches to many special functions. Experimental mathematics is dedicated to expanding mathematical knowledge rapidly, occasionally with a "temporary relaxation in rigor." As a sample, consider the Nilakantha-Gregory series

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots\right),$$

After 5 million terms, the series differs from π in the 7th decimal place and periodically thereafter. "Such anomalous behavior begs explanation," and the authors proceed to explain exactly why and how this behavior occurs. They also include challenge problems, Internet sites, and enough mathematical connections to astonish anyone. (Note: An insert advises that some citations point to the wrong items in the references; those and other errata, plus other useful information, can be found at `http://www.expmath.info`).

Campbell-Kelly, Martin, Mary Croarken, Raymond Flood, and Eleanor Robson (eds.), *The History of Mathematical Tables: From Sumer to Spreadsheets*, Oxford University Press, 2003; x + 361 pp, $89.50. ISBN 0–19–850841–7.

I suspect that my students in introductory statistics find it quite quaint to have to consult tables of the normal, t, and χ^2 distributions. Of course, specialized calculators ("tables in dynamic form") are available. Still, even today, despite the ubiquity of $10 scientific calculators, some trigonometry textbooks include trig tables in the back (current students tell of learning interpolation in high school). Well, even after 4,500 years, mathematical tables are still important. This book's dozen essays treat Babylonian tables, log tables, actuarial tables, astronomical tables, table-making with mechanical and electronic calculators, and finally spreadsheets. But there are no tables in the back of the book.

Fauvel, John, Raymond Flood, and Robin Wilson (eds.), *Music and Mathematics: From Pythagoras to Fractals*, Oxford University Press, 2003; vi + 189 pp, $74.50. ISBN 0–19–851187–6.

The 10 essays here treat tuning and temperament, Kepler's cosmology, the science of musical sound, the spacing of frets on stringed instruments, Helmholtz's physiological basis for music, frieze patterns in music, change ringing, mathematical ideas in composition, projective planes and the 7-tone system, and composing with fractals.

Berlinghoff, William P., and Fernando Q. Gouvêa, *Math through the Ages: A Gentle History for Teachers and Others*, expanded ed.. Oxton House Publishers and the MAA, 2004; x + 273 pp, $39.95 (minus discount to members). ISBN 0–88385–736–7.

This book begins with a sketch of the history of mathematics "in a large nutshell (56 pp but only 5 equations), followed by 25 sketches about common ideas in basic mathematics (whole numbers, numerals, zero, ..., noneuclidean geometries, statistics, computers, boolean algebra, infinite sets). The book was conceived as a textbook for the history of mathematics course commonly required for teacher certification in mathematics, hence the title and the concentration on basic mathematics. This expanded edition includes questions (which require thinking and occasionally research) and projects (which require research). The authors include print and Internet sources for further investigation.

Grattan-Guinness, I. (ed.), *Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*, 2 vols., Johns Hopkins Press, 2003; xiv + 1806 pp, $49.95 per volume (P). ISBN 0–8018–7396–7, 0–8018–7397–5.

This is the collection for those whose appetite for history of mathematics is far greater than can be satiated by *Math through the Ages*. Originally published in 1994 but now reissued in paperback, this encyclopedia contains 180 articles devoted primarily to the history of mathematics. Here, too, the density of equations is low.

Bekken, Otto B., and Reidar Mosvold, *Study the Masters: The Abel-Fauvel Conference*, Gimlekollen Mediacentre—Kristiansand, June 12–15, 2002, Nationellt Centrum för Matemaikutbildning (Göteborgs Universitet, Vera Sandbergs allé 5A, SE-412 96 Göteborg, Sweden), 2003; vii + 310 pp, Sw. kr. 250 (P). ISBN 91–85143–00–6.

In 1995, the MAA published *Learn from the Masters*, a book of conference proceedings about using ideas from the history of mathematics to motivate and enhance the understanding of mathematics. This book, which gives the proceedings of the second conference on that topic at the same location, takes its title from Abel's dictum "study the masters and not the pupils." Part of the volume is devoted to topics related to Abel (on the 200th anniversary of his birth), and part is a tribute to the late John Fauvel in which the themes range from Nordic topics to why history is important in mathematics teaching, from "back to Euclid" to 17th-century topics. (Thanks to David Pengelley.)

Crannell, Annalisa, Gavin LaRose, Thomas Ratliff, and Elyn Rykken, *Writing Projects for Mathematics Courses: Crushed Clowns, Cars, and Coffee to Go*, MAA, 2004; viii + 119 pp, $28.50 (minus discount to members) (P). ISBN 0–88385–735–9.

This book offers writing projects for mathematics courses in calculus, differential equations, and "survey of mathematics." Each project is cast as a letter from a client and specifies the concepts, student reactions, needed technology, degree of realism (from real applications to "completely fabricated"), and a solution. Once students get beyond the artificiality, they may enjoy these projects; but with the authors' solutions averaging half a dozen lines, where's all the writing to be done—apart from the much-needed learning of such economy of expression? Two full solution papers (3 pp and 5 pp) are given, one with a grading rubric and the other with a grading checklist.

Tomlinson, Tommy, A beautiful find, *Charlotte Observer* (16 November 2003), `http://www.charlotte.com/mld/charlotte/news/columnists/tommy_tomlinson/7274830.htm` .

This column by one of the country's top local columnists, which won a journalism award for 2003, relates the experience of mathematician John Swallow (Davidson University) over five years in arriving at a proof of a theorem about Brauer groups. The piece gives a good sense of what mathematicians do and how they work, in a novel format.

NEWS AND LETTERS*

32nd United States of America Mathematical Olympiad

April 29 and 30, 2003

edited by Titu Andreescu and Zuming Feng

Problems

1. Prove that for every positive integer n there exists an n-digit number divisible by 5^n all of whose digits are odd.
2. A convex polygon $\mathcal{P}$ in the plane is dissected into smaller convex polygons by drawing all of its diagonals. The lengths of all sides and all diagonals of the polygon $\mathcal{P}$ are rational numbers. Prove that the lengths of all sides of all polygons in the dissection are also rational numbers.
3. Let $n \neq 0$. For every sequence of integers

$$a = a_0, a_1, a_2, \ldots, a_n$$

satisfying $0 \leq a_i \leq i$, for $i = 0, \ldots, n$, define another sequence

$$t(a) = t(a)_0, t(a)_1, t(a)_2, \ldots, t(a)_n$$

by setting $t(a)_i$ to be the number of terms in the sequence a that precede the term a_i and are different from a_i. Show that, starting from any sequence a as above, fewer than n applications of the transformation t lead to a sequence b such that $t(b) = b$.
4. Let ABC be a triangle. A circle passing through A and B intersects segments AC and BC at D and E, respectively. Rays BA and ED intersect at F while lines BD and CF intersect at M. Prove that $MF = MC$ if and only if $MB \cdot MD = MC^2$.
5. Let a, b, c be positive real numbers. Prove that

$$\frac{(2a+b+c)^2}{2a^2+(b+c)^2} + \frac{(2b+c+a)^2}{2b^2+(c+a)^2} + \frac{(2c+a+b)^2}{2c^2+(a+b)^2} \leq 8.$$

6. At the vertices of a regular hexagon are written six nonnegative integers whose sum is 2003. Bert is allowed to make moves of the following form: he may pick a vertex and replace the number written there by the absolute value of the difference between the numbers written at the two neighboring vertices. Prove that Bert can make a sequence of moves, after which the number 0 appears at all six vertices.

Solutions

Note: For interested readers, the editors recommend the *USA and International Mathematical Olympiads 2003*. Many of the problems are presented there, together with a collection of remarkable solutions developed by the examination committees, contestants, and experts, during or after the contests.

*Additional News and Letters appear on p. 155.

1. We proceed by induction. The property is clearly true for $n = 1$. Assume that $N = a_1a_2 \ldots a_n$ is divisible by 5^n and has only odd digits. Consider the numbers

$$N_1 = 1a_1a_2 \ldots a_n = 1 \cdot 10^n + 5^n M = 5^n(1 \cdot 2^n + M),$$
$$N_2 = 3a_1a_2 \ldots a_n = 3 \cdot 10^n + 5^n M = 5^n(3 \cdot 2^n + M),$$
$$N_3 = 5a_1a_2 \ldots a_n = 5 \cdot 10^n + 5^n M = 5^n(5 \cdot 2^n + M),$$
$$N_4 = 7a_1a_2 \ldots a_n = 7 \cdot 10^n + 5^n M = 5^n(7 \cdot 2^n + M),$$
$$N_5 = 9a_1a_2 \ldots a_n = 9 \cdot 10^n + 5^n M = 5^n(9 \cdot 2^n + M).$$

The numbers $1 \cdot 2^n + M$, $3 \cdot 2^n + M$, $5 \cdot 2^n + M$, $7 \cdot 2^n + M$, $9 \cdot 2^n + M$ give distinct remainders when divided by 5. Otherwise the difference of some two of them would be a multiple of 5, which is impossible, because neither 2^n is a multiple of 5, nor is the difference of any two of the numbers 1, 3, 5, 7, 9. It follows that one of the numbers N_1, N_2, N_3, N_4, N_5 is divisible by $5^n \cdot 5$, and the induction is complete. With a bit more effort, one can show that this number is unique.

2. Let $\mathcal{P} = A_1A_2 \ldots A_n$, where n is an integer with $n \geq 3$. The problem is trivial for $n = 3$ because there are no diagonals and thus no dissections. We assume that $n \geq 4$. Our proof is based on the following Lemma.

LEMMA. *Let $ABCD$ be a convex quadrilateral such that all its sides and diagonals have rational lengths. If segments AC and BD meet at P, then segments AP, BP, CP, DP all have rational lengths.*

It is clear by the Lemma that the desired result holds when $\mathcal{P}$ is a convex quadrilateral. Let A_iA_j $(1 \leq i < j \leq n)$ be a diagonal of $\mathcal{P}$. Assume that $C_1, C_2, \ldots, C_m$ are the consecutive division points on diagonal A_iA_j (where point C_1 is the closest to vertex A_i and C_m is the closest to A_j). Then the segments $C_\ell C_{\ell+1}$, $1 \leq \ell \leq m - 1$, are the sides of all polygons in the dissection. Let C_ℓ be the point where diagonal A_iA_j meets diagonal A_sA_t. Then quadrilateral $A_iA_sA_jA_t$ satisfies the conditions of the Lemma. Consequently, segments A_iC_ℓ and $C_\ell A_j$ have rational lengths. Therefore, segments $A_iC_1, A_iC_2, \ldots, A_jC_m$ all have rational lengths. Thus, $C_\ell C_{\ell+1} = AC_{\ell+1} - AC_\ell$ is rational. Because i, j, ℓ are arbitrarily chosen, we proved that all sides of all polygons in the dissection are also rational numbers.

Now we prove the Lemma. We set $\angle DAP = A_1$ and $\angle BAP = A_2$. Apply the Law of Cosines to triangles ADC, ABC, ABD to show that angles A_1, A_2, $A_1 + A_2$ all have rational cosine values. By the Addition formula, we have

$$\sin A_1 \sin A_2 = \cos A_1 \cos A_2 - \cos(A_1 + A_2),$$

implying that $\sin A_1 \sin A_2$ is rational. Thus

$$\frac{\sin A_2}{\sin A_1} = \frac{\sin A_2 \sin A_1}{\sin^2 A_1} = \frac{\sin A_2 \sin A_1}{1 - \cos^2 A_1}$$

is rational. Note that the ratio between the areas of triangles ADP and ABP is equal to PD/BP. Therefore

$$\frac{BP}{PD} = \frac{[ABP]}{[ADP]} = \frac{\frac{1}{2}AB \cdot AP \cdot \sin A_2}{\frac{1}{2}AD \cdot AP \cdot \sin A_1} = \frac{AB}{AD} \cdot \frac{\sin A_2}{\sin A_1},$$

implying that PD/BP is rational. Because $BP + PD = BD$ is rational, both BP and PD are rational. Similarly, AP and PC are rational, proving the Lemma.

3. We prove that for $n \geq 2$, the claim holds without the initial condition $0 \leq a_i \leq i$. (Of course this does not prove anything stronger, but it's convenient.) We do this by induction on n, the case $n = 2$ being easy to check by hand.

   Note that if $c = (c_0, \ldots, c_n)$ is a sequence in the image of t, and d is the sequence $(c_1, \ldots, c_n)$, then the following two statements are true:

   (a) If e is the sequence obtained from d by subtracting 1 from each nonzero term, then $t(d) = t(e)$. (If there are no zero terms in d, then subtracting 1 clearly has no effect. If there is a zero term in d, it must occur at the beginning, and then every nonzero term is at least 2.)

   (b) One can compute $t(c)$ by applying t to the sequence $c_1, \ldots, c_n$, adding 1 to each nonzero term, and putting a zero in front.

   The recipe in (b) works for computing $t^i(c)$ for any i, by (a) and induction on i.

   We now apply the induction hypothesis to $t(a)_1, \ldots, t(a)_n$ to see that it stabilizes after $n - 2$ more applications of t; by the recipe above, this means that a stabilizes after $n - 1$ applications of t.

4. Extend segment DM through M to G such that $FG \parallel CD$. Then $MF = MC$ if and only if quadrilateral $CDFG$ is a parallelogram, or, $FD \parallel CG$. Hence $MC = MF$ if and only if $\angle GCD = \angle FDA$, that is, $\angle FDA + \angle CGF = 180°$.

   Because quadrilateral $ABED$ is cyclic, $\angle FDA = \angle ABE$. It follows that $MC = MF$ if and only if

$$180° = \angle FDA + \angle CGF = \angle ABE + \angle CGF,$$

   that is, quadrilateral $CBFG$ is cyclic, which is equivalent to

$$\angle CBM = \angle CBG = \angle CFG = \angle DCF = \angle DCM.$$

   Because $\angle DMC = \angle CMB$, $\angle CBM = \angle DCM$ if and only if triangles BCM and CDM are similar, that is

$$\frac{CM}{BM} = \frac{DM}{CM},$$

   or $MB \cdot MD = MC^2$.

5. By multiplying a, b, and c by a suitable factor, we reduce the problem to the case when $a + b + c = 3$. The desired inequality reads

$$\frac{(a+3)^2}{2a^2 + (3-a)^2} + \frac{(b+3)^2}{2b^2 + (3-b)^2} + \frac{(c+3)^2}{2c^2 + (3-c)^2} \leq 8.$$

   Set

$$f(x) = \frac{(x+3)^2}{2x^2 + (3-x)^2}$$

   It suffices to prove that $f(a) + f(b) + f(c) \leq 8$. Note that

$$\begin{aligned} f(x) &= \frac{x^2 + 6x + 9}{3(x^2 - 2x + 3)} = \frac{1}{3} \cdot \frac{x^2 + 6x + 9}{x^2 - 2x + 3} \\ &= \frac{1}{3}\left(1 + \frac{8x + 6}{x^2 - 2x + 3}\right) = \frac{1}{3}\left(1 + \frac{8x + 6}{(x-1)^2 + 2}\right) \\ &\leq \frac{1}{3}\left(1 + \frac{8x + 6}{2}\right) = \frac{1}{3}(4x + 4). \end{aligned}$$

Hence,

$$f(a) + f(b) + f(c) \leq \frac{1}{3}(4a + 4 + 4b + 4 + 4c + 4) = 8,$$

as desired, with equality if and only if $a = b = c$.

6. In the beginning, because $A + B + C + D + E + F$ is odd, either $A + C + E$ or $B + D + F$ is odd; assume without loss of generality it is the former. Perform the following steps repeatedly.

   a. *If A, C, E are all nonzero.* Suppose without loss of generality that $A \geq C \geq E$. Perform the sequence of moves

$$\begin{matrix} A & \begin{matrix} B & C \\ F & E \end{matrix} & D \end{matrix} \to \begin{matrix} A & \begin{matrix} (\boldsymbol{A-C}) & C \\ (\boldsymbol{A-E}) & E \end{matrix} & (\boldsymbol{C-E}) \end{matrix}$$

$$\to \begin{matrix} (\boldsymbol{C-E}) & \begin{matrix} (A-C) & C \\ (A-E) & (\boldsymbol{A-C}) \end{matrix} & (C-E), \end{matrix}$$

   which decreases the sum of the numbers in positions A, C, E while keeping that sum odd.

   b. *If exactly one among A, C, E is zero.* Assume without loss of generality that $A \geq C > E = 0$. Then, because $A + C + E$ is odd, A must be strictly greater than C. Therefore, $-A < A - 2C < A$, and the sequence of moves

$$\begin{matrix} A & \begin{matrix} B & C \\ F & 0 \end{matrix} & D \end{matrix} \to \begin{matrix} A & \begin{matrix} (\boldsymbol{A-C}) & C \\ \boldsymbol{A} & 0 \end{matrix} & \boldsymbol{C} \end{matrix}$$

$$\to \begin{matrix} \boldsymbol{C} & \begin{matrix} (A-C) & |\boldsymbol{A-2C}| \\ A & 0 \end{matrix} & C, \end{matrix}$$

   decreases the sum of the numbers in positions A, C, E while keeping that sum odd.

   c. *If exactly two among A, C, E are zero.* Assume without loss of generality that $A > C = E = 0$. Then perform the sequence of moves

$$\begin{matrix} A & \begin{matrix} B & 0 \\ F & 0 \end{matrix} & D \end{matrix} \to \begin{matrix} A & \begin{matrix} \boldsymbol{A} & 0 \\ \boldsymbol{A} & 0 \end{matrix} & \boldsymbol{0} \end{matrix} \to \begin{matrix} \boldsymbol{0} & \begin{matrix} A & 0 \\ A & 0 \end{matrix} & 0 \end{matrix} \to \begin{matrix} 0 & \begin{matrix} \boldsymbol{0} & 0 \\ \boldsymbol{0} & 0 \end{matrix} & 0. \end{matrix}$$

   By repeatedly applying step (a) as long as it applies, then doing the same for step (b) if necessary, and finally applying step (c) if necessary, $\begin{matrix} 0 & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & 0 \end{matrix}$ can eventually be achieved.

CONTENTS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, DC 20036