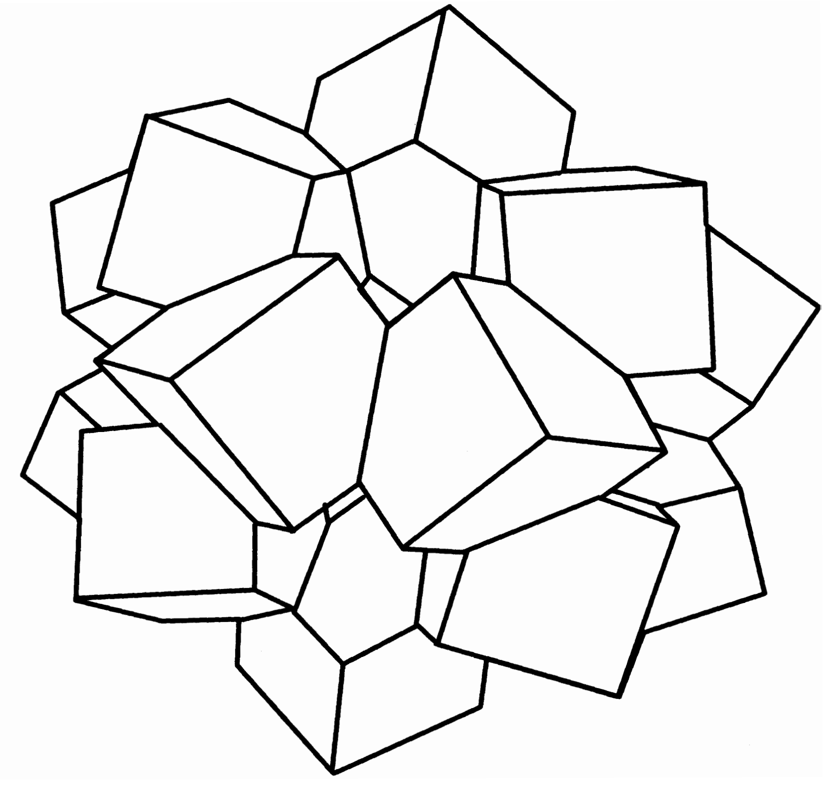
Vol. 79, No. 3, June 2006

MATHEMATICS MAGAZINE



Twenty-Cubes

- Compound Platonic Polyhedra in Origami
- Applying Burnside's Lemma to a One-Dimensional Escher Problem
- Dropping Lowest Grades

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 Allen Schwenk, Editor, Mathematics Magazine, Department of Mathematics, Western Michigan University, Kalamazoo, MI, 49008. 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: Imagine starting with a regular dodecahedron. At each of the twenty vertices we attach a cube with one corner indented (dimpled?) to fit the faces meeting at that vertex. The result is the Twenty-Cubes configuration invented by origamist David Mitchell and described, along with numerous variations, in the article by Zsolt Lengvarsky on page 190.

AUTHORS

Tomaž Pisanski has a joint position at the University of Ljubljana and University of Primorska, Slovenia. He is the founder of a strong research group in discrete mathematics in Slovenia. His current research interests include the interplay between graphs, groups,maps, and finite geometrical structures. His university teaching of mathematics and computer science includes visiting positions in Austria, Canada, Croatia, England, Italy, New Zealand, and USA, the latest in 2003 as Neil R. Grabois Professor of Mathematics at Colgate University.

Doris Schattschneider is Professor Emerita of Mathematics at Moravian College in Bethlehem, Pennsylvania. Escher's symmetry investigations have fascinated her for a long time; the study Escher's work was a natural outgrowth of her research interests in discrete geometry and tiling. In the Afterword of a new edition of *M.C. Escher: Visions of Symmetry,* she discusses several mathematical problems generated by Escher's work, including his combinatorial patterns, which provided the inspiration for this article.

Brigitte Servatius is a combinatorist whose main interest is matroid theory, in particular the study of rigidity matroids. She has been teaching at Worcester Polytechnic Institute since 1987. During a most enjoyable sabbatical year, hosted by Pisanski at the University of Ljubljana, she started to develop a bit into a geometer. She is the editor of the *Pi Mu Epsilon Journal* which focuses on articles for and by students. She also edits Student Research Projects for the *College Math Journal.*

Daniel Kane is an undergraduate student at the Massachusetts Institute of Technology. He was a gold medalist at the International Mathematics Olympiad in 2002 and 2003, a Putnam Fellow in 2003 and 2004, and a Fellow of the Davidson Institute for Talent Development in 2003. Daniel has written over a dozen papers in number theory, game theory, and combinatorics, which are at various stages of publication. The work for this paper grew out of a discussion about the subtleties of the dropping lowest grades problem which took place on a 1●K run with his father and co-author, Jonathan Kane.

Jonathan Kane earned a Ph.D. in several complex variables from the University of Wisconsin, Madison in 1980. He is now professor of mathematics and computer science at the University of Wisconsin, Whitewater. Dr. Kane is the creator of GRADE GUIDE, a shareware computer program which helps teachers store, analyze, and report students' grades, which first appeared in 1985. In addition to his continued interest in analysis, probability, combinatorics, and computer science, he is actively involved in high school mathematics competitions both as a member of the American Invitational Mathematics Examination Committee and as co-creator of the Purple Comet on-line competition.

Zsolt Lengvarszky was educated in Hungary (University of Szeged) and in the US (University of South Carolina). He earned graduate degrees in mathematics, statistics, and computer science. He has professional experience in both academia and industry. His research interests include lattice theory, universal algebra, discrete mathematics, and the mathematics of paper folding.

Vol. 79, No. 3, June 2006

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 Frank Peterson (*FPetersonj@aol.com*), 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

ARTICLES

Applying Burnside's Lemma to a One-Dimensional Escher Problem

TOMAŽ PISANSKI
University of Ljubljana and University of Primorska
Slovenia
Tomaz.Pisanski@fmf.uni-lj.si

DORIS SCHATTSCHNEIDER
Moravian College
Bethlehem, PA 18018-6650
schattdo@moravian.edu

BRIGITTE SERVATIUS
Worcester Polytechnic Institute
Worcester, MA 01609-2280
bservat@wpi.edu

Our point of departure is the paper **[7]** in which a problem of M. C. Escher is solved using methods of contemporary combinatorics, in particular, Burnside's lemma. Escher originally determined (by laboriously examining multitudes of sketches) how many different patterns would result by repeatedly translating a 2×2 square having its four unit squares filled with copies of an asymmetric motif in any of four rotated aspects. In this note we simplify the problem from two dimensions to one dimension but at the same time we generalize it from the case in which a 2×2 block stamps out a repeating planar pattern to the case in which a $1 \times n$ block stamps out a repeating strip pattern.

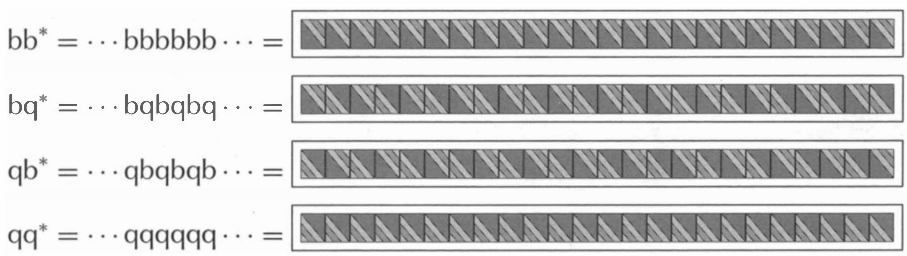
The 1×2 case

Suppose we are tiling a strip by a single rectangle containing an asymmetric motif, say ■, a motif taken from South African beadwork which is a rectangle divided by a diagonal into two triangles, one solid red, and the other yellow with a green stripe. The original motif has three additional aspects, namely the motif rotated by 180°, reflected in a vertical line and in a horizontal line. We note the motif by b and its other aspects as follows:

since the letters p, q and d are the corresponding aspects of the letter b under these transformations. This notation was first introduced in **[9]** to encode the symmetry groups of strip patterns.

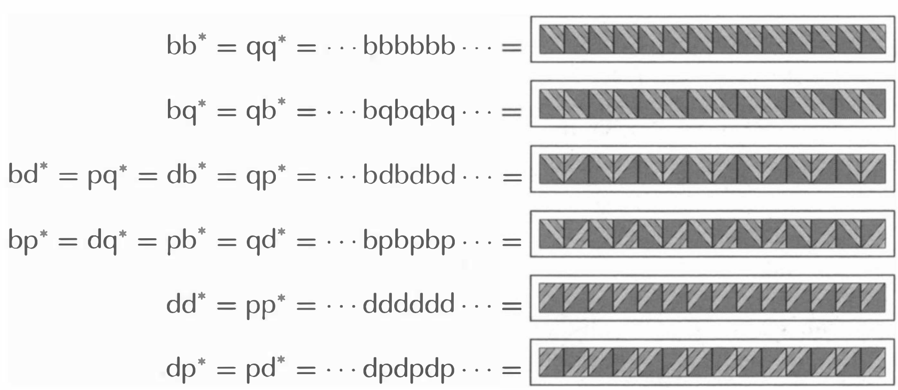
Aspects b and q are translated and rotated images of the original aspect b; we call these *direct aspects*. Assume that we may select any two direct aspects (with repetition allowed) to form a signature for a 1×2 block of two rectangles containing those aspects of the motif. There are four possible signatures:

By repeating a 1×2 block horizontally and removing the outline of the rectangles, each signature determines uniquely a 2–way infinite strip pattern:



The patterns bq* and qb* differ only by translation and so we write bq* = qb*. Similarly, the pattern bb* can be turned into qq* by rotating the strip by 180°, so we have as well bb* = qq* and thus there are only two different patterns.

If we repeat the above construction of the patterns, but allow the two-letter signature to be any ordered pair of aspects chosen from {b, q, d, p}, the number of possible signatures increases to 16. If we do not distinguish between patterns that can be obtained from each other by translations and rotations, we will find that there are six patterns.



If, however, we do not distinguish between patterns which are mirror images of one another, then the first four complete the list.

The key observation is that we do not actually have to construct the strip patterns and observe them in order to determine how many different ones there are. Since the patterns are determined by their signatures, the method is to study what permutations of signatures do not change the pattern. The general model can be set up as follows.

We are given a set of permutations P that generates a group $\langle P \rangle$ which acts on the set of signatures $S = \{w_1, w_2, \ldots\}$, where each permutation in P transforms each signature into one that produces the "same" strip pattern, with the choice of group determining the definition of sameness. To count how many different strip patterns there are, we have to determine the number of orbits under the action of $\langle P \rangle$ on S. The perfect tool for counting the number of orbits is Burnside's lemma: The number of orbits equals the average number of points fixed by the permutations in the group.

More precisely, Burnside's lemma says that the number of orbits N of the group $\langle P\rangle$ acting on S is

$$N = \frac{1}{|\langle P\rangle|}\sum_{p\in\langle P\rangle}|\mathrm{fix}(p)|$$

where $\mathrm{fix}(p)$ is the set of signatures fixed by the permutation p.

Burnside's lemma, also called the Cauchy-Frobenius lemma in the literature, has a long history, which can be found in **[5, 11]**, but still has its place in advanced texts, e.g., see **[10]**.

Suppose the group $\langle P\rangle$ is generated by two elements T and R. Here T interchanges the first and second elements of the signature, $T(XY) = YX$, and corresponds to a 1-unit horizontal translation of the strip pattern. R replaces each aspect with its rotated aspect and interchanges their order in the signature: $R(XY) = R(Y)R(X)$, where $R(\mathrm{b}) = \mathrm{q}$, $R(\mathrm{q}) = \mathrm{b}$, $R(\mathrm{p}) = \mathrm{d}$ and $R(\mathrm{d}) = \mathrm{p}$. R corresponds to a 180° rotation of the strip pattern.

The group $\langle P\rangle = \langle T, R \mid T^2 = R^2 = (TR)^2 = I\rangle$ is isomorphic to the Klein four group. Its Cayley graph is shown in Table 1 below: the group elements are represented as vertices. The horizontal edges correspond to multiplication by T, vertical edges to multiplication by R. In Table 1 we also show the action of the group $\langle P\rangle$ on the four signatures bb, bq, qb, qq, and see that there are a total of 8 signatures fixed by elements of $\langle P\rangle$ (boxed). Since the group $\langle P\rangle$ has 4 elements, Burnside's lemma confirms the number of distinct patterns for $\langle P\rangle$ acting on signatures with two aspects to be $8/4 = 2$.

TABLE 1: $\langle P\rangle$ and its action on four signatures

$$\begin{tikzcd}
R \arrow[r, dash] \arrow[d, dash] & TR \arrow[d, dash] \\
I \arrow[r, dash] & T
\end{tikzcd}$$

	I	R	T	TR
bb	$\boxed{\mathrm{bb}}$	qq	$\boxed{\mathrm{bb}}$	qq
bq	$\boxed{\mathrm{bq}}$	$\boxed{\mathrm{bq}}$	qb	qb
qb	$\boxed{\mathrm{qb}}$	$\boxed{\mathrm{qb}}$	bq	bq
qq	$\boxed{\mathrm{qq}}$	bb	$\boxed{\mathrm{qq}}$	bb

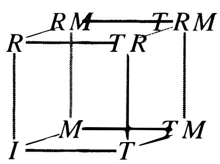
If we extend Table 1 to include 16 rows of signatures to account for all four aspects, we obtain a total of 24 signatures fixed by elements of $\langle P\rangle$, which are shown boxed in the first 4 columns of Table 2, and so the formula in Burnside's lemma gives the number of distinct patterns as $24/4 = 6$.

To regard the strips as identical even after orientation-reversing transformations, we extend the group $\langle P\rangle$ by adding another generator, the mirror M, where M acts on signatures by $M(XY) = M(X)M(Y)$, and on aspects by $M(\mathrm{b}) = \mathrm{p}$, $M(\mathrm{p}) = \mathrm{b}$, $M(\mathrm{q}) = \mathrm{d}$ and $M(\mathrm{d}) = \mathrm{q}$. This corresponds to taking the mirror image of the infinite strip in a horizontal axis, and, together with the transformations we already have, allows us to consider strip patterns as identical if they differ by orientation-preserving as well as orientation-reversing transformations. Let $P' = \{T, R, M\}$. The extended group

$$\langle P'\rangle = \langle T, R, M \mid T^2 = R^2 = M^2 = (TR)^2 = (TM)^2 = (RM)^2 = I\rangle$$

has 8 elements and is isomorphic to the direct product of three copies of the cyclic group on 2 elements. Table 2 shows also the Cayley graph of $\langle P' \rangle$ in which the three sets of mutually parallel edges correspond to multiplication by R, T and M, respectively. Table 2 shows the action of $\langle P' \rangle$ on the sixteen signatures; there are 32 signatures fixed by elements of $\langle P' \rangle$, which are boxed.

TABLE 2: $\langle P' \rangle$ and its action on 16 signatures



	I	R	T	TR	M	RM	TM	TRM
bb	$\boxed{\text{bb}}$	qq	$\boxed{\text{bb}}$	qq	pp	dd	pp	dd
bq	$\boxed{\text{bq}}$	$\boxed{\text{bq}}$	qb	qb	pd	pd	dp	dp
qb	$\boxed{\text{qb}}$	$\boxed{\text{qb}}$	bq	bq	dp	dp	pd	pd
qq	$\boxed{\text{qq}}$	bb	$\boxed{\text{qq}}$	bb	dd	pp	dd	pp
bp	$\boxed{\text{bp}}$	dq	pb	qd	pb	qd	$\boxed{\text{bp}}$	dq
bd	$\boxed{\text{bd}}$	pq	db	qp	pq	$\boxed{\text{bd}}$	qp	db
qp	$\boxed{\text{qp}}$	db	pq	bd	db	$\boxed{\text{qp}}$	bd	pq
qd	$\boxed{\text{qd}}$	pb	dq	bp	dq	bp	$\boxed{\text{qd}}$	pb
pb	$\boxed{\text{pb}}$	qd	bp	dq	bp	dq	$\boxed{\text{pb}}$	qd
pq	$\boxed{\text{pq}}$	bd	qp	db	bd	$\boxed{\text{pq}}$	db	qp
db	$\boxed{\text{db}}$	qp	bd	pq	qp	$\boxed{\text{dq}}$	pq	bd
dq	$\boxed{\text{dq}}$	bp	qd	pb	qd	pb	$\boxed{\text{dq}}$	bp
pp	$\boxed{\text{pp}}$	dd	$\boxed{\text{pp}}$	dd	bb	qq	bb	qq
pd	$\boxed{\text{pd}}$	$\boxed{\text{pd}}$	dp	dp	bq	bq	qb	qb
dp	$\boxed{\text{dp}}$	$\boxed{\text{dp}}$	pd	pd	qb	qb	bq	bq
dd	$\boxed{\text{dd}}$	pp	$\boxed{\text{dd}}$	pp	qq	bb	qq	bb

Note: M mirrors the aspects in a horizontal mirror. We could have, alternately, used a vertical mirror M_V which mirrors aspects b and d, p and q; however the three groups generated by $\{T, R, M\}$, $\{T, R, M_V\}$, and $\{T, R, M, M_V\}$ are all the same since $M_V = RM$, and $M = RM_V$. Try to draw the corresponding Cayley graphs!

From Table 2 and Burnside's lemma, we obtain the result of $32/8 = 4$ different strip patterns with four motif aspects, confirming our earlier observation for the 'beadwork' pattern.

In fact, from the first four columns of Table 2, we can determine the previously computed number of patterns up to rotation and translation, with either all four aspects, all 16 rows, or just the two direct aspects, the first 8 rows.

The main purpose of this note is to generalize the approach from the 1×2 case to the general case $1 \times n$, $n \geq 1$. The permutation groups become much more complicated and the sets of signatures on which they act grow much larger. To understand the general case it is enough to consider two relatively small representatives.

The 1 × 12 case

Let's compute the number of patterns arising from a strip of length 12 filled with choices from all four aspects, regarding patterns to be the same up to translation, rotation and reflection, that is, using the extended group, $\langle P' \rangle$.

To study the transformations of the signature, it is convenient to think of the signature as being drawn on the surface of a ring with 12 marked sections, see Figure 1, in which the initial point in the signature is marked with a small triangle.

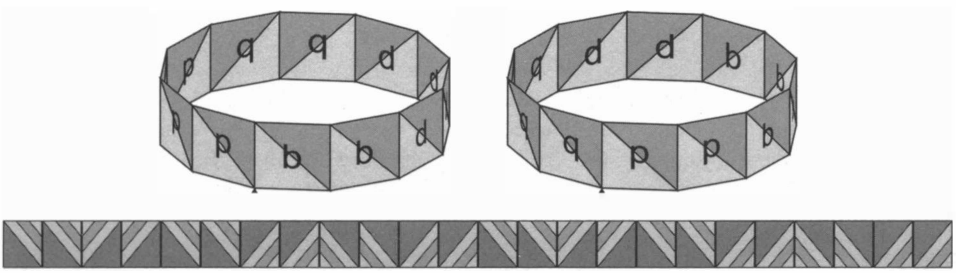


Figure 1 The signature $w =$ bbddbbppqqpp on a ring, and $T^4M(w)$, and their pattern. $T^6M(w) = w$, so $w \in \text{fix}(T^6M)$

In fact, this is how you can create the strip patterns in practice; by inking the ring and then rolling out the pattern!

Any symmetry of the ring clearly yields the same pattern. Rotationally, the ring has dihedral symmetry, and the rotation group is generated by two rotations. The first is a rotation of 30° about the vertical axis through the center of the ring and corresponds to a translation of the strip pattern by one unit. We denote it by T:

$$T(X_1X_2\ldots X_{12}) = X_2X_3\ldots X_{12}X_1$$

The second is a 180° rotation about the axis passing through the center of the ring and passing through the midpoint of the initial boundary of the first motif, and corresponds to a 180° rotation of of the strip pattern. We denote it by R and its action on the signature is

$$R(X_1X_2\ldots X_{11}X_{12}) = R(X_{12})R(X_{11})\ldots R(X_2)R(X_1).$$

See Figure 2.

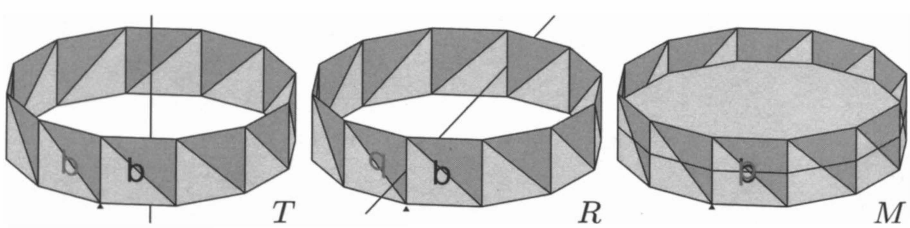


Figure 2 Transforming the signature on a 12-ring

The elements T and R generate the dihedral group D_{12} in the usual way:

$$\langle P \rangle = \langle T, R \mid R^2 = T^{12} = I, RTR = T^{-1} \rangle.$$

The orientation-reversing transformations can be added by adding the generator M, which is the reflection in the horizontal plane that bisects the ring,

$$M(X_1 \dots X_{12}) = M(X_1) \dots M(X_{12}),$$

and corresponds to a reflection of the strip pattern in a horizontal plane; see Figure 2. We get the following presentation for $\langle P' \rangle$,

$$\langle T, R, M \mid I = M^2 = R^2 = T^{12}, RTR = T^{-1}, TM = MT, RM = MR \rangle.$$

Of course, it is convenient to describe groups in terms of generators and relations, but that really doesn't help us in using Burnside's lemma, since we have to take the mean over *all* the elements of the group, not just the generators. Fortunately, at least for the dihedral group and its extension, we can easily visualize all the transformations. See Figure 3.

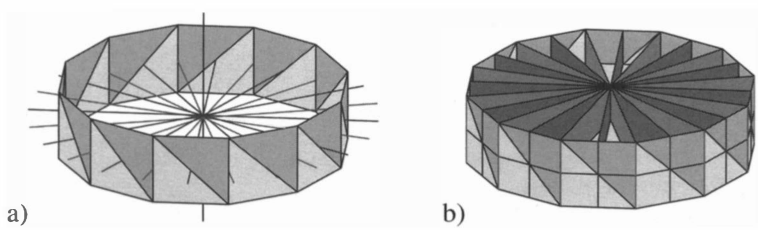


Figure 3 The axes of the rotational symmetries of the ring, and the planes of the mirror symmetries

All 24 transformations in $\langle P \rangle$ are rotational symmetries of the ring. There are 12 rotations of 180° around axes in the horizontal plane through the center of the ring, see Figure 3a. Of these, 6 have axes passing through the centers of two opposite motifs, and so fix no signatures since the motifs are asymmetric. The other six have axes on the midpoints of motif boundaries, with the motifs being divided into 6 pairs of orbits. So there are $6 \cdot 4^6$ fixed signatures for these transformations. See Figure 4, in which 6 independent choices (b, d, d, p, b, q) for the first six positions yield the fixed signature bddpbqbqdppq.

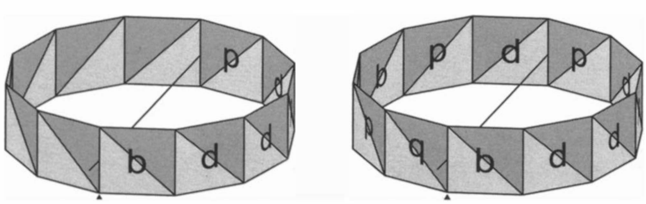


Figure 4 Creating a fixed signature for a horizontal axis rotation

The other 12 transformations in $\langle P \rangle$ are rotations about the vertical axis of $\frac{i}{12}360° = i \cdot 30°$, $i = 1, \dots, 12$. If i and 12 have a common divisor k, which means that there

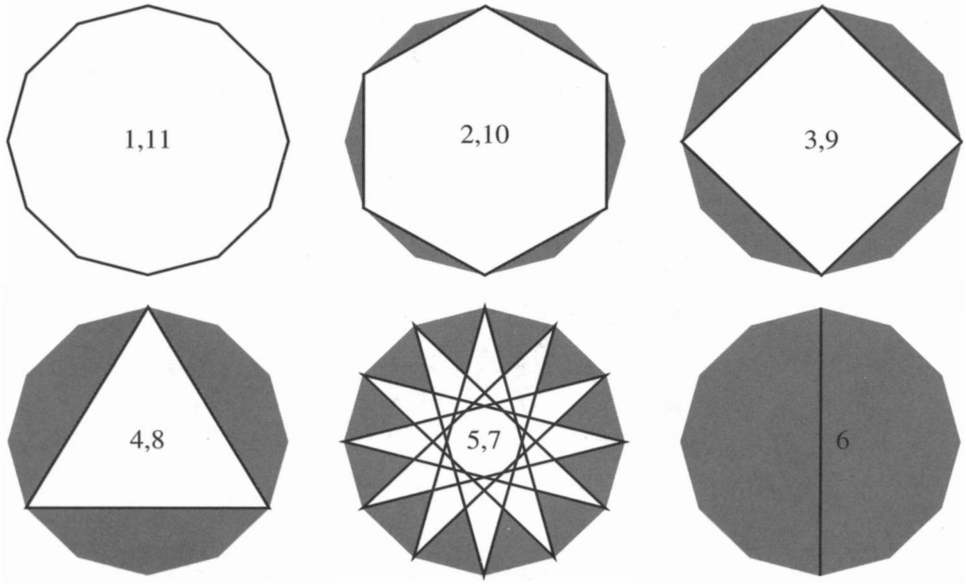


Figure 5 Orbits of rotations of a 12-gon

are integers p and q such that $i = pk$ and $12 = qk$, then $q \cdot (i \cdot 30°)$ is a multiple of 360° and so this rotation has motif orbits of size a divisor of q. In fact, it is easy to see that the orbits of $i \cdot 30°$ are of size $12/\gcd(i, 12)$, and there are $\gcd(i, 12)$ of them. See Figure 5. So, for each divisor k of 12 there are rotations with aspect orbit sizes k. Each of these will have $4^{12/k}$ fixed signatures, since we are free to choose any of the four aspects for each orbit. See Figure 6 in which four independent choices, (b, d, d, p) for the first four positions, and a $3 \cdot \frac{180°}{12}$ rotation, yield the signature bddpbddpbddp. Twelve has divisors 12, 6, 4, 3, 2, and 1. For 12 there will be 4 rotations with orbit size 12, corresponding to $i = 1, 5, 7, 11$, which is the number of positive integers less than 12 which are coprime to 12, giving $4 \cdot 4^1$ fixed signatures. For 6 there are two rotations of orbit size 6, corresponding to $i = 2 = 1 \cdot \frac{12}{6}$ and $i = 10 = 5 \cdot \frac{12}{6}$. Observe that 1 and 5 are the integers less than 6 relatively prime to 6; we get $2 \cdot 4^2$ fixed signatures. For 4 there are two rotations of orbit size 4, $i = 3 = 1 \cdot \frac{12}{4}$ and $i = 9 = 3 \cdot \frac{12}{4}$, with 1 and 3 being the integers less than 4 relatively prime to 4; we get $2 \cdot 4^3$ fixed signatures. For 3 there are two rotations of orbit size 3, $i = 4 = 1 \cdot \frac{12}{3}$ and $i = 8 = 2 \cdot \frac{12}{3}$, with 1 and 2 being the integers less that 3 relatively prime to 3; we get $2 \cdot 4^4$ fixed signatures. For 2 there is one rotation of orbit size 2, $i = 6 = 1 \cdot \frac{12}{2}$, with 1 the only integer less than 2 relatively prime to 2; we get $1 \cdot 4^6$ fixed signatures. For 1 there is one rotation of orbit size 1, $i = 12$; we get $1 \cdot 4^{12}$ fixed signatures.

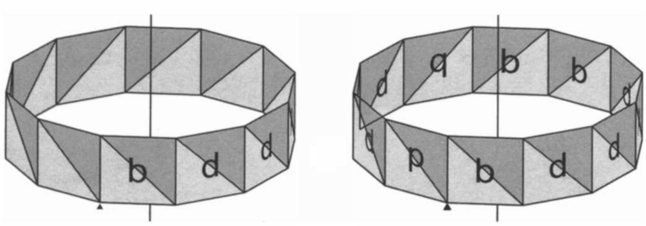


Figure 6 Creating a fixed signature for a vertical axis rotation

The reader may recall that for $k > 1$, the number of positive integers at most k which are relatively prime to k is denoted by $\varphi(k)$ and is called the *Euler phi function.* Note that $\varphi(1) = 1$ by definition. So we have for the rotations about the vertical axis

$$\varphi(12)\cdot 4^{12/12} + \varphi(6)\cdot 4^{12/6} + \varphi(4)\cdot 4^{12/4} + \varphi(3)\cdot 4^{12/3} + \varphi(2)\cdot 4^{12/2} + \varphi(1)\cdot 4^{12/1}$$

fixed signatures.

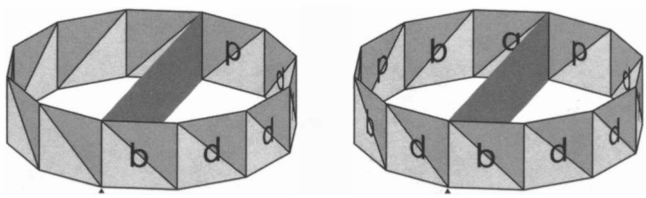
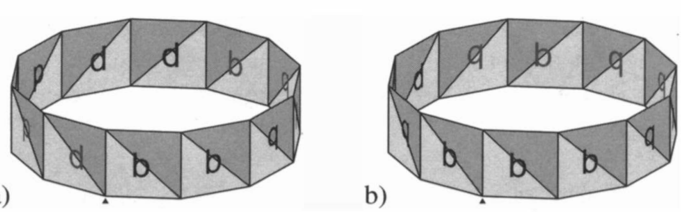


Figure 7 Creating a fixed signature for a reflection

For the 24 orientation-reversing symmetries, twelve are reflections in a vertical mirror, see Figure 3b, 6 of which pass through the center of a motif, and so have no fixed signatures, and 6 of which pass through the boundaries of the aspects, giving 6 aspect orbits of size 2 each, hence $6 \cdot 4^6$ fixed signatures, See Figure 7, in which 6 independent choices (b, d, d, p, b, q) for the first six positions yield the fixed signature bddpbqpdqbbd. The other 12 orientation-reversing symmetries are not reflections at all, but the product of one of the twelve rotations about the vertical axis with the reflection in the horizontal mirror, and are called rotary reflections. We have already analyzed the aspect orbits under these rotations. The only difference now is that, with the horizontal mirror, if the aspect orbit size is odd, specifically for $k = 3$ and 1, ($i = 4, 8, 12$) there will be no fixed signatures since, following the aspect through its orbit, the aspect would return to its original position on the ring as a reflected aspect. For example, in Figure 8a we have chosen aspects b, b, and q respectively for the first three positions of a rotatory reflection of angle 90°, one for each of the three orbits, yielding the fixed signature bbqppdbbqppd. Trying the same method, Figure 8b, and choosing b, b, q, and b for the first four positions with the rotary reflection of 120°, with rotational order 3, gives the signature bbqbppdpbbqb which is not a fixed signature under the 120° rotary reflection.



a) b)

Figure 8 A fixed signature for a rotary reflection

So, omitting the odd divisors, there are

$$\varphi(12)\cdot 4^{12/12} + \varphi(6)\cdot 4^{12/6} + \varphi(4)\cdot 4^{12/4} + \varphi(2)\cdot 4^{12/2}$$

fixed signatures of rotary reflections.

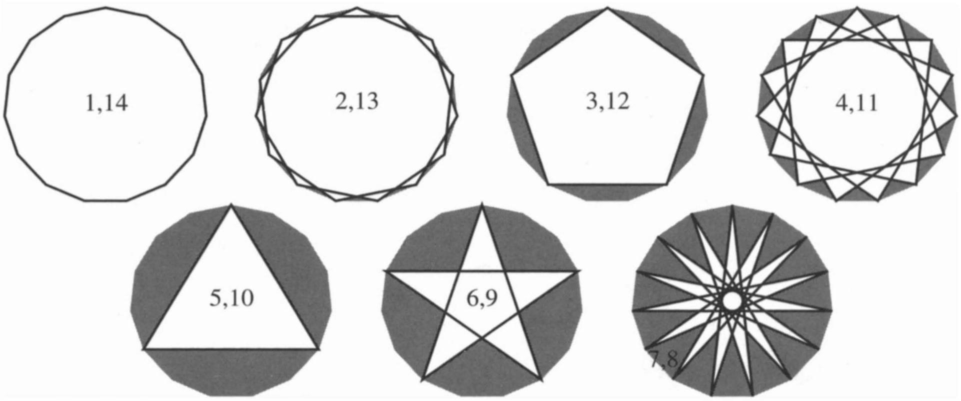


Figure 9 Orbits of rotations of a 15-gon

The 1 × 15 case

For a ring of size 15, the rotations with vertical axis, and their rotary reflections are analyzed just as before; see Figure 9. Thus there are

$$\varphi(15) \cdot 4^{15/15} + \varphi(5) \cdot 4^{15/5} + \varphi(3) \cdot 4^{15/3} + \varphi(1) \cdot 4^{15/1}$$

fixed signatures for the first kind and no signatures fixed by the second kind because 15 has no even divisors.

The main difference here is that, since the ring is of odd size, every 180° rotation about a horizontal axis has one pole of the axis passing through the midpoint of an aspect boundary and the other passing though the center of the aspect; see Figure 10. None of these will have fixed signatures, since the motif is assumed to be asymmetric. Similarly, each vertical mirror passes through an aspect boundary on one side of the ring, and passes through the middle of an aspect on the opposite side, so, since the motifs are asymmetric, there are no fixed signatures. Thus the total number of fixed signatures is $8 \cdot 4^{15/15} + 4 \cdot 4^{15/5} + 2 \cdot 4^{15/3} + 1 \cdot 4^{15/1}$.

The 1 × n case

In the general case, the group $\langle T, R, M\rangle$ has elements:

Vertical axis rotations	T^i
Horizonal axis rotations	$T^i R$
Vertical reflections	$T^i RM$
Rotary reflections	$T^i M.$

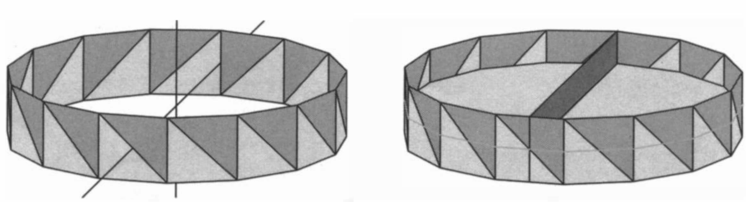


Figure 10 Axes and mirrors of a 15-ring

and acts on a signature $w = X_1X_2\cdots X_n$, $X_i \in \{\mathrm{b, q, d, p}\}$ via

Translation:	$T(X_1X_2\cdots X_n) = X_2X_3\cdots X_nX_1$.
Rotation:	$R(X_1X_2\cdots X_n) = R(X_n)\cdots R(X_2)R(X_1)$.
Mirror:	$M(X_1X_2\cdots X_n) = M(X_1)M(X_2)\cdots M(X_n)$.

For the vertical axis rotations the number of fixed signatures if there are only the two direct aspects is

$$v(n) = \sum_{k|n} \varphi(k)2^{n/k},$$

while if there are 4 aspects the number is

$$V(n) = \sum_{k|n} \varphi(k)4^{n/k}.$$

For horizontal axis rotations the number of fixed signatures if there are only the two direct aspects is

$$h(n) = \begin{cases} (n/2)2^{n/2} & \text{for } n \text{ even and} \\ 0 & \text{for } n \text{ odd} \end{cases},$$

while if there are four aspects the number is

$$H(n) = \begin{cases} (n/2)4^{n/2} & \text{for } n \text{ even and} \\ 0 & \text{for } n \text{ odd.} \end{cases}$$

For the orientation-reversing transformations, we are only considering the case with four aspects. For the vertical mirror reflections, there are $H(n)$ fixed signatures and, lastly, for the rotary reflections, there are

$$R(n) = \sum_{k|n,2|k} \varphi(k)4^{n/k}$$

fixed signatures.

So, if we consider only the two direct aspects and rotational symmetry we have by Burnside's lemma

$$f(n) = \frac{v(n) + h(n)}{2n}$$

patterns.

If we allow four aspects but only consider rotational symmetry we have

$$F(n) = \frac{V(n) + H(n)}{2n}$$

patterns.

If we allow four aspects and consider mirror symmetry as well, there are only

$$G(n) = \frac{V(n) + 2H(n) + R(n)}{4n}$$

patterns.

The number of orbits for each of the three cases, where $n = 1, \ldots, 30$ is given in Table 3. The sequence of numbers $f(n)$ appears as sequence with ID number A053656 in Sloane's On-Line Encyclopedia of integer sequences [**8**], where it is described as arising from the number of necklaces with n blue or red beads such that the beads switch color when the necklace is turned over, which is clearly equivalent to our situation. Our interpretation of $f(n)$ via the number of strip patterns is more naturally motivated than color-switching beads.

TABLE 3: Numbers of strip patterns under different notions of 'sameness'.

n	2 aspects $\langle P\rangle$ $f(n)$	4 aspects $\langle P\rangle$ $F(n)$	4 aspects $\langle P'\rangle$ $G(n)$
1	1	2	1
2	2	6	4
3	2	12	6
4	4	39	23
5	4	104	52
6	9	366	194
7	10	1172	586
8	22	4179	2131
9	30	14572	7286
10	62	52740	26524
11	94	190652	95326
12	192	700274	350738
13	316	2581112	1290556
14	623	9591666	4798174
15	1096	35791472	17895736
16	2122	134236179	67127315
17	3856	505290272	252645136
18	7429	1908947406	954510114
19	13798	7233629132	3616814566
20	26500	27488079132	13744183772
21	49940	104715393912	52357696956
22	95885	399823554006	199912348954
23	182362	1529755308212	764877654106
24	350650	5864066561554	2932035552786
25	671092	22517998136936	11258999068468
26	1292762	86607703209516	43303860638644
27	2485534	333599972407532	166799986203766
28	4797886	1286742822580254	643371447241598
29	9256396	4969489243995032	2484744621997516
30	17904476	19215358696480536	9607679491405864

From the numbers in Table 3, we can observe that $G(n) \approx 2F(n)$, which is to be expected, since the set of signatures is the same in both cases but the group that acts on it doubles in size, with equality occurring exactly when n is odd, in which case $R(n) = H(n) = 0$.

If n is large, then we expect that most signatures are asymmetric, and so will have orbit size $4n$. This would give us an approximate count of $G(n) \approx 4^n/(4n)$ which is necessarily an undercount since at least the signature bbb... is symmetric. If $n = p$, a prime, then this is the only signature which is not in an orbit of size $4p$, so rounding up the rough approximation will give the actual value

$$G(p)=\frac{4^p+(p-1)4}{4p}=\frac{4^p}{4p}+\frac{p-1}{p}.$$

In any case, these are large numbers. In the case of $G(12)$, to scan over all the distinct patterns would take at least 14 days at a rate of one pattern per second and working 8 hours a day. For $G(15)$, the other case we examined, it would take more than a year.

Note also that the results are valid only in the case the original motif is asymmetric: $R(X)$, $M(X)$ and $RM(X)$ are distinct from X.

There are several other problems that the reader is invited to explore; in several of these, other number sequences in Sloane's Encyclopedia **[8]** appear.

1. Discover the formulas (and possible patterns) for the cases when the original motif has 180° rotation symmetry, (motif N), or mirror symmetry – horizontal (motif E) or vertical (motif A).
2. Consider the problem for the two other $1 \times n$ cases for two aspects, namely, the original aspect and one of its reflected images: the sets {b, d} and {b, p}, and a suitable group of transformations.
3. Consider the case where there are m_a asymmetric motifs and m_s motifs with 180° rotation symmetry, and discover the formula $f(n, m_a, m_s)$.

The 1 × 4 case: Escher revisited

In Table 3, we see that $G(4) = 23$, which is exactly the number of different planar patterns that Escher found in answer to his original problem. The occurrence of the same numbers is not a coincidence; in fact, the 1×4 strip pattern problem corresponds exactly to Escher's 2×2 problem discussed in **[7]**.

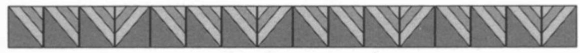
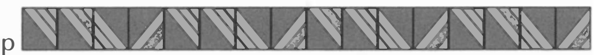
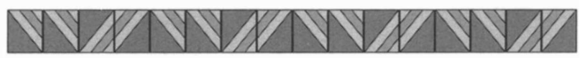
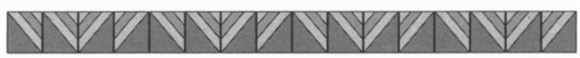
In each case, there are four units that are filled with aspects of an asymmetric motif chosen from a set of four aspects. In our 1×4 case, the aspects are all obtained from aspect b by the action of a Klein-four group generated by the 180° rotation R and reflection M in a horizontal axis. In Escher's case, the aspects were all obtained from aspect b by the action of a cyclic group of order 4, generated by a 90° rotation.

Also, in each case, the group that acts on signatures for the patterns has order 16; it is a semi-direct product of a cyclic group of order 4 and a Klein four-group. In our 1×4 case, the cyclic group is generated by T, and the Klein four-group is generated by R and M. In Escher's case, the cyclic group was generated by a permutation induced by a 90° rotation of the 2×2 block, and the Klein four-group was generated by permutations induced by horizontal and vertical unit translations of the 2×2 block.

Table 4 shows the $G(4) = 23$ strip patterns with their signatures ordered 'lexicographically' with respect to the order b < q < p < d. The $f(4) = 4$ patterns with two direct aspects are in rows 1, 2, 5, and 14.

It is well-known that there are exactly seven symmetry groups of strip patterns. The notation for these groups are: 11 (translations only—bb*); 12 (translations and 180° rotations—bq*); m1 (translations and vertical mirrors—bd*), 1g (translations and glide-reflections—bp*); mg (translations, 180° rotations, vertical mirrors, glide-reflections—bdpq*); 1m (translations, glide-reflections and horizontal mirror—$\frac{\text{bb}}{\text{pp}}{}^*$); and mm (translations, 180° rotations, vertical mirrors, glide-reflections, and horizontal mirror—$\frac{\text{bd}}{\text{pq}}{}^*$).

TABLE 4

1 bbbb	
2 bbbq	
3 bbbp	
4 bbbd	
5 bbqq	
6 bbqp	
7 bbqd	
8 bbpq	
9 bbpp	
10 bbpd	
11 bbdq	
12 bbdp	
13 bbdd	
14 bqbq	
15 bqbp	
16 bqbd	
17 bqpd	
18 bqdp	
19 bpbp	
20 bdbp	
21 bdbd	
22 bdqp	
23 bdpq	

All five without reflection symmetry in a midline mirror parallel to the edges of the strip occur in the patterns in Table 4. Most patterns have only translation symmetry. Here is the distribution of symmetry types:

Type 11	13 patterns
Type 12	3 patterns
Type m1	3 patterns
Type 1g	2 patterns
Type mg	2 patterns

Escher pursued several generalizations of his original problem, and these in turn have spawned many others: generalize to an $m \times m$ block with aspects chosen from a set of n aspects; generalize to higher dimensions; if the motif has under-over weave, allow inversion of over-under relationships to be a group operation; color the pattern so that overlapping strands do not share the same color; automate the pattern-creating process and pattern-coloring process. Many of these problems have been solved, and several are still under investigation. We list some published work on these problems in the references.

Acknowledgment. We would like to thank Marko Petkovšek for fruitful discussion. We also wish to thank a referee for a careful reading of an early draft, and the helpful suggestions provided.

REFERENCES

1. Dan Davis, On a tiling scheme from M.C. Escher, *Electronic Journal of Combinatorics*, R23, Volume 4(2), 1997.
`http://www.combinatorics.org/Volume_4/Abstracts/v4i2r23ab.html`
2. Ellen Gethner, Doris Schattschneider, Steve Passiouras, and J. Joseph Fowler, Combinatorial Enumeration of Escher's 2×2 Ribbon Patterns, to appear in the *European Journal of Combinatorics.*
3. Ellen Gethner, On a generalization of a combinatorial problem posed by M.C. Escher, In *Proceedings of the Thirty-second Southeastern International Conference on Combinatorics, Graph Theory and Computing* (Baton Rouge, LA, 2001), volume 153, pages 77–96, 2001.
4. R. Mabry, S. Wagon, and D. Schattschneider, Automating Escher's Combinatorial Patterns, *Mathematica in Education and Research*, v. 5, no. 4 (1996–97), pp. 38–52.
5. P.M. Neumann, A Lemma that is not Burnside's, *Math. Scientist* 4, 133–141, 1979.
6. S. Passiouras, web site `http://www.eschertiles.com`
7. Doris Schattschneider, Escher's combinatorial patterns, *Electronic Journal of Combinatorics*, R17 Volume 4(2), 1997.
`http://www.combinatorics.org/Volume_4/Abstracts/v4i2r17ab.html`
8. *On-Line Encyclopedia of Integer Sequences*
`http://www.research.att.com/cgi-bin/access.cgi/as/njas/sequences/eisA.cgi?Anum=A053656` Maintained by N.J.A. Sloane, njas@research.att.com
9. Brigitte Servatius, The geometry of folding paper dolls, *The Mathematical Gazette*, v. 81, no. 490 (1997), pp. 29–37.
`http://www.m-a.org.uk/docs/library/2062.pdf`
10. Richard P. Stanley, *Enumerative Combinatorics*, Volume 2, Cambridge Studies in Advanced Mathematics 62, Cambridge University Press 1999.
11. E.M. Wright, Burnside's lemma: A historical note, *J. Combinatorial Theory (B)*, 30 (1981), pp. 89–90.

Dropping Lowest Grades

DANIEL M. KANE
Massachusetts Institute of Technology
Cambridge, MA, 02139
dankane@mit.edu

JONATHAN M. KANE
University of Wisconsin-Whitewater
Whitewater, WI, 53190
kanej@uww.edu

Introduction

Many teachers allow students to drop the lowest score from a sequence of quizzes, tests, or homework assignments. When the number of grades is large, some teachers will even allow students to drop several of their lowest scores. A computer gradebook program would need to implement an algorithm to provide this feature (see one of many examples of computer gradebook software, for example, [**4**]). A natural criterion to decide which grades to drop would be to drop the set of grades that maximizes the student's final grade. In some circumstances, it can be non-trivial to determine the best grades to drop. Using natural brute force methods, the time needed to find this optimal set of grades to drop can grow exponentially as the number of grades involved grows making these methods impractical even on fast computers. We discuss some unexpected behavior exhibited by this problem and provide a simple and very efficient algorithm for finding the best set of grades to drop.

Grade dropping

Assume that a teacher has given a sequence of $k > 0$ quizzes and will allow each student to drop r of the quiz grades. Suppose that for $j = 1, 2, 3, \ldots, k$ a particular student has earned on quiz j a score of m_j points out of a possible n_j points. For simplicity assume earned scores are integers, and possible points are positive integers. Let N be an upper bound for the n_j. We will refer to the set of r grades that are dropped as the *deletion set*, and the set of $k - r$ grades that are not dropped as the *retained set*. The goal is to identify the deletion set which will result in the student receiving the highest possible final grade, the *optimal deletion set*.

If the teacher is only basing the student's final grade on the student's raw score, $\sum_{j=1}^{k} m_j$, then finding the best grades to drop is a simple matter of finding the r smallest m_j values and dropping them. For example, suppose that Alan has earned the quiz scores shown in Table 1. If the teacher wants to drop two quiz scores, this student

TABLE 1: Alan's Quiz scores

Quiz	1	2	3	4	5
Score	2	6	24	3	6
Possible	8	12	40	4	24
Percentage	25	50	60	75	25

does the best by dropping quizzes 1 and 4 since those are the two with the smallest number of points assigned, leaving the student with an accumulated quiz total of $6 + 24 + 6 = 36$, the largest possible sum of three scores. Notice that we dropped quiz 4, on which the student scored a higher percentage than on any other quiz.

On the other hand, if the teacher is basing the student's final grade on the ratio of total points earned to the total points possible, then the problem of finding the best set of r scores to drop is far more interesting. What we need is a subset $S \subset K = \{1, 2, 3, \ldots, k\}$ of $k - r$ retained grades so that the ratio $\sum_{j\in S} m_j / \sum_{j\in S} n_j$ is maximized. If all the quizzes are worth the same amount, that is, if all of the n_j are equal, then this reduces to finding the r smallest m_j values, just as it was in the above example.

Paradoxical behavior

Intuitively, one might suspect that a way to obtain an optimal solution would be to drop those quiz grades where the student performed the worst either by obtaining the smallest number of points or by obtaining the smallest percentage grade, $m_j(100\%)/n_j$. However, this is not always the case as the following examples illustrate. Consider Beth's quiz scores shown in Table 2. It is clear that Beth performed worst on quiz 3 where she obtained the smallest raw score (1) and the smallest percentage grade (5%). If that grade is dropped, Beth's remaining quiz grades would give a mean score of $(80 + 20)/(100 + 100) = 50\%$. On the other hand, if quiz 2 is dropped instead, she would receive a mean score of $(80 + 1)/(100 + 20) = 67.5\%$. The reason for this is that quiz 3 is not worth very many points, so its impact on the final score is much smaller than that of quiz 2.

TABLE 2: Beth's Quiz Scores

Quiz	1	2	3
Score	80	20	1
Possible	100	100	20
Percentage	80	20	5

One conclusion is clear. As long as the number of grades to drop is smaller than the total number of grades, the optimal retained set of grades will always contain the grade that has the largest percentage score. If more than one grade share the same largest percentage score, none of those grades will be dropped unless there are more of them than the number of retained grades. For example, with Beth's grades, quiz 1 will not be dropped. The reason for this is that if the retained set S contains any grade whose percentage is not the largest percentage, the average $\sum_{j\in S} m_j / \sum_{j\in S} n_j$ will be less than this largest percentage. S will then contain at least one grade whose percentage is less than or equal to the average of the grades in S. Removing that grade and replacing it with a grade with the largest percentage will raise the average since both the removal and the addition raise the average.

As seen with Beth's grades, the reverse argument does not work. That is, the grade with the smallest percentage score does not necessarily appear in the optimal deletion set. We can conclude that the grade with the smallest percentage will be among the grades retained if we want to get the smallest possible average score. But getting the smallest possible average score is not the goal.

One might hope that the best way to drop a set of r grades can be constructed inductively by finding the best one grade to drop, and then finding the best grade to drop from the remaining grades, and so on. This strategy turns out not to work. Consider Carl's quiz scores shown in Table 3. If we wish to drop just one grade, then the best score is obtained by dropping quiz 4 yielding an average of $(100 + 42 + 14)/(100 + 91 + 55) = 63.4\%$ as compared to 32.0% for dropping quiz 1, 60.6% for dropping quiz 2, and 63.3% for dropping quiz 3. If we need to drop two scores, it is best to drop quizzes 2 and 3 and retain quiz 4 to get the average $(100 + 3)/(100 + 38) = 74.6\%$ as compared to 74.3% for dropping quizzes 3 and 4, 73.5% for dropping quizzes 2 and 4, and 38.4% for dropping quizzes 1 and 4. Notice that the optimal deletion set of two grades does not include the best single grade to drop.

TABLE 3: Carl's Quiz Scores

Quiz	1	2	3	4
Score	100	42	14	3
Possible	100	91	55	38
Percentage	100	46	25	8

Also surprising is how slight changes to a problem can result in radically different results. To see this, consider Dale's eleven quiz grades displayed in Table 4. We consider several examples of Dale's quiz scores where c and each of the b_j are positive integers. Since quiz 0 is the only quiz with percentage over 50%, we would not want to drop quiz 0. If $A \subset \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ is the set of other quiz grades retained, the resulting average score is

$$\frac{20 + c + \frac{1}{2}\sum_{j\in A} n_j - \sum_{j\in A} b_j}{40 + \sum_{j\in A} n_j} = 0.5 + \frac{c - \sum_{j\in A} b_j}{40 + \sum_{j\in A} n_j}.$$

TABLE 4: Dale's Quiz Scores

Quiz	0	1	2	3	4	5
Score	$20 + c$	$21 - b_1$	$22 - b_2$	$23 - b_3$	$24 - b_4$	$25 - b_5$
Possible	40	42	44	46	48	50
Percentage	$50 + \frac{c}{.40}$	$50 - \frac{b_1}{.42}$	$50 - \frac{b_2}{.44}$	$50 - \frac{b_3}{.46}$	$50 - \frac{b_4}{.48}$	$50 - \frac{b_5}{.50}$
Quiz		6	7	8	9	10
Score		$26 - b_6$	$27 - b_7$	$28 - b_8$	$29 - b_9$	$30 - b_{10}$
Possible		52	54	56	58	60
Percentage		$50 - \frac{b_6}{.52}$	$50 - \frac{b_7}{.54}$	$50 - \frac{b_8}{.56}$	$50 - \frac{b_9}{.58}$	$50 - \frac{b_{10}}{.60}$

First, let us set $c = 4$ and each of the $b_j = 1$. If we drop five quiz grades, the average score will be

$$0.5 + \frac{4 - \sum_{j\in A} b_j}{40 + \sum_{j\in A} n_j} = 0.5 - \frac{1}{40 + \sum_{j\in A} n_j}.$$

So, to maximize this average we want A to represent the quizzes with the largest possible values, n_j, in order to make the denominator of the fraction as large as possible. Thus, the optimal deletion set is $\{1, 2, 3, 4, 5\}$.

But if we just change the value of c from 4 to 6, the average score becomes

$$0.5+\frac{6-\sum_{j\in A} b_j}{40+\sum_{j\in A} n_j}=0.5+\frac{1}{40+\sum_{j\in A} n_j}.$$

In this case we want A to represent the quizzes with the smallest possible values in order to make the denominator of the fraction as small as possible. Thus, the optimal deletion set is $\{6, 7, 8, 9, 10\}$. A slight change in c completely changed the optimal deletion set. Note that if $c = 5$, all deletion sets which do not include quiz 0 result in the same average of 50%, so every set A of size five gives the same optimal average.

Next, consider what happens with Dale's quiz scores when $c = 11$ and $b_j = 2$ for each j. If we drop four quiz grades, the set A will have six elements, and the average score will be

$$0.5+\frac{11-2\sum_{j\in A} b_j}{40+\sum_{j\in A} n_j}=0.5-\frac{1}{40+\sum_{j\in A} n_j}.$$

To maximize this average, we want to retain the quizzes with the largest possible scores, so the optimal deletion set is $\{1, 2, 3, 4\}$. If, on the other hand, we drop five quiz grades, the set A will have only five elements, and the average score becomes

$$0.5+\frac{11-2\sum_{j\in A} b_j}{40+\sum_{j\in A} n_j}=0.5+\frac{1}{40+\sum_{j\in A} n_j}.$$

To maximize this average, we want to retain the quizzes with the smallest possible scores, so the optimal deletion set is $\{6, 7, 8, 9, 10\}$ which has no elements in common with the optimal deletion set when we dropped only four grades.

Finally, Dale's quiz scores can be used to show that the optimal deletion set when dropping four grades can overlap with the optimal deletion set when dropping five grades to whatever extent we like. Indeed, let t represent the number of grades we wish the two optimal deletion sets to have in common where t is one of the numbers 1, 2, 3, or 4. Set $b_j = 3$ for each j from 1 to t, and set $b_j = 2$ for each $j > t$. Let $c = 11$. If we drop four quizzes, and s is the number of retained quizzes which have their $b_j = 3$, the set A will have six elements, and the average score will be

$$0.5+\frac{11-\sum_{j\in A} b_j}{40+\sum_{j\in A} n_j}=0.5-\frac{11-[s+2(6-s)]}{40+\sum_{j\in A} n_j}=0.5-\frac{1+s}{40+\sum_{j\in A} n_j}.$$

To maximize this average, s needs to be as small as possible (0), and we need to retain the quizzes with the largest possible score. This means the optimal deletion set is $\{1, 2, 3, 4\}$.

Now, if we drop five quiz scores, and s is the number retained quizzes which have their $b_j = 3$, the set A will have five elements, and the average score becomes

$$0.5+\frac{11-\sum_{j\in A} b_j}{40+\sum_{j\in A} n_j}=0.5+\frac{11-[s+2(5-s)]}{40+\sum_{j\in A} n_j}=0.5+\frac{1-s}{40+\sum_{j\in A} n_j}.$$

To maximize this average, s needs to be 0 or else the numerator $1 - s$ will be less than or equal to zero, and the average will not exceed 50%. Thus, this average will be maximized only when we drop all the quizzes with $b_j = 3$ and retain quizzes with the smallest possible score. This means the optimal deletion set is the set containing the quizzes with $b_j = 3$ and as many of the high numbered quizzes as needed. Thus, the overlap between the optimal deletion set when dropping four grades and the optimal deletion set when dropping five grades will be exactly the set of t grades with the $b_j = 3$.

Note that it is easy to construct examples similar to Dale's grades which include a very large number of quiz scores that exhibit the same paradoxical behavior as in the examples just given. Even though such examples exist only when the possible scores, the n_j values, are not all the same, paradoxical examples can still be constructed where the n_j value are all very close to each other, for example, within 1 of a fixed value. These examples are reminiscent of Simpson's Paradox (see [**6**]) which also deals with creating ratios by combining the numerators and denominators of other fractions.

Algorithms for finding the optimal deletion set

We return to the question of how one can identify the optimal deletion set when we want to drop r grades from a list of k quiz scores. One *brute force algorithm* would have us calculate the average grade for each possible set of $k - r$ retained grades. There are several well-known algorithms for enumerating all such subsets (see virtually any book on combinatorics, for example [**1**]). The arithmetic for calculating each average grade is straightforward. Unfortunately, even though checking any one average is very fast, the number of average grades which need to be calculated is given by the binomial coefficient $\binom{k}{r}$ which grows at a rate of $k^r/r!$. For small k and r, these calculations pose no problems. However, if a teacher wanted to drop just 10 grades from a list of 100 grades, even on a computer this algorithm would take far too long to be of any practical value.

The examples of the last section suggest that small changes in a problem can result in completely different optimal deletion sets. This indicates that we would run into difficulties by trying to implement either a *greedy algorithm* or a *dynamic programming algorithm.* These standard approaches to developing algorithms attempt to find solutions to problems by constructing an array of solutions to smaller problems which, in our case, have little bearing on the results of the original problem (see [**3**] for a discussion of how these methods are used to generate algorithms).

The optimal drop function

Our goal is to find the retained set $S \subset K = \{1, 2, 3, \ldots, k\}$ of size $k - r$ so that the ratio

$$\frac{\sum_{j\in S} m_j}{\sum_{j\in S} n_j} = q \tag{1}$$

is maximized. For each j define $f_j(q) = m_j - qn_j$. Then equation (1) is equivalent to

$$\sum_{j\in S} f_j(q) = 0. \tag{2}$$

Notice the the left-hand side of equation (1) is greater than q if and only if the left-hand side of equation (2) is greater than 0.

Since each $f_j(q)$ is a linear, decreasing function of q, for any given set S, $\sum_{j\in S} f_j(q)$ is also a linear, decreasing function of q. For a particular selection of retained grades, S, the equation $\sum_{j\in S} f_j(q) = 0$ is satisfied by the value of q which represents the average of the quizzes in S. We will have found the optimal set of retained quizzes, S_{best}, when we find the S where the associated average, q_{best}, is as large as possible. Define the *optimal drop function* to be

$$F(q) = \max\left\{\sum_{j\in S} f_j(q) : S \subseteq K, |S| = k - r\right\}. \tag{3}$$

Since F is the maximum of a finite number of linear, decreasing functions, it must be a piecewise linear, decreasing, concave up function. Moreover, $F(q_{\text{best}}) = 0$ since $\sum_{j\in S_{\text{best}}} f_j(q_{\text{best}}) = 0$ while for any other $S \subseteq K$ with $|S| = k - r$, it follows that $\sum_{j\in S} f_j(q_{\text{best}}) \le 0$.

Consider, for example, Carl's quiz scores from Table 2 where we drop two of four quizzes. There are six possible sets S and six associated sums shown in Figure 1.

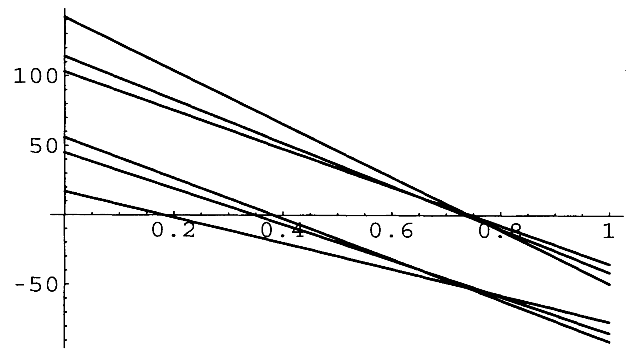


Figure 1 The six possible sums of two f_j

The function F in this case has the graph

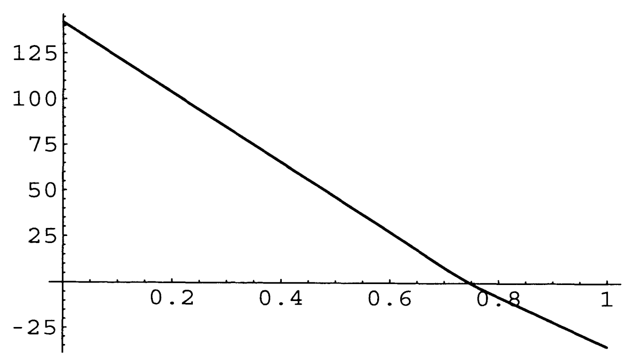


Figure 2 The graph of F

The problem of determining the best set of r grades to drop is now equivalent to finding the subset $S \subset K$ with $|S| = k - r$ and a rational number q, so that $F(q) = \sum_{j\in S} f_j(q) = 0$. The advantage of considering the function F is that it is a simple matter to evaluate $F(q)$ for any given q. Indeed, given a list of k grades $m_1, m_2, \ldots, m_k$ and k maximum possible scores $n_1, n_2, \ldots, n_k$, a number, r, of grades to drop, and a real number q, one merely has to evaluate each $f_j(q) = m_j - qn_j$ for each $j = 1, 2, \ldots, k$. Then one identifies the $k - r$ largest values among the $f_j(q)$ values. The set S becomes the set of j values corresponding to the largest $f_j(q)$ values. Finally, $F(q)$ is calculated as $\sum_{j\in S} f_j(q)$. Since there are well-known efficient algorithms for identifying the largest values out of a collection of numbers (see virtually any book about data structures or algorithms, for example, [**3**] or [**5**]), $F(q)$ can be calculated efficiently.

It remains to find the value of q where $F(q) = 0$. Since for a given S, $\sum_{j\in S} f_j(q)$ is linear, the graph of F can change slope at a value of q only if the associated set S changes at this value of q. For each q we can consider the collection of the k values of $f_j(q)$ for $j = 1, 2, 3, \ldots, k$. We can order these values in decreasing order. As the value of q changes, the values of the $f_j(q)$ change, and their order changes. Notice that the values of S depend only on the order of the $f_j(q)$, and hence the set S changes only when the order of $f_j(q)$ changes. Since each f_j is a continuous function, the order of $f_i(q)$ and $f_j(q)$ can change only for values of q where $f_i(q) = f_j(q)$. Notice that since each f_j is linear, this occurs at most once for every pair of i and j. Therefore, the set S cannot change at more than $\binom{k}{2}$ values of q since there are only that many pairs of i and j.

The condition $f_i(q) = f_j(q)$ occurs when $m_i - qn_i = m_j - qn_j$, or when

$$q = \frac{m_i - m_j}{n_i - n_j}.$$

Thus, if the graph of F changes slope at some value q, q has to be a rational number with denominator bounded by N (recall that N is an upper bound for all the n_i). Since

$$\frac{\sum_{j\in S_{\text{best}}} m_j}{\sum_{j\in S_{\text{best}}} n_j} = q_{\text{best}},$$

q_{best} is a rational number with denominator no larger than $(k - r)N$.

This can be used to find S_{best} and q_{best}. One could identify all the values of q where $f_j(q) = f_i(q)$ for some two values i and j. Then, by evaluating $F(q)$ at each of those points, the function F can be constructed since it is linear between each of those values of q. From this, one can easily find where $F(q) = 0$. But there are more efficient ways to find where $F(q) = 0$.

The bisection algorithm

An even more efficient algorithm is obtained by approximating the q where $F(q) = 0$ using the *bisection method* (see virtually any book about numerical analysis, for example, [**2**]). Since we know that q_{best} must lie in the interval between the minimum and maximum values of m_j/n_j, we begin by setting

$$q_{\text{high}} = \max_j \left\{\frac{m_j}{n_j}\right\}, \quad q_{\text{low}} = \min_j \left\{\frac{m_j}{n_j}\right\}, \quad \text{and} \quad q_{\text{middle}} = \frac{q_{\min} + q_{\max}}{2}.$$

Then we calculate $F(q_{\text{middle}})$ and its associated set S. If $F(q_{\text{middle}}) < 0$, we reset q_{low} to q_{middle}. Otherwise we reset q_{high} to q_{middle}. Finally, we reset q_{middle} to $(q_{\min} + q_{\max})/2$. We repeatedly calculate q_{middle}, $F(q_{\text{middle}})$, S, and reset q_{high}, q_{low}, and q_{middle} until

$$q_{\text{high}} - q_{\text{low}} < \frac{1}{2(k - r)N^2}.$$

At that point the value of S is S_{best}. Then, q_{best} can be calculated from S_{best}.

How do we know that this final set S is S_{best}? To answer this, we carefully consider the function F. Recall that F is piecewise linear, decreasing, and concave up. If F is linear in a neighborhood of q_{best}, then the distance between q_{best} and the next q where F changes slope is the distance between a rational number with denominator at most N and a rational number with denominator at most $(k-r)N$, which must be at least $1/[(k-r)N^2]$. So our approximation to q_{best} must be closer to q_{best} than to this closest point of slope change. Thus, the set S associated with this approximation is S_{best}.

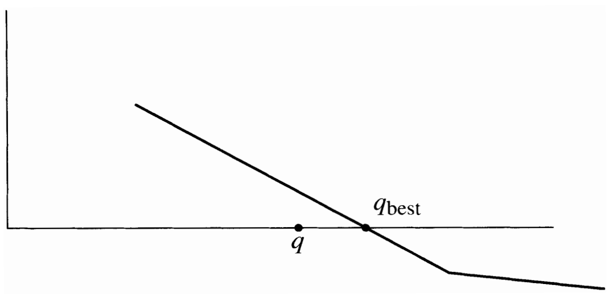


Figure 3 q and q_{best} when F is linear near q_{best}

If F were to change slope at q_{best}, then our approximation to q_{best} would be associated with one of two different sets of grades where both of these sets are associated with the average q_{best} and, thus, are equally good sets of grades to drop.

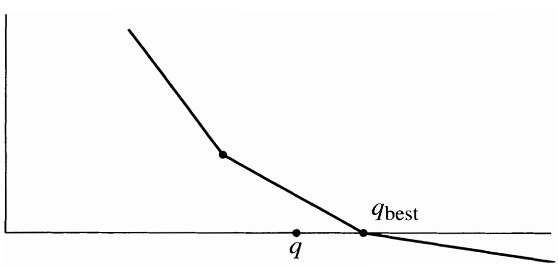


Figure 4 q and q_{best} when F is not linear at q_{best}

A more efficient algorithm

An improvement can be found in the bisection algorithm by considering the geometry of the graph of F.

Figure 5 shows several of the linear pieces which form the graph of F. Suppose the value $q_1 < q_{\text{best}}$ is chosen at random, and $F(q_1)$ is calculated yielding the associated set S_1 of grades to keep. Consider the linear piece of the graph of F passing through the point $(q_1, F(q_1))$. Let q_2 be the location where this linear piece crosses the x-axis. This q_2 is the average of the grades of S_1. Since the graph of F is concave up, q_2 lies strictly between q_1 and q_{best}. Iterating this process will yield a sequence of q_j which reach q_{best} after finitely many steps. At that point, $F(q)$ will be 0. If the value of q_1 happened to be larger than q_{best}, one iteration of this process will yield a q_2 less than or equal to q_{best}.

Note that the determination of the point where $F(q)$ is 0 poses no problem. Each q_j used in this algorithm will be a rational number. In practice, rather than calculating

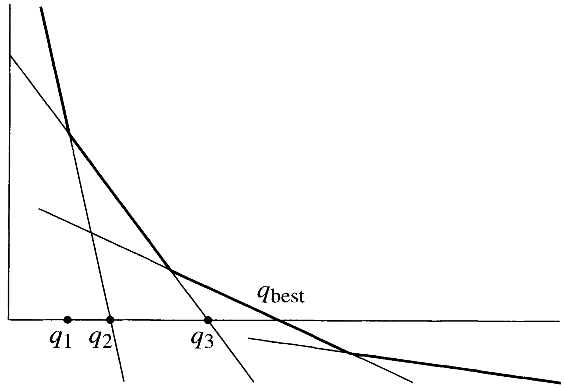


Figure 5 Sequence of q_j's approaching q_{best}

$F(q)$, one would calculate $F(q)$ multiplied by the denominator of q. Doing this allows $F(q)$ to be calculated using integer (fixed point) arithmetic which is not subject to the round-off error and inaccuracy problems common when using real (floating point) arithmetic.

Although it is not clear from the above discussion that this algorithm will run any faster than the bisection method algorithm, extensive running of simulations suggest that the algorithm always converges very rapidly requiring only a very small number of iterations to solve the most complicated problems. For example, we randomly generated many sets of quiz grades each containing 1,000 grades. Using this algorithm to drop 300 of the 1,000 grades, we never found a case where more than five iterations were needed to identify the optimal deletion set. This makes the algorithm particularly well suited for implementation in a computer gradebook program. Why does this algorithm converge so rapidly? Perhaps it is because it is essentially Newton's method applied to the piecewise linear function F.

Acknowledgment. We would like to thank the referees for their helpful suggestions in the preparation of this paper.

REFERENCES

1. Richard A. Brualdi, *Introductory Combinatorics* 3rd ed., Prentice-Hall, 1999.
2. Richard L. Burden & J. Douglas Faires, *Numerical Analysis* 8th ed., Brooks Cole, 2004.
3. T.H. Cormen, C.E. Leiserson, R.L. Rivest, & C. Stein, *Introduction to Algorithms* 2nd ed., MIT Press, 2001.
4. Jon Kane, *GRADE GUIDE* ver. 5.5, gradebook software, http://www.gradeguide.com, 2005.
5. Donald E. Knuth, *The Art of Computer Programming*, Addison-Wesley Reading, 1998.
6. Ronald Meester, *A Natural Introduction to Probability Theory*, Birkhäuser Verlag, 2003.

Compound Platonic Polyhedra in Origami

ZSOLT LENGVARSZKY
Claflin University
Orangeburg, SC 29115
zsolt@claflin.edu

Introduction

Origami, the fascinating art of paper folding, is a great source of both recreational and research mathematics. There is no textbook on the subject, however, there are conference proceedings [**1**], [**3**], [**8**], popular books [**4**], and web sites [**2**] addressing the various mathematical aspects of origami. In this article, we consider a problem that was motivated by a particular compound polyhedral design.

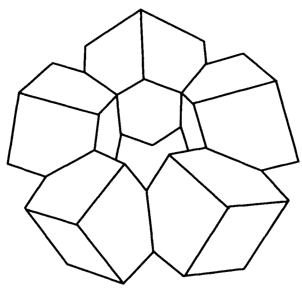


Figure 1 Ring-Of-Five-Cubes

Internationally known origamist David Mitchell invented the model "Ring-Of-Five-Cubes" (see [**6**]), and the model "Twenty-Cubes" (see [**5**], [**7**]). As the name suggests, the Ring-Of-Five-Cubes consists of five cubes, arranged in a ring so that one corner of every cube is inserted in the dimpled corner of one of its neighbors (see Figure 1). Alternatively, we may truncate two opposite corners on one face of each cube, and glue them together along the truncated vertices (which are now triangle faces) to form a ring. The Twenty-Cubes is an extension of this ring to what may be described as a dodecahedron of cubes. The cubes are centered at the vertices of a dodecahedron whose edges are represented by adjacent pairs of intersecting cubes (see Figure 2).

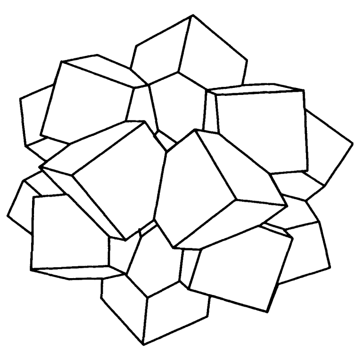


Figure 2 Twenty-Cubes

For both of these models the corners are dimpled/truncated in such a way that an equal portion of each edge at a particular vertex is cut off. It turns out that these designs are "nonmathematical", or "invalid", that is, if made with mathematical precision, the ring could not be completed since the angle formed by two adjacent cubes would be just a bit larger than the required 108° for a regular pentagon. This is not uncommon in origami: a model that is only an approximation of a valid, mathematically correct configuration can often be assembled without much difficulty due to the inherent imprecision in the paper folding process.

Jeannine Moseley [**9**] provided a short argument proving the invalidity of Ring-Of-Five-Cubes (and by extension the invalidity of the Twenty-Cubes), however, no further exploration was carried out concerning these designs.

In this paper we describe the correct geometry for these models, and investigate similar configurations as well. First let us formulate the undertaking in a somewhat more general context.

Generalization

In the Twenty-Cubes model either the cubes or the dodecahedron or both could be replaced by other regular polyhedra. Thus, we arrive at the following generalization of the Twenty-Cubes model:

> *Let A and B be Platonic solids. For each vertex V of B, place a copy A_V of A in space so that A_V is centered at V, and for adjacent vertices U and V, the copies A_U and A_V intersect. Describe the geometry of the compound that consists of the collection of A_V.*

We require that the intersection of two adjacent copies of A be *simple*, by which we mean that the truncation cuts off only one corner of A, and not more, the truncated portion is completely absorbed by the adjacent copy of A, and that every vertex of each copy of A is involved in at most one intersection.

In addition, we wish to consider compounds that exhibit a maximum amount of symmetry. Let us assume we are directly above a particular vertex V of B looking down on B and the copy of A at this vertex; we refer to this as the overhead view of the vertex V. We would like to see the same symmetric configuration of vertices, edges, and faces for every possible choice of V. This can happen in one of the following two ways:

Type 1. A vertex W of A coincides with V in the overhead view.

Type 2. The center C of one of the faces F of A coincides with V in the overhead view.

For both Type 1 and Type 2, we require rotational symmetry. More symmetry will be present when we have one of the following subtypes (all visualized in the overhead view):

Type 1.a. Edges of A with an endpoint at W lie on edges of B with an endpoint at V.

Type 1.b. Edges of A with an endpoint at W bisect the angles between adjacent edges of B with an endpoint at V.

Type 2.a. Edges of B with an endpoint at V go in the direction of vertices of F.

Type 2.b. Edges of B with an endpoint at V go in the direction of edge-midpoints of F.

Figure 3 illustrates the four subtypes. Solid lines are used to represent the edges of A and dashed lines to represent the edges of B.

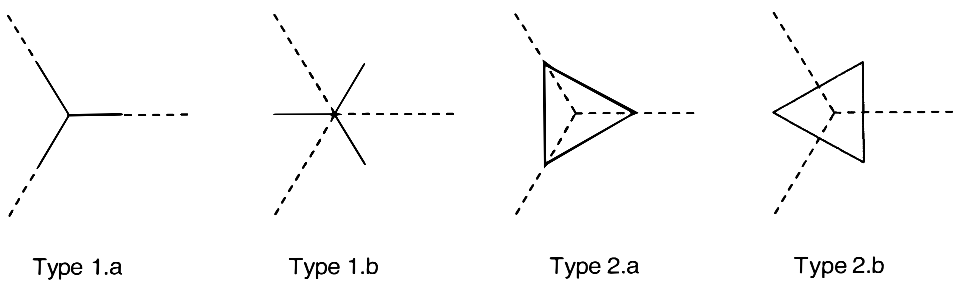


Figure 3 The four symmetry types

David Mitchell's Twenty-Cubes is of Type 1.b. We will see that there exists a Type 1.a Twenty-Cubes as well.

In some cases two types or subtypes may be regarded as one. For example, if A is a tetrahedron, for each vertex W there is a face F opposite W so that both W and the center C of F lie on the line connecting the center of A and the center of B. Thus, a Type 1.a configuration is also a Type 2.a configuration and vice versa. However, it seems natural to use Type 1.a if the vertex W is closer to the viewer and Type 2.a if the opposite face F is closer to the viewer. In the case when A is a cube, two opposite vertices can play the role of W and the configuration may be viewed as either Type 1.a or Type 1.b. Again, it is natural to base the type on the vertex that is closer to the viewer.

Not all possible combinations of A and B are meaningful. The octahedron has four edges at each of its vertices while this number is five for the icosahedron and three for the rest of the Platonic solids. For example, it is not possible to place tetrahedra at the vertices of an octahedron and to preserve the rotational symmetry.

We shall focus on the subtypes allowing for maximum symmetry. An important consequence is that the intersection of two adjacent copies of A is regular in the sense that edges of one copy of A intersect edges of the other copy. We describe the intersection using the truncating ratio. By truncating ratio we mean a triplet (for the most part) corresponding to the portions that are cut off from the edges at the dimpled vertex. For example, in the case of the Ring-Of-Five-Cubes and Twenty-Cubes, the ratio $1/2 : 1/2 : 1/2$, or equivalently, $1 : 1 : 1$, was used (as an approximation to the true ratio, as it turns out). For a given edge length of A, the actual values of a, b, and c are determined by the edge length of B, and we consider these values unimportant. Instead, we are interested in their relative proportions and thus consider two ratios $a : b : c$ and $ka : kb : kc$ equivalent.

The importance of the truncating ratio is signified by the fact that it describes all the essential aspects of the intersection geometry for the subtypes. In addition, the values in the truncating ratio are used directly in the folding pattern when origami models of these configurations are made. If we relax the symmetry requirements and consider Type 1 and 2 in general, the geometry will become more complex and the truncating ratio alone will not be sufficient to describe how two copies of A intersect.

Twenty-Cubes

Let us consider a Type 1.b configuration that is suggested by David Mitchell's design. Figure 4 shows two adjacent intersecting cubes in the dodecahedron of cubes (compare with Figure 2). Assuming unit edge length for the cubes, let a be the length of $\overline{OP}$ and b the length of $\overline{OQ}$. Then the length of $\overline{OR}$ is also a due to symmetry.

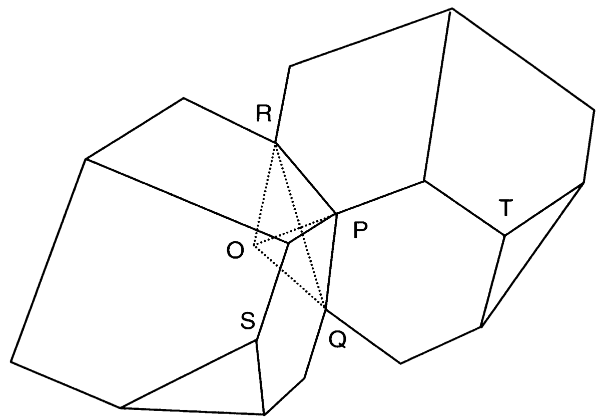


Figure 4 Two intersecting cubes

In fact, we can exploit symmetry further. The points S, P, and T, together with two more points not shown in the diagram form a regular pentagon providing one way to calculate d, the half-length of $\overline{ST}$:

$$d = (1-a)\sqrt{2}\sin 54^\circ = (1-a)\sqrt{2}\frac{\sqrt{5}+1}{4}$$

For a second way to calculate d, note that the points S and T are mirror images of each other with respect to the plane through P, Q, and R, also because of symmetry. Hence, d is the distance between the point T and the plane PQR. Choose a suitable coordinate system, say, with the origin at O and with P, Q, and R on the x-, y-, and z-axis, respectively. Then we have the following coordinates:

$$P(a,0,0);\ Q(0,b,0);\ R(0,0,a);\ T(1,1-a,0)$$

The equation of the plane through P, Q, and R is

$$bx + ay + bz - ab = 0$$

and d, the distance between this plane and the point T is

$$d = \frac{|b(1) + a(1-a) + b(0) - ab|}{\sqrt{b^2 + a^2 + b^2}} = \frac{|(a+b)(1-a)|}{\sqrt{a^2 + 2b^2}}$$

Comparing the two expressions for d, we obtain the following quadratic equation:

$$(\sqrt{5}-1)a^2 - 8ab + 2(\sqrt{5}+1)b^2 = 0$$

The solutions are

$$a_1 = (2+\sqrt{2})\frac{\sqrt{5}+1}{2}b \approx 5.5243,$$

$$a_2 = (2-\sqrt{2})\frac{\sqrt{5}+1}{2}b \approx 0.9478$$

The solution a_2 is the one we expect for the Twenty-Cubes design; the correct truncating ratio in this case is

$$a : a : b = 1 : 1 : b/a = 1 : 1 : \frac{(2+\sqrt{2})(\sqrt{5}-1)}{4} \approx 1 : 1 : 1.055$$

In a paper-folding project, particularly when the sizes are smaller, this is hardly distinguishable from the ratio 1 : 1 : 1.

What about a_1? It is the Type 1.a solution; Figure 5 shows this Twenty-Cubes design with truncating ratio

$$1 : 1 : \frac{(2-\sqrt{2})(\sqrt{5}-1)}{4} \approx 1 : 1 : 0.181$$

As we mentioned earlier, certain configurations may be associated with two types, and this is the reason why we have obtained the Type 1.a solution also, when in fact we assumed a Type 1.b configuration.

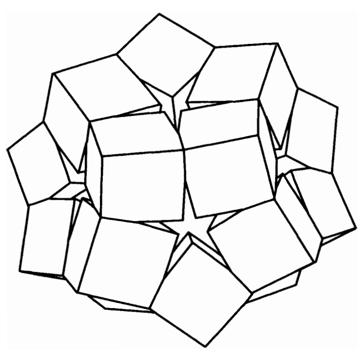


Figure 5 Twenty-Cubes, Type 1.a

Platonic solids built from cubes

The calculations of the previous section can be extended in an obvious manner. Replace 54° by 45° to get a Cube-Of-Cubes (Eight-Cubes) compound. Replace 54° by 30° to get a Tetrahedron-Of-Cubes (Four-Cubes) configuration. In the former case this is the equation:

$$2ab - b^2 = 0$$

Since for a simple intersection we expect positive lengths, only the Type 1.b solution, $b = 2a$, implying a truncating ratio 1 : 1 : 2, is of interest.

The equation for the Tetrahedron-Of-Cubes is

$$a^2 + 4ab = 0$$

This equation has no positive solution, implying a configuration that is non-simple (which we do not consider). Indeed for four cubes to be positioned at the vertices of a tetrahedron, they would have to intersect one another at more than one corner.

The remaining two cases are the octahedron of cubes and the icosahedron of cubes. The latter is not possible because of the incompatible symmetry groups. On the other hand, it is possible to build an octahedron from cubes. The Type 2.a configuration has the truncating ratio $1 : \sqrt{2} : \sqrt{2}$; we leave the calculations to the reader.

All configurations

Using an approach similar to that in the previous two sections, it is possible to determine the truncating ratios for all choices of A and B. However, some of the configurations are difficult to visualize and/or their computation may be rather tedious.

A better approach is to use an appropriate computer algebra system, both for visualization and computation. This method has its own challenges but it appears to be easier than doing the calculations for each case analytically.

For visualization we define and display B centered at O, and A_V, a copy of A, centered at a vertex V of B. It is unlikely that A_V is correctly oriented by default, thus it needs to be rotated. A rotation can be applied to make certain that one of the vertices (or face centers) of A_V lie on the line through O and V. Then a second rotation may be applied to obtain a particular symmetry type. Once A_V is correctly oriented, the remaining copies of A can be created by applying suitable transformations to A_V, or alternatively, the above procedure could be repeated for each vertex of B and the entire compound can be displayed.

To compute the truncating ratio, two correctly oriented adjacent copies of A are needed. First we determine the intersection point for each corresponding pairs of edges. Then the distances between these intersection points and the vertex of one of the copies of A that corresponds to the corner involved in the intersection can be computed. One of the challenges we encountered was to deal with the pages-long formulas that may occur for, say, the coordinates of a cube that was rotated a few times. If no explicit simplification is requested, the complexity of some of the expressions may overwhelm the computer algebra system or the computed values in the truncating ratio may not be in a desirable compact form.

The table below shows the complete list of Type 1.a, 1.b, 2.a, and 2.b configurations. Most of the results were obtained by using the computer algebra system Maple. Some were obtained analytically, and some were obtained both ways.

In each cell the ratios are preceded with the type in bold face. If a cell is empty, the configuration is not possible either because of incompatible symmetries or adjacent copies of A intersect in a non-simple fashion. When 4 (A = octahedron) or 5 (A = icosahedron) edges are involved in the intersection, the corresponding values occur cyclically in the truncating ratio.

Folding patterns

From a paper folder's point of view, the Twenty-Cubes, and the other configurations in this paper, belong to the field of modular (or macro modular) origami. In modular

		B				
		Tetrahedron	Hexahedron	Octahedron	Dodecahedron	Icosahedron
A	Tetrahedron	**1.a** $1:2:2$	**1.a** $3:4:4$		**1.a** $1:1:\frac{2}{3}+\sqrt{2}-\frac{\sqrt{2}\sqrt{5}}{3}$	
	Hexahedron		**1.b** $1:1:2$	**2.a** $1:\sqrt{2}:\sqrt{2}$	**1.a** $1:1:\frac{(2-\sqrt{2})(\sqrt{5}-1)}{4}$ **1.b** $1:1:\frac{(2+\sqrt{2})(\sqrt{5}-1)}{4}$	
	Octahedron	**2.b** $1:1:\frac{3}{5}:\frac{3}{5}$	**2.b** $1:1:1:1$		**2.b** $1:1:\sqrt{2}-\frac{\sqrt{10}-1}{3}:$ $\sqrt{2}-\frac{\sqrt{10}-1}{3}$	
	Dodecahedron	**1.b** $1:1:\frac{1}{2}+\frac{\sqrt{10}+\sqrt{2}}{4}$	**1.b** $1:1:\frac{\sqrt{10}-\sqrt{2}}{4}$		**1.b** $1:1:\frac{5-\sqrt{5}}{20}$	
	Icosahedron	**2.b** $\frac{5-\sqrt{10}+3\sqrt{2}-\sqrt{5}}{4}:$ $1:\frac{4\sqrt{2}+3-\sqrt{5}}{6}:$ $\frac{4\sqrt{2}+3-\sqrt{5}}{6}:1$	**2.a** $\frac{\sqrt{2}}{2}:1:\frac{1+\sqrt{5}+2\sqrt{2}}{2}$ $:\frac{1+\sqrt{5}+2\sqrt{2}}{2}:1$ **2.b** $\frac{\sqrt{2}}{2}+1:1:\frac{3-\sqrt{5}+2\sqrt{2}}{6}$ $:\frac{3-\sqrt{5}+2\sqrt{2}}{6}:1$		**2.a** $\frac{9-\sqrt{5}}{8}:1:\frac{29+\sqrt{5}}{22}$ $:\frac{29+\sqrt{5}}{22}:1$	**1.b** $\frac{\sqrt{5}}{2}:1:\frac{3\sqrt{5}-5}{2}:$ $\frac{3\sqrt{5}-5}{2}:1$

origami, several pieces of paper are used to fold individual, often identical, components that make up the final model. There are differing opinions about what should be allowed in true origami; some insist on square-shaped paper, others would allow any shape and even the use of scissors and tape or glue. The following example will perhaps interest most readers and satisfy the majority of paper folders.

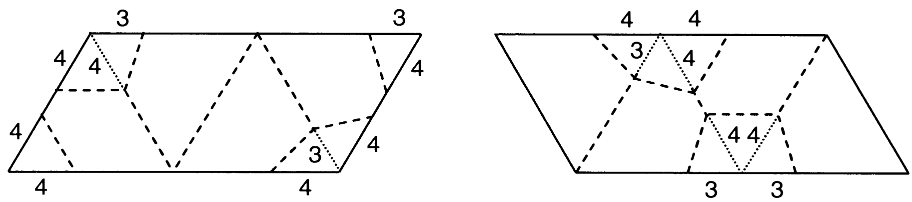


Figure 6 Folding pattern for the Eight-Tetrahedra

Figure 6 shows the crease pattern for an Eight-Tetrahedra (Cube-Of-Tetrahedra) compound using two parallelograms each consisting of four equilateral triangles. The valley- and mountain folds are shown with dashed and dotted lines, respectively. The 3s and 4s in the diagram correspond to the values in the truncating ratio 3 : 4 : 4.

Note that the size of the truncations does not matter so long as they are all the same and the two adjacent vertices truncations do not intersect. Hence, assuming the smaller side of the parallelogram is 1, the line segments marked by 4 could have a length between 0 and $1/2$.

Both pieces are needed for a tetrahedron; with glue one could use one of the two pieces only. For proper assembly, one piece must be left-handed and the other right-handed as shown in the picture.

The easiest way to make this model is to photocopy Figure 6 (perhaps after enlarging it), cut out the parallelograms, and fold them up along the crease lines. The two pieces wrap around each other so that the leftmost (large) equilateral triangle of the left parallelogram, rotated by 60 degrees counterclockwise, lies on the third (large) equilateral triangle of the right parallelogram.

A folding sequence, that would follow strict origami rules, can be found as well. David Mitchell [**6**] has a method to fold the parallelograms in Figure 6 from a rectangular piece of paper (see also T. S. Row [**10**] on how to fold equilateral triangles and other geometric objects). Adding the dimples would require finding $3/4$ of a given distance, which can be done by two midpoint constructions. For an optimal solution one would try to minimize the number of folds and avoid creases on the outside surface of the model.

In all, four tetrahedra with two dimpled corners, and four tetrahedra with one dimpled corner will be needed. To make a tetrahedron with just one dimpled corner, ignore the folds for any one of the two dimpled corners in Figure 6. Once the tetrahedra are made, they can be assembled without using tape or glue; resistance will hold them together. (It may help if one uses construction paper—less slippery than computer paper, and if the dimples are at least as large as suggested by Figure 6.) The Eight-Tetrahedra model is shown in Figure 7.

The assembly process is an entertaining puzzle in itself: a corner of a tetrahedron may be inserted into the dimpled corner of another in three different ways but at most one of the three possibilities will work. We encourage the reader to make his or her own Eight-Tetrahedra and to experiment with the other configurations in this paper as well.

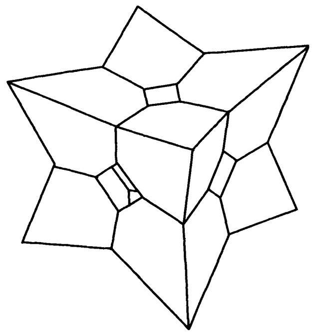


Figure 7 Eight-Tetrahedra

REFERENCES

1. T. Hull ed., *Origami3*, Third International Meeting of Origami Science, Mathematics, and Education, 2002, A. K. Peters Ltd.
2. T. Hull, *Origami Mathematics*, `http://www.merrimack.edu/~thull/OrigamiMath.html`
3. H. Huzita ed., Proceedings of the First International Meeting of Origami Science and Technology, Ferrara, 1989.
4. K. Kasahara, *Origami Omnibus: Paper-folding for Everybody*, Kodansha International, 1998.
5. D. Mitchell, Diagrams for 20-Cubes, *British Origami* (146), February, 1991.
6. D. Mitchell, *Mathematical Origami*, Tarquin Publications, Norfolk, England, 1997.
7. D. Mitchell, Origami Heaven, `http://www.mizushobai.freeserve.co.uk/20cubes.htm`
8. K. Miura ed., *Origami Science and Art*: Proceedings of the Second International Meeting of Origami Science and Scientific Origami, Seian University of Art and Design, Otsu, Japan, 1997.
9. Jeannine Mosely, *The Validity of the Orb, an Origami Model*, in Origami3, T. Hull, ed, Third International Meeting of Origami Science, Mathematics, and Education, 2002, A. K. Peters Ltd., 75–82.
10. T. S. Row, *Geometric Exercises in Paper Folding*, Dover Publications (New York), 1966.

Erratum

In the February 2006 issue of this MAGAZINE, in the note The Cross Ratio Is the Ratio of Cross Products! *by Leah Wrenn Berman, Gordon Ian Williams, and Bradley James Molnar*, 54–59, the passage below was inadvertently deleted from page 58 during the editing process. This quotation from Eves forms the basis for the discussion in the final three paragraphs of the note. The editor apologizes for this omission.

> Essentially the notation (AB, CD) was introduced by Möbius in 1827. He employed the term *Doppelschnitt-Verhältniss*, and this was later abbreviated by Jacob Stenier to *Doppelverhältniss*, the English equivalent of which is *double ratio*. Chasles used the expression *rapport anharmonique (anharmonic ratio)* in 1837, and William Kingdon Clifford coined the term *cross ratio* in 1878.

NOTES

The Bernoulli Trials 2004

CHRISTOPHER G. SMALL
Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, ON N2L 3G1
cgsmall@uwaterloo.ca

IAN VANDERBURGH
Centre for Education in Mathematics and Computing
University of Waterloo
Waterloo, ON N2L 3G1
iwtvande@uwaterloo.ca

On Saturday, March 13, 2004, a competition called *The Bernoulli Trials* was held at the University of Waterloo for the nth time, where $n = 8$. A total of 55 undergraduate students participated, all sharing insatiable appetites for mathematical problems and freshly baked croissants.

The Bernoulli Trials consists of a number of rounds, each of which involves contestants having to decide whether a mathematical statement is TRUE or FALSE, without the aid of a calculator. Rounds usually last 10 minutes, unless the organizers feel particularly malicious or hungry, in which case the round may be shortened (for example to 2 or 5 minutes). At the end of the round, students submit only their decision of TRUE or FALSE, obtained by some combination of skill, intuition and chance. Contestants can continue until they have made two mistakes, whereupon they are eliminated. The competition continues until one contestant is left standing and is declared champion.

The 2004 competition lasted a total of 14 rounds, and produced quite an exciting finish. After 8 rounds, only six contestants remained standing, each of whom had answered one problem incorrectly. So it was sudden death! In round 9, Iouri Khramtsov and Marcin Mika succumbed to sigma-itis and were eliminated. Round 10 saw the elimination of Yuli Ye, leaving only Raymond Chiu, Ralph Furmaniak, and Feng Tian, all of whom then answered correctly in round 11.

As round 12 closed, Ralph hastily changed his answer, leaving all three remaining contestants with an incorrect answer. After much deliberation and consultation of the BT Rule Book, the organizers determined that they could not eliminate all three remaining contestants and thus keep the prizes for themselves, so Raymond, Ralph, and Feng continued on.

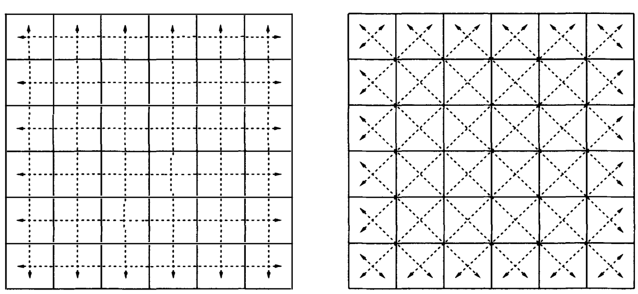
Ralph redeemed himself in round 13 by being the only contestant to answer correctly, and so was declared champion. In March 2004, Ralph Furmaniak (from London, ON) was a first-year student and a verteran of the IMO. It would later be known that Ralph was a Putnam Fellow in the 2003 Putnam Competition.

Round 14 was a two-minute tie-breaker for second place, which saw Raymond Chiu prevail as second place finisher, leaving Feng Tian in third place. (Our top-secret proof-reader did assure us that he managed to differentiate the given function 45 times in 1 minute 38 seconds, so 2 minutes seemed a reasonable length of time.) Thankfully, the organizers were prepared with enough problems for the 14 Rounds here. In fact, the organizers always bring infinitely many rounds to the Contest. (They could always make each round half as long as the previous round and still be done before lunch.)

The winners received a variety of prizes, from indescribable medals to mathematical T-shirts to stone plates. Congratulations go out to all competitors for a thrilling competition.

Here are the problems.

1. TRUE or FALSE?
   In the following list of statements, there is exactly one false statement:
   (a) $2004^3 - 2004$ is divisible by 3;
   (b) $2004^5 - 2004$ is divisible by 5;
   (c) $2004^7 - 2004$ is divisible by 7;
   (d) $2004^9 - 2004$ is divisible by 9;
   (e) $2004^{11} - 2004$ is divisible by 11.
2. TRUE or FALSE?
   The equation
   $$\sin^4 x - \sin x = \cos x - \cos^4 x$$
   has no solutions in the interval $(0, \pi/2)$.
3. A regular polygon has 2004 sides, each of which has length 1.
   TRUE or FALSE?
   The area between the circumscribed and inscribed circles of the polygon is greater than 1.
4. TRUE or FALSE?
   $$\left\lfloor \left(\tan\left(\tan^{-1}(987^{-1}) - \tan^{-1}(2584^{-1})\right)\right)^{-1} \right\rfloor = 1597.$$
5. We are given 36 squares arranged on a board in a 6×6 array. Therefore, there are 6 rows, 6 columns, and 22 diagonals, as shown in the figure.

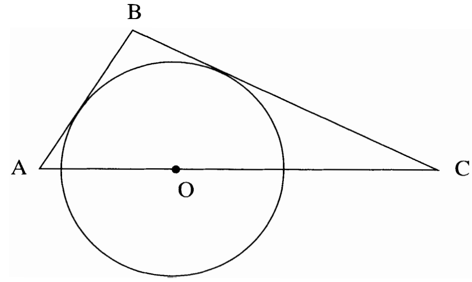


   TRUE or FALSE?
   It is possible to place 12 counters on the squares of this board—no 2 on the same square—so that each of the rows, columns, and diagonals of the board has at most 2 counters.
6. TRUE or FALSE?
   The number of pairs of positive integers (x, y), each less than 2004, whose arithmetic mean exceeds their geometric mean by 1 is 62.
7. Let a_n be the nth positive integer whose digits do not include 9 when written in base 10.

TRUE or FALSE?

$$\sum_{n=1}^{\infty} \frac{1}{a_n} \quad \text{diverges.}$$

8. In the figure, the circle shown has its center O on AC, and is tangent to AB and to BC. We also have $AO = 3$, $AB = 4$, and $BC = 8$.



TRUE or FALSE?
The area of $\triangle ABC$ is greater than or equal to $12\sqrt{2}$.

9. Let

$$S = 1 \cdot 3 \cdot 5 + 2 \cdot 4 \cdot 6 + 3 \cdot 5 \cdot 7 + \cdots + 2004 \cdot 2006 \cdot 2008$$

TRUE or FALSE?

$$4S = 2004 \cdot 2005 \cdot 2008 \cdot 2009$$

10. Let $A = 2004^2 + 17$.
TRUE or FALSE?
Among the five numbers

$$A,\ A+1,\ A+2,\ A+3,\ A+4$$

is a number that is relatively prime to the product of the other four.

11. TRUE or FALSE?

$$\lim_{k\to\infty} (2^k)\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2 + \cdots + \sqrt{2}}}}} = 2\sqrt{4 - \sqrt{2}}.$$

(There is a total of k 2s under the outermost square root.)

12. TRUE or FALSE?

$$\int_0^1 \frac{x^{2004} - 1}{\ln x}\, dx \le \ln 2004.$$

13. By an *ellipsoidal ball* in $\mathbb{R}^3$, we shall mean an ellipsoid in $\mathbb{R}^3$ together with its interior. Let a set of ellipsoidal balls $D_1, \cdots, D_n$ be given, as well as a plane π. Suppose that D_i and D_j have nonempty intersection whenever $i \neq j$.
TRUE or FALSE?
It is always possible to find a plane π' parallel to π that intersects each ellipsoidal ball.

14. TRUE or FALSE?
The coefficient of x^{45} in the power series expansion of

$$(1-x)^{-1}(1-x^3)^{-1}(1-x^9)^{-1}(1-x^{15})^{-1}$$

is 88.

Here are the solutions to the problems.

1. TRUE.
The divisibility statements for 3, 5, 7, and 11 are special cases of Fermat's Little Theorem. Since 9 is not prime that case does not follow from this theorem. It can be checked that 2004 is congruent to 6 mod 9. Since 2004 is divisible by 3 it follows that 2004^9 is divisible by 9.

2. TRUE.
In the interval $(0, \pi/2)$, $\sin x$ and $\cos x$ belong to $(0, 1)$. Thus, $0 < \sin^4 x < \sin x < 1$ and $0 < \cos^4 x < \cos x < 1$, so

$$\sin^4 x - \sin x < 0 < \cos x - \cos^4 x.$$

3. FALSE.
If the radius of the inscribed circle is r, then the radius of the circumscribed circle is $\sqrt{r^2 + (1/4)}$ by the Pythagorean Theorem.
So the area between the circles is

$$\pi\,(r^2 + 1/4) - \pi\, r^2 = \pi/4 < 1\,.$$

4. TRUE.
Since 987, 1597, and 2584 are Fibonacci numbers, and

$$(F_{2n-1})^2 = F_{2n-2}F_{2n} + 1$$

(where the numbering starts with $F_0 = 0$ and $F_1 = 1$),

$$\begin{aligned}\frac{1}{\tan\left(\tan^{-1}(987^{-1}) - \tan^{-1}(2584^{-1})\right)} &= \frac{1 + \frac{1}{987}\frac{1}{2584}}{\frac{1}{987} - \frac{1}{2584}} \\ &= \frac{2584(987) + 1}{1597} \\ &= 1597.\end{aligned}$$

5. TRUE.

●		●			
		●		●	
●	●				
				●	●
	●		●		
			●		●

6. FALSE.
We must have

$$\frac{1}{2}(x+y) = \sqrt{xy} + 1.$$

Rewriting this as $x + y - 2 = 2\sqrt{xy}$ and squaring both sides, we get

$$x^2 + y^2 + 4 - 2xy - 4x - 4y = x^2 - 2(y+2)x + (y-2)^2 = 0$$

By the quadratic formula, $x = (y+2) \pm \sqrt{8y}$. We will count the number of pairs with $x < y$, and double this total. If $x < y$, then $x = (y+2) - 2\sqrt{2y}$. Since x and y are integers, then y must be twice a perfect square. Since $y < 2004$, then $y = 2(1^2), 2(2^2), 2(3^2), \ldots, 2(31^2) = 1922$. Each gives a positive integer for x except for $y = 2$. So there are 30 values of y giving 30 pairs in this case, and 60 pairs in total.

7. FALSE.
There are $9(10^{k-1})$ positive integers with k digits. The reciprocal of each number is at most $1/10^{k-1}$. Of these integers, there are $8(9^{k-1})$ that do not use the digit 9. (There are 8 possibilities for the first digit and 9 for each of the remaining.) Therefore,

$$\sum_{n=1}^{\infty} \frac{1}{a_n} \leq \sum_{k=1}^{\infty} 8(9^{k-1}) \left(\frac{1}{10^{k-1}}\right) = 80$$

8. FALSE.
Since O is centre of the circle tangent to AB and BC it follows that OB bisects $\angle B$. Therefore

$$\frac{AO}{AB} = \frac{OC}{BC},$$

implying that $OC = 6$. Thus $AC = 9$. The semiperimeter of $\triangle ABC$ is therefore $21/2$.

Using Heron's formula for the area of a triangle we get

$$\begin{aligned} A &= \sqrt{\frac{21}{2} \cdot \left(\frac{21}{2} - 4\right) \cdot \left(\frac{21}{2} - 8\right) \cdot \left(\frac{21}{2} - 9\right)} = \frac{3}{4}\sqrt{455} \\ &< \sqrt{\frac{9}{16}(464)} = \sqrt{261} \\ &< \sqrt{288} = 12\sqrt{2}. \end{aligned}$$

9. TRUE.

$$\begin{aligned} S &= \sum_{n=3}^{N} (n-2)(n)(n+2) = \left(\sum_{n=3}^{N} n^3\right) - 4\left(\sum_{n=3}^{N} n\right) \\ &= \left(\sum_{n=1}^{N} n^3\right) - 4\left(\sum_{n=1}^{N} n\right) - 1^3 - 2^3 + 4(1) + 4(2) \\ &= \frac{N^2(N+1)^2}{4} - 2N(N+1) + 3 \\ &= \frac{1}{4}[N^2(N+1)^2 - 8N(N+1) + 12] \\ &= \frac{1}{4}[N(N+1) - 2][N(N+1) - 6] \end{aligned}$$

$$= \frac{1}{4}(N+2)(N-1)(N+3)(N-2)$$
$$= \frac{1}{4}(N-1)(N-2)(N+2)(N+3)$$

as required.

10. TRUE.
The result is true for any positive integer A, not just the one given. Two numbers that differ by at most four can have only factors ≤ 4 in common. Thus the only prime factors that any two can have in common are 2, 3. There are at least two odd numbers in any consecutive set of five numbers. Taking two consecutive odd numbers from the list, say B, $B+2$, we observe that one of them will not be divisible by 3. As this number is not divisible by 2, it must be relatively prime to all of the other numbers. Thus it is relatively prime to the product of the other numbers.

11. FALSE.
Using the fact that $\cos(\frac{1}{2}x) = \sqrt{\frac{1+\cos x}{2}}$, we can determine that

$$\cos\left(\frac{\pi}{2^k}\right) = \sqrt{\frac{2+\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}}{4}},$$

where there are $k-1$ 2s in the numerator.
Using the fact that $\sin\left(\pi/2^k\right) = \sqrt{1-\cos^2\left(\pi/2^k\right)}$, we get that

$$\sin\left(\frac{\pi}{2^k}\right) = \sqrt{\frac{2-\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}}{4}}$$

Using the fact that $\lim_{x\to 0}\frac{\sin \pi x}{x} = \pi$, we get that $\lim_{k\to\infty} 2^k \sin\left(\pi/2^k\right) = \pi$, so

$$\lim_{k\to\infty} 2^{k-1}\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}}} = \pi$$

where there are $k-1$ 2s under the outermost square root.

12. FALSE.
Let

$$g(\alpha) = \int_0^1 \frac{x^\alpha - 1}{\ln x}\,dx\,.$$

Then

$$g'(\alpha) = \int_0^1 x^\alpha\,dx = \frac{1}{\alpha+1},$$

provided that $\alpha > -1$. Integrating we get

$$g(\alpha) = \ln(\alpha+1) + C.$$

Checking $\alpha = 0$ we see that $C = 0$.

13. TRUE.
    Given such a set of ellipsoidal balls, project the ellipsoidal balls orthogonally onto a line ℓ perpendicular to π. By Helly's theorem in dimension one, these projections must have a common point. Take a plane π' parallel to π through this common point.
14. FALSE.

$$(1-x)^{-1}(1-x^3)^{-1}(1-x^9)^{-1}(1-x^{15})^{-1}$$
$$= \left(\sum_{a=0}^{\infty} x^a\right)\left(\sum_{b=0}^{\infty} x^{3b}\right)\left(\sum_{c=0}^{\infty} x^{9c}\right)\left(\sum_{d=0}^{\infty} x^{15d}\right),$$

so we need to count the number of 4-tuples (a, b, c, d) of nonnegative integers such that $a + 3b + 9c + 15d = 45$. First, a must be a multiple of 3, say $a = 3A$, which leads to $A + b + 3c + 5d = 15$.

$d = 3$ gives $A = b = c = 0$. (1 way.)
$d = 2$ gives $c = 1$ or 0 giving 3 and 6 possibilities for the pair (A, b). (9 ways.)
$d = 1$ gives $c = 3, 2, 1$ or 0, giving 2, 5, 8, and 11 possibilities. (26 ways.)
$d = 0$ gives $c = 5, 4, 3, 2, 1,$ or 0, giving 1, 4, 7, 10, 13, and 16 possibilities. (51 ways.)

There are 87 ways, so the coefficient of x^{45} is 87.

Since the initial writing of this Note, the Bernoulli Trials were held in 2005 for the $(n+1)$st time. After 9 rounds of competition, only 7 students remained standing, and of these only 1 had made no mistakes to that point. Round 10 saw the elimination of all but one of the remaining competitors, leaving the winner, Lino Demasi. But then ties needed to be broken so these 6 eliminated competitors were brought back in for Round 11, which eliminated all but two of them: Ian Baillargeon and Ralph Furmaniak. In a valiant effort to break this tie for 2nd/3rd place, the organizers tried to get them to crack under pressure. After two rounds with equal responses, a tie was declared for 2nd/3rd place.

The Comparison Test— Not Just for Nonnegative Series

MICHELE LONGO
Università Cattolica del Sacro Cuore
Milan, Italy
michele.longo@unicatt.it

VINCENZO VALORI
Università degli Studi di Firenze
Florence, Italy
vincenzo.valori@dmd.unifi.it

Is it possible to generalize the Comparison Test to generic real series? More precisely, is it true that, given $a_n \le b_n \le c_n$ for all sufficiently large natural numbers n, the convergence of $\sum b_n$ follows from the convergence of $\sum a_n$ and $\sum c_n$? At first glance, many of us (certainly the authors) could argue something like "*If it were true then it would certainly appear in some of the books standing on the shelves in my office.*" As a matter of fact, all the books on the authors' shelves state the test only for nonnegative series. In particular, Hardy [**5**, p. 376, lines 1–3] considers possible extensions involving the comparison of only two series and affirms: "*... there are no comparison tests for convergence of conditionally convergent series.*"

A generalization of the comparison test The Comparison Test is usually stated only for nonnegative real series, both in calculus books [**1, 4, 5, 8**] and in more specialized texts [**2, 3, 6, 7**]. There could be many reasons for this; however, the most immediate generalization can be used to establish the convergence or divergence of many series for which standard tests do not apply. Also, the proof is so straightforward that, at least, it could be considered as an exercise in calculus courses. The Comparison Test can be generalized as follows.

THEOREM. *Let $\sum a_n$, $\sum b_n$, and $\sum c_n$ be three real series such that $a_n \leq b_n \leq c_n$ for all sufficiently large $n \in \mathbb{N}$, then:*

(i) *the series $\sum b_n$ converges if $\sum a_n$ and $\sum c_n$ converge;*

(ii) *the series $\sum b_n$ diverges to $+\infty$ if $\sum a_n$ diverges to $+\infty$; and*

(iii) *the series $\sum b_n$ diverges to $-\infty$ if $\sum c_n$ diverges to $-\infty$.*

Proof. Assume $a_n \leq b_n \leq c_n$ for all natural numbers $n \geq N$. (i) If $\sum a_n$ and $\sum c_n$ converge, then $\sum (c_n - a_n)$ converges. The relation $0 \leq b_n - a_n \leq c_n - a_n$ and the usual Comparison Test imply that $\sum (b_n - a_n)$ converges. Since $\sum b_n = \sum (b_n - a_n) + \sum a_n$, we conclude that $\sum b_n$ converges. Parts (ii) and (iii) follow from the relation

$$\sum_{n=N}^{M} a_n \leq \sum_{n=N}^{M} b_n \leq \sum_{n=N}^{M} c_n,$$

which is true for each $M \geq N$. ∎

In the Theorem, the only requirement on $\sum a_n$ and $\sum c_n$ is that they converge. Clearly, the case of interest is that of conditionally convergent series (that is, convergent but not absolutely convergent, see Hardy [**5**, p. 375]). It is worth noting that the proof of our Theorem can be adapted to show that if $a_n \leq b_n$ and one of the two series $\sum a_n$ or $\sum b_n$ converges then the other must either converge or else have an infinite limit. We give now an example.

EXAMPLE 1. Consider the alternating real series $\sum b_n$ where the generic term

$$b_n = \ln\left(1 + \frac{(-1)^n}{n^\gamma}\right)$$

depends on the positive real parameter γ. Normally, we would apply the Alternating Series Test (also known as Leibniz Test) which states that an alternating series $\sum b_n$ converges if $|b_n|$ is monotonic decreasing and $b_n \to 0$ [**3**, p. 55]. However, the test applies in this case if and only if $\gamma \geq 1$, because otherwise the sequence $|b_n|$ is not decreasing. To see this, observe that, if n is even,

$$\left|\ln\left(1 + \frac{(-1)^n}{n^\gamma}\right)\right| = \ln\left(1 + \frac{1}{n^\gamma}\right) = \ln\left(1 + \frac{2}{2n^\gamma - 1 + (-1)^n}\right);$$

whereas, if n is odd,

$$|b_n| = -\ln\left(1 - \frac{1}{n^\gamma}\right) = \ln\left(1 + \frac{1}{n^\gamma - 1}\right) = \ln\left(1 + \frac{2}{2n^\gamma - 1 + (-1)^n}\right).$$

Therefore, $|b_n| = \ln\left(1 + 2/\left(2n^\gamma - 1 + (-1)^n\right)\right)$ for every $n \in \mathbb{N}$ and

$$|b_n| \geq |b_{n+1}| \Leftrightarrow n^\gamma \leq (n+1)^\gamma + (-1)^{n+1}.$$

The second inequality is verified for all positive γ if n is odd, but holds only for $\gamma \geq 1$ if n is even. Furthermore, $\sum b_n$ is absolutely convergent if and only if $\gamma > 1$, as follows. The power series expansion of $\ln(1+x)$ allows us to assert that $|x|/2 \leq |\ln(1+x)| \leq 2|x|$, as long as $|x|$ is sufficiently small; this in turn implies that

$$0 < 1/2n^\gamma < \left|\ln\left(1 + (-1)^n/n^\gamma\right)\right| < 2/n^\gamma$$

for sufficiently large n and the usual Comparison Test yields the result.

This analysis shows that we cannot use standard tests to study $\sum b_n$ when $0 < \gamma < 1$, so let us proceed as follows: First observe that

$$x - 3x^2/4 \leq \ln(1+x) \leq x - x^2/4$$

in a neighborhood of zero (again from the power series expansion), so the substitution $x = (-1)^n/n^\gamma$ yields

$$\frac{(-1)^n}{n^\gamma} - \frac{3}{4n^{2\gamma}} \leq \ln\left(1 + \frac{(-1)^n}{n^\gamma}\right) \leq \frac{(-1)^n}{n^\gamma} - \frac{1}{4n^{2\gamma}}$$

for sufficiently large n. Second, notice that by the Alternating Series Test, $\sum (-1)^n/n^\gamma$ converges for all $\gamma > 0$ whereas $\sum 1/n^{2\gamma}$ converges if and only if $\gamma > 1/2$. Finally, by setting $a_n = (-1)^n/n^\gamma - 3/4n^{2\gamma}$ and $c_n = (-1)^n/n^\gamma - 1/4n^{2\gamma}$, an application of our Theorem enables us to conclude that

$$\text{if } \gamma > \frac{1}{2}, \text{ then } \sum b_n \text{ converges, and if } 0 < \gamma \leq \frac{1}{2}, \text{ then } \sum b_n \text{ diverges to } -\infty.$$

The previous example suggests a general procedure that can be useful to decompose the general term of the series into simpler quantities, which can be analyzed separately. Our generalized Comparison Test then applies to give information about the original series. This procedure is summarized in the following proposition.

PROPOSITION. *Let $\sum a_n$ be convergent and suppose f is a real valued function such that, in a neighborhood of* 0,

$$f(x) = \alpha x + \beta x^{2k} + o\left(x^{2k}\right), \quad \beta \neq 0, k \in \mathbb{N}.$$

Then $\sum f(a_n)$ converges if and only if $\sum (a_n)^{2k}$ converges.

Proof. By assumption there exists $\varepsilon > 0$ such that for $|x| < \varepsilon$ we have

$$\alpha x + \left(\beta - \frac{|\beta|}{2}\right)x^{2k} \leq f(x) \leq \alpha x + \left(\beta + \frac{|\beta|}{2}\right)x^{2k}.$$

Hence, there exists N_ε such that, for each $n > N_\varepsilon$,

$$\alpha a_n + \left(\beta - \frac{|\beta|}{2}\right)(a_n)^{2k} \leq f(a_n) \leq \alpha a_n + \left(\beta + \frac{|\beta|}{2}\right)(a_n)^{2k}.$$

Now, convergence of $\sum f(a_n)$ follows from convergence of $\sum (a_n)^{2k}$ and part (*i*) of the Theorem. Conversely, if $\sum (a_n)^{2k}$ diverges then parts (*ii*) and (*iii*) imply that also $\sum f(a_n)$ diverges. ■

REMARK. The Proposition cannot be extended to the case where the expansion ends with an odd power (Examples 4 and 5 show this point). Nevertheless, if $\sum a_n$ converges and

$$f(x) = \alpha x + \beta x^{2k+1} + o\left(x^{2k+1}\right), \quad \beta \neq 0, k \in \mathbb{N},$$

then

$$\sum |a_n|^{2k+1} \text{ converges} \Rightarrow \sum f(a_n) \text{ converges.} \tag{1}$$

Finally, notice that a sufficient condition for the convergence of $\sum |a_n|^{2k+1}$ is the convergence of $\sum (a_n)^{2i}$ for some $i \in \{1, 2, \ldots, k\}$.

This Remark gives only a sufficient condition for the convergence of $\sum f(a_n)$. For this reason, it may be possible to obtain more information about the convergence of the series by refining the Taylor expansion of f. For example, consider the series $\sum \arctan\left((-1)^n / \sqrt[4]{n}\right)$. In this case $\sum \left((-1)^n / \sqrt[4]{n}\right)^{2k+1}$ converges for all natural numbers k whereas it converges absolutely if and only if $k \geq 2$; so if we use $\arctan x = x - x^3/3 + o\left(x^3\right)$, we can say nothing about the convergence of the series since $\sum \left|(-1)^n / \sqrt[4]{n}\right|^3$ does not converge; on the contrary, if we consider the expansion up to the fifth order we can conclude that the series is convergent. A less trivial example of this occurrence is the following.

EXAMPLE 2. Consider the alternating real series $\sum b_n$, whose generic term, depending on the positive real parameter γ, is defined by

$$b_n = \tan\left(\frac{(-1)^n}{n^\gamma} + \frac{1}{n^{\gamma+1}}\right).$$

The Taylor expansion of $\tan x$ up to the third order is $\tan x = x + x^3/3 + o\left(x^3\right)$. The series $\sum a_n = \sum \left((-1)^n / n^\gamma + 1/n^{\gamma+1}\right)$ converges for all $\gamma > 0$, whereas

$$\sum |a_n|^3 = \sum \left|\frac{(-1)^n \left(n + (-1)^n\right)}{n^{\gamma+1}}\right|^3 = \sum \left(\frac{1}{n^\gamma} + \frac{(-1)^n}{n^{\gamma+1}}\right)^3$$

converges if and only if $\gamma > 1/3$. Hence, by the Remark we conclude that $\sum b_n$ converges if $\gamma > 1/3$. It is worth noting that, on the basis of the Remark, we can say nothing about the behavior of the series when $0 < \gamma \leq 1/3$. Nevertheless, the series converges for all $\gamma > 0$ as can be proved by considering more terms from the Taylor expansion of $\tan x$ and observing that the series $\sum_n \left((-1)^n / n^\gamma + 1/n^{\gamma+1}\right)^k$ converges for all $\gamma > 0$ if k is odd and $\sum \left|(-1)^n / n^\gamma + 1/n^{\gamma+1}\right|^k$ converges if and only if $\gamma > 1/k$. Notice that $|x| < |\tan x| < 2|x|$ for x small enough, hence

$$0 < \frac{1}{n^\gamma} + \frac{(-1)^n}{n^{\gamma+1}} < \left|\tan\left(\frac{(-1)^n}{n^\gamma} + \frac{1}{n^{\gamma+1}}\right)\right| < \frac{2}{n^\gamma} + \frac{2(-1)^n}{n^{\gamma+1}}$$

and so the series is absolutely convergent if and only if $\gamma > 1$. Finally, the Alternating Series Test applies if and only if $\gamma > 2$. Indeed, since $\tan x$ is increasing and

$|\tan x| = \tan |x|$ for $|x| < \pi/2$, it follows that

$$|b_n| \geq |b_{n+1}| \Leftrightarrow \frac{1}{n^\gamma} + \frac{(-1)^n}{n^{\gamma+1}} \geq \frac{1}{(n+1)^\gamma} + \frac{(-1)^{n+1}}{(n+1)^{\gamma+1}}$$
$$\Leftrightarrow \frac{n + (-1)^n}{(n+1) + (-1)^{n+1}} \geq \left(\frac{n}{n+1}\right)^{\gamma+1}.$$

The last inequality is verified for all positive γ if n is even, whereas it is true only for $\gamma > 2$ if n is odd and sufficiently large.

The following example concerns a series with no regularity in sign.

EXAMPLE 3. Consider the real series $\sum b_n$, whose generic term, depending on the parameters $\alpha, \gamma \in \mathbb{R}$, $\gamma > 0$, is defined by

$$b_n = \exp\left(\frac{\sin \alpha n}{n^\gamma}\right) - 1.$$

If α is multiple of π then the series is identically zero. For $\alpha \neq k\pi$, consider the Taylor expansion, $e^x - 1 = x + x^2/2 + o\left(x^2\right)$. The series $\sum (\sin \alpha n)/n^\gamma$ converges for all $\alpha, \gamma \in \mathbb{R}$, $\gamma > 0$ by Dirichlet's Test [7, Theorem 3, p. 137] whereas $\sum \left(\sin^2 \alpha n\right)/n^{2\gamma} = \sum (1 - \cos 2\alpha n)/2n^{2\gamma}$ converges if and only if $\gamma > 1/2$. Hence, by the Proposition, we have that

$$\sum b_n \text{ converges if } \gamma > \frac{1}{2},$$
$$\sum b_n \text{ diverges to } +\infty \text{ if } 0 < \gamma \leq \frac{1}{2}.$$

Finally, observe that $\sum b_n$ is absolutely convergent if and only if $\gamma > 1$ [3, p. 60].

Counterexamples We now provide counterexamples to show why results cannot be extended in certain ways.

The Proposition cannot be extended to odd powers much more than we have done in our Remark. Indeed, if $\sum a_n$ converges and the function f satisfies

$$f(x) = \alpha x + \beta x^{2k+1} + o\left(x^{2k+1}\right), \quad \beta \neq 0, k \in \mathbb{N},$$

then no implications can be drawn in general between the convergence of $\sum (a_n)^{2k+1}$ and $\sum f(a_n)$; that is

$$\sum (a_n)^{2k+1} \text{ converges } \begin{array}{c}\not\Rightarrow \\ \not\Leftarrow\end{array} \sum f(a_n) \text{ converges.}$$

EXAMPLE 4. ($\not\Rightarrow$) Consider the real series $\sum a_n$, where

$$a_n = \frac{(-1)^n}{\sqrt[4]{n}}$$

and the function $f(x) = x + x^3 + x^4$.

EXAMPLE 5. ($\not\Leftarrow$) Consider the real series $\sum a_n$, where

$$a_{2k} = \frac{(-1)^k}{\sqrt[4]{2k}} + \frac{1}{3\sqrt{2k}} \quad \text{and} \quad a_{2k+1} = -\frac{1}{3\sqrt{2k}}$$

and the function $f(x) = x + x^3 - x^4$. In this case $\sum a_n$ and $\sum f(a_n)$ both converge, but $\sum a_n^3$ and $\sum a_n^4$ do not.

The following example shows that the converse of (1) is also not true; that is, if $\sum a_n$ converges and $f(x) = \alpha x + \beta x^{2k+1} + o\left(x^{2k+1}\right)$ then

$$\sum |a_n|^{2k+1} \text{ converges } \not\Leftarrow \sum f(a_n) \text{ converges.}$$

EXAMPLE 6. Consider the function $f(x) = \sin x = x + x^3/6 + o\left(x^3\right)$ and the series $\sum (-1)^n / \ln n$. Then $\sum \sin((-1)^n / \ln n)$ converges (Alternating Series Test) but $\sum \left|(-1)^n / \ln n\right|^3 = \sum 1/\ln^3 n$ does not.

We conclude by observing that the Limit Comparison Test, which, for series of positive terms, states that $\sum a_n$ and $\sum b_n$ behave the same if $a_n \sim b_n$ [**5**, p. 342, §173 D], does not extend to series without restrictions on sign. Take, for instance, b_n as in Example 1 and $a_n = (-1)^n / n^\gamma$.

Acknowledgment. We thank Giulio Pianigiani and Sergio Vessella for many valuable suggestions. The authors would also like to thank the anonymous referees for their comments. The usual disclaimer applies. The first author acknowledges Università Cattolica di Milano for financial support.

REFERENCES

1. T.M. Apostol, *Calculus Vol. 1*, John Wiley & Sons, New York, 1967.
2. G. Arfken, *Mathematical Methods for Physicists*, Academic Press, Orlando, 1985.
3. T.J.I'a. Bromwich and T.M. MacRobert, *An Introduction to the Theory of Infinite Series*, McMillan & Co., London, 1926.
4. R. Courant and F. John, *Introduction to Calculus and Analysis*, Springer, New York, 1989.
5. G.H. Hardy, *A Course of Pure Mathematics*, Cambridge University Press, Cambridge, 1993.
6. K. Knopp, *Theorie und Anwendung der unendlichen Reihen*, Springer, Berlin, 1922.
7. K. Knopp, *Infinite Sequences and Series*, Dover, New York, 1956.
8. W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, Auckland, 1976.

From the Cauchy-Riemann Equations to the Fundamental Theorem of Algebra

ALAN C. LAZER
University of Miami
Coral Gables, FL 33124
a.lazer@math.miami.edu

Beginning with the doctoral dissertation of C.F. Gauss, there have been at least eleven proofs of the Fundamental Theorem of Algebra: For every nonconstant complex polynomial $p(z)$, there is a complex number z_0 such that $p(z_0) = 0$ [**3, 7, 8**]. Why then would anybody be interested in another proof?

Our principal reason is that it gives an application of the Cauchy-Riemann equations which is usually taught in the first two weeks of an undergraduate complex variable course.

Our proof has three main ingredients. The first reduces to first-semester calculus.

LEMMA 1. *If $w(x, y)$ is a real-valued function on $\mathbb{R}^2$ with second-order partial derivatives and $w(x, y)$ attains its maximum at (x_0, y_0), then*

$$\Delta w(x_0, y_0) \equiv \frac{\partial^2 w}{\partial x^2}(x_0, y_0) + \frac{\partial^2 w}{\partial y^2}(x_0, y_0) \leq 0.$$

Proof. The two functions of one variable $x \mapsto w(x, y_0)$, $\mathbb{R} \mapsto \mathbb{R}$, and $y \mapsto w(x_0, y)$, $\mathbb{R} \mapsto \mathbb{R}$, attain their maxima at x_0 and y_0 respectively. From the second derivative test of single-variable calculus it follows that $\frac{\partial^2 w}{\partial x^2}(x_0, y_0) \leq 0$ and $\frac{\partial^2 w}{\partial y^2}(x_0, y_0) \leq 0$, which proves Lemma 1. ∎

Our second ingredient is a property of complex polynomials.

LEMMA 2. *Let $Q(z)$ be a nonzero complex polynomial. There exists a number $d > 0$ such that $Q(a) = 0$ implies $Q(a + d) \neq 0$.*

Proof. From the theorem about factoring polynomials, it follows that $Q(z) = 0$ has a finite number of solutions. The assertion of the lemma follows by taking d to be a positive number that is less than the distance between any two distinct roots of Q. This proves Lemma 2. ∎

Our third ingredient is an identity, which we will prove after a brief review.

Let D be an open subset of $\mathbb{C}$. Recall that a complex-valued function defined on D is holomorphic in D if it is differentiable at each point of D. We identify the complex numbers $\mathbb{C}$ with $\mathbb{R}^2$ in the usual way and recall that if $f(z) = u(x, y) + iv(x, y)$, where u and v are real, is holomorphic in D, the well-known Cauchy-Riemann equations imply that

$$f'(z) = u_x(x, y) + iv_x(x, y) = v_y(x, y) - iu_y(x, y).$$

It follows that if f' is also holomorphic in D, then

$$f''(z) = u_{xx}(x, y) + iv_{xx}(x, y) = -u_{yy}(x, y) - iv_{yy}(x, y).$$

Equating real and imaginary parts in the above equation, we find that $\Delta u(x, y) = 0$ and $\Delta v(x, y) = 0$.

The following facts are usually assumed without proof in a complex variable course, since the proofs are the same as in a single-variable calculus course [**1, 2**].

(i) The sum and product of two holomorphic functions in D are holomorphic in D and the usual formulas hold for derivatives of sums and products.

(ii) If g is holomorphic in D and $g(z) \neq 0$ for every $z \in D$, then $1/g$ is holomorphic in D and

$$\left(\frac{1}{g}\right)' = -g'/g^2.$$

(iii) If

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n,$$

where $a_0, a_1, \ldots, a_n$ are complex constants, then p is holomorphic in C and

$$p'(z) = a_1 + \cdots + (n-1)a_{n-1}z^{n-2} + na_n z^{n-1}.$$

We now prove the identity referred to above.

LEMMA 3. *If $f(z)$ and $f'(z)$ are both holomorphic in D, then*

$$\Delta(|f(z)|^2) = 4|f'(z)|^2.$$

Proof. Let $f(z) = u(x, y) + iv(x, y)$. Then $|f(z)|^2 = (u(x, y))^2 + (v(x, y))^2$. It follows that if $w(x, y) = |f(z)|^2$, then $w_x = 2uu_x + 2vv_x$, $w_{xx} = 2uu_{xx} + 2(u_x)^2 + 2vv_{xx} + 2(v_x)^2$, and, similarly, $w_{yy} = 2uu_{yy} + 2(u_y)^2 + 2vv_{yy} + 2(v_y)^2$. Since $\Delta u = \Delta v = 0$ in D,

$$\begin{aligned}\Delta w &= 2(u_x)^2 + 2(v_x)^2 + 2(u_y)^2 + 2(v_y)^2,\\ &= 4\,|f'(z)|^2,\end{aligned}$$

where the last equality follows from the Cauchy-Riemann equations. This proves Lemma 3. ∎

Proof of the Fundamental Theorem of Algebra. Let $p(z)$ be a nonconstant complex polynomial. Then $p'(z)$ is a nonzero complex polynomial. Therefore, there exists a number d, $d > 0$, such that $p'(z) = 0$ implies $p'(z + d) \neq 0$.

We claim that there exists z_0 such that $p(z_0) = 0$. Assuming the contrary, the function $f(z) = 1/p(z)$ and $f(z + d)$ are holomorphic on $\mathbb{C}$. Therefore,

$$w(x, y) = |f(z)|^2 + |f(z + d)|^2$$

is continuous and has continuous second-order partial derivatives. It is easy to see that $w(x, y) \to 0$ as $x^2 + y^2 \to \infty$.

Let w attain its maximum at (x_0, y_0). From Lemma 1, $\Delta w(x_0, y_0) \leq 0$. From Lemma 3, $\Delta w(x, y) = 4(|f'(z)|^2 + |f'(z + d)^2|$. Therefore, letting $z_0 = x_0 + iy_0$, we have $f'(z_0) = 0$ and $f'(z_0 + d) = 0$. But $f'(z_0) = -p'(z_0)/p(z_0)^2$ and $f'(z_0 + d) = -p'(z_0 + d)/p(z_0 + d)^2$. Therefore $p'(z_0) = 0 = p'(z_0 + d)$ which is a contradiction. This proves the Fundamental Theorem of Algebra. ∎

We would like to point out that for functions f holomorphic in an open set, the identity

$$\Delta|f|^2 = 4|f'|^2 \qquad (*)$$

is well known. It is an exercise in Nehari [**4**, p. 64] and also in the classic reference, Titchmarsh [**6**, p. 7]. However, to derive it from the Cauchy-Riemann equations, one must first establish that the real and imaginary parts of f have second-order partial derivatives. For general holomorphic functions this can't be done without the Cauchy theory of integration or other advanced theory. For our application of $(*)$, with $f = 1/p$, this was circumvented by use of the elementary properties (i), (ii), and (iii).

Finally, we indicate another proof based on Lemma 3. A function w that is twice continuously differentiable in a connected, bounded, open set D, continuous on $\bar{D}$, and satisfies $\Delta w \geq 0$ is called *subharmonic* in D. The Maximum Principle [**5**] for subharmonic functions shows that $\max_{\bar{D}} w = \max_{\partial D} w$, where ∂D is the boundary of D. Lemma 3 shows that if $p(z)$ is a nonconstant polynomial without zeros, then $w(x, y)$ defined as

$$w(x, y) = \frac{1}{|p(z)|}, \quad \text{where} \quad z = x + iy,$$

satisfies $\Delta w \geq 0$ on R^2. Taking D to be a disc of radius r centered at 0 and using the fact that $w \to 0$ as $r \to \infty$, the Maximum Principle shows that $w \equiv 0$, contradicting $w(x, y) > 0$.

REFERENCES

1. Joseph Bak and Donald J. Newman, *Complex Analysis*, 2nd edition, Springer Verlag, New York, 1997.
2. William R. Derrick, *Introductory Complex Analysis, and Applications*, Academic Press, New York and London, 1972.
3. Benjamin Fine and Gerard Rosenberger, *The Fundamental Theorem of Algebra*, Springer Verlag, New York, 1997.
4. Z. Nehari, *Conformal Mapping*, McGraw-Hill, New York, 1952.
5. M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations*, Prentice Hall, Englewood Cliffs, N.J., 1967.
6. E.C. Titchmarsh, *The Theory of Functions*, 2nd ed., Oxford University Press, London, 1939.
7. Gabor Toth, *Glimpses of Algebra and Geometry*, Springer-Verlag, New York, 1997.
8. J.V. Uspensky, *Theory of Equations*, Mc-Graw Hill, New York, 1948.

Path Representation of a Free Throw Shooter's Progress

CHRISTOPHER L. BOUCHER
Salem State College
Salem, MA 01970
cboucher@salemstate.edu

A problem on the Putnam exam given December 7, 2002 involved a basketball player (Shanille O'Keal) taking a sequence of free throws. The player makes the first shot, misses the second, and makes each subsequent shot with probability equal to the fraction of successful shots prior to that point. Thus, Shanille makes her third shot with probability $1/2$. If she makes her third shot, she makes her fourth shot with probability $2/3$ and so forth. The exam asked for the probability that she made 50 of her first 100 shots. We are interested in the probability that Shanille ever finds herself having missed k more shots than she has made.

Shanille's "state" after n shots can be represented by a pair (x, y) where x is the number of successful shots to that point in the sequence and y is the number of unsuccessful shots to that point in the sequence (so $n = x + y$). If Shanille's current state is (x, y), her history can be represented by a lattice path from $(1, 1)$ to (x, y) involving only rightward and upward steps where a rightward step represents a successful shot and an upward step an unsuccessful shot. Each edge in the lattice is naturally associated with a conditional probability . The horizontal edge connecting (x, y) and $(x + 1, y)$ has probability equal to the probability that the Shanille, having made x shots and missed y shots, makes her next shot. According to our rule, this probability is $x/(x + y)$. Similarly, the vertical edge connecting (x, y) and $(x, y + 1)$ has probability $y/(x + y)$. The probability of her history following any particular lattice path is the product of the probabilities associated with the edges of the path.

THEOREM 1. *The two lattice paths connecting (x, y) and $(x + 1, y + 1)$ are equiprobable.*

Proof. If we write RU for the path from (x, y) to $(x + 1, y + 1)$ consisting of a rightward step followed by an upward step and UR for the path from (x, y) to $(x + 1, y + 1)$ consisting of an upward step followed by a rightward step, then

$$P(RU) = \frac{x}{x+y} \cdot \frac{y}{x+y+1}$$

$$= \frac{y}{x+y} \cdot \frac{x}{x+y+1}$$
$$= P(UR). \qquad \blacksquare$$

It follows from Theorem 1 that if one path can be obtained from another by transposing two steps, then the two paths are equiprobable. Since any permutation (i.e., a reordering of the steps) of a path can be accomplished by a sequence of transpositions, all permutations of a path are equiprobable, a fact we record in this corollary.

COROLLARY 1. *Let α be a lattice path from $(1, 1)$ to (x, y) consisting of $x - 1$ rightward and $y - 1$ upward steps and let $\sigma(\alpha)$ be a permutation of α. Then α and $\sigma(\alpha)$ are equiprobable.*

In particular, Corollary 1 implies that any two lattice paths with the same number of rightward and upward steps are equiprobable. Thus, to calculate the probability that Shanille arrives at any point (x, y) we need only calculate the probability of any one path consisting of $x - 1$ rightward and $y - 1$ upward steps from $(1, 1)$ and count how many such paths there are. The path consisting of $x - 1$ consecutive rightward steps from $(1, 1)$ to $(x, 1)$ has probability

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots \frac{x-1}{x} = \frac{1}{x}.$$

The path consisting of $y - 1$ consecutive upward steps from $(x, 1)$ to (x, y) has probability

$$\frac{1}{x+1} \cdot \frac{2}{x+2} \cdots \frac{y-1}{x+y-1} = \frac{x!\,(y-1)!}{(x+y-1)!}.$$

We record these calculations in the following theorem.

THEOREM 2. *The lattice path from $(1, 1)$ to (x, y) consisting of $x - 1$ rightward steps followed by $y - 1$ upward steps has probability*

$$p(x, y) \doteq \frac{(x-1)!(y-1)!}{(x+y-1)!}.$$

The number of lattice paths from $(1, 1)$ to (x, y) is

$$\binom{x+y-2}{x-1}$$

and so the probability that Shanille winds up at state (x, y) is, remarkably,

$$\binom{x+y-2}{x-1}\frac{(x-1)!\,(y-1)!}{(x+y-1)!} = \frac{1}{x+y-1}.$$

In particular, the $n - 1$ states in which Shanille could be after taking n shots are all equiprobable and the expected number of shots made in n attempts is $n/2$. (Thus the solution to the Putnam question is $1/(100 - 1) = 1/99$.)

The use of lattice paths to visualize the history of our basketball player caused me to read with interest [**1**], which calculates the probability that a one dimensional random walk returns to the origin given that the walker starts at $x = k$ by viewing the walker's progress as a lattice path. A similar analysis can be applied to our basketball player. For $k \in \mathbb{N}$, let P_k be the probability that our shooter ever finds herself having missed k more shots than she has made, that is, the probability that she ever finds herself at a state of the form $(n, n+k)$ for some $n \in \mathbb{N}$. This is the analog of the probability addressed in [**1**]. The number of $2n+k-2$-length lattice paths from $(1,1)$ to $(n, n+k)$ that intersect the line $y = x+k$ only at the point $(n, n+k)$ is the same as the number of $2n+k-2$-length lattice paths from $(0,0)$ to $(n-1, n-1+k)$ that intersect the line $y = x+k$ only at the point $(n-1, n-1+k)$. The article [**1**] denotes this number $C_k(n-1)$, where

$$C_1(n-1) = \frac{1}{n}\binom{2n-2}{n-1}$$

is the $(n-1)$st Catalan Number, $C_2(n-1) = C_1(n)$, and for $k \geq 3$, the numbers $C_k(n)$ satisfy the following recurrence relation (see Theorem 2 in [**1**]):

$$C_k(n) = C_{k-1}(n+1) - C_{k-2}(n+1). \tag{1}$$

The probability we seek is

$$P_k = \sum_{n=1}^{\infty} C_k(n-1)p(n, n+k). \tag{2}$$

The analogous sum in [**1**] is different because for the walker, all steps up are equiprobable and all steps right are equiprobable, while for the basketball player it is paths with the same initial and terminal point that are equiprobable. If $k = 1$, then using

$$C_1(n-1) = \frac{1}{n}\binom{2n-2}{n-1},$$

we have

$$\begin{aligned} P_1 &= \sum_{n=1}^{\infty} C_1(n-1)p(n, n+1) = \sum_{n=1}^{\infty} \frac{1}{(2n)(2n-1)} \\ &= \sum_{n=1}^{\infty}\left(\frac{1}{2n} - \frac{1}{2n-1}\right) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log 2. \end{aligned} \tag{3}$$

So the probability that, at some point, Shanille has missed one more shot that she has made is, remarkably, $\log 2$. (Conversely, the probability that she has always made at least as many as she has missed is $1 - \log 2$.) For $k = 2$, we have

$$P_2 = \sum_{n=1}^{\infty} C_2(n-1)p(n, n+2) = \sum_{n=1}^{\infty} \frac{1}{n(2n+1)} = 2 - 2\log 2. \tag{4}$$

An evaluation of the sum in (2) for general k, however, requires a closed form expression for $C_k(n)$, which is given in the following theorem.

THEOREM 3. *If* $n, k \in \mathbb{N}$ *then*

$$C_k(n) = \frac{k}{2n+k}\binom{2n+k}{n}. \tag{5}$$

Proof. For $n = 1, 2 \ldots$ and $k = 0, 1, \ldots$, let

$$B_k(n) = \frac{k}{2n+k}\binom{2n+k}{n},$$

and let $C_k(n)$ be the numbers that satisfy (1) with $C_0(n) = 0$ and

$$C_1(n) = \frac{1}{n+1}\binom{2n}{n}.$$

(Note that this implies that $C_2(n) = C_1(n+1)$). For $n \in \mathbb{N}$, $B_0(n) = C_0(n)$ and

$$B_1(n) = \frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{2n+1}\cdot\frac{(2n+1)!}{n!\,(n+1)!} = \frac{1}{n+1}\binom{2n}{n} = C_1(n).$$

Using similarly straightforward algebra, one can verify that for all $n = 1, 2, \ldots$ and $k = 0, 1, \ldots$ the numbers $B_k(n)$ satisfy:

$$B_k(n) = B_{k-1}(n+1) - B_{k-2}(n+1). \tag{6}$$

This is the same recurrence relation satisfied by the numbers $C_k(n)$. It follows that the numbers $B_k(n)$ are identical to the numbers $C_k(n)$, which is what the theorem asserts. ■

Having a closed form expression for the path counts doesn't appear to be much use in computing the probability that the walker ever finds himself k steps to one side of where he started, however it does allow us to solve the analogous problem for Shanille, i.e., to compute (2) for general k.

$$\begin{aligned}
P_k &= \sum_{n=1}^{\infty} C_k(n-1)p(n, n+k) \\
&= \sum_{n=1}^{\infty} \frac{k}{2(n-1)+k}\binom{2(n-1)+k}{n-1}\frac{(n-1)!\,(n+k-1)!}{(2n+k-1)!} \\
&= \sum_{n=1}^{\infty} \frac{k}{(2n+k-2)(2n+k-1)} \\
&= \sum_{n=1}^{\infty} k\left(\frac{1}{2n+k-2} - \frac{1}{2n+k-1}\right) \\
&= k\left[\frac{1}{k} - \frac{1}{k+1} + \frac{1}{k+2} - \frac{1}{k+3} + \cdots\right] \\
&= k\cdot(-1)^{k+1}\left(\log 2 - \sum_{j=1}^{k-1}\frac{(-1)^{j+1}}{j}\right).
\end{aligned} \tag{7}$$

The last line in this display uses the series

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

Note that the sums in (3) and (4) are computed in the same manner as we compute P_k here. The first ten values of P_k are displayed below in Table 1.

The numerical evidence suggests that these probabilities decline from $\log 2$ to $1/2$. This is in fact the case. Since the event that Shanille ever finds herself having missed k more shots than she has made is a superset of the event that she ever finds herself having missed $k+1$ more shots than she has made, the probabilities certainly decline.

TABLE 1: Some values of P_k

k	P_k
1	.693147
2	.613706
3	.579442
4	.560745
5	.549069
6	.541117
7	.535364
8	.531013
9	.527610
10	.524877

Thus, $\lim_{k\to\infty} P_k$ surely exists. Determining the value of this limit is a nice exercise in several techniques from first-year calculus. First, we note that if

$$f_k(x) = \frac{k}{(2x+k-2)(2x+k-1)},$$

then

$$f_k'(x) = -\frac{2k(4x+2k-3)}{(2x+k-2)^2(2x+k-1)^2}$$

and for any $k = 1, 2, \ldots$, f_k is a decreasing function on $[1, \infty)$. Thus, we can use integrals to bound the series (7) above and below:

$$\int_1^\infty f_k(x)\,dx \le \sum_{n=1}^\infty f_k(n) \le f_k(1) + \int_1^\infty f_k(x)\,dx. \tag{8}$$

Next, we calculate the improper integrals above:

$$\int_1^\infty f_k(x)\,dx = \frac{k}{2}\log\left(1+\frac{1}{k}\right)$$

So, we have for any $k = 1, 2, \ldots$,

$$\frac{k}{2}\log\left(1+\frac{1}{k}\right) \le \sum_{n=1}^\infty f_k(n) \le \frac{k}{k(k+1)} + \frac{k}{2}\log\left(1+\frac{1}{k}\right).$$

Finally, an application of L'Hôpital's Rule shows that

$$\lim_{k\to\infty} \frac{k}{2}\log\left(1+\frac{1}{k}\right) = \frac{1}{2}.$$

This proves our claim that the probability that Shanille ever finds herself having missed k more shots than she has made approaches $1/2$ as $k \to \infty$.

Acknowledgment. The author is grateful to the referees whose comments materially improved the paper.

REFERENCE

1. Bolina, O. "Path Representation of One-Dimensional Random Walks," this MAGAZINE **77** 218–225, 2004.

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; BYRON WALDEN, Santa Clara University; PAUL ZEITZ, The University of San Francisco

Proposals

To be considered for publication, solutions should be received by November 1, 2006.

1746. *Proposed by Stephen J. Herschkorn, Highland Park, NJ.*

Alice and Bob play a game in which they alternately flip a (biased) coin that has probability p of coming up heads when tossed. Alice goes first. With one possible exception, each player flips the coin once per turn. The first player to have cumulatively flipped k heads is the winner. To compensate for Alice's advantage in going first, Bob gets a second flip on his first turn if his first flip turns up tails; this is the exception. (Note that if $k = 1$, Bob may not get to flip at all.)

For each of the cases $k = 1$ and $k = 2$, determine the value of p for which the game fair, and calculate the expected value and variance of the number of flips in the game when p takes on this value.

1747. *Proposed by Stephen J. Herschkorn, Highland Park, NJ.*

Does there exist a Hausdorf space with a countably infinite topology?

1748. *Proposed by Anonymous*

Let m and n be positive integers such that mn is a triangular number. Prove that there exists an integer k such that the sequence $\{R_j\}$ generated by

$$R_0 = m, \qquad R_1 = n, \qquad R_j = 6R_{j-1} - R_{j-2} + k, \quad j \geq 2,$$

has the property that $R_j R_{j+1}$ is a triangular number for all integers $j \geq 0$.

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 reader's name, full address, and an e-mail address and/or FAX number on every page.

1749. *Proposed by Mihàly Bencze, Nègygalu, Romania.*

Let r and R be, respectively, the inradius and circumradius of a triangle with sides of length a, b, and c and let n be a positive integer. Prove that

$$\frac{R(R^n - r^n)}{(R-r)r^n} \geq \frac{(a^n - b^n)(a^{n+1} + b^{n+1})}{(a-b)a^n b^n} + 2^{n+1} - 2(n+1).$$

1750. *Proposed by Christopher J. Hillar, Texas A&M University, College Station, TX.*

Let $p > 3$ be prime. Define a sequence $x_1,\ x_2,\ x_3$ of integers to be a 3-progression if they are in arithmetical progression modulo p. If $A_1,\ A_2, \dots, A_n \subseteq \mathbb{Z}/p\mathbb{Z}$ is a collection of sets such that each three progression is contained in at least one of the A_ks, then the collection $\{A_1,\ A_2, \dots, A_n\}$ is called a 3-covering of $\mathbb{Z}/p\mathbb{Z}$. Find the minimum over all 3-coverings of the quantity

$$\sum_{i=1}^{n} |A_i|^2.$$

Quickies

Answers to the Quickies are on page 226.

Q961. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Suppose that $(X,\ d)$ is a compact metric space and that $f : X \to X$ is a function. Suppose that ϕ is a lower semicontinuous real-valued function defined on X and that

$$d\,(x,\ f(x)) \leq \phi(x) - \phi\,(f(x))\,,$$

for all $x \in X$. Prove that f has a fixed point. (To say that ϕ is lower semicontinuous at x_0 means that for each $\epsilon > 0$, there exists a $\delta > 0$ such that $\phi(x) > \phi(x_0) - \epsilon$ when $d(x,\ x_0) < \delta$.)

Q962. *Proposed by Murray Klamkin (deceased), University of Alberta, Edmonton, Alberta, Canada.*

Let a, b, and c be the lengths of the sides of a triangle. Prove that

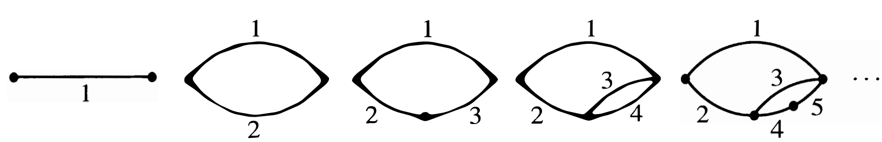
$$2(b^3c + bc^3 + c^3a + ca^3 + a^3b + ab^3) \geq (a^2 + b^2 + c^2)^2 + abc(a+b+c).$$

Solutions

Fibonacci Graphs **June 2005**

1721. *Proposed by Donald Knuth, Stanford University, Stanford, CA.*

The "Fibonacci graphs"



are defined by successively replacing the edge with maximum label n by two edges n and $n+1$, in series if n is even, and in parallel of n is odd. Prove that the Fibonacci graph with n edges has exactly F_{n+1} spanning trees, where $F_1 = F_2 = 1$ and $F_{n+1} = F_n + F_{n-1}$. Show also that these spanning trees can be listed in such a way that some edge k is replaced by $k \pm 1$ as we pass from one tree to the next. For example, for $n = 5$ the eight spanning trees can be listed as 125, 124, 134, 135, 145, 245, 235, 234.

Solution by Michel Bataille, Rouen, France.

Denote by $\mathcal{G}_n$ the Fibonacci graph with n edges, and by t_n the number of spanning trees of $\mathcal{G}_n$. We use induction to prove that $t_n = F_{n+1}$. It is readily checked that $t_1 = 1$ and $t_2 = 2$. Let $n \geq 3$.

If n is odd, then $\mathcal{G}_n$ has one more vertex than $\mathcal{G}_{n-1}$ and this vertex is an endpoint of edges $n-1$ and n only. First count the spanning trees of $\mathcal{G}_n$ that contain the edge n. These are obtained by adding the edge n to a spanning tree of $\mathcal{G}_{n-1}$ that does not contain edge $n-1$ or, in a spanning tree of $\mathcal{G}_{n-1}$ that contains edge $n-1$, by dividing the existing edge $n-1$ into two edges $n-1$ and n. Thus there are t_{n-1} spanning trees of $\mathcal{G}_n$ that contain edge n. The spanning trees of $\mathcal{G}_n$ that do not contain edge n are obtained by adding edge $n-1$ to a spanning tree of $\mathcal{G}_{n-2}$.

If n is even, then $\mathcal{G}_n$ has the same number of vertices as $\mathcal{G}_{n-1}$. It follows from the construction of $\mathcal{G}_n$ that every one of the t_{n-1} spanning trees of $\mathcal{G}_{n-1}$ is a spanning tree of $\mathcal{G}_n$. If a spanning tree of $\mathcal{G}_n$ contains the edge n, then it cannot contain edge $n-1$. It follows that such spanning trees are obtained by adding edge n to any of the t_{n-2} spanning tree of $\mathcal{G}_{n-1}$.

In both cases we find that $t_n = t_{n-1} + t_{n-2}$, and the result $t_n = F_{n+1}$ follows.

We now construct inductively a suitable list L_n of the spanning trees of $\mathcal{G}_n$. We take $L_1 = 1$ and $L_2 = 2; 1$. Now let $n \geq 3$. If n is odd, we take the terms of L_{n-1} in the order given, and append n to the right end of each element of the list. We then continue the list with the terms of L_{n-2} in opposite order, with $n-1$ appended to the right end of each element. If n is even, first take the terms of L_{n-2} in reverse order and append n to the right end of each element of the list, then continue the list with the terms of L_{n-1} in order. This process results in the lists

$$L_3 = 23\,;\,13\,;\,12$$

$$L_4 = 14\,;\,24\,;\,23\,;\,13\,;\,12$$

$$L_5 = 145\,;\,245\,;\,235\,;\,135\,;\,125\,;\,124\,;\,134\,;\,234$$

It follows from our argument that $t_n = F_{n+1}$ that the list L_n contains all spanning trees for $\mathcal{G}_n$. The list also satisfies the desired criteria because each list L_{n-1} and L_{n-2} are assumed to do so, and because in each case the last spanning tree in the first part of the list and the first spanning tree of the second part differ only in the last edge.

Also solved by Con Amore Problem Group (Denmark), Dave Cromley, A. K. Desai (India), G.R.A.20 Math Problems Group (Italy), Enkel Hyselaj (Australia), Houghten College Problem Solving Group, Richard F. McCoart, Christopher N. Swanson, Paul K. Stockmeyer, Paul Weisenhorn (Germany), Xinyi Zhang, and the proposer.

Permutations with a Subsequence **June 2005**

1722. *Proposed by Emeric Deutsch, Polytechnic University, Brooklyn, NY.*

Let k and n be positive integers with $k \leq n$. Find the number of permutations of $\{1, 2, \ldots, n\}$ in which $1, 2, \ldots, k$ appears as a subsequence but $1, 2, \ldots, k, k+1$ does not.

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

Let $s(n, k)$ denote the number of permutations of $\{1, 2, \ldots, n\}$ in which $1, 2, \ldots, k$ occurs as a subsequence. Let $a(n, k)$ denote the number of permutations in which $1, 2, \ldots, k$ occurs as a subsequence but $1, 2, \ldots, k, k+1$ does not. Then

$$a(n, k) = s(n, k) - s(n, k+1).$$

To compute $s(n, k)$, note that there are $\binom{n}{k}$ ways to select the positions for $1, 2. \ldots, k$ to appear in their natural order, then $(n-k)!$ ways to order the elements $k+1, k+2, \ldots, n$ in the remaining positions. Thus

$$s(n, k) = \binom{n}{k}(n-k)!.$$

It follows that

$$a(n, k) = \binom{n}{k}(n-k)! - \binom{n}{k+1}(n-k-1)! = \frac{k \cdot n!}{(k+1)!}.$$

Also solved by Anurag Agarwal, Michel Bataille (France), Tom Beatty, J. C. Binz (Switzerland), Marc Brodie, Robert Calcaterra, John Christopher, Con Amore Problem Group (Denmark), Toni Davies and Lauren McMullen and Michelle Pullman and Anna Wilkins, A. K. Desai (India), Fejéntaláltuka Szeged Problem Group (Hungary), Dmitry Fleishman, Marty Getz and Dixon Jones, G.R.A.20 Math Problems Group (Italy), Ralph P. Grimaldi, Arup Guha, Peter Hohler (Switzerland), Houghten College Math Club, Jerry G. Ianni, Kathleen E. Lewis, Peter W. Lindstrom, Marvin Littman, Auturo Magidin, William Moser (Canada), José H. Nieto (Venezuela), Northwestern University Math Problem Solving Group, Angel Plaza de la Hoz (Spain), Rob Pratt, Adriana Rivera and Cecilia Greene and Farley Mawyer, Arthur J. Rosenthal, Nicholas C. Singer, Paul K. Stockmeyer, Christopher N. Swanson, Li Zhou, and the proposer. There was one solution with no name and eight incorrect submissions.

An Area Formula — June 2005

1723. *Proposed by Herb Bailey, Rose Hulman Institute of Technology, Terre Haute, IN.*

Let I be the incenter of triangle ABC with BC tangent to the incircle at D. Let E be the intersection of the extension of $\overline{ID}$ with the circle through B, I, and C. Prove that

$$DE = \frac{T}{s-a},$$

where T and s are, respectively, the area and semiperimeter of triangle ABC, and $a = BC$.

Many readers submitted a solution along the following lines.

Let Γ be the circle through B, I, and C. Note that $\overline{IE}$ and $\overline{BC}$ are chords of Γ meeting at D, with $DB = s-b$ and $DC = s-c$. Then $DE \cdot DI = DB \cdot DC = (s-b)(s-c)$. Solving for DE we find

$$DE = \frac{(s-b)(s-c)}{DI} = \frac{s(s-a)(s-b)(s-c)}{rs(s-a)} = \frac{T^2}{T(s-a)} = \frac{T}{s-a}.$$

Solved by Michel Bataille (France), J. C. Binz (Switzerland), Bruce S. Burdick, Robert Calcaterra, Minh Can, Adam Coffman, Miguel Amengual Covas (Spain), Prithwijit De (Ireland), Emeric Deutsch, Habib Y. Far, Fejéntaláltuka Szeged Problem Group (Hungary), John Ferdinands, Dmitry Fleischman, Ovidiu Furdui, Marty Getz and Dixon Jones, Michael Goldenberg and Mark Kaplan, Peter Gressis, John G. Heuver (Canada), Peter Hohler (Switzerland), Houghton College Problem Solving Group, Enkel Hysnelaj (Australia), Victor Y. Kutsenok, Elias Lampakis (Greece), Kim McInturff, Juniad N. Mansuri, Dao T. Nguyen, José H. Nieto (Venezuela), Northwestern University Math Problem Solving Group, Peter E. Nüesch (Switzerland), Thomas Peter and Yuguang

Bai, Richard E. Pfiefer, Raúl A. Simón (Chile), Albert Stadler (Switzerland), Man Kam Tam, R. S. Tiberio, Paul Weisenhorn (Germany), Yan-loi Wong (Singapore), Charles Worrall, John Zacharias, Tom Zerger, Li Zhou, and the proposer.

Sharpening the AM-GM Inequality **June 2005**

1724. *Proposed by Mihály Bencze, Săcele-Négyfalu, Romania.*

Let $x_1, x_2, \ldots, x_n$ be positive real numbers. Prove that

$$\frac{1}{n}\sum_{k-1}^{n} x_k - \left(\prod_{k=1}^{n} x_k\right)^{1/n} \leq \frac{1}{n}\sum_{1\leq j<k\leq n} (\sqrt{x_j}-\sqrt{x_k})^2. \tag{1}$$

I. *Solution by Li Zhou, Polk Community College, Winter Haven, FL.*

In [1] it is shown that if f is convex on an interval I and $a_1, a_2, \ldots, a_n \in I$, then

$$(n-2)\sum_{k=1}^{n} f(a_k) + nf\left(\frac{1}{n}\sum_{k=1}^{n} a_k\right) \geq 2\sum_{1\leq j<k\leq n} f\left(\frac{a_j+a_k}{2}\right).$$

Applying this with $f(x) = e^x$ on $(-\infty, \infty)$ with $a_k = \ln x_k$ for all k we obtain

$$(n-2)\sum_{k=1}^{n} x_k + n\left(\prod_{k=1}^{n} x_k\right)^{1/n} \geq 2\sum_{1\leq j<k\leq n} \sqrt{x_j x_k},$$

which is equivalent to the desired inequality.

1. Vasile Cîrtoaje, Two generalizations of Popoviciu's inequality, *Crux Math.*, **31.5** (2005), p 313–318.

II. *Solution by Peter W. Lindstrom, Saint Anselm College, Manchester, NH.*

Because (1) is homogeneous in $x_1, x_2, \ldots, x_n$, we may assume that $\prod_{k=1}^{n} x_k = 1$. Then, with some algebraic manipulation, it can be shown that (1) is equivalent to

$$(n-2)S_n \geq 2Q_n - n, \tag{2}$$

where

$$S_n = \sum_{k=1}^{n} x_k \quad \text{and} \quad Q_n = \sum_{1\leq j<k\leq n} \sqrt{x_j x_k}.$$

We prove (2) for $n \geq 2$ by induction. It is easily checked that (2) holds with equality when $n = 2$. Now let $n \geq 2$. Assume that (2) holds for any n positive real numbers whose product is 1, and assume that $x_1, x_2, \ldots, x_n, x_{n+1}$ are positive real numbers with $\prod_{k=1}^{n+1} x_k = 1$. Because both sides of (2) are symmetric in the x_ks we may assume that $x_n \leq 1$ and $x_{n+1} \geq 1$ are, respectively, the minimum and maximum of the numbers $x_1, x_2, \ldots, x_n, x_{n+1}$. For these numbers, with n replaced by $n+1$, (2) becomes

$$(n-1)S_n + (n-1)x_{n+1} \geq 2Q_n + 2T_n\sqrt{x_{n+1}} - (n+1), \tag{3}$$

where $T_n = \sum_{k=1}^{n} \sqrt{x_k}$. We establish this inequality be adding two other inequalities. The first inequality arises from (2) using the n positive numbers $x_1, x_2, \ldots, x_n x_{n+1}$:

$$(n-2)S_{n-1} + (n-2)x_n x_{n+1} \geq 2Q_{n-1} + 2T_{n-1}\sqrt{x_n}\sqrt{x_{n+1}} - n. \tag{4}$$

The second inequality we shall need to show is

$$(n-1)x_n + S_{n-1} + x_{n+1}\left((n-1)-(n-2)x_n\right) \\ \geq 2T_{n-1}\sqrt{x_n} + 2\sqrt{x_{n+1}}\left(T_{n-1}(1-\sqrt{x_n}) + \sqrt{x_n}\right) - 1. \tag{5}$$

Note that adding (4) and (5) gives (3). To prove (5), let $x = \sqrt{x_n} \leq 1$ and $y = \sqrt{x_{n+1}} \geq 1$. Then (5) can be rearranged to give

$$(n-2)\left(1+(1-x^2)(y^2-1)\right) + (y-x)^2 + 1 \geq -(S_{n-1} - \beta T_{n-1}), \tag{6}$$

where $\beta = 2\left(1+(1-x)(y-1)\right)$. On completing the square on the right, (6) becomes

$$(n-2)\left(1+(1-x^2)(y^2-1)\right) + (y-x)^2 + 1 \geq -\sum_{k=1}^{n-1}\left(\sqrt{x_k} - \frac{\beta}{2}\right)^2 + (n-1)\frac{\beta^2}{4}$$

Thus we can establish (5) by showing that

$$(n-2)\left(1+(1-x^2)(y^2-1)\right) + (y-x)^2 + 1 \geq (n-1)\frac{\beta^2}{4},$$

or equivalently,

$$(n-2)\left(1+(1-x^2)(y^2-1)\right) + (y-x)^2 + 1 - (n-1)\left(1+(1-x)(y-1)\right)^2 \\ = 1 + \left((y-1)^2 - 1\right)\left(1-(1-x)^2\right) + 2(n-2)xy(1-x)(y-1) \geq 0. \tag{7}$$

The last inequality holds because $\left((y-1)^2-1\right)\left(1-(1-x)^2\right) \geq -1$ and $xy(1-x)(y-1) \geq 0$. This completes the proof.

Also solved by Michel Bataille (France), Robert Calcaterra, Ovudui Furdui, Eugene A. Herman, Chi-Kwong Li, Albert Stadler (Switzerland), Henry Ricardo, Xiaoshen Wang, and the proposer.

Parallelograms about an Ellipse **June 2005**

1725. *Proposed by Michel Bataille, Rouen, France.*

Let $\mathcal{E}$ be the ellipse with equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

where a and b are positive integers. Find the number of parallelograms with vertices at integer lattice points and sides tangent to $\mathcal{E}$ at their midpoints.

Solution by Eugene A. Herman, Grinnell College, Grinnell, IA.

We refer to a parallelogram with vertices at integer lattice points and sides tangent to $\mathcal{E}$ at their midpoints as a *fitting* parallelogram, and we denote the number of fitting parallelograms by $N(a, b)$. We will show that $N(a, b)$ can be computed as follows. Let $g = \gcd(a, b)$, and let $g = p_1^{e_1} \cdots p_k^{e_k} h$ be the factorization of g into primes, where $p_1, \ldots, p_k$ are the distinct prime factors of g that are congruent to 1 modulo 4. Then

$$N(a, b) = \prod_{j=1}^{k}(1 + 2e_j)$$

LEMMA 1. *Suppose we have any circumscribed parallelogram. If the points of tangency of two adjacent sides are denoted* (x_1, y_1) *and* (x_2, y_2)*, then the four points of tangency are the midpoints of the four sides if and only if*

$$\frac{x_1x_2}{a^2} + \frac{y_1y_2}{b^2} = 0$$

Proof. Note that the other two points of tangency are the points $(-x_1, -y_1)$ and $(-x_2, -y_2)$. If the four points of tangency are the midpoints of the four sides, then the side tangent at (x_1, y_1) is parallel to the line joining the midpoints of the two adjacent sides. That line has direction vector (x_2, y_2), and the side tangent at (x_1, y_1) has normal vector $(x_1/a^2, y_1/b^2)$. These vectors are orthogonal, which means that $x_1x_2/a^2 + y_1y_2/b^2 = 0$.

Conversely, if $x_1x_2/a^2 + y_1y_2/b^2 = 0$, then the line joining (x_2, y_2) and $(-x_2, -y_2)$ is parallel to the side tangent at (x_1, y_1). Likewise, the line joining (x_1, y_1) and $(-x_1, -y_1)$ is parallel to the side tangent at (x_2, y_2). Since $(0, 0)$ is the midpoint of the lines joining opposite points of tangency, the points of tangency are the midpoints of their sides. ■

Note that the parallelogram with $(x_1, y_1) = (a, 0)$ and $(x_2, y_2) = (0, b)$ is fitting. By Lemma 1, there are no other fitting parallelograms for which some point of tangency lies on a coordinate axis. We refer to this parallelogram as *special*, and we must now count the number of fitting parallelograms that are not special.

LEMMA 2. *If* $g = gcd(a, b)$*, the number of nonspecial fitting parallelograms,* $N(a, b) - 1$*, equals the number of primitive Pythagorean triples* (r, s, t) *such that* t *divides* g.

Proof. Suppose we are given a fitting nonspecial parallelogram. Since the vertices of the parallelogram are integers, we can write the midpoints of each side in the form $(m/2, n/2)$, where m and n are nonzero integers. Since $(m/2, n/2)$ is a point on the ellipse, we have

$$m^2b^2 + n^2a^2 = 4a^2b^2$$

Therefore, there is a primitive Pythagorean triple (r, s, t) and a nonzero integer k such that

$$mb = kr, \qquad na = ks, \qquad 2ab = kt \tag{1}$$

Hence, we can write $2mnab$ two ways, which yields $2krs = mnt$. Since $\gcd(r, s, t) = 1$, there exists a nonzero integer u such that

$$mn = urs \text{ and } 2k = ut$$

Since t is odd, u must be even. Writing $u = 2v$, we have

$$mn = 2vrs \text{ and } k = vt$$

Next, we combine these two equations with the second of the equations (1). Writing mna in two ways, we obtain $2ra = tm$. Hence m is even and there exists an integer α such that

$$\frac{m}{2} = \alpha r \text{ and } a = \alpha t$$

Similarly, n is even and there exists an integer β such that

$$\frac{n}{2} = \beta s \text{ and } b = \beta t$$

Therefore $t|g$, and one point on the ellipse is $(\alpha r, \beta s)$.

Conversely, suppose there is a primitive Pythagorean triple (r, s, t) such that $t|g$. Hence there exist integers α and β such that $a = \alpha t$ and $b = \beta t$. Note that all the points $(\pm\alpha r, \pm\beta s)$ and $(\pm\alpha s, \pm\beta r)$ lie on the ellipse. In particular, this yields two nonspecial fitting parallelograms. In the notation of Lemma 1, these have the following points of tangency on adjacent sides:

$$(x_1, y_1) = (\alpha r, \beta s), \qquad (x_2, y_2) = (\alpha s, -\beta r)$$

$$(x_1, y_1) = (\alpha s, \beta r), \qquad (x_2, y_2) = (\alpha r, -\beta s)$$

Note that each of these pairs of points satisfies Lemma 1 and that the permuted primitive Pythagorean triple (s, r, t) yields the same two pairs of points. Furthermore, the following computation shows that the two points of tangency can only arise from the same primitive Pythagorean triple and its permuted triple. Suppose (r', s', t') is another primitive Pythagorean triple and that α' and β' are integers such that $a = \alpha' t'$ and $b = \beta' t'$. Let $(x_1, y_1) = (\alpha r, \beta s)$ and $(x_2, y_2) = (\alpha' r', \beta' s')$. Then

$$\frac{x_1 x_2}{a^2} + \frac{y_1 y_2}{b^2} = \frac{\alpha r \alpha' r'}{\alpha t \alpha' t'} + \frac{\beta s \beta' s'}{\beta t \beta' t'} = \frac{1}{tt'}\left(rr' + ss'\right)$$

Thus $x_1 x_2/a^2 + y_1 y_2/b^2 = 0$ if and only if $rr' = -ss'$. This equation is equivalent to the pair of equations $r' = \pm s$ and $s' = \mp r$, since $\gcd(r, s) = 1$ and $\gcd(r', s') = 1$; hence, by Lemma 1, our assertion is proved. We have therefore established a one-one correspondence between the nonspecial fitting parallelograms and the primitive Pythagorean triples (r, s, t) such that $t|g$.

To complete the solution, we define the following two functions on the set of positive integers:

$$f(n) = \text{number of primitive Pythagorean triples } (r, s, t) \text{ such that } t|n$$

$$R(n) = \text{number of non-negative primitive solutions of } u^2 + v^2 = n$$

Note that $f(n) = 0$ for $n < 5$, since 5 is the smallest possible value of t. Also, (r, s, t) is a primitive Pythagorean triple if and only if there exist positive integers u and v of opposite parity such that $\gcd(u, v) = 1$ and $t = u^2 + v^2$. Hence

$$f(n) = \sum_{t|n, t>1, t \text{ odd}} R(t)$$

It is well known that

$$R(n) = \begin{cases} 0 & \text{if } 4|n \text{ or } p|n \text{ for some prime } p \text{ such that } p \equiv 3 \pmod 4 \\ 2^j & \text{otherwise, where } j = \text{ number of distinct odd primes } p \text{ such that } p|n \end{cases} \tag{2}$$

(See, for example, E. Landau, *Elementary Number Theory*, Chelsea, 1958, p. 136.) Therefore, if $p_1^{e_1} \cdots p_k^{e_k} h$ is the factorization of n into primes, where $p_1, \ldots, p_k$ are the distinct prime factors of n that are congruent to 1 modulo 4, then

$$f(n) = \sum_{t|p_1^{e_1} \cdots p_k^{e_k}, t>1} R(t)$$

In this sum, consider all the terms for which the divisor t is a product of exactly j distinct primes. According to formula (2), the sum of these terms is 2^j times the number of such terms. In fact, the number of such terms is the elementary symmetric function of degree j in $e_1, \ldots, e_k$, since $e_{i_1} \cdots e_{i_j}$ is the number of factors of the form $p_{i_1}^{k_1} \cdots p_{i_j}^{k_j}$ such that $1 \leq k_1 \leq e_{i_1}, \ldots, 1 \leq k_j \leq e_{i_j}$. Therefore

$$f(n) = \sum_{j=1}^{k} 2^j \sum_{i_1 < \cdots < i_j} e_{i_1} \cdots e_{i_j} = \sum_{j=1}^{k} \sum_{i_1 < \cdots < i_j} 2e_{i_1} \cdots 2e_{i_j} = \prod_{j=1}^{k} (1 + 2e_j) - 1$$

Combining this formula with Lemma 2 shows that $N(a, b) = \prod_{j=1}^{k} (1 + 2e_j)$. ■

Note: Thanks to Arnold Adelberg for assistance with the number theory aspects of this solution.

Also solved by the proposer. There were several incorrect submissions.

Answers

Solutions to the Quickies from page 219.

A961. It is well known that if (X, d) is a compact metric space and ϕ is lower semi-continuous on X, then ϕ takes on its minimum value on X. Thus, there exists an $m \in X$ such that $\phi(m) \leq \phi(x)$ for all $x \in X$. By the condition given in the problem statement we have

$$0 \leq d\,(m,\ f(m)) \leq \phi(m) - \phi\,(f(m)) \leq 0.$$

Thus $d\,(m,\ f(m)) = 0$, and it follows that $f(m) = m$.

A962. After some rearrangement, we see that the inequality is equivalent to

$$2 \sum b^2c^2 - 2abc \sum a \geq 2 \sum a^4 - 4 \sum (b^3c + bc^3) + 6 \sum b^2c^2,$$

where the sums are symmetric over a, b, and c. This inequality is equivalent to

$$\sum a^2(b-c)^2 \geq \sum (b-c)^4 \qquad \text{or} \qquad \sum \left(a^2 - (b-c)^2\right)(b-c)^2 \geq 0.$$

It is easy to check that this last inequality is true, with equality if and only if the triangle is equilateral or degenerate with one side of length 0.

Fun, Fun, Functions
(with apologies to the Beach Boys)

Well, she got the calculator and she headed straight away for the beach now
Seems she forgot all about the graph paper like she told her ol' Teach now
She has the buttons all flying, always trying just to keep it in reach now
And she'll graph fun, fun, functions 'til her teacher takes the TI away
(fun, fun, functions 'til her teacher takes the TI away)

Well, the class can't stand her 'cause she DRAWs, MODEs, and GRAPHs like an ace now
(you DRAW like an ace now, you DRAW like an ace)
She switches Function to Polar to Parametric at a really fast pace now
(you MODE like an ace now, you MODE like an ace)
A lot of guys try to catch her but she knows how to ZOOM and to TRACE now
(you GRAPH like an ace now, you GRAPH like an ace)
And she'll graph fun, fun, functions 'til her teacher takes the TI away
(fun, fun, functions 'til her teacher takes the TI away)

Well, you knew all along that the Teach was gettin' wise to you now
(you shouldn't have π'ed now, you shouldn't have π'ed)
And since she took your batteries you've been thinking that your functions are through now
(you shouldn't have π'ed now, you shouldn't have π'ed)
But you can come along with me 'cause we've got a lot of graphing to do now
(you shouldn't have π'ed now, you shouldn't have π'ed)
And we'll graph fun, fun, functions now that Teacher took the TI away
(fun, fun, functions now that Teacher took the TI away)
And we'll graph fun, fun, functions now that Teacher took the TI away
(fun, fun, functions now that Teacher took the TI away)

Brian D. Beasley
Presbyterian College
Clinton, SC 29325
bbeasley@mail.presby.edu

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles, books, and other materials 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.

Devlin, Keith, Math back in forefront, but debate lingers on how to teach it, *San Jose Mercury News* (19 February 2006) 4P; How do we learn math?, `http://www.maa.org/devlin/devlin_03_06.html` .

According to Devlin, math is hard, but humans can do it because of our abilities for language, ascribing meaning, and learning new skills. Learning mathematics is like learning chess, skiing, driving, or playing an instrument, or even using a VCR or a computer. His claim is that for those as well as for mathematics, mechanical rule-following and "procedural practice" *must* precede conceptual understanding. Consequently, high school mathematics education—and college calculus—should concentrate on fostering the ability "to learn and apply rule-based symbolic processes *without understanding them*" [his emphasis]. Explanations should be provided "as a matter of intellectual courtesy." Devlin is correct in operational terms, particularly for concrete learners: Whatever we may try to teach, much of what students learn in mathematics and other subject areas is merely to run rudimentary algorithms. Training students to such "initial mastery of use" may well equip them for [warning: cliché follows] daily "life in the highly technological world they will live in"—but it is indeed training, not education, and certainly not intellectual. Devlin's argument also disregards the role of motivation in [warning: another cliché ahead] a pragmatic world driven by immediate feedback and satisfaction. To use Devlin's examples, people learn chess, skiing, driving, or an instrument, or to use a VCR or a computer, for the pleasure, utility, or convenience that the skill provides. How many high school students want to learn math for the enjoyment of it, or have a problem that they look forward to using math to solve, or find that it affords them great convenience and pleasure in their life? But perhaps the discouraging answers to those questions support Devlin's argument in a backhanded way: We can make no apologia on behalf of mathematics that will thrill most students, so we can foster only preconceptual learning of ununderstood technique. Coda: If "understanding can come only later, as an emergent consequence of use," then we should endeavor to make the "later" sooner, and teach "rule-based symbolic processes" while students are young enough to enjoy learning them.

Ball, Deborah Loewenberg, Imani Masters Goffney, and Hyman Bass, The role of mathematics instruction in building a socially just and diverse democracy, *The Mathematics Educator* (Athens, Georgia) 15 (1) (2005) 2–6, `http://math.coe.uga.edu/TME/Issues/v15n1/V15N1_Ball.pdf` .

The authors, two elementary school teachers and a distinguished professor of mathematics, see mathematics not as "culturally neutral, politically irrelevant, and mainly a matter of innate ability," but as "a critical lever for social and educational progress." They urge changes in teaching practice: listen to students' ideas and use of terminology, avoid "real world" problems with cultural settings that favor middle-class students, and don't force students to construct knowledge (only some will, with class and ethnic differences). They cite the usefulness of tools from mathematics in analysis for social change, its "setting for developing cultural knowledge and appreciation," and its emphasis on reasoning and alternative solutions, which helps in learning "the value of others' perspectives and ideas." They conclude by urging a change in the population of teachers, to one more "diverse in race, culture and ethnicity, and linguistic resources."

Gold, Lauren, Physicist's algorithm simplifies biological imaging—and also solves Sudoku puzzles, Cornell Chronicle Online `http://www.news.cornell.edu/stories/Feb06/Elser.sudoku.lg.html`. Elser, Veit, Reconstruction of an object from its symmetry-averaged diffraction pattern, `http://arxiv.org/abs/physics/0505174`; The Mermin fixed point, *Foundations of Physics* 33 (11) 2003: 1691–1698, `http://arxiv.org/abs/nlin.CG/0206025`.

Promulgate an algorithm for imaging, and the world yawns; tell them that it can solve Sudoku puzzles, too, and your campus public relations office may make you and your algorithm famous. Veit Elser (Cornell University) has developed an algorithm for X-ray diffraction microscopy, a technique that uses "soft" X-rays that do not damage the specimen. The algorithm constructs the image from the diffraction pattern, using Fourier synthesis and two constraints: a clearly defined boundary, and matching wave amplitudes in the synthesis to those measured in the experiment (a nonconvex constraint). (For Sudoku, the two constraints are that each digit appears only once per row and column, and all nine digits appear in each subblock.) The principle behind Elser's *difference-map algorithm* is the same as with Newton's method for root-finding: "to construct an iterated map whose fixed points are by design the problem's solution." Elser claims that his algorithm is superior to the naive alternating projection map (special cases of which are known in other contexts as biproportional scaling or the balancing algorithm). Although the difference map can be computed efficiently ("in a time that grows only quasi-linearly with the number of pixels"), there is no theory yet about the number of iterations required for convergence.

Crilly, Tony, *Arthur Cayley: Mathematician Laureate of the Victorian Age*, John Hopkins Press, 2006; xxiii + 610 pp, $69.95. ISBN 0–8018–8011–4.

This is a thorough biography, by a mathematician, of Arthur Cayley, England's leading mathematician of the nineteenth century. He was "driven by the beauty of mathematics to the point of obsession" and "constantly diverted from subject to subject as he pursued the mathematical zeitgeist." We learn tantalizingly that his closest colleagues admired him for his character, though it is hard to a get a feel for that side of the man from this biography; given the thorough research that went into this volume, the shortcoming is no doubt in the information available.

Wapner, Leonard M., *The Pea and the Sun: A Mathematical Paradox*, A K Peters, 2005; xiv + 218 pp, $34. ISBN 1–56881–213–2.

This book is devoted to a "journalistic, (as opposed to mathematically) intensive, look" at the Banach-Tarski paradox, which in generalized form states that "a solid of any shape and volume can be decomposed and reassembled to form another solid of any specified shape and volume." To be accurate, however, volumes are not involved, since the decomposition involves nonmeasurable sets; naturally, the proof is nonconstructive. This book is entertaining and instructive, and its paradox may "hook" readers into an appreciation of mathematics.

Franzén, Torkel, *Gödel's Theorem: An Incomplete Guide to Its Use and Abuse*, A K Peters; x + 172 pp, $24.95 (P). ISBN 1–56881–238–8.

At last, a book, devoid of all but the most essential mathematical symbolism, to help set nonmathematical colleagues and friends straight about what Gödel's results do and (mostly) don't say—about human thought, theological "applications," randomness, infinity, or Roger Penrose's arguments for a "science of consciousness." Students of mathematics, whether they have had a course in mathematical logic or not, will find the book illuminating and highly readable.

Baker, Stephen (with Bremen Leak), Math will rock your world, *Business Week* (23 January 2006), `http://www.businessweek.com/magazine/content/06_04/b3968001.htm`.

This cover article in *Business Week* touts the growing use of mathematics to mine databases for business and social purposes, notes the increased demand for "luminary quants," and joins the call for training more "math entrepeneurs." It concludes, in the words of James R. Schatz (NSA), that "There has never been a better time to be a mathematician."

NEWS AND LETTERS

Almost Square? Bubba Majors in Business

Editor:

Bubba is a member of a hypothetical Calculus I class where the instructor has been motivated by the recent article [**1**] which reviewed *almost square* rectangles that had originally appeared in [**2**]. This hypothetical instructor decides to give a supplementary lecture on Farmer Ted's construction of a rectangular chicken coop of 190 square feet with the restriction of the pen having integer sides. The solution presented in class (and in these papers) is that Ted constructs an 11 foot by 17 foot pen reducing the area to 187 square feet. (A 10 by 19 pen is poor use of fencing.)

Bubba is perplexed. He reasons that if he used the same amount of fence, $2 \times (11 + 17) = 56$ feet and divides this by 4, he can have a square pen of 14 by 14 with an area of 196 square feet. Bubba decides that the kind of mathematics presented in this class is nonsense and decides to major in business.

If Bubba had had the nerve and patience to present his solution to the instructor, the instructor would argue that the pen must be less than or equal to 190 square feet. Bubba might argue that the instructor changed the original problem from exactly 190. Why can't the area be increased a little? With the same amount of fence we get 9 more square feet of area.

The above confusion (and the potential loss of a math major) results from us dealing with four different problems where all the variables are integers:

1. Find the smallest perimeter with integer dimensions for a rectangle of exactly A square feet.
2. In a prudent and economic manner find the integer dimensions for a rectangle of approximately A square feet.
3. Find the smallest perimeter with integer dimensions for the largest rectangle not exceeding A square feet whose area to perimeter ratio is maximized.
4. Find the largest integer area rectangle with perimeter not exceeding L feet.

(The reader is urged to stop reading and find the relatively easy solution to 4) which will be presented later.)

Problem 1) is based on factoring A, but when A is prime, the one solution has a large perimeter. Even when A contains a large prime factor, such as when A is 190, the best solution could be "unsatisfactory" in a larger context. Problem 2) is this broader context. Unfortunately, 2) does not have a precise mathematical interpretation. Martin in [**2**] uses problem 3) to create his concept of *almost square* rectangles along with some nice results in composite number theory.

Going back to 2), there is no inherent reason why A cannot be made a little larger. For example if A is 19 then increasing A to 20 produces a 4×5 rectangle rather than 3)'s solution of 18 giving a 3×6 rectangle. Both rectangles have a perimeter of 18.

Problem 4) is an alternate mathematically precise interpretation of 2). With a little prompting students can solve problem 4) themselves. The solution to 4) gives areas of the form $c \times c$ or of the form $c \times (c + 1)$. An integer L is one of the following forms: $4c$, $4c + 1$, $4c + 2$, or $4c + 3$. The second and fourth of these are odd and do not produce an integer sided rectangle. Length $4c$ gives an area of $c \times c$ and $4c + 2$ gives an area of $c \times (c + 1)$. There is an intuitive interpretation for this solution. Suppose one has a square that needs to be a little larger. Increase one of the dimensions by 1. If later we need something a little larger still, increase the other dimension by 1, getting the next square. We get the following areas:

$$\{1, 2, 4, 6.9, 12, 16, 20, 25, 30, 36, 42, 49, \ldots\}$$

If an instructor is motivated to stretch students' minds with integer area rectangles, approach 4) is probably more satisfying and less confusing to a typical undergraduate student than farmer Ted's approach in 3). Bubba might major in math!

Alice is another member of this hypothetical Calculus I class. She is perplexed by rectangles with dimensions 2 by 1, 3 by 1, and 5 by 3 being called *almost square*. The ratio of the length to width of these rectangles is high. She also hears the instructor indicate that an integer of the form $n^2 + 1$ for n greater than one is never the area of an almost square rectangle. The first few examples: 5, 10, 17, 26, just about convince her. She decides to challenge her boy friend, Joe, with this problem. Joe, who is a computer science major, uses his calculator to discover the example: $57^2 + 1 = 3250 = 65 \times 50$ that has a length to ratio of 1.3. Alice now has a 5 by 3 rectangle with a ratio of 1.67 being called *almost square* and a 65 by 50 rectangle with ratio 1.3 that is not *almost square*. Alice decides to major in art.

Joe on the other hand is hooked by the problem. He writes a computer program to generate more numbers of the form $n^2 + 1$ that can be factored into 2 integers whose ratio is smaller still: $73^2 + 1 = 65 \times 82$ and $91^2 + 1 = 82 \times 101$. He asks himself the question: Can the length to width ratio be made as small as one likes? A different math professor asks Joe to see if he can discover a pattern in these factors as it relates to n. Joe discovers the identity:

$$[m(m + 1) + 1]^2 + 1 = [m^2 + 1][(m + 1)^2 + 1]$$

Joe decides to major in math!

If we want to inspire new students to pursue mathematics, confusing terminology needs to be avoided. Instead we should have a collection of counter intuitive questions to "hook" their interest. One such question: Do there exist integers n so that $n^2 + 1 = jk$ where the ratio of j to k can be made close to 1? (For children this could be asked in terms of rearranging tiles.)

REFERENCES

1. S. Alspaugh, *Farmer Ted Goes 3D*, this MAGAZINE, **78** (2005) 192–204.
2. G. Martin, *Farmer Ted Goes Natural*, this MAGAZINE, **72** (1999) 259–276.

James Kropa
Marietta, GA
jkropa@spsu.edu
(received 2-15-06)

Response:

I must disagree with James Kropa's critique of [**1**] and [**2**] in which he claims that these types of articles would drive students away from mathematics. Rather, Martin's article was the basis for research completed by two undergraduates, which further motivated them to attend summer REUs and to pursue graduate degrees in mathematics. In that sense, these are exactly the types of articles that this magazine should publish.

One of the strengths of the problem is that there are many generalizations of what it means to be almost a square (or a cube). While some students may dispute the definition of terms in [**1**] or [**2**], these were chosen by the authors as interesting and worthy of study. These students are encouraged to create their own definitions and begin their own studies of these objects. This is the type of activity which helps potential math majors become actual math majors, even encouraging them to make mathematics a career.

Shawn Alspaugh
Indiana University
shalspau@indiana.edu
(received 3-7-06)

Response from the editor:

I am shocked that Bubba, budding business major, is unaware that the federal government imposes a steep tax and an onerous paperwork burden on chicken coops that exceed 190 square feet. His proposed solution of 14×14 would spell disaster for poor Farmer Ted. While the motivation for the conditions restricting the problem were not fully explained, we mathematicians accept such conditions on faith and let others worry about why these conditions are the proper restrictions.

And I am surprised that artist Alice has forgotten what Humpty Dumpty told her: "When I use a word it means just what I choose it to mean—neither more nor less." Alice wants to impose an extra condition on "almost square" that was not part of the original definition. I am not overly fond of the phrase chosen by the author, but, like Humpty Dumpty, he gets to choose the meaning of the words.

Mathematical problems like these are not posed as practical real world exercises intended to attract majors by demonstrating the remarkable utility of our subject. Instead they are intended to lay down the rules of an arcane game or puzzle. Is the Four Color Theorem a practical application? Is Fermat's Last Theorem useful? Do we urgently need to connect three house and three utilities with noncrossing lines?

The appeal of our subject is the strange, beautiful, and unanticipated consequences of these seemingly innocent restrictions. Real mathematicians, the ones I want to draw into the major, respond to the wonder and beauty of such problems. We chuckle about the strange conclusions, and pity those who do not love the mystery of it all. Practicality? Utility? That must be some other subject.

Allen Schwenk
Western Michigan University
schwenk@wmich.edu

46th International Mathematical Olympiad Mérida, México July 13 and 14, 2005

Edited by Zuming Feng, Cecil Rousseau, and Melanie Wood

PROBLEMS

1. Six points are chosen on the sides of an equilateral triangle ABC: A_1 and A_2 on BC, B_1 and B_2 on CA, and C_1 and C_2 on AB. These points are vertices of a convex equilateral hexagon $A_1A_2B_1B_2C_1C_2$. Prove that lines A_1B_2, B_1C_2, and C_1A_2 are concurrent.
2. Let $a_1, a_2, \ldots$ be a sequence of integers with infinitely many positive and infinitely many negative terms. Suppose that for each positive integer n, the numbers $a_1, a_2, \ldots, a_n$ leave distinct remainders upon division by n. Prove that every integer occurs exactly once in the sequence.
3. Let x, y, and z be positive real numbers such that $xyz \geq 1$. Prove that
$$\frac{x^5 - x^2}{x^5 + y^2 + z^2} + \frac{y^5 - y^2}{y^5 + z^2 + x^2} + \frac{z^5 - z^2}{z^5 + x^2 + y^2} \geq 0.$$
4. Consider the sequence $a_1, a_2, \ldots$ defined by
$$a_n = 2^n + 3^n + 6^n - 1.$$
for all positive integers n. Determine all positive integers that are relatively prime to every term of the sequence.
5. Let $ABCD$ be a given convex quadrilateral with sides BC and AD equal in length and not parallel. Let points E and F lie on sides BC and AD, respectively, such that $BE = DF$. Lines AC and BD meet at P, lines BD and EF meet at Q, and lines EF and AC meet at R. Consider all the triangles PQR as E and F vary. Show that the circumcircles of these triangles have a common point other than P.
6. In a mathematical competition 6 problems were posed to the contestants. Each pair of problems was solved by more than $\frac{2}{5}$ of the contestants. Nobody solved all 6 problems. Show that there are at least 2 contestants who each solved exactly 5 problems each.

SOLUTIONS

Note: For interested readers, the editors recommend the *USA and International Mathematical Olympiads 2005*. There many of the problems are presented together with a collection of remarkable solutions developed by the examination committees, contestants, and experts, during or after the contests.

1. Set $x = AB$ and $s = A_1A_2$. We construct an equilateral triangle $A_0B_0C_0$ with $A_0B_0 = x - s$. Points C_4, A_4, and B_4 lie on sides A_0B_0, B_0C_0, and C_0A_0, respectively, satisfying $C_4B_0 = C_2B$, $A_4C_0 = A_2C$, and $B_4A_0 = B_2A$. Then it is easy to obtain that $B_0A_4 = BA_1$, $C_0B_4 = CB_1$, and $A_0C_4 = AC_1$. We obtain three pairs of congruent triangles, namely, AB_2C_1 and $A_0B_4C_4$, BC_2A_1 and $B_0C_4A_4$, and CA_2B_1 and $C_0A_4B_4$. (Indeed, we are sliding the three *corner* triangles together.)

   It follows that $A_4B_4 = B_4C_4 = C_4A_4 = s$; that is, triangle $A_4B_4C_4$ is equilateral, implying that $\angle B_4C_4A_4 = \angle C_4A_4B_4 = \angle A_4B_4C_4 = 60°$. Hence $\angle A_0B_4C_4 + \angle A_0C_4B_4 = \angle B_0C_4A_4 + \angle A_0C_4B_4 = 120°$, and so $\angle A_0B_4C_4 = \angle B_0C_4A_4$. Hence

$\angle AB_2C_1 = \angle BC_2A_1$, or $\angle B_1B_2C_1 = \angle C_1C_2A_1$. Since the vertex angles of the isosceles triangles $B_1B_2C_1$ and $C_1C_2A_1$ are equal, then two triangles are similar and hence congruent to each other, implying that $C_1B_1 = C_1A_1$. Since $C_1B_1 = C_1A_1$ and $A_2B_1 = A_2A_1$, line C_1A_2 is a perpendicular bisector of triangle $A_1B_1C_1$. Likewise, so are lines A_1B_2 and B_1C_2. Therefore, lines C_1A_2, A_1B_2, and B_1C_2 concur at the circumcenter of triangle $A_1B_1C_1$.

It is not difficult to see that triangle $A_0B_4C_4$ is congruent to triangle $B_0C_4A_4$ (and to triangle $C_0A_4B_4$).

2. The conditions of the problem can be reformulated by saying that for every positive integer n, the numbers $a_1, a_2, \ldots, a_n$ form a complete set of residues modulo n. We proceed our proof as the following.

   (1) First, we claim that the sequence consists of distinct integers; that is, if $1 \le i < j$, then $a_i \neq a_j$. Otherwise the set $\{a_1, a_2, \ldots, a_j\}$ would contain at most $j-1$ distinct residues modulo j, violating our new formulation of the conditions of the problem.

   (2) Second, we show that numbers in the sequence are fairly close to each other. More precisely, we claim that if $1 \le i < j \le n$, then $|a_i - a_j| \le n-1$. For if $m = |a_i - a_j| \ge n$, then the set $\{a_1, a_2, \ldots, a_m\}$ would contain two numbers congruent modulo m, violating our new formulation of the conditions of the problem.

   (3) Third, we show that the set $\{a_1, a_2, \ldots, a_n\}$ contains a block of consecutive numbers. Indeed, for every positive integer n, let i_n and j_n be the indices such that a_{i_n} and a_{j_n} are respectively the smallest and the largest number among $a_1, a_2, \ldots, a_n$. By (2), we conclude that $a_{j_n} - a_{i_n} = |a_{j_n} - a_{i_n}| \le n-1$. By (1), we conclude that $\{a_1, a_2, \ldots, a_n\}$ consists of all integers between a_{i_n} and a_{j_n} (inclusive).

   (4) Finally, we show that every integer appears in the sequence. Let x be an arbitrary integer. Because $a_k < 0$ for infinitely many indices k and the terms of the sequence are distinct, it follows that there exists i such that $a_i < x$. Likewise, there exists j such that $x < a_j$. Let n be an integer with $n \ge \max\{i, j\}$. By (3), we conclude that every number between a_i and a_j, including x in particular, is in $\{a_1, a_2, \ldots, a_n\}$. Our proof is thus complete.

3. Note that

$$\frac{x^5 - x^2}{x^5 + y^2 + z^2} \ge \frac{x^5 - x^2}{x^3(x^2 + y^2 + z^2)}$$

is equivalent to

$$\frac{(x^3 - 1)^2(y^2 + z^2)}{x(x^5 + y^2 + z^2)(x^2 + y^2 + z^2)} \ge 0,$$

which is true for all positive x, y, z. Hence

$$\frac{x^5 - x^2}{x^5 + y^2 + z^2} \ge \frac{x^2 - \frac{1}{x}}{x^2 + y^2 + z^2}.$$

Summing the above inequality with its analogous cyclic inequalities, we see that the desired result follows from

$$x^2 + y^2 + z^2 - \frac{1}{x} - \frac{1}{y} - \frac{1}{z} \ge 0.$$

Since $xyz \geq 1$,

$$x^2 + y^2 + z^2 - \frac{1}{x} - \frac{1}{y} - \frac{1}{z} = x^2 + y^2 + z^2 - \frac{yz + xz + xy}{xyz}$$

$$\geq x^2 + x^2 + z^2 - yz - xz - xy = \frac{(x-y)^2 + (y-z)^2 + (z-x)^2}{2} \geq 0,$$

so we are done.

4. The answer is that 1 is the only such number. It suffices to show that every prime p divides a_n for some positive integer n. Note that both $p = 2$ and $p = 3$ divide $a_2 = 2^2 + 3^2 + 6^2 - 1 = 48$.

   Now we assume that $p \geq 5$. By **Fermat's Little Theorem**, we have $2^{p-1} \equiv 3^{p-1} \equiv 6^{p-1} \equiv 1 \pmod{p}$. Then

$$3 \cdot 2^{p-1} + 2 \cdot 3^{p-1} + 6^{p-1} \equiv 3 + 2 + 1 \equiv 6 \pmod{p},$$

   or, $6(2^{p-2} + 3^{p-2} + 6^{p-2} - 1) \equiv 0 \pmod{p}$; that is, $6a_{p-2}$ is divisible by p. Because p is relatively prime to 6, a_{p-2} is divisible by p, as desired.

5. Applying the Law of Sines to triangles ARF and CRE gives

$$\frac{AR}{RC} = \frac{AR}{AF} \cdot \frac{CE}{CR} = \frac{\sin\angle AFR}{\sin\angle ARF} \cdot \frac{\sin\angle CRE}{\sin\angle CER} = \frac{\sin\angle AFR}{\sin\angle CER},$$

   as $\angle ARF = \angle CRE$. Likewise,

$$\frac{DQ}{QB} = \frac{\sin\angle DFQ}{\sin\angle BEQ} = \frac{\sin\angle AFR}{\sin\angle CER} = \frac{AR}{RC},$$

   by noting that $\angle DFQ + \angle AFR = 180°$ and $\angle BEQ + \angle CER = 180°$. Let Y be the center of the spiral similarity (denoted by $\mathbf{S}_1$) that sends segment BD to CA. (The existence of this center is to be explained later). Then $\mathbf{S}_1(Q) = R$. Then we have $\angle BPC = \angle QYR$, because both are the angle of rotation of $\mathbf{S}_1$. Hence $RPQY$ is cyclic; that is, the circumcircle of triangle PQR always passes through Y.

   Now we consider the existence of point Y. For any two nonparallel segments AD and BC (not necessarily having equal length), let Z be the intersection of lines AD and BC. Then Y is the second intersection of circumcircles of triangles ACZ and BDZ. (Because these two circles clearly are not tangent at Z, point Y exists.) Indeed, from the cyclic quadrilaterals $BYDZ$ and $AZCY$, we have $\angle CBY = \angle ZBY = \angle ADY$ and $\angle YCB = \angle YAZ = \angle YAD$, implying that triangle ADY is similar to CBY; that is, Y is the center of spiral similarity that sends triangle ADO to triangle CBO.

6. Suppose that there were n contestants. Let p_{ij}, with $1 \leq i < j \leq 6$, be the number of contestants who solved problems i and j, and let n_r, with $0 \leq r \leq 6$, be the number of contestants who solved exactly r problems. Clearly, $n_6 = 0$ and $n_0 + n_1 + \cdots + n_5 = n$.

   By the given condition, $p_{ij} > \frac{2n}{5}$, or $5p_{ij} > 2n$. Hence $5p_{ij} \geq 2n + 1$, or $p_{ij} \geq \frac{2n+1}{5}$. We define the set

$$U = \{(c, \{i, j\}) \mid \text{contestant } c \text{ solved problems } i \text{ and } j\}.$$

   If we compute $|U|$, the number of elements in U, by summing over all pairs $\{i, j\}$, we have

$$|U| = \sum_{1 \leq i < j \leq 6} p_{ij} \geq 15 \cdot \frac{2n+1}{5} = 6n + 3 = 6(n_0 + n_1 + \cdots n_5) + 3.$$

A contestant who solved exactly r problems contributes a "1" to $\binom{r}{2}$ summands in this sum (where $\binom{r}{2} = 0$ for $r < 2$). Therefore,

$$|U| = \sum_{r=0}^{6} \binom{r}{2} n_r = n_2 + 3n_3 + 6n_4 + 10n_5.$$

It follows that $n_2 + 3n_3 + 6n_4 + 10n_5 \geq 6(n_0 + n_1 + \cdots + n_5) + 3$, or

$$4n_5 \geq 3 + 6n_0 + 6n_1 + 5n_2 + 3n_3 \geq 3,$$

implying that $n_5 \geq 1$. We need to show that $n_5 \geq 2$. We approach indirectly by assuming that $n_5 = 1$. We call this person the *winner* (denote by W), and without loss of generality, we may assume that the winner failed to solve problem 6. Then $n_0 = n_1 = n_2 = n_3 = 0$. Hence $n_4 = n - 1$, and so

$$|U| = n_2 + 3n_3 + 6n_4 + 10n_5 = 6n + 4 > 6n + 3 = 15 \cdot \frac{2n+1}{5}.$$

Let $m = \frac{2n+1}{5}$. It follows that $p_{ij} = m$ for 14 out of the 15 total pairs (i, j) with $1 \leq i < j \leq 15$, and for the remaining pair (s, t), $p_{st} = m + 1$.

Let

$$d_i = \sum_{j \neq i} p_{ij}, \quad i = 1, 2, \ldots, 6.$$

We have just seen that $d_s = d_t = 5m + 1$ and $d_i = 5m$ otherwise. On the other hand, consider what happens if we build up the 6-tuple $(d_1, d_2, \ldots, d_6)$ one contestant at a time, starting with W. Thus we start with $(4, 4, 4, 4, 4, 0)$, and every subsequent contestant adds a permutation of $(3, 3, 3, 3, 0, 0)$. Thus

$$(d_1, d_2, \ldots, d_6) \equiv (1, 1, 1, 1, 1, 0) \pmod 3,$$

contradicting the earlier conclusion that $d_s = d_t = 5m + 1$ and $d_i = 5m$ otherwise. Hence there were are least two persons to solve five problems.

2005 Olympiad Results

The top twelve students on the 2005 USAMO were (in alphabetical order):

Robert Cordwell	Manzano High School	Albuquerque, NM
Zhou Fan	Parsippany Hills High School	Parsippany, NJ
Sherry Gong	Phillips Exeter Academy	Exeter, NH
Rishi Gupta	Henry M. Gunn High School	Palo Alto, CA
Hyun Soo Kim	Academy of Advancement in Science and Technology	Hackensack, NJ
Brian Lawrence	Montgomery Blair High School	Silver Spring, MD
Albert Ni	Illinois Math and Science Academy	Aurora, IL
Natee Pitiwan	Brooks School	North Andover, MA
Eric Price	Thomas Jefferson High School of Science and Technology	Alexandria, VA
Peng Shi	Sir John A. MacDonald Collegiate Institute	Toronto, ON
Yi Sun	The Harker School	San Jose, CA
Yufei Zhao	Don Mills Collegiate Institute	Don Mills, ON

Brian Lawrence, was the winner of the Samuel Greitzer-Murray Klamkin award, given to the top scorer(s) on the USAMO. Brian Lawrence and Eric Price placed first and second, respectively, Peng Shi and Yufei Zhao tied for third, on the USAMO. They were awarded college scholarships of \$20000, \$15000, \$5000, and \$5000, respectively, by the Akamai Foundation. The Clay Mathematics Institute (CMI) award, for a solution of outstanding elegance, and carrying a \$5000 cash prize, was presented to Sherry Gong for her solution to USAMO Problem 3.

The USA team members were chosen according to their combined performance on the 34th annual USAMO, and the Team Selection Test that took place at the Mathematical Olympiad Summer Program (MOSP), held at the University of Nebraska-Lincoln, June 12–July 2, 2005. Members of the USA team at the 2005 IMO (Mérida, México) were Robert Cordwell, Sherry Gong, Hyun Soo Kim, Brian Lawrence, Thomas Mildorf, and Eric Price. Zuming Feng (Phillips Exeter Academy) and Melanie Wood (Princeton University) served as team leader and deputy leader, respectively. The team was also accompanied by Steven Dunbar (University of Nebraska-Lincoln), as observers of the deputy leader.

There were 513 contestants in the 2005 IMO. The average score is 13.97 (out of 42) points. Gold medals were awarded to students scoring between 35 and 42 points, silver medals to students scoring between 23 and 34 points, and bronze medals to students scoring between 12 and 22 points. There were 42 gold medalists, 79 silver medalists, and 122 bronze medalists. Brian submitted one of the 16 perfect papers. Moldovian contestant Iurie Boreico's elegant solution on problem 3 (presented in this article) won a special award in the IMO, the first time this award is given in the past 10 years. The team's individual performances were as follows:

Cordwell	GOLD Medallist	Gong	SILVER Medallist
Kim	SILVER Medallist	Lawrence	GOLD Medallist
Mildorf	GOLD Medallist	Price	GOLD Medallist

In terms of total score (out of a maximum of 252), the highest ranking of the 93 participating teams were as follows:

China	235	Iran	201	Taiwan	190	Ukraine	181
USA	213	Korea	200	Japan	188	Bulgaria	173
Russia	212	Romania	191	Hungary	181	Germany	163

Problems for the 2005 USAMO were chosen by the USAMO Committee [Steve Blasberg, Steven Dunbar, Gregory Galperin, Elgin Johnston, Kiran Kedlaya, Cecil Rousseau (chair), Richard Stong, Zoran Šunik, and David Wells]. Proposals were made by members of the committee and other highly experienced individuals [Titu Andreescu, Gabriel Dospinescu, Zuming Feng, Răzvan Gelca, Gerald Heuer, Alex Saltman, and Melanie Wood]. The Team Selection Test (TST) was prepared by Zuming Feng and Melanie Wood, with gracious help from Kiran Kedlaya, Richard Stong, and Ricky Liu.

The MOSP was held at the University of Nebraska-Lincoln. Because of a generous gift from the Akamai Foundation, the 2005 MOSP expanded from the usual 24–30 students to 55. Titu Andreescu, Reid Barton, Zuming Feng (Academic Director), Chris Jewell, Ian Le, Josh Nichols-Barrer, and Melanie Wood served as instructors. Ricky Liu and Po-Ru Loh were junior instructors. Oleg Golberg, Anders Kaseorg, Mark Lipson, Tiankai Liu, and Tony Zhang were graders.

For more information about the USAMO or the MOSP, contact Steven Dunbar at `sdunbar@math.unl.edu`.

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

THE MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth Street, NW
Washington, DC 20036