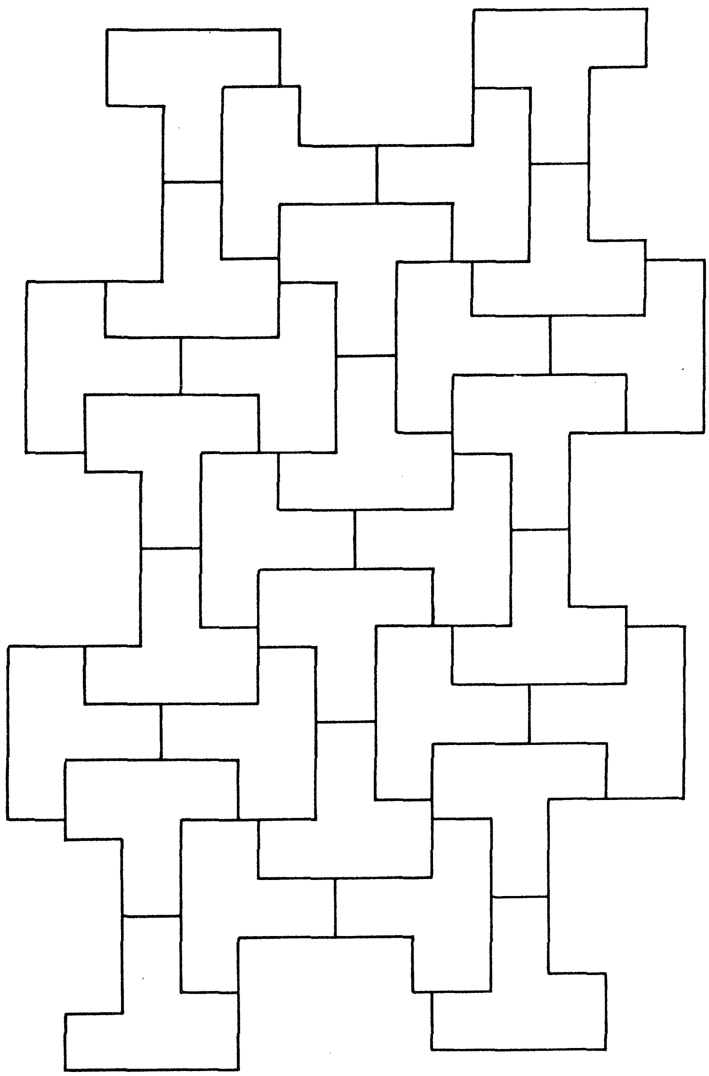
MATHEMATICS MAGAZINE



Vol. 57 No. 5
November 1984

PARLIAMENTARY COALITIONS • POLYOMINO TILES
APPROXIMATING INTEGRALS • TOURING CHECKERBOARDS

New from The Mathematical Association of America

VOLUME 6 — DOLCIANI MATHEMATICAL EXPOSITIONS

MAXIMA AND MINIMA WITHOUT CALCULUS,

by Ivan Niven
xv + 323 pp. Hardbound
List: $24.50   MAA Member: $18.50

Ivan Niven, distinguished author of several books on number theory and probability, has compiled the basic elementary techniques for solving maxima and minima problems. Since many books and courses already cover calculus and linear programming techniques, the author deliberately omits these areas from his discussions and concentrates instead on methods in algebra and geometry not so widely known. These methods are organized according to the mathematical ideas used, and many chapters can be read independently without reference to what precedes or follows. Some of the problems presented are left for the reader to solve with sketches of solutions given in the later pages.

The book is written for an audience at or near the maturity level of second- or third-year college students with a good working knowledge of precalculus mathematics. Although calculus is not a prerequisite, a prior knowledge of that subject will enhance the reader's comprehension. An excellent sourcebook in the area of maxima and minima, the book will serve as a textbook or as enrichment material for the talented undergraduate.

The main topics covered are:

- **Simple Algebraic Results**
- **Elementary Geometric Questions**
- **Isoperimetric Results**
- **Basic Trigonometric Inequalities**
- **Polygons Inscribed and Circumscribed**
- **Ellipses**
- **The Bees and Their Hexagons**
- **Further Geometric Results**
- **Applied and Miscellaneous Problems**
- **Euclidean Three-Space**
- **Isoperimetric Results not Assuming Existence**
- **Postscript on Calculus**

Order from:
The Mathematical Association of America
1529 Eighteenth St., NW
Washington, DC 20036

MATHEMATICS MAGAZINE

Volume 57, Number 5, November 1984
(ISSN— 0025- 570X)

COVER: Tiling by a decamorphic "T" polyomino. See pp. 275–283.

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS

INDEX TO VOLUME 57

EDITORIAL POLICY

Mathematics Magazine is a journal which aims to provide inviting, informal mathematical exposition. Manuscripts for the *Magazine* should be written in a clear and lively expository style and stocked with appropriate examples and graphics. Our advice to authors is: say something new in an appealing way or say something old in a refreshing way. The *Magazine* is not a research journal and so the style, quality, and level of articles should realistically permit their use to supplement undergraduate courses. The editor invites manuscripts that provide insight into the history and application of mathematics, that point out interrelationships between several branches of mathematics and that illustrate the fun of doing mathematics.

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

Send new manuscripts to: Doris Schattschneider, Editor, Mathematics Magazine, Moravian College, Bethlehem, PA 18018. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and one copy and keep one copy. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added.

AUTHORS

Philip D. Straffin Jr. ("Parliamentary Coalitions: A Tour of Models") is Professor of Mathematics and Computer Science at Beloit College. He first became interested in the problems and challenges of mathematical modeling in social science through teaching in Beloit's interdisciplinary division. He has written a number of articles on game theory and mathematical political science and a UMAP monograph *Topics in the Theory of Voting*. Work on this article was done during a 1983 sabbatical term at the Kennedy School of Government at Harvard University. His current project is a book on *Game Theory and Social Science*.

Co-author **Bernard Grofman** is Professor of Political Science and Social Psychology, School of Social Sciences, University of California, Irvine. He is a specialist in mathematical models of collective decision-making and the political consequences of electoral laws, with over two dozen published articles on topics such as jury verdict choice, reapportionment and voter turnout, and coalition formation models. Recently he has served in five states as an expert witness in redistricting litigation or as a court-appointed reapportionment expert.

ILLUSTRATIONS

David Logothetti portrays the non-mathematical aspects of politics for "Parliamentary Coalitions: A Tour of Models."

Ryland Loos provided the illustrations for "Polymorphic Polyominoes," pp. 275–283.

All other illustrations were provided by the authors.

The MATHEMATICS MAGAZINE (ISSN 0025–570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, five times a year: January, March, May, September, and November.
The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $11 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $22. Student, unemployed and emeritus members receive a 50% discount; new members receive a 30% dues discount for the first two years of membership.) The non-member/library subscription price is $28 per year. Bulk subscriptions (5 or more copies) are available to colleges and universities for classroom distribution to undergraduate students at a 41% discount ($6.50 per copy—minimum order $32.50).
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. Back issues may be purchased, when in print, from P. and H. Bliss Company, Middletown, CT 06457. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.
Advertising correspondence should be addressed to Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.

Second class postage paid at Washington, D.C. and additional mailing offices.
Postmaster: Send address changes to Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.
PRINTED IN THE UNITED STATES OF AMERICA

Parliamentary Coalitions: A Tour of Models

Mathematics explores the relationship between politics and rationality.

PHILIP D. STRAFFIN JR.
Beloit College
Beloit, WI 53511

BERNARD GROFMAN
University of California
Irvine, CA 92717

The study of government and politics, like all the other social sciences, has made increasing use of mathematics in the last thirty years. This has involved not only the use of statistics and computers to increase the scope and thoroughness of empirical work, but also the development of mathematical models in political theory. Through the use of models, mathematics has come to play an important role in political science at the conceptual level. This development has not been greeted with universal enthusiasm by practitioners in the field. An academic colleague of one of the authors, in a department of Government, has been heard to complain with some bitterness that he can no longer read the *American Political Science Review* because it is full of obscure mathematical symbols. Non-enthusiasts have their point. Politics often seems far removed from the domain of precise, rational thought which mathematics epitomizes. Nevertheless, mathematics *has* been able to contribute insights to the study of political systems. In this article we explore the use of mathematical models in various attempts to understand one of the central concerns of politics—the ways in which politicians go about forming coalitions with other politicians.

In the first section we introduce the problem of coalition formation in the context of parties trying to form a coalition government in a parliamentary system. This will be our central focus in the article. We start by considering some of the earliest and simplest models, which lead to a geometric formulation of the problem. In the second section we discuss two more complex models from the tradition of mathematical game theory. We have tried to illustrate the methods of thinking involved in these models without presenting all of the formal details (for which we give references). In the next section we return to a simpler "dynamic" model of coalition formation and explore its properties. Finally, we have some general words to say about the problems of testing models in the social sciences, and a "coda" in which we invite you to think about how a coalition model might be useful outside of political science.

The problem of coalition formation and early models

The motivating problem for much of the mathematical work on coalition theory, and the data base for most of the empirical work until quite recently, is the problem of understanding the formation of governing coalitions in parliamentary democracies. Most of the data is from western Europe. Let's look at a particular situation, in order to have a specific example before us.

After the 1965 parliamentary elections in Norway, the five major parties had won the following numbers of seats:

Party	Number of seats
A. Labor	68
B. Christian	13
C. Liberal	18
D. Center	18
E. Conservative	31
	148

Since no party had the necessary majority of 75, a majority government could be formed only by a coalition of two or more parties. Which coalition would we expect to form? The goal in positive political theory (see, for example, [**23**], [**24**], or [**1**]) is to build a model to explain which coalition should form in this and similar situations, on the basis of some kind of rational choices by the actors (here parties) involved. We would then test the model against the large amount of data on twentieth century parliamentary coalitions. If the fit is good, we think we have a deeper and more general understanding of parliamentary coalition formation than would be obtained by, say, case-by-case study or exhaustive descriptive classification.

To build a rational choice model, we need to posit some kind of goals for the actors involved. We think they might then act to try to reach those goals. The earliest models posited the following goal:

Goal P. *There is a positive payoff associated to the act of governing. This payoff will be shared in some way among the members of a governing coalition. Parties wish to maximize their share of this payoff.*

Notice that we need not be specific about the nature of the payoff. It might be something like "power," that nebulous entity which *homo politicus* might value for its own sake, as opposed to

"There is a positive payoff associated to the act of governing."

valuing it for the sake of attaining other goals. It might be some kind of "spoils" of governing, say in the form of patronage appointments. However, it is something divisible which can be shared. It isn't related to political ideology or, at least directly, to the ability to implement certain governing policies. This goal has the advantage of being minimalist. The idea is to see how well we can explain coalition formation using only this goal. If we fail, we will want to consider other possible goals.

The most direct consequence of this model was most clearly enunciated by William Riker [**23**] in the late 1950s. If the payoff is to be shared among the members of a governing coalition with each member wanting as much as possible, then the coalition which forms should have no superfluous members—no members whose deletion would leave the coalition still of winning size. If there were such a member, all of the other members would be better off expelling him and dividing his share of the payoff among themselves. In other words, we have the

Riker Minimal Winning Coalition Principle. *Only a minimal winning coalition should form.*

In the Norwegian example there are sixteen winning coalitions, but only five minimal winning coalitions: AB, AC, AD, AE, and $BCDE$. The minimal winning coalition principle says that only one of these five should form. It thus makes a testable prediction, but the prediction is fairly weak, since it doesn't say which one of the five should form. If we make stronger assumptions about how the sharing of the spoils might be done, we can make stronger predictions.

In the late 1950s sociologists experimenting with three-person coalition formation found that people often acted as if they believed that payoffs might be divided in proportion to the resources which each person brought to the coalition [**6**], [**12**]. In our situation, the resources are parliamentary votes. We thus might posit that payoffs will be divided in proportion to the number of parliamentary seats each party has. Riker used this idea to make a more specific prediction:

Riker Least Resources Principle. *The coalition which forms should be the winning coalition with the smallest total number of votes.*

The idea here is that all members of this coalition will prefer it to any other winning coalition because it will give them a larger share of the spoils. For example, consider our Norwegian parties.

Minimal winning coalition	Number of votes
AB	81
AC	86
AD	86
AE	99
$BCDE$	80

The least resources principle predicts that coalition $BCDE$ will form. Party B will prefer this coalition, which will give it $13/80$ of the spoils, to coalition AB, which will give it only $13/81$. Similarly, C, D and E will all prefer this coalition to any other winning coalition. Party A prefers AB ($68/81$ of the spoils), but can't implement this coalition because B prefers $BCDE$.

Of course, we could make other assumptions about how spoils would be divided. For instance, if we assumed they would be divided *equally*, on the grounds that all members of any minimal winning coalition are equally important (since the defection of any one of them would cause the coalition to lose), we would get the

Fewest Actor Principle [**18**]. *The coalition which forms should have the fewest members of all winning coalitions.*

In the Norwegian example, this principle would predict AB, AC, AD or AE instead of $BCDE$. The fewest actor principle might also be attractive if there are large **transaction costs** in forming coalitions. It costs more time, negotiating energy, and uncertainty to put together a four-member coalition than to agree on a two-member coalition.

It's time to come down to earth and look at the facts. When we look at the data on coalition formation in parliamentary systems, we find that ([3], [10], [11], [19])

i) although there are a number of countries which are exceptions, in most countries minimal winning coalitions form much of the time;
ii) neither the least resources principle nor the fewest actor principle predicts with better than chance accuracy.

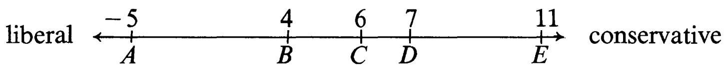
Parties often form minimal winning coalitions, but not necessarily least resource or fewest actor coalitions.

We might note that there are situations in which not even the minimal winning coalition theory is supported. On the one hand, it may not be necessary for a governing coalition to be winning: minority coalition governments have not been uncommon in western democracies [19]. On the other hand, winning coalitions on legislative bills are often larger than minimal winning. One reason seems to be *uncertainty*. In building a large, temporary coalition, you are never quite sure who is with you and who is not, so you need to allow a margin for error (see [23] or [24] for a discussion).

Apparently we need to bring in other goals if we want to build a model which will make more exact predictions than the minimal winning coalition principle. The most important factor which we have not yet considered, and one of the staples of political analysis, is the role of political ideology. Beyond the desire for power or patronage, it is generally believed that political parties do have policies which they would like to see implemented:

Goal I. *A party would like to see its policies implemented. Hence it wants to join a coalition with other parties whose values and ideological positions are close to its own.*

The first problem is how we are to model things like values and ideological positions. The most familiar model dates back to 1789, when parties in the French National Assembly with "radical," "liberal," "moderate," "conservative," and "reactionary" ideologies were seated from left to right in the chamber. In common political analysis it seems to be believed that we can often place parties on a one-dimensional left-right continuum based on their views about the extent to which government should intervene in the economy, redistributional issues, the desirability of legislating social behavior, and other issues. If we think this can be done, perhaps we can assign parties to points on the real line, representing their ideological position in this "ideological space." For instance, if we place the Norwegian parties by their stands on economic issues, we might get something like this [8]:



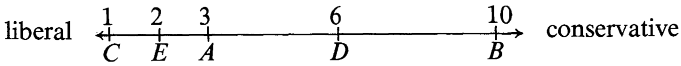
The weak (**ordinal**) assumption is that only the order of the parties along the line is meaningful. The stronger (**cardinal**) assumption is that relative distances between parties is meaningful: D is twice as close to C as B is, E is three times as close to D as A is, and so forth (see, for example, Chapter 8 of [25]).

In 1970 Robert Axelrod [3] noted that even with just the ordinal assumption, we can say that coalitions which are **connected** will have more commonality of values among their members than coalitions which are not connected. Here a coalition S is connected if whenever $X, Z \in S$ and $X < Y < Z$, then $Y \in S$. In the Norwegian example the connected winning coalitions are AB, ABC, $ABCD$, $BCDE$ and $ABCDE$.

Connected Coalition Principle. *The coalition which forms should be connected.*

We can sharpen this principle, as Axelrod did, by combining it with the minimal winning coalition principle to predict that a *connected minimal winning* coalition should form: AB or $BCDE$ in the Norwegian example.

The connected coalition principle and its minimal winning strengthening are both supported fairly well by parliamentary data [**3**], [**10**], [**19**]. For example, Abram DeSwann found that in a sample of 108 coalition governments, 55 were connected minimal winning. However, we can note two problems. The first is that there may be a number of connected minimal winning coalitions, and we might like to make a more precise prediction. The second is that sometimes it is quite difficult to place parties on a one-dimensional continuum. What do you do with a party which is "conservative" on government spending but "liberal" on social issues? As an example, cultural issues are important to many Norwegian voters, and if we locate parties by their stands on cultural issues we get something like [**8**]:



Now the connected minimal winning coalitions are AE and AD.

The most natural way out of this difficulty, for a mathematician at least, is to plot positions on both scales at once, i.e., to represent parties' ideological positions as points in a coordinate plane (see FIGURE 1). If there were three salient kinds of issues, we would plot parties as points in three-dimensional space. Notice that we now need a new definition of when a coalition is connected. The natural definition is that a coalition is connected if it includes any party within the **convex hull** of its members. (The convex hull of a set of points is the smallest convex set containing those points.) In FIGURE 1, AE or CDE or $BCDE$ are connected, but BCE is not connected since it doesn't include D, which is in its convex hull. For this definition to work, the relevant scales must be *cardinal* scales: if D were moved to $(8,6)$ the ordinal relations would be unchanged, but now BCE would be connected. In fact, it is traditional in this kind of "spatial modeling" in political science to make a much stronger assumption: that the relevant scales are **comparable** in the sense that *Euclidean distance* measures ideological proximity (see for example [**24**] or [**1**]). In FIGURE 1, $d(B,D)=5$ and $d(A,B)=11.4$, so A is more than twice as far from B as D is. This quite strong assumption could of course be relaxed, for instance, by developing the theories of the next sections in other metrics, but we will accept this assumption for our discussion.

Finally, notice that, with the expansion into two or more dimensions, the connected minimal winning coalition theory becomes distinctly less helpful. In the Norwegian example of FIGURE 1, *all* minimal winning coalitions are connected, and we would want to narrow down this class.

Party	Weight	Position
A. Labor	68	$(-5, 3)$
B. Christian	13	$(4,10)$
C. Liberal	18	$(6, 1)$
D. Center	18	$(7, 6)$
E. Conservative	31	$(11, 2)$
	148	

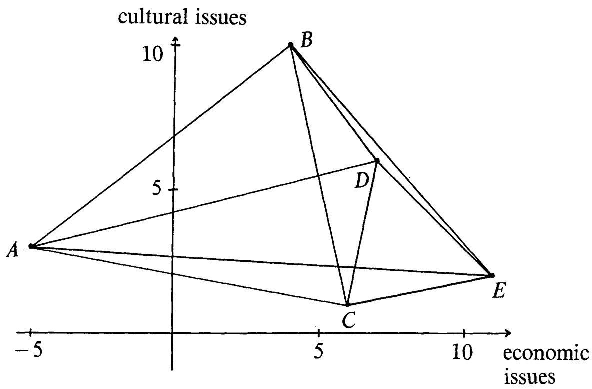


FIGURE 1. **Parties in Norway, 1965.**

Two game theoretic approaches to spatial coalitions

Our model is now as follows. We are given an n-dimensional Euclidean **ideological space**. Parties are represented as points in that space, and the distance between points represents the ideological proximity of the parties. Each party has a **weight** (its number of seats) and winning coalitions are those with large enough total weights (e.g., a majority of seats in the parliament). We can now be more specific about the goal that parties want to join a coalition which will pursue policies similar to theirs. We can think of any coalition, when it forms or thinks about forming, as choosing a point in the ideological space which represents the policies it would follow. Parties want to join coalitions which will choose points close to theirs. For example, suppose that in FIGURE 2 the five parties are approximately equally weighted, so that any three would be a winning coalition. If coalition GHJ formed and chose policy point α, its members would be fairly pleased, since α is fairly close to each of their policy points.

Now the question. In this situation, can we predict which coalition would form? This **spatial coalition theory** has been actively studied for the past ten years, and a number of different solutions have been proposed. We would like to describe two of them, to illustrate the kinds of thinking involved and the variety of possible results. Both solution ideas come from the perspective of the mathematical theory of cooperative games.

In 1964, game theorists Robert Aumann and Michael Maschler proposed a general solution concept for cooperative games which is now known as the **Aumann-Maschler Bargaining Set**. (Actually, there are several different Bargaining Sets. We describe one of the simplest.) For a general elementary introduction, see [**9**]. In the 1970s this idea was applied to the spatial coalition problem ([**26**], [**20**]) and it is this application we will discuss. The Bargaining Set approach centers around the problem of a winning coalition which is considering forming: it must choose a policy point which is *fair* to all its members. In fact, any member which considers the selected point unfair will be likely to refuse to join the coalition. Thus if a potential coalition cannot find a policy point which all its members consider fair, it probably will not form.

We will illustrate the way that Aumann and Maschler make the idea of fairness operational by an example. Suppose that in the spatial situation of FIGURE 2, GHJ are considering forming a winning coalition with policy point α. Although α may seem fair (by symmetry) to us as outsiders, G thinks the policy point should be closer to him and further from J. He **objects** against J. The nature of his objection is that he can do better in another coalition:

> *G's objection against J*: "I can go to F and K and propose forming the winning coalition FGK with policy point β. That is better for me than α. Furthermore, it is also better than α for both F and K, so they would be happy to join me."

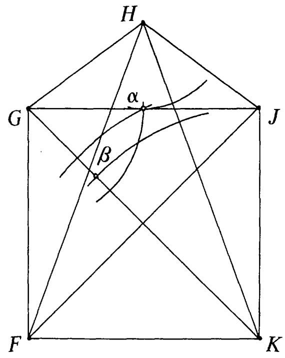


FIGURE 2. **Bargaining set theory. Point α is not a stable proposal for coalition *GHJ*.**

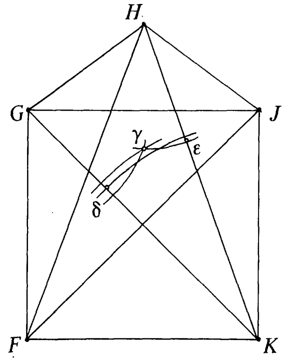


FIGURE 3. **Bargaining set theory. Point γ is a stable proposal for coalition *FHK*.**

G's objection has force, but Aumann and Maschler argue that we cannot consider α unfair, or unstable, just because one coalition member can object against it in this way. The reason is that there could then be *no* fair policy proposals. A little experimentation with the picture should convince you that *any* policy point proposed by any coalition gives rise to this kind of objection by at least one member. (Game theorists express this fact by saying that the **core** of this game is empty.) Aumann and Maschler propose that G's objection should not invalidate α if J can come up with a reasonable **counter-objection**, i.e., a convincing plan of action for what she would do if G tried to carry out the plan in his objection. This should have the following form.

J's attempted counter-objection against G: "I will form a winning coalition which includes me but not you, with a policy point which is
i) at least as good as α for everyone in my coalition, and
ii) at least as good as β for F or K, if they are involved in my coalition."

Condition ii) is necessary in order to convince G that J can compete with him for F or K if she needs them.

Unfortunately, in this case J cannot make such a counter-objection. Any point which is as close as α to J is *not* as close as β to F, and any point which is as close as α to H is *not* as close as β to K. J cannot offer enough to two parties simultaneously to convince them to join her. Since J has no convincing counter-objection to G's objection, we say that α is an **unstable** policy proposal for *GHJ*. In fact, *GHJ* has *no* policy proposal which is stable in the Aumann-Maschler sense. G could successfully object against any point to the right of α, and J could successfully object against any point to the left.

Are there coalitions which do have stable policy proposals? The answer is yes. For example, consider coalition *FHK* proposing policy point γ in FIGURE 3. Suppose F objects against H:

F's objection against H: "I can propose *FGK* with policy point δ. It is closer than γ to me, and also to G and K."

This objection can be countered.

H's counter-objection against F: "If you do, I can propose *HJK* with policy point ε. That's as good as γ was for me and J, and it's as good as your offer of δ for K."

In fact, H can counter any objection by F, using either *GHJ* or *HJK*. In a similar way, any member of *FHK* can counter any objection by any other member. The point γ (which lies on a line segment defined by inequalities which say that any objection can be countered) is a stable policy proposal for *FHK*.

In this example, it turns out that five three-party coalitions have Aumann-Maschler stable proposals: *FHK*, *GJK*, *FHJ*, *GHK* and *FGJ*. These are exactly the coalitions which "split the opposition," in the sense that the convex hull of such a coalition intersects the convex hull of its opposition (for example, the convex hull of *FHK* intersects the convex hull of *GJ*). We'll call them **internal** coalitions [**20**]. All of their stable policy points are inside the central pentagon, reasonable compromise policies which are about as good for nonmembers of the coalition as they are for members. The coalitions which do not have stable policy proposals are the coalitions like *GHJ* whose convex hulls are disjoint from the convex hulls of their opposition. We'll call them **external** coalitions. They are more extremist coalitions which would choose policies quite unfavorable to their nonmembers. The intuitive idea behind the instability of these coalitions is that they give their opposition so little that it is easy for a member to threaten to defect and offer the opposition more.

The prediction by the Aumann-Maschler Bargaining Set Theory is that one of the five internal coalitions (the theory doesn't say *which* one) should form, and pursue a policy represented by a stable policy point. We have seen that such a policy would always be "centrist." Do you think this always happens? Many political scientists are pleased with the kind of reasoning in the Aumann-Maschler argument. Its idea of objections being met by counter-objections seems to capture the give and take of hard bargaining, threats and counterthreats in the political arena. Nevertheless, they are uneasy with the predicted results. In the real world, external coalitions pursuing off-center policies do occur.

In 1978 political scientists Richard McKelvey, Peter Ordeshook and Mark Winer proposed an alternative solution for the spatial coalition problem which they called the **competitive solution** [**20**]. They try to capture the bargaining process in a different way. Instead of members within a coalition trying to get what is fair to them by threatening to go off and form other coalitions, McKelvey, Ordeshook and Winer think of alternative coalitions which include some common members as bidding for the support of those members. They bid by making policy proposals which are as attractive as possible to those pivotal members. A coalition can only form if it can compete successfully.

In FIGURE 4 (which is borrowed from [**20**]) the five points α, β, γ, δ and ε on the central pentagon are determined by five equidistance relationships, shown by the circular arcs. For instance, α and δ are equidistant from G. Consider potential coalition *FGK* offering δ and potential coalition *GHJ* offering α. These two coalitions are competing for G, and their proposals are equally attractive to G. Their offers to G are **balanced**. Now consider *FGK* offering δ and *FGH* offering ε. These two coalitions are competing for both F and G. The offer δ of *FGK* is more attractive to F, but the offer ε of *FGH* is more attractive to G. These offers have a kind of balance too, since neither one is more attractive to *both* parties for which the coalitions are competing. In fact, if we consider the collection of all five external coalitions *GHJ*, *HJK*, *FJK*, *FGK* and *FGH* with respective offers α, β, γ, δ and ε, all of these are balanced against each

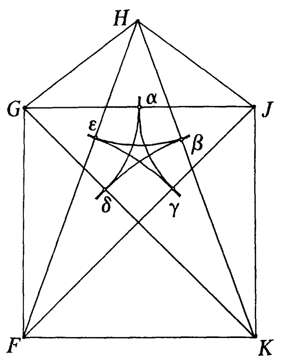


FIGURE 4. **The competitive solution.**

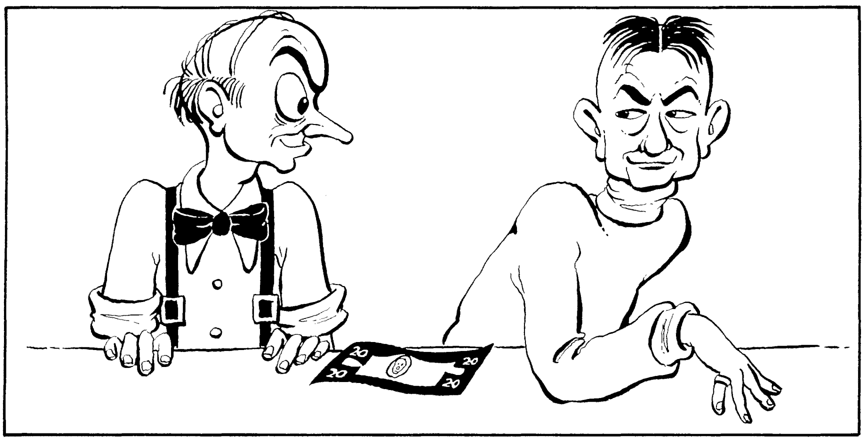
other. McKelvey, Ordeshook and Winer call a collection of coalitions and offers like this **competitively balanced**.

The five internal coalitions *FHK*, *GJK*, *FHJ*, *GHK* and *FGJ* are not represented here. Consider one of them, say *FHK*. Can *FHK* offer a proposal which would compete against all the proposals in our set of five? No, it can't. If *FHK* is to compete with *GHJ* for *H*, it must make a proposal which is at least as attractive to *H* as α. But then its offer will be less attractive than γ to both *F* and *K*, so it will lose *F* and *K* to *FJK*. The collection of five competitively balanced offers by the external coalitions cannot be invaded by other coalitional offers. Such a balanced, uninvadable collection is called a **competitive solution**. It is fairly easy to check that the five internal coalitions are not part of any competitive solution: any balanced set of proposals they might make can be disrupted by a more competitive proposal from an external coalition. McKelvey, Ordeshook and Winer predict that the coalition which forms will be one which is contained in a competitive solution. Here that means an external coalition.

This picture of coalitions trying to make offers which will hold their pivotal members is more complicated to think about than the Aumann-Maschler bargaining procedure. The problem is that a coalition is not a single entity (after all, it hasn't formed yet) but a collection of individual parties, all of whom are simultaneously offerers trying to hold pivotal members, and pivotal members seeking offers themselves. And they are doing this in many coalitions. After a while, your head begins to hurt. However, there is no question that the theory offers a prediction which is testably different from the Aumann-Maschler Bargaining Set. In the pentagon example it predicts that an external coalition will form with a policy proposal on the boundary of the central pentagon, while the Bargaining Set predicts an internal coalition with a policy point inside the central pentagon.

McKelvey, Ordeshook and Winer tested their theory by running a series of eight experimental five-person pentagon bargaining games. The subjects played for money and could win up to about \$20, so they were well motivated to bargain. The results were clearcut. In all eight cases an external coalition formed, none of the policy points were inside the central pentagon, and six of the eight points were close to the competitive solution prediction [**20**]. The bargaining procedure of Aumann and Maschler is a lovely abstraction of considerable intuitive appeal, but it just isn't what people seem to do in this kind of situation. It doesn't seem to work very well for parliamentary coalitions either [**26**]. Sometimes even the nicest model has to be abandoned when reality won't cooperate.

By the way, in our original Norwegian example, the bargaining set predicts *AB*, *AC*, *AD* or *AE*, while the competitive solution predicts *AB*, *AC* or *BCDE*.

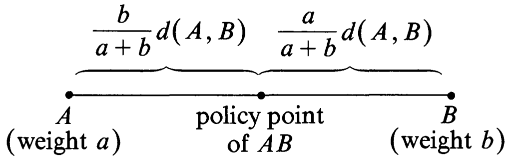


"The subjects played for money and could win up to about \$20, so they were well motivated to bargain."

A dynamic model of spatial coalition formation

The complexity of the game-theoretic solutions for spatial coalition formation comes from the fact that they try to model a negotiation process in which the participants keep in mind a number of possible final results and play them off against each other. What would happen if we didn't assume such farsightedness in the parties? Remember our original assumption that each party wants to join with other parties of similar values. One short-sighted way to do this would be step by step. Go to a nearby party and offer to form a coalition with it. If it accepts, form the coalition and go together with an offer to another nearby party or coalition. Continue until your coalition is winning, or until other parties following the same strategy beat you to it.

To model this idea, we start with our parties as points in some n-dimensional Euclidean space, each party weighted by its number of seats. Consider two parties, A and B, with weights a and b, respectively, thinking about whether to form a coalition. If they do, the coalition will have to adopt a policy point, probably somewhere between A and B. We'll adopt the sociologists' "resource theory" (see the first section) and assume that the most natural policy point would be the weighted average $(aA + bB)/(a+b)$:



A would want to join with B if this weighted average policy is close to A. In fact, A would look at all possible partnerships with others and make an offer to the party whose weighted average with A would be *closest to* A. If each party does this, each party will make an offer to some other party, its preferred coalition partner. We can represent such offers as a directed graph, the **preference digraph**, in which each vertex (party) has a directed edge leading from it to its preferred coalition partner.

We will assume that two parties form a coalition if and only if their offers are reciprocal—each is the other's preferred partner. If that happens, the two original parties, A and B, say, are replaced by their coalition AB, a new actor with weight $(a+b)$ at point $(aA + bB)/(a+b)$. In this first stage, several pairs of parties may form coalitions. If none of these coalitions is of winning size, the process continues to a second stage with the new actors. As soon as a winning coalition forms, the process terminates.

Let's try this process out on the Norwegian example of FIGURE 1. First we compute A's preferred partner. The distance from A to B is 11.40. Since A's weight is 68 and B's weight is 13, the distance from A to the weighted average of A and B would be $(13/(68+13))(11.40) = 1.83$. Other distances are calculated similarly:

	B	C	D	E
Distance from A	11.40	11.18	12.37	16.03
Weighted distance from A	1.83	2.34	2.59	5.02

Since the smallest of these weighted distances from A is to B, A prefers to join with B. The complete preference digraph for stage one is shown in FIGURE 5. We invite you to verify it. Since B and D prefer each other, they join in a coalition of weight 31 at point $(13/31)(4,10) + (18/31)(7,6) = (5.74, 7.68)$. Since no coalition is yet winning, we proceed to stage two, with a new preference digraph shown in FIGURE 5. Now C and E join to form a coalition of weight 49 at $(9.16, 1.63)$. From the stage three digraph of FIGURE 5 we conclude that the winning coalition $BCDE$ will form.

In studying this model, we must first show that it always produces an answer. We must show that at every stage there is at least one pair of coalitions which prefer each other as partners, for if there were not, the process would grind to a halt. What we need to show is that any digraph

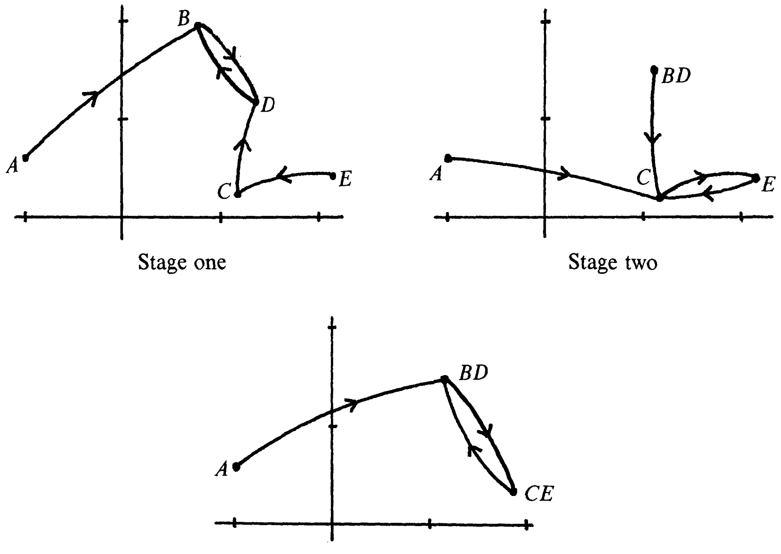


FIGURE 5. **Preference digraphs for the dynamic model.**

arising in the process must have a *cycle of length two*. A nice argument for this is in two steps.

i) *The digraph must have a cycle*. Start at any vertex. Each vertex has a directed edge leading away from it. Follow these edges in sequence. Since the number of vertices is finite, we must eventually return to a vertex we have already visited, at which point we have traversed a cycle.

ii) *No cycle can have length greater than two*. Suppose $A_1, A_2, \dots, A_n$ $(n \geqslant 3)$ were such a cycle, where A_i has weight a_i.

Since A_1 prefers A_2 to A_n, $$\frac{a_2}{a_1 + a_2} d(A_1, A_2) < \frac{a_n}{a_n + a_1} d(A_n, A_1).$$

Since A_2 prefers A_3 to A_1, $$\frac{a_3}{a_2 + a_3} d(A_2, A_3) < \frac{a_1}{a_1 + a_2} d(A_1, A_2).$$

$$\vdots$$

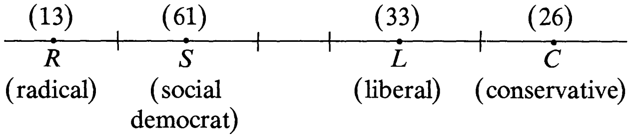
Since A_n prefers A_1 to A_{n-1}, $$\frac{a_1}{a_n + a_1} d(A_n, A_1) < \frac{a_{n-1}}{a_{n-1} + a_n} d(A_{n-1}, A_n).$$

But if we multiply all of these inequalities together, noting that all of the terms are positive, we find that the left and right sides contain exactly the same terms, which is a contradiction.

The argument is worth contemplating. In particular, you should be sure to see how ii) depends on the condition that $n \geqslant 3$.

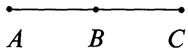
In the presentation so far, we haven't dealt with the case where some party might have a tie for its preferred coalition partner. The preference digraph would then have two directed edges from that vertex. You might like to check that the presence of ties introduces some indeterminancy into the sequence of coalition formation, but produces no serious problems. In particular, there is still always a cycle of length two at each stage.

Recall from the first section that there is some empirical evidence that governing parliamentary coalitions will be i) minimal winning and ii) connected. We should investigate how the dynamic model performs with respect to these properties. For the first, consider the following one-dimensional example, which is abstracted from the April 1953 elections in Denmark. To win, 76 votes are needed (this is larger than a majority because we have omitted some minor parties).



The dynamic process gives coalitions *RS* and *LC* in stage one. Neither wins, and in the second stage we get the grand coalition *RSLC*, which is clearly not minimal winning. The dynamic process can produce nonminimal winning coalitions. Before we get too discouraged, we might consider what happened in Denmark in 1953, which was that *no* governing coalition formed and a minority caretaker government ruled until new elections were held in September. Grofman [**15**] has proposed that if the dynamic process produces a nonminimal winning coalition, it may be a sign that it will be difficult for a governing coalition to form, or be stable if it does form. The idea is that in this case the strategic desire to expel extraneous members conflicts with primary ideological alliances.

As for connected coalitions, the dynamic process will always produce connected coalitions in a *one-dimensional* policy space. To see why, consider



and note that it will never be possible for A to prefer C *and* C to prefer A. For if the weighted average of A and C is at or to the left of B, then C will prefer B to A (since the weighted average of B and C is to the right of B). Similarly, if the weighted average of A and C is to the right of B, then A will prefer B to C. But these cases exhaust the possibilities. Hence a disconnected coalition will never form at any stage of the dynamic process.

We thought for a while that this result might be true in higher dimensions as well, but it is not. A counterexample in two dimensions is shown in FIGURE 6. In stage one, with preference digraph as shown, *EF* forms with weight 10 at (10,10), and *AB* forms with weight 10 at (2.8,0). In stage two *ABC* forms at $(-3.26,0)$. In stage three the winning coalition *DEF* forms at (10,0). The coalition *DEF* is not connected, since it doesn't include party A, whose policy point is in the convex hull of its members. The situation of poor A in this example is worth thinking about. Because he shortsightedly let himself be enticed away first by B and then by C, he is left out of a winning coalition which forms precisely at his preferred policy point.

We do not view the possibility of disconnected coalitions in two or higher dimensions as a defect of the dynamic model. First, the care it took to construct the counterexample hints that it is

Party	Weight	Location
A	1	(10, 0)
B	9	(2, 0)
C	9	$(-10, 0)$
D	10	$(10, -10)$
E	5	(9, 10)
F	5	(11, 10)
	39	

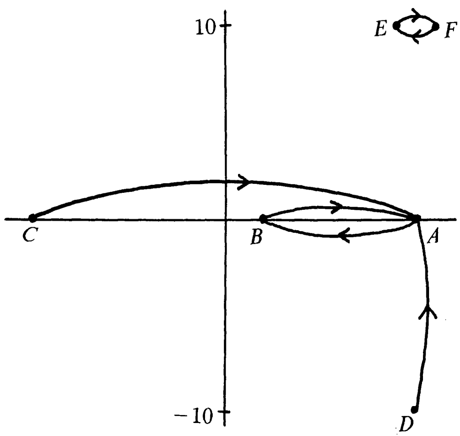


FIGURE 6. **The dynamic process forms a disconnected winning coalition.**

not easy for disconnected coalitions to arise by the dynamic process. Second, the empirical support for connected coalitions, as cited by Axelrod, for instance, comes from one-dimensional models.

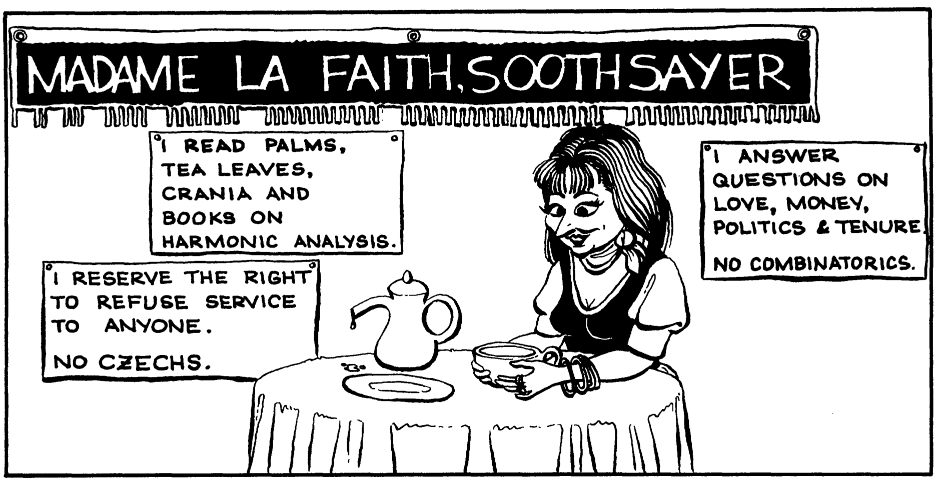
The dynamic model has not yet been tested against the full range of twentieth-century data on parliamentary coalitions, but Grofman [14] has gathered some preliminary evidence that, for all its simplicity and shortsightedness, it predicts well the formation of governing coalitions in Norway, Denmark and Germany. For example, here are the predictions of the models we have considered for the Norwegian situation of FIGURE 1:

Theory	Prediction
Minimal winning	$AB, AC, AD, AE, BCDE$
Least resources	$BCDE$
Fewest actor	AB, AC, AD, AE
Bargaining set	AB, AC, AD, AE
Competitive solution	$AB, AC, BCDE$
Dynamic model	$BCDE$

The coalition which actually formed was $BCDE$. In Denmark in 1973 the dynamic model predicts the six-party, nonminimal winning coalition which actually formed and which is predicted by no other model. Grofman [15] reviews evidence that the dynamic model makes 16 correct predictions out of 18 predictions for Denmark in the period 1913–1973. This kind of performance is especially impressive since the dynamic model makes a *unique* prediction, whereas except for the least resources theory, the other theories generally do not. If you would enjoy running further tests, there is plenty of data in [5], [10], [11], [13] and [17].

Comparing and testing models

We have considered a number of models of coalition formation, some naive, some quite sophisticated. Some predictions are weak, others are very precise. How do we go about deciding which model is the best one? We might gather all the available data, compute the predictions of all the models, and choose the model with the highest percentage of correct predictions. Of course we would have to figure out a way to compensate for the varying degrees of precision with which the models make their predictions (see [27] for a statistical technique to do this). However, in the domain of social science, things are not generally that easy. There are at least two factors which make the decision among competing models a much more subtle process.



"We... choose the model with the highest percentage of correct predictions."

The first factor involves the nature of data in social science. It is hardly ever just *there*. It has to be carefully gathered, organized and put in suitable form before it can be used to test models. In spatial coalition theory we see this problem in a particularly basic form. Recall that we have assumed as given that parties are located as points in an n-dimensional ideological space. In the literature, the methods used to get such a placement have varied from individual "expert judgment" through quantitative analysis based on public perceptions or previous appearances in coalitions. At the very least, the inexactness of placement methods should caution us against claiming precision in prediction. There are a number of additional problems, for instance

i) there may be, and are, inconsistencies in the spatial placements obtained by different analysts using the same or different methods;
ii) although Euclidean distance has usually been used as the measure of ideological proximity, other metrics (for instance the "taxicab metric" $d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$) might have an equal or better claim to validity;
iii) there may be some logical circularity in deriving likely coalitions from spatial placements, since parties might be perceived as similar *because* they have often been in coalitions together.

With enough caution, sensitivity analysis and modesty for claimed results, we think these problems can be controlled, but they illustrate the subtle relationship between data and models in the social sciences.

The second factor complicating the process of deciding among coalition models is that we are trying to understand a process which is so complicated that it probably does not make sense to ask for a unique *best* model. Some of the models have parties concerned mainly with power, some mainly with ideology. The game theory models posit careful rational consideration of alternatives; the dynamic model relies on more shortsighted behavior. We will likely find that different models work best in different situations. For instance, the political cultures of various countries may result in parties balancing differently concerns for power and ideology. Some systems may be conducive to hard rational bargaining of the kind that McKelvey, Ordeshook and Winer's experimental subjects were motivated and able to do, and the competitive solution may be applicable. Other systems may make flexible offers and counteroffers harder to make, and the step-by-step process of the dynamic model will be more suitable. In fact, if different models predict well for different cultures, that could indicate interesting things about differing motivations and concepts of bargaining. This kind of comparative richness would be lost if we insisted on one best model.

In general, the role of mathematical models in the social sciences is more flexible than in the physical sciences. We rarely expect that a model will embody the "correct" understanding of a phenomenon, only that it will give insight, lead to new and interesting questions and directions for further study. Models are not an end of thought, but an aid to thought. If you would like to look at other ways in which mathematical models are being used in political science, we would recommend **[4]**, **[21]**, **[24]** and **[1]**.

Coda: Applying the dynamic model outside of politics

We would like to close by illustrating one way in which the dynamic model of coalition formation might be useful outside of political science. Consider the geographical problem of the formation of a trading network among spatially separated communities. We think of the communities starting off isolated. Each community decides on a preferred trading partner, which in the simplest case might be just the *closest* other community. Two communities form a **trade link** between them if and only if they are each other's preferred trading partner. Once a trade link is established, communities connected by it are considered together as a **trading coalition**, and the

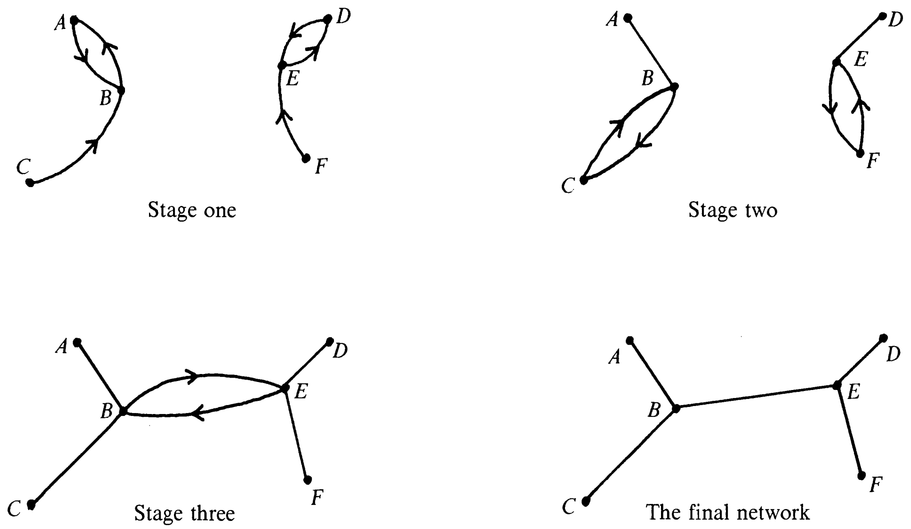


FIGURE 7. **The dynamic model of the formation of a trading network.**

distance between two trading coalitions is simply the distance between their closest members. The process continues until all of the communities are connected by a trading network. FIGURE 7 shows an example of the process. Notice the changes we have made from the political model. First, the communities don't amalgamate at their center of gravity, but stay fixed and are joined by a trade link. Second, the distance between coalitions is measured from the closest members, and that is where the trade link is drawn. For this modification, the proof that at least one trade link must form at each stage is trivial: the two coalitions which are closest together must prefer each other.

In this model, two things might be of interest to geographers—the sequence in which trade routes are built up over time (could we observe it?) and the final trading network. For the latter, a little thought should convince you of the following:

i) the final network will always be a **tree** (a connected graph with no cycles) which spans the communities, and
ii) this tree is the same tree as would be obtained if you started by drawing the shortest possible edge, and at each step added the next shortest edge which would not complete a cycle. (The order in which the edges come in may be different, but the final result will be the same.)

Now it is well known that the algorithm in ii) generates the **minimal spanning tree** of the communities, i.e., the tree which has the smallest total length among all possible spanning trees (see [7] or [**22**]). We thus have the result that communities which shortsightedly build up a trading network by this dynamic process will efficiently solve the global problem of joining themselves by a minimal length trading network.

Grofman and Landa [**16**] have used this model to investigate the development of the "Kula ring" trading network in the Melanesian Islands, which was first described by Malinowski in 1922.

One useful modification of the model might be to weight communities by their value as trading partners if distance were not a consideration. We might, for instance, weight them by the wealth of their economies. A community A might then judge the attractiveness of another community B by $b/d(A,B)$, where b is the economic weight of B. A trading link would be established if two communities were most attractive to each other, and the weight of a coalition would be the sum of the weights of its members. For this modification, we would once again have to prove that at each stage there is a pair of mutually most attractive coalitions. Fortunately, the argument in the third section goes through *verbatim*, except that the inequalities which must be multipled together now have the form

$$\frac{a_{i+1}}{d(A_i, A_{i+1})} > \frac{a_{i-1}}{d(A_{i-1}, A_i)}.$$

In fact, the argument will work when the attractiveness of A_j for A_i is measured by a wide variety of functions of a_i, a_j and $d(A_i, A_j)$. Modifications of the dynamic model of coalition formation might apply in many cases where political, economic or social agents join with other agents whose attractiveness as partners depends on their "weight" and their "location."

References

[1] Peter Aranson, American Government: Strategy and Choice, Winthrop Publishers, Cambridge, 1981.

[2] Robert Aumann and Michael Maschler, The bargaining set for cooperative games, in Dresher, Shapley and Tucker, eds., Advances in Game Theory, Princeton U. Press, 1964, pp. 443–476.

[3] Robert Axelrod, Conflict of Interest, Markham Publishing Co., Chicago, 1970.

[4] S. Brams, W. Lucas and P. Straffin, eds., Political and Related Models, Springer-Verlag, New York, 1982.

[5] E. Browne and J. Dreijmanis, eds., Government Coalitions in Western Democracies, Longman, New York, 1982.

[6] T. Caplow, Further developments of a theory of coalitions in the triad, Amer. J. of Sociology, 66 (1959) 488–493.

[7] Gary Chartrand, Graphs as Mathematical Models, Prindle, Weber and Schmidt, Boston, 1977, pp. 83–85.

[8] P. Converse and H. Valen, Dimensions of change and perceived party distance in Norwegian voting, Scandinavian Political Studies, 6 (1971) 107–151.

[9] Morton Davis, Game Theory: A Non-technical Introduction, Basic Books, New York, 1970, chapter 6.

[10] A. DeSwann, Coalition Theories and Cabinet Formation, Elsevier, Amsterdam, 1973.

[11] L. Dodd, Coalitions in Parliamentary Government, Princeton U. Press, 1976.

[12] W. Gamson, A theory of coalition formation, Amer. Sociological Review, 26 (1961) 373–382.

[13] S. Groennings, E. Kelley and M. Leiserson, eds., The Study of Coalition Behavior, Holt, Rinehart and Winston, New York, 1970.

[14] Bernard Grofman, A dynamic model of protocoalition formation in ideological n-space, Behavioral Science, 27 (1982) 77–90.

[15] ______, Inferring party distances from cabinet coalition patterns, School of Social Sciences, University of California, Irvine, 1982.

[16] B. Grofman and J. Landa, The development of trading networks among spatially separated traders as a process of protocoalition formation: the Kula trade, Social Networks, 5 (1983) 347–365.

[17] M. Holler, ed., Coalition Theory, Physica-Verlag, Wurzburg, 1983.

[18] M. Leiserson, Power and ideology in coalition behavior, in reference [13], pp. 323–335.

[19] G. Luebbert, Coalition theory and government formation in multiparty democracies, Comparative Politics (January 1983) 235–249.

[20] R. McKelvey, P. Ordeshook and M. Winer, The competitive solution for n-person games without transferable utility, with an application to committee games, Amer. Political Sci. Rev., 72 (1978) 599–615.

[21] P. Ordeshook, ed., Game Theory and Political Science, NYU Press, 1978.

[22] Oystein Ore, Graphs and Their Uses, New Mathematical Library #10, MAA, 1963.

[23] William Riker, The Theory of Political Coalitions, Yale Univ. Press, 1962.

[24] W. Riker and P. Ordeshook, An Introduction to Positive Political Theory, Prentice-Hall, Englewood Cliffs, 1973.

[25] F. S. Roberts, Discrete Mathematical Models, Prentice-Hall, Englewood Cliffs, 1976.

[26] N. Schofield, The kernel and payoffs in European government coalitions, Public Choice, 26 (1976) 29–49.

[27] M. Winer, Cabinet coalition formation: a game theoretic analysis, in Brams, Schotter and Schwodiauer, eds., Applied Game Theory, Physica-Verlag, Vienna, 1979, pp. 133–151.

Polymorphic Polyominoes

ANNE FONTAINE
The College of Saint Rose
Albany NY 12203

GEORGE E. MARTIN
State University of New York at Albany
Albany NY 12222

Some polygons tile the plane. For such a polygon this means that the plane can be covered with congruent copies of the polygon in such a way that no two copies have overlapping interiors. Ordinary graph paper suggests that a square tiles the plane. Once it has been determined that a given polygon does tile the plane, we can ask the question: *How many different (noncongruent) ways are there to tile the plane with this polygon*? A polygon admitting one and only one tiling is said to be **monomorphic**. A regular hexagon jumps to mind; once one hexagon is placed in the plane, every position of the congruent copies required for a tiling is determined. Actually, there are many monomorphic polygons. Squares and triangles are at the other end of the spectrum however. With either of these, we can form an uncountable infinity of tilings by sliding tiled parallel strips back and forth. In this Note, we are interested only in polygons that tile the plane in finitely many ways.

The problem of finding tiles that admit exactly r tilings of the plane for a given value of r was posed in 1977 by Grünbaum and Shephard in [7]. In that paper, they exhibited **dimorphic** tiles, that is, polygons admitting exactly two tilings up to congruence. A footnote added in proof announced that they had also found a polygon that tiles in exactly three ways. This **trimorphic** tile is presented in [**11**], where Harborth gives a pair of polygons that together tile the plane in exactly r different ways. For a given r, by cutting one of the pair into pieces such that the only way the pieces can be reassembled is by reconstructing the original polygon, it is then easy to obtain a set of k polygons that together tile the plane in exactly r ways when $k > 1$. These results then focused attention on the problem of tiling with copies of a single polygon. We define a polygon to be **r-morphic** if congruent copies of the polygon tile the plane in exactly r different ways. (See [**8**], [**9**], or [**10**] for an introduction to tiling.)

In 1981, Grünbaum and Shephard reproduced their polymorphic tiles in [**9**] and posed the problem of finding additional, essentially different dimorphic and trimorphic tiles. They also asked whether r-morphic tiles exist for values of r with $r > 3$. We have responded to these problems in [**1**], [**2**], and [**3**]. In particular, we have shown that there exist r-morphic polygons for $r < 11$. The existence of r-morphic polygons for $r > 10$ remains open. We indicate below how we constructed the tiles announced in the cited papers but, with one exception, present entirely different polygons that prove these results. The exception is the decamorphic tile given later, which is the only known polygon up to similarity that tiles the plane in exactly ten ways.

A **polyomino** is a set of rookwise connected unit squares, a generalization of the case where two squares are connected edge-to-edge to form a domino. Golomb introduced polyominoes in 1954 in [**4**]. Although he discusses tiling the plane with polyominoes in [**6**], it is from his book [**5**], a classic in recreational mathematics, that most readers will enjoy learning about polyomino puzzles. More by accident than by design, all examples of polymorphic polygons given below are polyominoes.

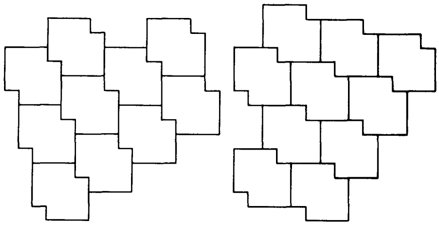


FIGURE 1

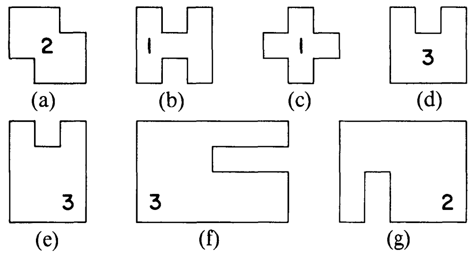


FIGURE 2

We are going to state many results but prove only one. The reason for this is that the proofs depend simply on exhausting all the possible combinations. You are invited, encouraged, even urged to verify some of the stated results for yourself. Although some take only a few minutes or less, others will exhaust most readers before the reader has exhausted all the possibilities. Some results take the authors over six hours to check carefully just once; it is by making many such checks that we are convinced of our results. If you like jigsaw puzzles, then this is for you. Cut out many copies of a tile and go at it for yourself. Perhaps only one more warning is not out of place: playing with polymorphic polyominoes can be habit-forming.

The Z's

One infinite family of monomorphic polygons is constructed as follows. Start with a square polyomino with sides of length n where $n > 3$ and at each of two diagonally opposite corners cut out a unit square. The resulting figure tiles the plane in the two ways illustrated in FIGURE 1. However, these two tilings are congruent; in fact, the patches shown in the figure are themselves congruent. With copies of the tile in hand, it is very easy to see that the tile is indeed monomorphic. Actually, we can cut out a square with sides of length k from each of two diagonally opposite corners of a square with sides of length n in the construction above, provided that $2k < n$. With one exception, the tile formed is monomorphic. The interesting exception is where we start with a square having sides of length $3k$ and cut out squares with sides of length k. This tile is shown in FIGURE 2(a); as indicated there by the "2" and as we shall prove below, this tile is dimorphic.

FIGURE 2 shows some r-morphic polyominoes for $r = 1,2,3$. The trimorphic tile of FIGURE 2(d) can be augmented to form an infinite family of dimorphic polyominoes. These and a comparable family of trimorphic polyominoes are indicated in FIGURE 3. Two copies of any one of these tiles fit together to fill the slot in each in an obvious way and form what we shall call a Z. By $Z(a,b,c,d)$ we shall mean the polygon obtained by starting with an $a+c$ by $b+d$ rectangle and removing the $b \times c$ rectangle from each of two diagonally opposite corners such that the resulting polygon has consecutive sides of lengths a,b,c,d,a,b,c,d (see FIGURE 4). Polygon $Z(d,c,b,a)$ is congruent to $Z(a,b,c,d)$. In the construction we must have $a > c$ or $d > b$ in order to have a polygon left after removing the corner rectangles. For example, our first monomorphic Z mentioned above is $Z(a,1,1,a)$ with $a > 2$; FIGURE 1 shows the case $a = 3$. The

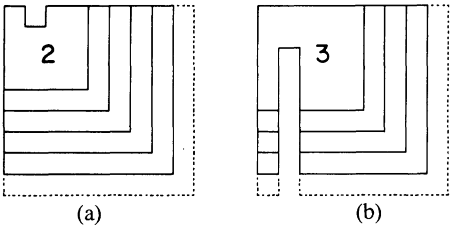


FIGURE 3

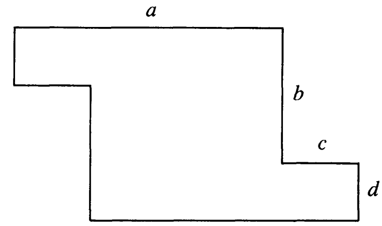


FIGURE 4

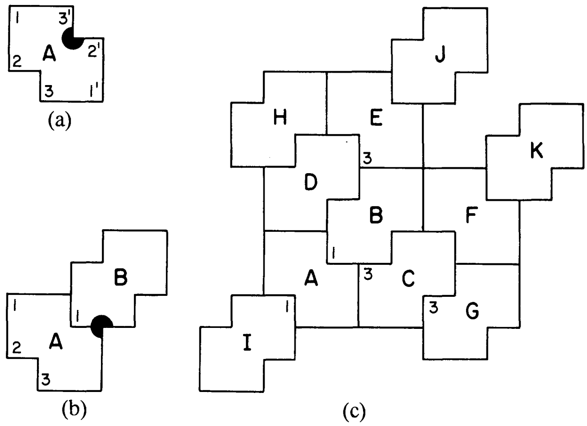


FIGURE 5

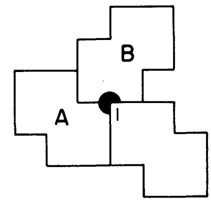


FIGURE 6

dimorphic Z of FIGURE 2(a) is $Z(2,1,1,2)$. Two copies of a dimorphic tile from FIGURE 3(a) fit together to form a dimorphic $Z(x, x-1, x-2, x)$ with $x > 3$, and two copies of a trimorphic tile from FIGURE 3(b) form a trimorphic $Z(x, 2, x-2, x)$ with $x > 4$.

To give some indication of the method used to prove our results, we choose $Z(2,1,1,2)$ and argue that this tile is dimorphic. We first label the projecting corners of a copy of the tile in question; see FIGURE 5(a). Since $Z(2,1,1,2)$ has a line of symmetry, turning over a copy produces another tile directly congruent to the original. Hence, we need not distinguish "front" and "back" here, as is necessary with most tiles. Further, since a copy of $Z(2,1,1,2)$ rotated 180 degrees about its center coincides with its original position, the projecting corners $1'$, $2'$, and $3'$ of copy A of the tile in FIGURE 5(a) need not be distinguished from those of 1, 2, and 3, respectively. So, of the twelve projecting corners of a copy of $Z(2,1,1,2)$ we need label only three. All copies are assumed to have the labels 1, 2, and 3, as copy A does. Next we select a slot (concave corner) of copy A, darkened in FIGURE 5(a), and begin placing in this slot the three labeled corners of a copy of $Z(2,1,1,2)$ in turn. Thus, we begin with the configuration of copies A and B as in FIGURE 5(b). In this configuration of two tiles, we select a slot, darkened in FIGURE 5(b), and begin placing in this slot the three labeled corners of a copy of $Z(2,1,1,2)$ in turn. Placing corner 1 in the slot leads to the contradiction of a "hole" that cannot be filled, as shown in FIGURE 6. Placing corner 2 in the slot leads to a forbidden overlapping of the tiles. Placing corner 3 in the slot is, however, possible. So the position of copy C is uniquely determined; see FIGURE 5(c). By symmetry, the position of copy D in the figure is also uniquely determined. Next, we select a slot in the configuration of copies A, B, C, and D and continue as above. Although the selecting of slots is somewhat arbitrary, practice and good luck distinguish those that most quickly produce results. We see corner 3 of copy E in FIGURE 5(c) is uniquely determined in its slot, since each of corners 1 and 2 in that slot leads to a contradiction. The position of copy F is uniquely determined by symmetry. Likewise, the positions of copies G and H are uniquely determined. Now, we consider filling the open slot in copy A with copy I. Corner 1 does fit. At this point, it is advantageous to determine that corner 1 is not only compatible but forced. In other words, instead of continuing on and then later backtracking to consider the compatibility of corners 2 and 3, we try them now and exclude those that lead to a contradiction. In this special example, both of these choices do lead to a contradiction and the position of copy I is thus uniquely determined. With two or more labeled corners compatible with selected slots at several stages, you can imagine the rapid growth of the number of possible branches, each of which must be followed to a contradiction, a tiling, or another branch. In our example, we now have observed the configuration of A and B determines the position of I. Hence, the configuration of E and D determines the position of J, and the configuration of F and C determines the position of K. We note that the initial configuration of A and B is now repeated in two independent directions with the determined configuration of E and J

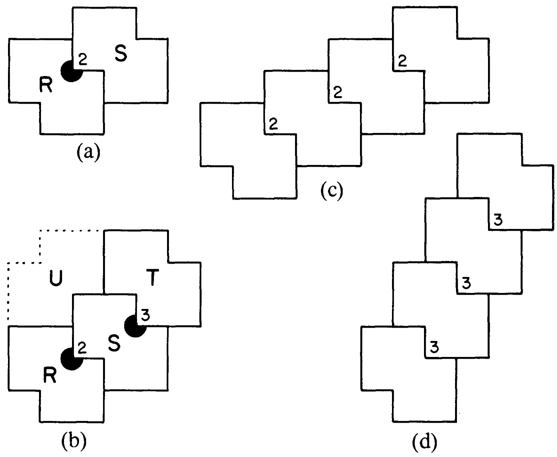
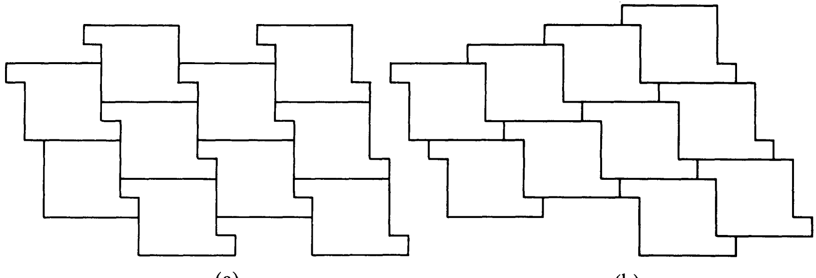


FIGURE 7

and the determined configuration of F and K. It follows easily that the configuration of A and B determines a tiling and, more importantly, that this tiling is unique. In general, the difficult task is to determine the exact number of tilings compatible with a given configuration rather than to observe only that the given configuration can be completed to a tiling. We are nearly done with our example. We can now eliminate the occurrence of the configuration of A and B in FIGURE 5(b) in all further tilings by $Z(2,1,1,2)$, since we know that this configuration leads to a unique tiling.

Now consider the configuration of R and S in FIGURE 7(a), which corresponds to putting corner 2 in the original slot of copy A. The addition of a third tile T to form the configuration in FIGURE 7(b) does not produce a new tiling because this forces the configuration of U and S, which is congruent to the configuration of A and B. Thus the configuration of R and S in FIGURE 7(a) can lead only to the infinite strip suggested by FIGURE 7(c). Likewise, the configuration of S and T, which corresponds to placing corner 3 in the original slot of copy A, can lead only to the infinite strip suggested by FIGURE 7(d). However, these two infinite strips are congruent and the copies fit together in only one way, forming the single tiling analogous to that illustrated in FIGURE 1. We conclude that $Z(2,1,1,2)$ tiles the plane in exactly two ways.

The Z's are the key to all the polymorphic tiles we shall present. Copies of $Z(a,b,c,d)$ always stack in the ways shown in FIGURE 8: first with sides of length a adjacent, as in FIGURE 8(a), and second, with sides of length d adjacent, as in FIGURE 8(b). Except when $a=d$ and $b=c$, these two ways of stacking the Z's will give noncongruent tilings. Of course, a Z may give rise to an infinite number of tilings. For example, $Z(x+y, 2x, 2y, x+y)$ always tiles the plane in infinitely many ways, and you might like to find out for yourself just why this is so. In any case, dimorphic Z's are easy to come by. Trimorphic Z's are less abundant. The five families of trimorphic Z's we have found are:



(a) FIGURE 8 (b)

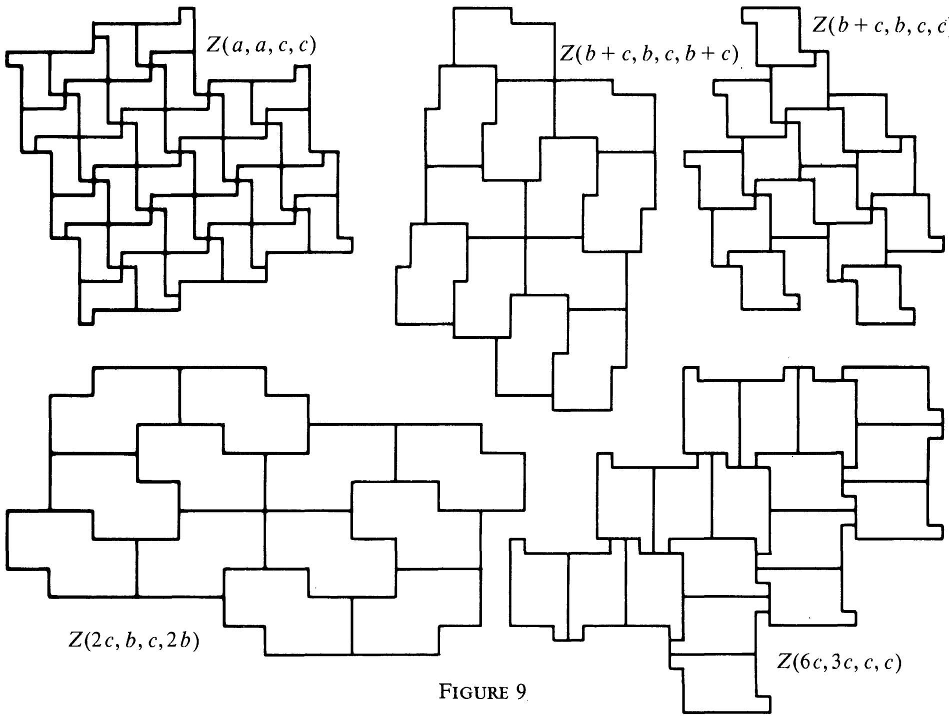


FIGURE 9

(1) $Z(a,a,c,c)$ with $a \neq c$, $a \neq 2c$, and $c \neq 2a$.

(2) $Z(b+c,b,c,b+c)$ with $b \neq c$.

(3) $Z(b+c,b,c,c)$ with $b \neq c$ and $b \neq 2c$.

(4) $Z(2c,b,c,2b)$ with $b \neq c$, $b \neq 2c$, $b \neq 3c$, $c \neq 2b$, and $c \neq 3b$.

(5) $Z(6c,3c,c,c)$.

An example for each of these families is illustrated in FIGURE 9, where the tiling given is the one not obtained by stacking the tiles in the two ways shown in FIGURE 8.

The nice thing about the Z's is that they have a point of symmetry; in other words, they are invariant under a rotation of 180 degrees. Thus we can easily bisect a Z into two congruent pieces, and the resulting tiles sometimes can be put together to form a different Z. This process is the key to the discovery of all of the polymorphic tiles we shall mention. Bisecting a Z of FIGURE 4 with a horizontal or vertical line through the center yields polygons that tile in infinitely many ways. Our most successful bisection gives the tile in FIGURE 10, where two copies form three different Z's. The tile in FIGURE 10 has consecutive sides of lengths 4,3,4,3,5,3,3,9 and up to

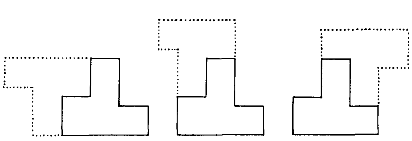


FIGURE 10

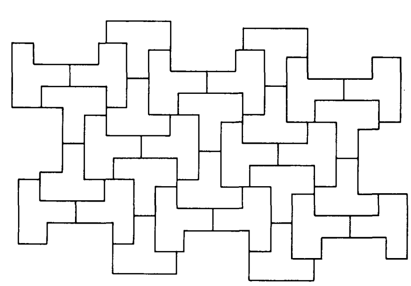


FIGURE 11

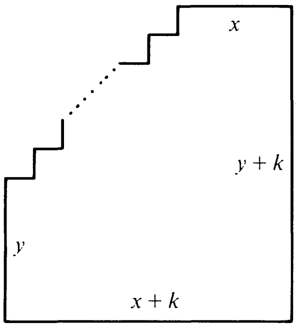


FIGURE 12

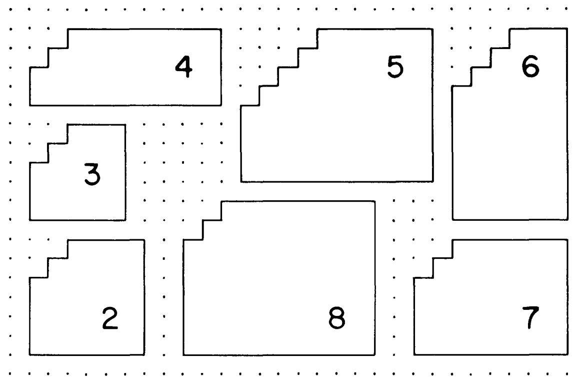


FIGURE 13

similarity is the only known decamorphic tile. That this smallest member of an infinite family of tiles that were designed to be heptamorphic actually turned out to be decamorphic is admittedly serendipitous. Nine of the ten tilings admitted by this decamorphic polyomino are formed by various combinations of the Z's of FIGURE 10. The tenth tiling, shown in FIGURE 11, contains no Z. Although quite similar to a common tiling, this intriguing tiling seems to be just a little off everywhere. The discovery of the others is left to the delight of the reader; you may check your results with those illustrated in [**2**].

Stairs

Tiles we call **stairs** and denote by $S(x, y, k)$ are formed by deleting k steps from a corner of an $x+k$ by $y+k$ rectangle. See FIGURE 12. Examples of polymorphic stairs are shown in FIGURE 13, while TABLE 1 gives an infinite family of r-morphic stairs for each value of r from 2 through 8. Of the polyominoes in FIGURE 13, only the trimorphic tile also appears in TABLE 1.

r-morphic stairs

r	x	y	k	
2	2	2	k	$(k>1)$
3	x	x	$x-1$	$(x>2)$
4	2	5	k	$(k>3)$
5	2	4	k	$(k>3)$
6	2	$k+2$	k	$(k>2)$
7	$2k+1$	$3k+2$	k	$(k>1)$
8	x	$2x$	x	$(x>2)$

TABLE 1

A few observations about how copies of a stairs naturally fit together will simplify the discovery of the various tilings determined by the tiles in FIGURE 13 and TABLE 1. Two copies of a stairs can be put together to form $Z(x+k, y, x-1, y+k)$ or to form $Z(x+k, y-1, x, y+k)$. The two Z's coincide when $x=y$. Turning over one of the two copies of a stairs in a Z and then placing the steps together gives a polygon that, for obvious reasons, we shall call a W. The W has consecutive sides of lengths $x, x-1, y+k, x+k, y, y-1, x+k, y+k$. The W duplicates the Z when $x=y$. However, not all the W's tile the plane. For example, the pentamorphic polyomino in FIGURE 13 yields a trimorphic Z, a dimorphic Z, and a W that does not tile. In contrast, a pentamorphic polyomino from TABLE 1 yields two dimorphic Z's and a monomorphic W.

With $r>5$, the r-morphic stairs from FIGURE 13 and those from TABLE 1 have the property that the W's do tile. In fact, four copies of a stairs can be put together to form a big Z in one of the two ways illustrated in FIGURE 14. Note each of these big Z's is made up of two W's. The W for the octamorphic stairs from FIGURE 13 tiles in exactly two ways, while a W for the

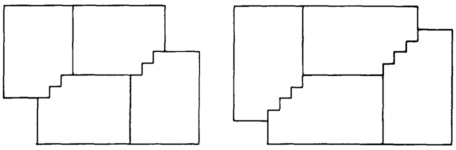


FIGURE 14

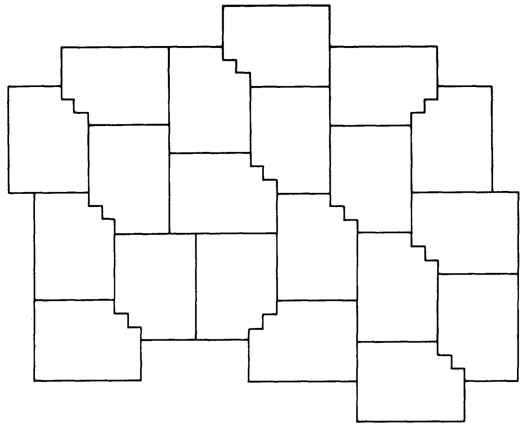


FIGURE 15

octamorphic stairs from TABLE 1 tiles in exactly three ways. However, perhaps the most interesting of the polyominoes from FIGURE 13 is the heptamorphic stairs. Here, in addition to the six tilings produced by the three dimorphic Z's, there is a seventh tiling that requires twenty copies of the stairs to form a patch that can be translated to cover the plane. Such a patch consisting only of Z's and W's is shown in FIGURE 15, and the tiling is obtained by repeating copies of this pattern under translation. You might try to pick out ten contiguous copies of the stairs from FIGURE 15 that would form the pattern for a linoleum block that theoretically could be used to print the tessellation design. Only ten copies of the stairs are required for the linoleum block because such a block can be rotated in making a print. Can you find a patch of twenty copies that has a point of symmetry and that gives the same tiling when translated?

We end our discussion of stairs with two problems. One is very, very easy and one is not. What can be said about the infinite family of **stairs with landing** illustrated in FIGURE 16(a)? What can be said about the infinite family of **stairs with stoop** illustrated in FIGURE 16(b)?

The L's

The last class of polyominoes we consider are the L's of FIGURE 17. Two copies of an L polyomino generally fit together to form Z's in two different ways. These ways coincide for the dimorphic and trimorphic L's. For the heptamorphic L both of these Z's are trimorphic. However, it is the enneamorphic L that is most interesting. For this 9-morphic polyomino, six copies can be put together to form a third Z. These three Z's are shown in FIGURE 18; each is dimorphic. The two tilings suggested by FIGURE 8 for each of these three Z's produce six tilings for the enneamorphic L. The three additional tilings are displayed in FIGURE 19. For each of the three tilings in the figure, it is rather fun to find a patch that can be translated to create the tessellation.

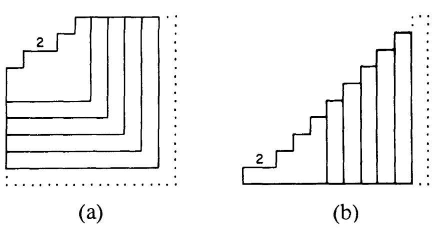


FIGURE 16

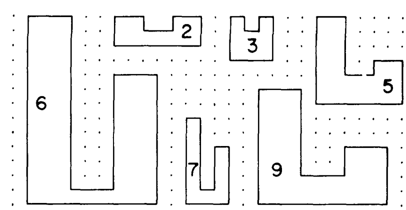


FIGURE 17

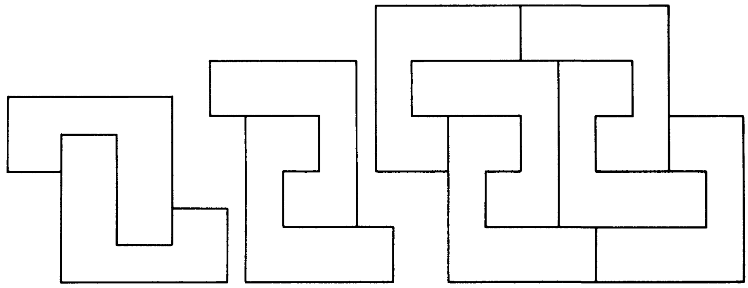


FIGURE 18

A careful examination of the Z's in FIGURE 18 and the tilings in FIGURE 19 shows that all of the L's, whether reversed or not, have parallel axes in any one of the nine tilings by the enneamorphic polyomino. Therefore, stretching the enneamorphic L along its axis should give rise to another enneamorphic polygon. Indeed, the stretch does give a polygon that tiles the plane in at least nine ways. The problem is with the "at least." We generalize as follows. The particular L defined by FIGURE 20 is denoted by L_{ij}. The 9-morphic polyomino from FIGURES 17, 18, and 19 is $L_{2,3}$. Copies of L_{ij} always form the three Z's of FIGURE 18, namely $Z(2i, j, 3i, 3j)$, $Z(4i, 3j, i, j)$, and $Z(8i, 2j, 2i, 3j)$, and always form the three tilings shown in FIGURE 19. You may check this last statement for yourself by applying the Conway criterion of **[12]**. Hence, L_{ij} always tiles the plane in at least nine ways. Both $L_{3,2}$ and $L_{3,4}$ are also enneamorphic polyominoes, while $L_{4,3}$ tiles the plane in infinitely many ways. Although $L_{2,3}$ is 9-morphic, an $L_{2k,3}$ tiles in infinitely many ways for $k=2$, for $k=3$, and for $k=6$. An L_{ij} will tile in infinitely many ways if i/j is any one of $1/8, 1/5, 1/4, 1/3, 1/2, 3/5, 1, 4/3, 2, 5/2, 3$, and 4. Beyond this we know little, although we claim $L_{k,1}$ is 9-morphic for $k>4$.

Questions to explore

Several problems have already been suggested. In another direction, you may wish to search for polymorphic polygons among the polyiamonds. A **polyiamond** is formed by connecting equilateral triangles edge-to-edge, analogous to the forming of polyominoes. Just as a polyomino is a generalization of a domino, so a polyiamond is a generalization of a diamond. Perhaps among the polyominoes or the polyiamonds the answers to the following questions can be found: Are there other decamorphic polygons? Are there r-morphic polygons for $r>10$? Is there a polygon that tiles in only a countably infinite number of ways?

References

[1] A. Fontaine and G. E. Martin, Tetramorphic and pentamorphic prototiles, J. Combin. Theory Ser. A, 34 (1983) 115–118.

[2] ______, Polymorphic prototiles, J. Combin. Theory Ser. A, 34 (1983) 119–121.

[3] ______, An enneamorphic prototile, J. Combin. Theory Ser. A, to appear.

[4] S. W. Golomb, Checkerboards and polyominoes, Amer. Math. Monthly, 61 (1954) 675–682.

[5] ______, Polyominoes, Scribner's, New York, 1965.

[6] ______, Tiling with polyominoes, J. Combin. Theory, 1 (1966) 280–296.

[7] B. Grünbaum and G. C. Shephard, Patch-determined tilings, Math. Gaz., 61 (1977) 31–38.

[8] ______, Tilings by regular polygons, this MAGAZINE, 50 (1977) 227–247 and 51 (1978) 205–206.

[9] ______, Some problems on plane tilings, The Mathematical Gardner (D. A. Klarner, ed.), Wadsworth, Belmont, 1981, pp. 167–196.

[10] ______, Tilings and Patterns, Freeman, New York, to appear.

[11] H. Harborth, Prescribed numbers of tiles and tilings, Math. Gaz., 61 (1977) 296–299.

[12] D. Schattschneider, Will it tile? Try the Conway criterion!, this MAGAZINE, 53 (1980) 224–233.

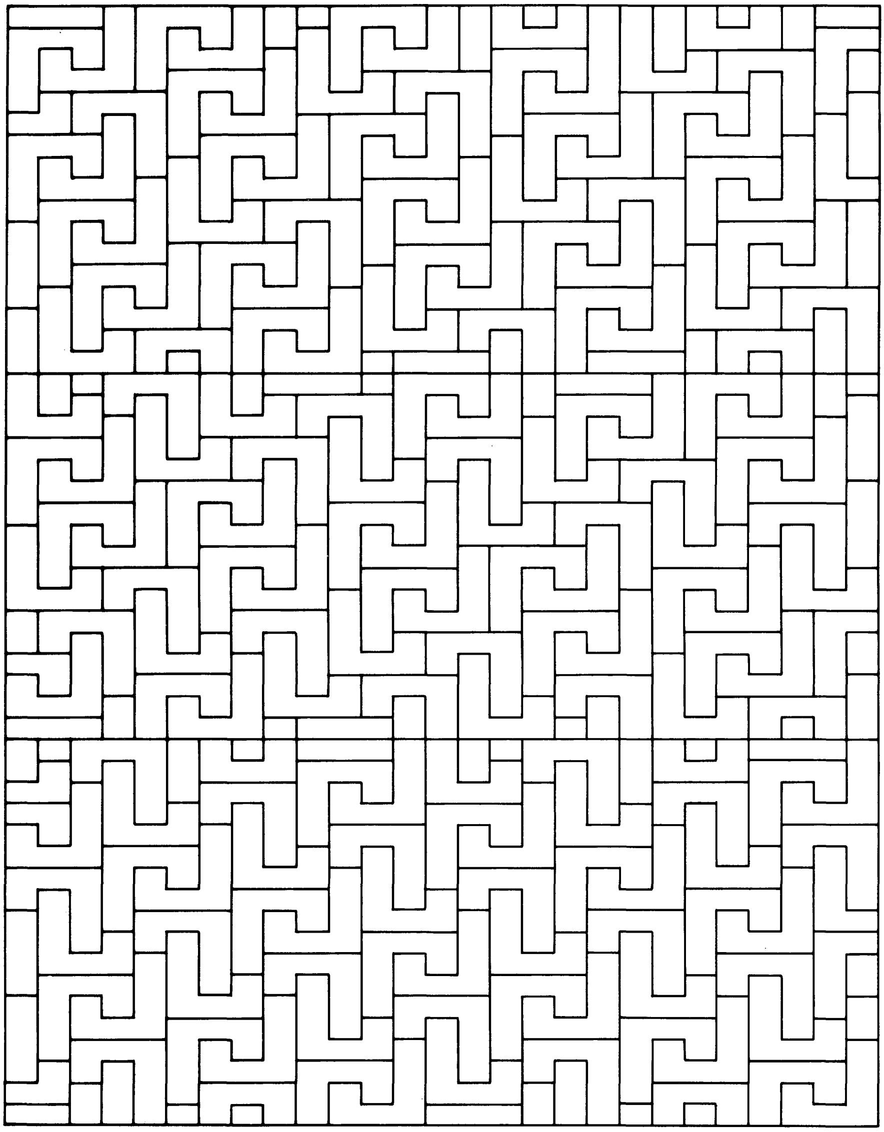


FIGURE 19

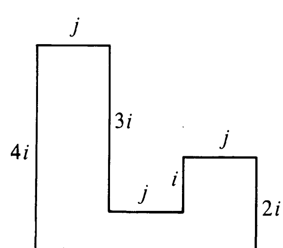


FIGURE 20

Optimal Quadrature Points for Approximating Integrals When Function Values are Observed with Error

DARRYL KATZ
California State University
Fullerton, CA 92634

A discussion of numerical integration in an introductory calculus course typically includes discussion of the trapezoidal rule and Simpson's rule. Attention is usually restricted to the case where the interval of integration is subdivided equally; this leads to simpler formulas but may also lead to very poor approximations to the definite integral when compared with other subdivision strategies.

Consider approximation of the definite integral

$$IEXACT = \int_{x_0}^{x_{m+1}} f(x)\, dx \tag{1}$$

by the trapezoidal rule

$$TR = 0.5\left[f(x_0)(x_1 - x_0) + \sum_{i=1}^{m} f(x_i)(x_{i+1} - x_{i-1}) + f(x_{m+1})(x_{m+1} - x_m)\right], \tag{2}$$

where $x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots \leqslant x_m \leqslant x_{m+1}$. This general form of the trapezoidal rule does not assume that the x_i are equally spaced, and is obtained by summing the terms $0.5(f(x_{i-1}) + f(x_i))(x_i - x_{i-1})$, $1 \leqslant i \leqslant m+1$; these are the areas of the approximating trapezoids. Equation (2) reduces to the familiar form of the trapezoidal rule when equal subdivisions are used. The example in FIGURE 1 shows that equal spacing of $x_0, x_1, \ldots, x_{m+1}$ can lead to very inaccurate estimates of the definite integral in (1), whereas a different choice of the same number of subdivision points x_i can give a much better estimate. In fact, FIGURE 1 suggests the adoption of a strategy for choosing $x_1, \ldots, x_m$ that results in larger subintervals where $f(x)$ is nearly linear (i.e., where $|f''(x)|$ is small) and smaller ones where $f(x)$ is "wiggly" (i.e., has large $|f''(x)|$).

The absolute error in approximating

$$\int_{x_{i-1}}^{x_i} f(x)\, dx$$

by the corresponding trapezoidal approximation is

$$\frac{(x_i - x_{i-1})^3}{12}|f''(u)|$$

for some u in the interval $[x_{i-1}, x_i]$ (see, e.g., [**1**]). Examination of this error term also suggests that the points x_i be selected to yield small differences $x_i - x_{i-1}$ on intervals in which $|f''(x)|$ may be large and vice versa. Methods based on this general idea are discussed in the numerical analysis literature (see, e.g., [**3**]) and are used in widely distributed software packages. (For example, an analogous approach is used in the IMSL adaptive Simpson's rule subroutine [7].)

The arguments presented above are based on the assumption that the values $f(x_i)$ can be calculated exactly. In many applications, however, the function values used in formulas such as (2) are obtained through experimental observation and thus have measurement error. In addition, expense may limit the number of observations available and, consequently, add importance to the choice of where to make these observations. These facts can substantially affect the appropriate strategies for choosing $x_1, \ldots, x_m$ and can lead to results that are surprising and very nonintuitive.

An important area of application that fits this description involves the estimation of the "relative bioavailability" of a proposed new form of a drug and some standard form. These estimates are used in deciding whether a new drug formulation is suitable and in determining appropriate doses. They are required by the Food and Drug Administration and usually eliminate the need to treat each new formulation as though it were an entirely new drug. This saves time and money and exposes research subjects to less risk [**10**].

We define the **relative bioavailability** of two forms of a drug as the ratio of the amounts of the administered doses that reach the general circulation. Suppose that

$f_1(x) =$ concentration of drug in the blood x hours after the dose of the new form,
$f_2(x) =$ concentration of drug in the blood x hours after the dose of the standard form.

It can be shown ([4]), under fairly general assumptions (which are made in practice), that

$$RB = \text{Relative Bioavailability} = \frac{\int_0^\infty f_1(x)\,dx}{\int_0^\infty f_2(x)\,dx}.$$

Since, for most drugs, $\lim_{x\to\infty} f(x) = 0$, and the above improper integrals exist,

$$RB \doteq \frac{\int_0^X f_1(x)\,dx}{\int_0^X f_2(x)\,dx} \tag{3}$$

for sufficiently large X. In typical drug kinetics applications, the integrands are sums of decaying exponential functions and, consequently, (3) often holds for small to moderately sized X. The integrals in (3) may be estimated by numerical quadrature and the trapezoidal rule is probably the most common choice. The measurements, taken at times $0 = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_{m+1} = X$, consist

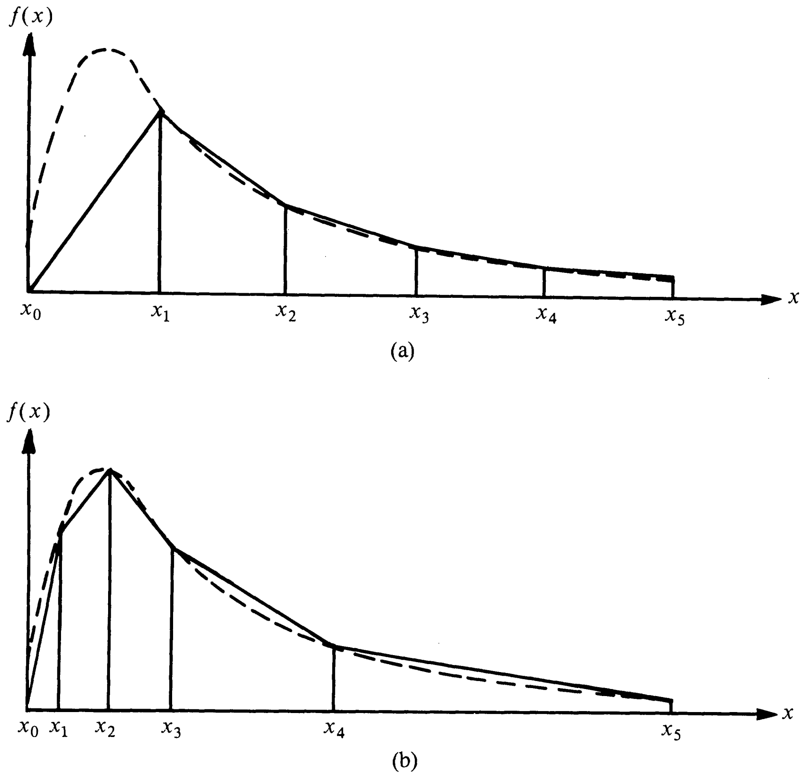


FIGURE 1. **Trapezoidal rule approximation to $\int_{x_0}^{x_5} f(x)\,dx$ using six points and (a) equal subintervals or (b) "optimal" subintervals.**

of laboratory estimates of the $f_1(x_i)$ and $f_2(x_i)$:

$$\hat{f}_j(x_i) = f_j(x_i) + e_{ij}, \qquad j = 1,2 \tag{4}$$

where the e_{ij} are assumed to be additive errors attributed to imprecision in the laboratory assay procedures. For each of the required integral estimates we would like to choose observation times that minimize some measure of error in the approximations. This problem is discussed more generally in the next section and, in a special case, is solved.

The general problem

Since we have the same problem to solve for f_1 and f_2, we will suppress the function subscript j in the following discussion and refer to either function as f. Suppose $\{e_i\}$ are independent random variables with mean 0 and variances $\{\sigma_i^2\}$ representing the error process in observing $f(x_i)$. (This is regarded as a reasonable assumption in many applications, including the assay error example discussed above. See, for example, [**13**].) In this case, the integral $IEXACT$ in (1) is approximated by

$$TR_2 = 0.5\Bigg[(f(x_0)+e_0)(x_1-x_0) + \sum_{i=1}^{m}(f(x_i)+e_i)(x_{i+1}-x_{i-1}) + (f(x_{m+1})+e_{m+1})(x_{m+1}-x_m)\Bigg]. \tag{5}$$

On the average, we would like TR_2 to be a good approximation to $IEXACT$ and will address that goal by choosing $x_1,\dots,x_m$ to minimize the objective function

$$OBJ = E(TR_2 - IEXACT)^2, \tag{6}$$

where the expectation is taken with respect to the measurement errors $\{e_i\}$.

By expanding $(TR_2 - IEXACT)^2$, regrouping the terms, and taking the expectation (recalling that $E(e_i)=0$ and $E(e_i^2)=\sigma_i^2$, $i=0,1,\dots,m+1$), we obtain

$$OBJ = R^2 + Q^2, \tag{7}$$

where

$$R^2 = (0.5)^2\Bigg[\sigma_0^2(x_1-x_0)^2 + \sum_{i=1}^{m}\sigma_i^2(x_{i+1}-x_{i-1})^2 + \sigma_{m+1}^2(x_{m+1}-x_m)^2\Bigg] \tag{8}$$

and

$$Q^2 = (TR - IEXACT)^2. \tag{9}$$

Thus, OBJ decomposes into the sum of Q^2, the square of the error in approximating $IEXACT$ due to use of the trapezoidal rule with no measurement error, and R^2, which depends only on the e_i and the x_i and not on $f(x)$.

If a specific form is assumed for $f(x)$ and for the e_i, OBJ can often be minimized using some nonlinear optimization algorithm, such as a Newton method or the Nelder-Mead simplex procedure [**3**], [**9**].

Application of (7) to a specific problem will frequently show that R^2 is substantially larger than Q^2 [**8**]. For example, if $f(x)$ is a sum of exponential functions (a typical assumption in practice), m is of reasonable size (eight or more, usually) and the σ_i's are representative of assay error (e.g., 10% of typical $f(x)$ values), any "reasonable" choice of $x_1,\dots,x_m$ will result in R^2 that is much larger than Q^2. (In an actual application, this assertion can be examined directly.) If, in (3), the upper limit of integration, X, is close enough to the lower limit to provide reasonable trapezoidal approximations for the particular integrands and numbers of subdivisions, we are likely to have OBJ approximately equal to R^2. In the extreme case in which $f(x)$ is linear, Q^2 is identically 0. In light of the above observations, it is of interest to consider the specific problem of selecting $x_1,\dots,x_m$ to minimize R^2 in (8).

If we assume that the standard deviations of the measurement errors are the same for all values of $f(x)$, i.e., $\sigma_i^2=\sigma^2$ for all i and, without loss of generality, set $x_0=0$ and $x_{m+1}=1$, then the problem of minimizing R^2 is equivalent to minimizing

$$S=x_1^2+\sum_{i=1}^{m}(x_{i+1}-x_{i-1})^2+(1-x_m)^2. \tag{10}$$

An example of such an application is provided by electroimmunodiffusion assays. This method involves placing an antigen on a gel containing an antibody and applying an electric potential to accelerate precipitation along the direction of the current. The length of the resulting immunoprecipitate is a direct function of the concentration of the reactant (antigen). The measurement error is considered to be additive and to have a constant standard deviation [**12**].

We can write S in (10) as the quadratic form

$$S=X^TAX-2B^TX+1 \tag{11}$$

where

$$A=\begin{bmatrix} 2 & 0 & -1 & 0 & 0 & & \cdots & 0\\ 0 & 2 & 0 & -1 & 0 & & \cdots & 0\\ -1 & 0 & 2 & 0 & -1 & & \cdots & 0\\ 0 & -1 & 0 & 2 & 0 & & \cdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & \cdots & & & & 2 & 0 & -1\\ 0 & \cdots & & & & 0 & 2 & 0\\ 0 & \cdots & & & & -1 & 0 & 2 \end{bmatrix},\qquad B=\begin{bmatrix}0\\0\\ \vdots\\0\\1\\1\end{bmatrix} \tag{12}$$

and $X=(x_1,\dots,x_m)^T$. It can be readily verified that

$$\det(A)=\prod_{i=1}^{m}\frac{\left[\frac{i+1}{2}+1\right]}{\left[\frac{i+1}{2}\right]}, \tag{13}$$

where $[x]=$ the greatest integer in x. It follows that the leading principal minors of A are positive and hence A is positive definite [**5**]. From this fact, we know that S has a unique global minimum, X^*, which is the solution to

$$AX=B. \tag{14}$$

This can be established by "completing the square." Briefly, if we write $A=Q^TQ$ and let $D=Q^{-T}B$, then (11) is

$$S=(QX-D)^T(QX-D)+C,$$

which has a unique global minimum when $QX-D=0$ or, equivalently, $AX=B$.

The solutions of (14) for small values of m suggest the following general solution:

If m is even,

$$x_i=\left\{\frac{\left[\frac{i+1}{2}\right]}{\frac{m}{2}+1},\qquad i=1,\dots,m.\right. \tag{15}$$

For example, if $m=8$, two observations should be taken at each of the points $1/5$, $2/5$, $3/5$, and $4/5$ of the way from x_0 to x_{m+1}. This result is quite strange, at least upon initial consideration. Every other trapezoid used in the approximation has a base of length zero and thus contributes nothing to TR_2!

If m is odd, and we set $N = (m+3)/2$,

$$x_i = \begin{cases} \dfrac{i+1}{2N} & \text{for } i \text{ odd}, \\ \dfrac{i}{2(N-1)} & \text{for } i \text{ even}. \end{cases} \tag{16}$$

For example, if $m = 11$, observations should be taken at points 1/7, 1/6, 2/7, 2/6, 3/7, 3/6, 4/7, 4/6, 5/7, 5/6, and 6/7 of the way from x_0 to x_{m+1}. If we treat (15) and (16) as hypotheses, they can be proven by verifying that they satisfy (14) and thus, as noted above, they give the minimizers of S.

Substituting (15) and (16) into (8), we see that the minimum R^2, with constant σ^2, is proportional to $4/(m+2)$ for m even and to $2/(m+1)+2/(m+3)$ for m odd. It should be noted that equal spacing ($x_i = i/(m+1)$, $i = 1,\dots,m$) leads to $(4m+2)/(m+1)^2$ as the proportionality factor. This is an intuitively appealing choice for many people, agrees with typical calculus text examples and is not substantially worse than the optimal solution.

EXAMPLE. The above results involve minimizing an expectation and, hence, cannot be properly illustrated by a single instance. Instead, as an example, we will compare the average results of 1,000 approximations to the same integral in each of two ways: using three "optimal" subdivisions given by (15) and (16), and using three equispaced subdivisions. We consider the integral of

$$f(x) = x^{1/2}$$

on the interval [1, 4] with $x_0 = 1$, $x_3 = 4$, and subdivide the interval with x_2 and x_3 first at 2 and 3, respectively (equispaced choice), then both at 2.5, the choice indicated by (15). Using a digital computer, we add a simulated measurement error to each $f(x_i)$ under each subdivision scheme and calculate the corresponding trapezoidal rule approximation using (2). The errors are chosen independently from a normal distribution with mean 0.0 and standard deviation 0.2 (see FIGURE 2). This process is then repeated an additional 999 times, using new independent errors each time, and the average squared deviation between the approximation and the exact integral, 14/3, is calculated for each subdivision strategy.

For this example, these results, together with the corresponding values of R^2 from (8) and Q^2 from (9) are shown in TABLE 1.

	Equispaced x_i	Optimal x_i
Average of $(IEXACT - TR_2)^2$	0.1004	0.0920
Q^2 (error due to trap. approx.)	0.0004	0.0020
R^2 (error due to measurement)	0.1000	0.0900

TABLE 1

TABLE 1 shows that Q^2 is smaller when equal subdivisions are used but that $R^2 \gg Q^2$ and thus the "optimal" strategy, which minimizes the average value of R^2, yields estimates that are, overall, superior. As mentioned earlier, this is intuitively unexpected in that the superior strategy incorporates a base of zero length in one of its three trapezoids (see FIGURE 2).

Discussion and other applications

The curves in FIGURE 1 are typical of concentration-time curves following a single oral dose of a drug [4]. As noted in the earlier discussion, R^2 is generally larger than Q^2 and thus, if σ^2 is

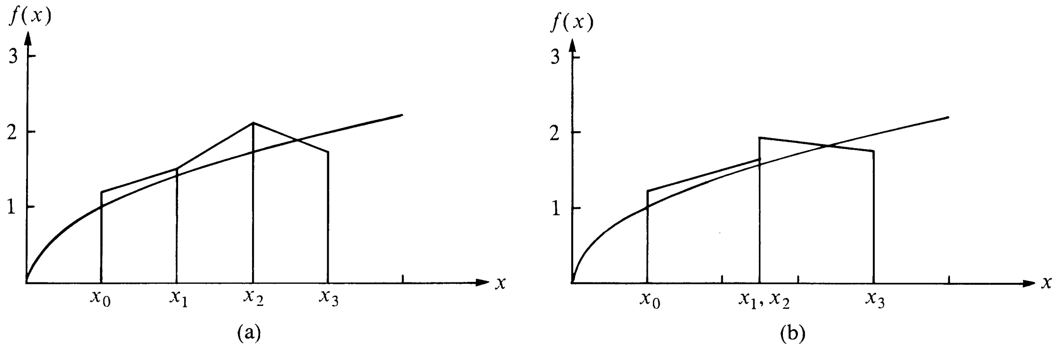


FIGURE 2. **Trapezoidal rule approximation to $\int_{x_0}^{x_3} x^{1/2}\,dx$, where $x^{1/2}$ is observed with error, using (a) equal subintervals or (b) "optimal" subintervals.**

constant, (15) and (16) may provide "approximately optimal" sampling schedules. Conventional practice, on the other hand, is to use geometric sampling schedules (e.g., $x = 1, 2, 4, 8, 16, \ldots$), somewhat similar to those in FIGURE 1(b). We have seen that this would be reasonable for $Q^2 \gg R^2$ but that this assumption is likely to be contrary to the truth in actual applications. Thus, conventional practice seems to result from ignoring the effect of measurement error, although this error may be the dominant influence on the accuracy of the resulting estimates.

In many cases, σ^2 is not constant. For example, radial immunodiffusion assays are similar to the electroimmunodiffusion assays described earlier but they do not employ electric potentials to accelerate precipitation. In this case, the diffusion is radial instead of linear and the standard deviations of the measurement errors tend to be proportional to the magnitude of the quantity being measured [**12**]. In this case, we could assume a form for $f(x)$ and a model for the error e and minimize OBJ numerically. The results of such analyses tend to be similar to those obtained by application of (15) and (16) and are also fairly insensitive to assumptions made concerning $f(x)$ and e [**8**].

It is interesting to consider the relative merits of using an even m (15) versus an odd m (16) in the constant standard deviation case considered above. Although the optimum R^2 is always reduced by increasing m ("the more points the better"), for any given m the inequality

$$\frac{4}{m+2} < \frac{2}{m+1} + \frac{2}{m+3}$$

indicates more "information" per observation for even m. (In addition, it is often much less expensive to take two observations at the same time.) The apparent advantages of using an even m are especially noteworthy when we recall that this strategy involves ignoring every other trapezoid in approximating $IEXACT$.

There are many other applied problems that require the estimation of one or more definite integrals from a limited number of noisy observations. One such application is the estimation of cardiac output, a measure of heart function, by the dye dilution process [**6**]. Here, a dose D of dye is input to the heart and $f(x)$ is the concentration of dye, as a function of time, leaving the heart. The **cardiac output** is the quotient

$$\frac{D}{\int_0^\infty f(x)\,dx}$$

and, as in the drug examples, the integral is generally estimated by numerical integration.

Another interesting application to drug kinetics involves estimation of the average amount of time a molecule of a drug spends in the body (mean residence time) following administration. Under some general assumptions [**11**], it can be shown that the **mean residence time** (MRT) is the x-coordinate of the centroid of the area under the curve $f(x)$:

$$MRT=\frac{\int_0^\infty xf(x)\,dx}{\int_0^\infty f(x)\,dx}. \tag{17}$$

Again, people conventionally use strategies that would be appropriate if there was no measurement error. Here, the effect of measurement error is profound as can be seen by considering the numerator in (17). Errors in $f(x)$ are multiplied by x, which goes to infinity. Consequently, errors can have a large effect on the resulting estimate of MRT [**2**].

Different relationships between the magnitude of these errors and the corresponding levels of measurements were illustrated earlier by two specific assay procedures. In practice, it is common to perform replicated assays at each of two or more levels, selected to cover the range of interest. The empirical standard deviations are examined and used to infer an approximate relationship between error standard deviation and concentration. Typically, the standard deviations are constant or proportional to the concentration (see, e.g., [**11**]).

In summary, many applied problems require integral estimates from noisy data. Mathematics provides strategies for selecting observation points that are appropriate but these strategies can be quite different for the noisy case than for the case in which the integrand can be calculated exactly. Conventional practice is often based on the latter assumption but applied to the former situation. In fact, results such as those in (15) and (16) are so unappealing to the intuitions of many researchers that they are often met with resistance and even hostility.

This work was supported in part by NIH Grant RR01629.

References

[1] R. L. Burden, J. D. Faires and A. C. Reynolds, Numerical Analysis, Prindle, Weber and Schmidt, Boston, 1981.

[2] D. Z. D'Argenio and D. Katz, Sampling strategies for noncompartmental estimation of mean residence time, J. Pharmacokin. Biopharm., 11 (1983) 437–448.

[3] V. A. Dixon, Numerical Quadrature: A Survey of the Available Algorithms, in D. J. Evans (Ed.), Software for Numerical Mathematics, Academic Press, London, 1974.

[4] M. Gibaldi and M. Perrier, Pharmacokinetics, Dekker, New York, 1975.

[5] F. E. Hohn, Elementary Matrix Algebra, Macmillan, New York, 1964.

[6] B. Horelick and S. Koont, Measuring Cardiac Output, UMAP Module (Unit 71), Birkhäuser Boston, Inc., Cambridge, 1980.

[7] International Mathematical and Statistical Libraries, IMSL Library Reference Manual, Volumes 2 and 3, Edition 8, Houston, 1980.

[8] D. Katz and D. Z. D'Argenio, Experimental design for estimating integrals by numerical quadrature, with applications to pharmacokinetic studies, Biometrics, 39 (1983) 621–628.

[9] J. Nelder and R. Mead, A Simplex method for function minimization, Comput. J., 4 (1965) 308–313.

[10] E. Purich, Bioavailability/Bioequivalency Regulations: An FDA Perspective, in K. S. Albert (Ed.), Drug Absorption and Disposition: Statistical Considerations, American Pharmaceutical Association, Washington, 1980.

[11] R. J. Sawchuk, S. M. Pepin, I. E. Leppik and R. J. Gumnit, Rapid and slow release Phenytoin in epileptic patients at steady state: comparative plasma levels and toxicity, J. Pharmacokin. Biopharm., 10 (1982) 365–391.

[12] M. Werner, Immunoprecipitation Assays in Stabilizing Media, in M. Werner (Ed.), Microtechniques for the Clinical Laboratory, John Wiley and Sons, New York, 1976.

[13] M. Werner, S. H. Brooks and L. B. Knott, Additive, multiplicative, and mixed analytical error, Clin. Chem., 21 (1978) 1895–1898.

[14] J. Yamaoka, T. Nakagawa and T. Uno, Statistical moments in pharmacokinetics, J. Pharmacokin. Biopharm., 6 (1978) 547–558.

Hamiltonian Checkerboards

JOSEPH A. GALLIAN
University of Minnesota, Duluth
Duluth, MN 55812

DAVID WITTE
University of Chicago
Chicago, IL 60637

Suppose we put a checker in the square at the Southwest corner of a standard 8×8 checkerboard. Let's assume that the checker cannot step off the edge of the board and that it can take a step (a distance of one square) in any of the four compass directions (North, South, East, West)—in particular, it does not move diagonally. Is it possible to move the checker in this manner through a *tour* of the board? That is, *can the checker travel a route which passes through each square exactly once before returning to the starting place*? The answer is "yes." (See FIGURE 1.)

It's trivial to find a tour on an 8×8 board, but what if we pose the problem for a checkerboard of another size—or perhaps a rectangular board (rather than square)? As long as the board has an even number of squares, it's still easy to plan the tour. In fact, a route like that of FIGURE 1 is suitable. (We can assume the number of rows is even by rotating the board if necessary.) On the other hand, if we use a board with an odd number of squares, then no tour is possible. The reason is just simple counting. Since the checker would start and end at the same place, for each step North, there must be one step South; for each step East, a single step West. Thus, there would be an even number of steps in the tour. This would entail visiting an even number of squares; whereas, the board has an odd number of squares.

If we change the constraints on the movement of the checker and allow the checker to step off the edge of the board (like Pac-Man), it becomes possible to tour a board with *any* number of squares—even or odd—and any rectangular shape. When the checker steps off the East edge, it moves to the West-most square of the same row. (Think of the row as "wrapping around.") The other three edges also wrap around. This can be described by saying the board is not flat; it is in the shape of a torus or doughnut. A tour of this type on a board with an odd number of squares is shown in FIGURE 2.

We need to be more sophisticated to deal with our next problem: *find a route which always travels North or East*. (We allow no steps to the South or West, but the checker may step off the edge of the board as described before.) This is easy to do on the 8×8 checkerboard. (Try it; there is a solution in FIGURE 3.) Not every rectangular board admits such a tour; a 3×5 checkerboard provides an example. The proof is by contradiction. Since the board has 15 squares, the tour must have a total of 15 steps. Let us say that E of these are Easterly and N are Northerly. Then, $E + N = 15$ and, since the checker must begin and end in the Southwest corner, it must be true that E is divisible by 5 and N is divisible by 3. Clearly, E can't be 0 or 15; otherwise, the checker

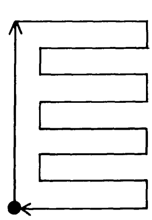


FIGURE 1. **A tour of a checkerboard.**

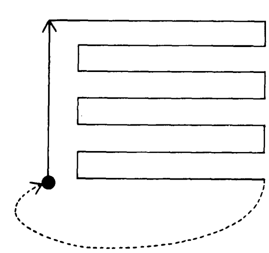


FIGURE 2. **A tour of a 7×9 checkerboard.**

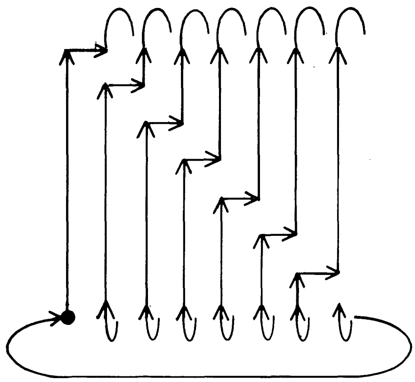


FIGURE 3. **A North-East tour of a checkerboard.**

would remain exclusively in one column or one row of the board. Thus, $E=5$ or 10; and, consequently, N must be 10 or 5. But neither 10 nor 5 is divisible by 3. This contradiction completes the proof that there is no tour on the 3×5 board.

For historical reasons, we say that a checkerboard is **hamiltonian** if it admits a tour of this type: the checker travels only North and East, a step off the edge of the board is permitted, the tour visits every square exactly once, and it ends in the square where it began. In general, it can be shown that *an $r\times c$ rectangular checkerboard is hamiltonian if and only if there is a pair of relatively prime positive integers a and b such that $ar+bc=rc$.* (The argument in the above paragraph for a 3×5 board can be generalized: take $a=N/r$ and $b=E/c$.)

Having determined which checkerboards are hamiltonian, it is natural to consider the analogous problem in three dimensions. Here, we replace the rectangular checkerboard with a rectangular parallelopiped, made of stacked cubes. In addition to steps East or North, the checker can also move up (but not down) one level in the configuration. (Of course, a move up from the top level wraps around, taking the checker to the bottom level of that same stack.) Now we ask: *which three-dimensional checkerboards are hamiltonian*? Although it is difficult to prove, it turns out that with 3- or higher-dimensional checkerboards, the extra freedom of movement always permits a hamiltonian tour [**2**].

Now that we've considered checkerboards of every dimension, it might seem that there is nothing left to do. Not so! The hamiltonian checkerboard problem has many natural generalizations. One such generalization can be obtained by representing a checkerboard schematically as a grid of dots, one dot for each square on the checkerboard. Then, each permissible move the checker can make can be denoted by drawing arrows between dots. The resulting **directed graph**, or **digraph**, looks like FIGURE 4. Each dot is called a **vertex**, and the arrows are **arcs**.

It is easy to describe the digraphs that arise in the checkerboard problem. Let $Z_r\times Z_c$ be the direct product of two cyclic groups—one of order r (for rows) and the other of order c (for columns). We may construct the digraph corresponding to an $r\times c$ rectangular checkerboard by drawing a vertex for each element of $Z_r\times Z_c$, and drawing an arc from every vertex (m,n) to each of $(m+1,n)$ and $(m,n+1)$. The same idea can be used to construct the digraph for a multidimensional checkerboard from an abelian group $Z_{n_1}\times Z_{n_2}\times\cdots\times Z_{n_d}$. In this way, our problem about a hamiltonian tour of a d-dimensional checkerboard corresponds to a problem about the digraph arising from the group $Z_{n_1}\times\cdots\times Z_{n_d}$ with the natural generating set. The checkerboard is hamiltonian if and only if there is a **hamiltonian tour** of the digraph, that is, a path which follows the directed arcs and goes through each vertex exactly once, returning to the vertex at which it began. This translation of our original problem to one of deciding if a digraph is hamiltonian leads to other interesting questions.

In general, given any finite group G and a generating set S for G, we can construct the so-called **Cayley digraph** on G. The set of vertices corresponds to the elements of G. The

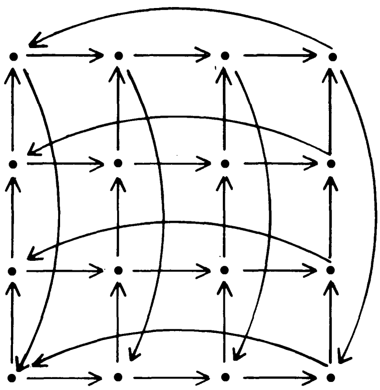


FIGURE 4. **The digraph for a 4 × 4 checkerboard which permits only Northerly and Easterly moves.**

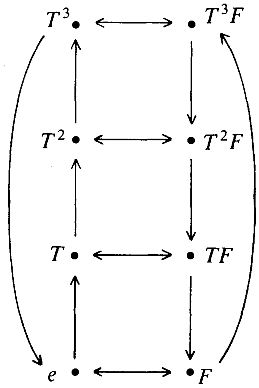


FIGURE 5. **The Cayley digraph on D_8.**

generating set S tells you where to put arcs: for each vertex g and each generator s, draw an arc from g to gs. When G is the abelian group $Z_{n_1} \times Z_{n_2} \times \cdots \times Z_{n_d}$ and S is the obvious generating set, this is the same as the construction above for (multidimensional) checkerboards.

Other interesting digraphs are obtained when G is a nonabelian group. Perhaps the least complicated nonabelian examples are the dihedral groups. The **dihedral group** D_{2m} of order $2m$ is the plane symmetry group of a regular m-gon ($m \geqslant 3$). It is generated by two elements: a rotation T of 360°/m about the center of the m-gon and a reflection F across an axis of symmetry. Algebraically, the group can also be described as being generated by T and F, subject only to the relations $T^m = F^2 = e$ and $FTF = T^{-1}$. FIGURE 5 shows the Cayley digraph on D_8 (with generating set $\{T, F\}$ of course). It is obvious that the Cayley digraph on D_{2m} is hamiltonian.

With some nonabelian groups, it is not easy to tell whether or not the Cayley digraph is hamiltonian. For example, consider the Cayley digraph of the direct product $D_{2m} \times Z_r$ of a dihedral group with a cyclic group. (The obvious generating set is $\{(T,0),(F,0),(e,1)\}$.) You can easily show the Cayley digraph is hamiltonian if r is even (see FIGURE 6 for an example). On the other hand, it is quite a challenge to show that the Cayley digraph on $D_8 \times Z_3$, for instance, is hamiltonian (see FIGURE 7). Even so, it has been proved that the Cayley digraph on $D_{2m} \times Z_r$ is hamiltonian for all m and r (see [**11**]).

There have been a number of applications of hamiltonian tours in Cayley digraphs and hamiltonian paths in Cayley digraphs (a **hamiltonian path** is the same as a hamiltonian tour, except

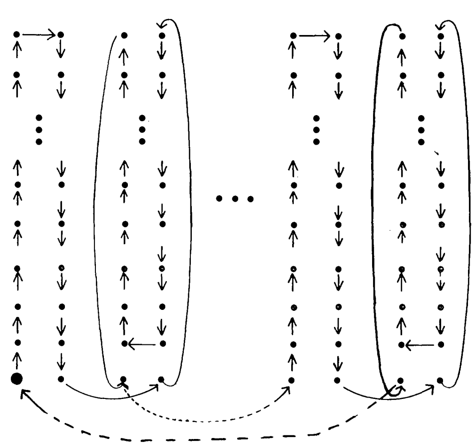


FIGURE 6. **A hamiltonian tour of $D_{2m} \times Z_r$ when r is even.**

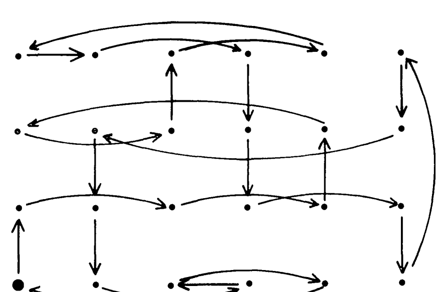


FIGURE 7. **A hamiltonian tour of $D_8 \times Z_3$.**

one need not return to the starting point). Rankin **[10]** was the first to formalize the intimate connection between hamiltonian tours on the symmetric groups and change-ringing (see also **[1]**, **[3]**, and **[5]**). The design of post office machinery discussed in **[6]** and **[7]** employs hamiltonian paths in Cayley digraphs on dihedral groups. Nijenhuis and Wilf **[9]** use hamiltonian paths in symmetric groups to generate all the permutations of a set. Another recent application utilizes hamiltonian paths in Cayley digraphs on infinite symmetry groups of hyperbolic space to produce patterns like those of the late graphic artist M. C. Escher **[4]**.

For further information and a comprehensive bibliography, see **[12]**. Reference **[8]** gives a nice introduction to Cayley digraphs.

References

[1] F. J. Budden, The Fascination of Groups, Cambridge University Press, Cambridge, 1972.

[2] Stephen J. Curran and David Witte, Hamiltonian paths in Cartesian products of directed cycles, Ann. Discrete Math., to appear.

[3] D. J. Dickinson, On Fletcher's paper campanological groups, Amer. Math. Monthly, 64 (1957) 331–332.

[4] Douglas Dunham, John Lindgren, and David Witte, Creating repeating hyperbolic patterns, Computer Graphics, 15 (1981) 215–223.

[5] T. J. Fletcher, Campanological groups, Amer. Math. Monthly, 63 (1956) 619–626.

[6] Joseph A. Gallian, Group theory and the design of a letter facing machine, Amer. Math. Monthly, 84 (1977) 285–287.

[7] Joseph A. Gallian and Charles A. Marttila, Reorienting regular *n*-gons, Aequationes Math., 20 (1980) 97–103.

[8] I. Grossman and W. Magnus, Groups and Their Graphs, MAA, Washington, D.C., 1964.

[9] A. Nijenhuis and H. S. Wilf, Chapter NEXPER in Combinatorial Algorithms, 2nd ed., Academic Press, New York, 1978.

[10] R. A. Rankin, A campanological problem in group theory, Proc. Camb. Phil. Soc., 44 (1948) 17–25.

[11] David Witte, Gail Letzter, and Joseph A. Gallian, On hamiltonian circuits in cartesian products of Cayley digraphs, Discrete Math., 43 (1983) 297–307.

[12] David Witte and Joseph A. Gallian, A survey: Hamiltonian cycles in Cayley graphs, Discrete Math., to appear.

A Property of Triangles

LEON BANKOFF
6360 Wilshire Blvd.
Los Angeles, CA 90048

CHARLES W. TRIGG
2404 Loring Street
San Diego, CA 92109

Blundon **[1]**, with some neat but complicated trigonometric manipulations, has established that the identity

$$(\cos A-\cos\theta)(\cos B-\cos\theta)(\cos C-\cos\theta)=0$$

leads to

$$s=2R\sin\theta+r\cot(\theta/2) \tag{1}$$

one need not return to the starting point). Rankin [**10**] was the first to formalize the intimate connection between hamiltonian tours on the symmetric groups and change-ringing (see also [**1**], [**3**], and [**5**]). The design of post office machinery discussed in [**6**] and [**7**] employs hamiltonian paths in Cayley digraphs on dihedral groups. Nijenhuis and Wilf [**9**] use hamiltonian paths in symmetric groups to generate all the permutations of a set. Another recent application utilizes hamiltonian paths in Cayley digraphs on infinite symmetry groups of hyperbolic space to produce patterns like those of the late graphic artist M. C. Escher [**4**].

For further information and a comprehensive bibliography, see [**12**]. Reference [**8**] gives a nice introduction to Cayley digraphs.

References

[1] F. J. Budden, The Fascination of Groups, Cambridge University Press, Cambridge, 1972.
[2] Stephen J. Curran and David Witte, Hamiltonian paths in Cartesian products of directed cycles, Ann. Discrete Math., to appear.
[3] D. J. Dickinson, On Fletcher's paper campanological groups, Amer. Math. Monthly, 64 (1957) 331–332.
[4] Douglas Dunham, John Lindgren, and David Witte, Creating repeating hyperbolic patterns, Computer Graphics, 15 (1981) 215–223.
[5] T. J. Fletcher, Campanological groups, Amer. Math. Monthly, 63 (1956) 619–626.
[6] Joseph A. Gallian, Group theory and the design of a letter facing machine, Amer. Math. Monthly, 84 (1977) 285–287.
[7] Joseph A. Gallian and Charles A. Marttila, Reorienting regular *n*-gons, Aequationes Math., 20 (1980) 97–103.
[8] I. Grossman and W. Magnus, Groups and Their Graphs, MAA, Washington, D.C., 1964.
[9] A. Nijenhuis and H. S. Wilf, Chapter NEXPER in Combinatorial Algorithms, 2nd ed., Academic Press, New York, 1978.
[10] R. A. Rankin, A campanological problem in group theory, Proc. Camb. Phil. Soc., 44 (1948) 17–25.
[11] David Witte, Gail Letzter, and Joseph A. Gallian, On hamiltonian circuits in cartesian products of Cayley digraphs, Discrete Math., 43 (1983) 297–307.
[12] David Witte and Joseph A. Gallian, A survey: Hamiltonian cycles in Cayley graphs, Discrete Math., to appear.

A Property of Triangles

Leon Bankoff
6360 Wilshire Blvd.
Los Angeles, CA 90048

Charles W. Trigg
2404 Loring Street
San Diego, CA 92109

Blundon [**1**], with some neat but complicated trigonometric manipulations, has established that the identity

$$(\cos A - \cos\theta)(\cos B - \cos\theta)(\cos C - \cos\theta) = 0$$

leads to

$$s = 2R\sin\theta + r\cot(\theta/2) \tag{1}$$

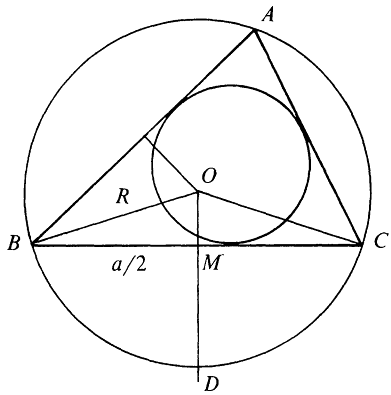


FIGURE 1

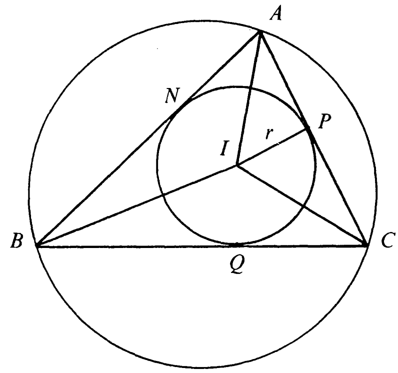


FIGURE 2

as a necessary and sufficient condition that θ be an angle of the triangle ABC where s, R, and r are the semiperimeter, circumradius, and inradius, respectively.

This result can be obtained more expeditiously by first blending three well-known properties of a triangle. Consider the two aspects of triangle ABC in FIGURES 1 and 2. Since the perpendicular bisectors of the sides of a triangle intersect in its circumcenter O, then, in FIGURE 1, $OB = R$. Also, OD bisects the arc BDC. It follows that $\angle BOD = \angle A$, and from triangle BOM, $(a/2)/R = \sin A$, or

$$a = 2R\sin A. \qquad (2)$$

In FIGURE 2, where the intersection of the angle bisectors is the incenter I, from triangle AIP, we see that $AP/r = \cot(A/2)$, or

$$AP = r\cot(A/2). \qquad (3)$$

Furthermore, in FIGURE 2, the tangents $BQ = BN$, $CQ = CP$, and $AP = AN$. Thus the perimeter, $2s = 2BQ + 2CQ + 2AP$, or

$$s = BQ + CQ + AP = a + AP.$$

Whereupon, substituting from (2) and (3),

$$s = 2R\sin A + r\cot(A/2).$$

In like manner, $s = 2R\sin B + r\cot(B/2)$ and $s = 2R\sin C + r\cot(C/2)$ can be derived, all in the form of (1).

When the trigonometric functions in (1) are expressed in terms of $\cos(\theta/2) = x$, and radicals are eliminated, we have

$$16R^2x^6 - 8R(4R+r)x^4 + \left[(4R+r)^2 + s^2\right]x^2 - s^2 = 0. \qquad (4)$$

By Descartes' rule of signs, this equation can have no more than three positive roots. Now, A, B, and C have been shown to satisfy (1), and since each of their three half-angles is $< 90°$, it follows that $\cos(A/2)$, $\cos(B/2)$, and $\cos(C/2)$ are the three positive roots of (4). Thus (1) holds if and only if θ is an angle of a triangle with parameters s, R, and r.

For certain specific values of θ, we have:

$$\begin{aligned} \theta &= 30°, & s &= R + (2+\sqrt{3})\,r; \\ \theta &= 60°, & s &= \sqrt{3}\,(R+r); \\ \theta &= 90°, & s &= 2R + r; \text{ and} \\ \theta &= 120°, & s &= \sqrt{3}\,(R + r/3). \end{aligned}$$

It follows, by equating the appropriate different expressions for s, that in a 30°-60°-90° triangle, $R/r=\sqrt{3}+1$, and in a 30°-30°-120° triangle, $R/r=2+4/\sqrt{3}$.

Reference

[1] W. J. Blundon, Generalization of a relation involving right triangles, Amer. Math. Monthly, 74 (1967) 566–567.

An Algorithm for Sums of Integer Powers

CLIVE KELLY
University of Hull,
Hull, HU6 7RX, England

Let $S_k(n)=1^k+2^k+\cdots+n^k$, where k, n are positive integers. The usual proofs of the evaluation of $S_k(n)$ for small $k=1$, 2 or 3, use induction. The drawback of these is, of course, that you need to know the answer in order to prove it.

A proof of the identity

$$1+\sum_{i=0}^{k}\binom{k+1}{i}S_i(n)=(n+1)^{k+1} \tag{1}$$

that involves generating functions is in Riordan [**2**, p. 160]; a combinatorial proof is in Paul [**1**]; and here is one that is neither, but is very easy. (As usual, $\binom{k}{i}$ denotes the binomial coefficient $k!/i!(k-i)!$.) A simple consequence of (1) is of course that on rewriting it in the form

$$(k+1)S_k(n)=(n+1)^{k+1}-1-\sum_{i=0}^{k-1}\binom{k+1}{i}S_i(n), \tag{2}$$

each $S_k(n)$ can be calculated in terms of the previous sums $S_0(n), S_1(n),\ldots, S_{k-1}(n)$, and induction is not needed.

Thus we get, using (2),

$$\begin{aligned}
k&=1, & 2S_1(n)&=(n+1)^2-(n+1)=n(n+1).\\
k&=2, & 3S_2(n)&=(n+1)^3-(n+1)-\frac{3n}{2}(n+1)=\frac{n}{2}(n+1)(2n+1).\\
k&=3, & 4S_3(n)&=(n+1)^4-(n+1)-6S_2(n)-4S_1(n)=n^2(n+1)^2.\\
k&=4, & S_4(n)&=\frac{n}{30}(n+1)(2n+1)(3n^2+3n-1),
\end{aligned}$$

but the manipulation begins to get tedious at this stage. Jakob Bernoulli in his book *Ars Conjectandi* on probability investigated the values of $S_k(n)$, presenting them as polynomials in n and giving exact formulas for the coefficients in terms of the well-known Bernoulli numbers. (See, for instance Struik [**3**, p. 317] and the references given there.)

We give a very elementary proof of (1), using the binomial expansion:

$$(r+1)^{k+1}=1+(k+1)r+\cdots+\binom{k+1}{i}r^i+\cdots+(k+1)r^k+r^{k+1}. \tag{3}$$

It follows, by equating the appropriate different expressions for s, that in a 30°-60°-90° triangle, $R/r=\sqrt{3}+1$, and in a 30°-30°-120° triangle, $R/r=2+4/\sqrt{3}$.

Reference

[1] W. J. Blundon, Generalization of a relation involving right triangles, Amer. Math. Monthly, 74 (1967) 566–567.

An Algorithm for Sums of Integer Powers

CLIVE KELLY
University of Hull,
Hull, HU6 7RX, England

Let $S_k(n)=1^k+2^k+\cdots+n^k$, where k, n are positive integers. The usual proofs of the evaluation of $S_k(n)$ for small $k=1$, 2 or 3, use induction. The drawback of these is, of course, that you need to know the answer in order to prove it.

A proof of the identity

$$1+\sum_{i=0}^{k}\binom{k+1}{i}S_i(n)=(n+1)^{k+1} \tag{1}$$

that involves generating functions is in Riordan [**2**, p. 160]; a combinatorial proof is in Paul [**1**]; and here is one that is neither, but is very easy. (As usual, $\binom{k}{i}$ denotes the binomial coefficient $k!/i!(k-i)!$.) A simple consequence of (1) is of course that on rewriting it in the form

$$(k+1)S_k(n)=(n+1)^{k+1}-1-\sum_{i=0}^{k-1}\binom{k+1}{i}S_i(n), \tag{2}$$

each $S_k(n)$ can be calculated in terms of the previous sums $S_0(n), S_1(n),\dots,S_{k-1}(n)$, and induction is not needed.

Thus we get, using (2),

$$k=1, \qquad 2S_1(n)=(n+1)^2-(n+1)=n(n+1).$$
$$k=2, \qquad 3S_2(n)=(n+1)^3-(n+1)-\frac{3n}{2}(n+1)=\frac{n}{2}(n+1)(2n+1).$$
$$k=3, \qquad 4S_3(n)=(n+1)^4-(n+1)-6S_2(n)-4S_1(n)=n^2(n+1)^2.$$
$$k=4, \qquad S_4(n)=\frac{n}{30}(n+1)(2n+1)(3n^2+3n-1),$$

but the manipulation begins to get tedious at this stage. Jakob Bernoulli in his book *Ars Conjectandi* on probability investigated the values of $S_k(n)$, presenting them as polynomials in n and giving exact formulas for the coefficients in terms of the well-known Bernoulli numbers. (See, for instance Struik [**3**, p. 317] and the references given there.)

We give a very elementary proof of (1), using the binomial expansion:

$$(r+1)^{k+1}=1+(k+1)r+\cdots+\binom{k+1}{i}r^i+\cdots+(k+1)r^k+r^{k+1}. \tag{3}$$

Setting $r = n, n-1, n-2, \ldots$ successively in (3), and going as far as $r = 1$, we get n equations, the last of which is

$$(1+1)^{k+1} = 1 + (k+1) + \cdots + \binom{k+1}{i} + \cdots + (k+1) + 1.$$

Adding these n equations, each term on the left, except for $(n+1)^{k+1}$, cancels with the last term on the right in the preceding equation. Thus

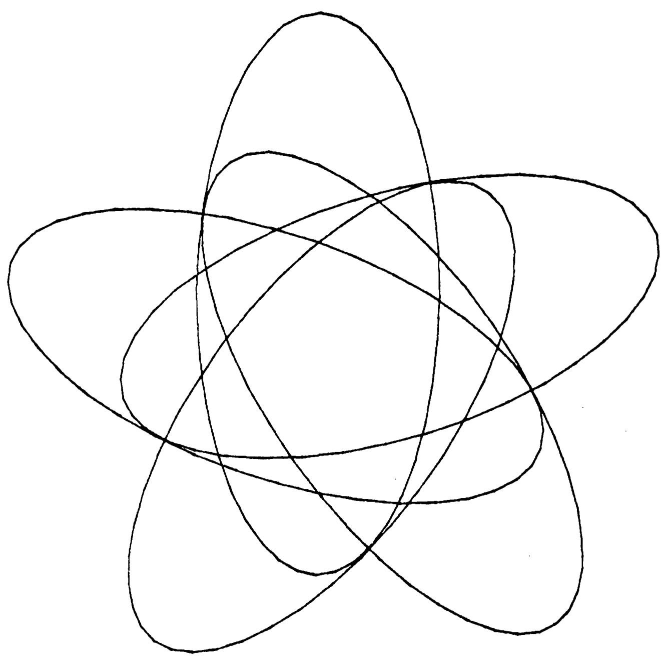
$$(n+1)^{k+1} = 1 + S_0(n) + (k+1)S_1(n) + \cdots + \binom{k+1}{i}S_i(n) + \cdots + (k+1)S_k(n)$$
$$= 1 + \sum_{i=0}^{k} \binom{k+1}{i} S_i(n),$$

which is identity (1).

References

[1] J. L. Paul, On the sum of the kth powers of the first n integers, Amer. Math. Monthly, 78 (1971) 271–272.
[2] J. Riordan, Combinatorial Identities, Wiley, 1968.
[3] D. J. Struik, A Source Book in Mathematics 1200–1800, Harvard U. Press, 1969.

Venn Diagram for Five Sets



—Allen J. Schwenk
U.S. Naval Academy

Editor's note: Readers may be interested in the article "The Construction of Venn Diagrams" by B. Grünbaum, College Math. J. 15 (1984) 238–247.

Setting $r = n, n-1, n-2, \ldots$ successively in (3), and going as far as $r = 1$, we get n equations, the last of which is

$$(1+1)^{k+1} = 1 + (k+1) + \cdots + \binom{k+1}{i} + \cdots + (k+1) + 1.$$

Adding these n equations, each term on the left, except for $(n+1)^{k+1}$, cancels with the last term on the right in the preceding equation. Thus

$$\begin{aligned}(n+1)^{k+1} &= 1 + S_0(n) + (k+1)S_1(n) + \cdots + \binom{k+1}{i}S_i(n) + \cdots + (k+1)S_k(n)\\ &= 1 + \sum_{i=0}^{k}\binom{k+1}{i}S_i(n),\end{aligned}$$

which is identity (1).

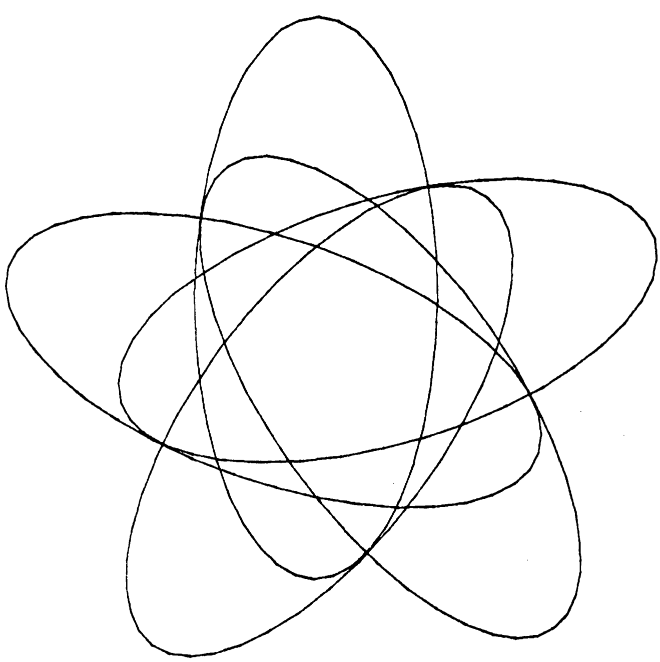
References

[1] J. L. Paul, On the sum of the kth powers of the first n integers, Amer. Math. Monthly, 78 (1971) 271–272.

[2] J. Riordan, Combinatorial Identities, Wiley, 1968.

[3] D. J. Struik, A Source Book in Mathematics 1200–1800, Harvard U. Press, 1969.

Venn Diagram for Five Sets



—ALLEN J. SCHWENK
U.S. Naval Academy

Editor's note: Readers may be interested in the article "The Construction of Venn Diagrams" by B. Grünbaum, College Math. J. 15 (1984) 238–247.

PROBLEMS

LEROY F. MEYERS, Editor
G. A. EDGAR, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before April 1, 1985.

1201. Let

$$S_n = \sum_{k=0}^{n} 2^k \tan\frac{x}{2^k} \tan^2\frac{x}{2^{k+1}}.$$

Find simple expressions for S_n and $\lim_{n\to\infty} S_n$. [*Themistocles M. Rassias, Athens, Greece.*]

1202. Let m and n be coprime positive integers with $n > m$. Let θ be a real number. Consider the following algorithm:

(i) Initialize (A, B, C, d, e) to $(2\cos\theta, 2\cos\theta, 2, n-m, m)$.

(ii) If $d > e$, replace (B, C, d) by $(AB - C, B, d-e)$; otherwise replace (A, C, e) by $(AB - C, A, e - d)$.

(iii) If $e = 0$, terminate the algorithm; otherwise return to step (ii).

Prove that $A = 2\cos(n\theta)$ when the algorithm terminates. [*Peter L. Montgomery, System Development Corporation, Santa Monica, California.*]

1203. Let $p(x) = ax^2 + bx + c$, where a, b, and c are integers with $a \neq 0$. If n is an integer such that $n < p(n) < p(p(n))$, show that $p(p(n)) < p(p(p(n)))$ if and only if $a > 0$. [*B. Landman and J. Layman, Virginia Polytechnic Institute and State University, and B. Klein, Davidson College.*]

1204. Euler's constant γ is defined by

$$\gamma = \lim_{n\to\infty}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n\right). \qquad (1)$$

Prove the following generalization of (1):

$$\gamma = \lim_{n\to\infty}\left[\left(\frac{2^m}{1} + \frac{3^m}{2} + \cdots + \frac{n^m}{n-1}\right) + m - \left(S_n^0 + S_n^1 + \cdots + S_n^{m-1}\right) - \ln(n-1)\right],$$

where $S_n^m = 1^m + 2^m + \cdots + n^m$. [*H. Roelants, Hoger Instituut voor Wijsbegeerte, Leuven, Belgium.*]

---

*We invite readers to submit problems believed to be new. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) will be placed next to a problem number to indicate that the proposer did not supply a solution.*

Solutions should be written in a style appropriate for Mathematics Magazine. Each solution should begin on a separate sheet containing the solver's name and full address. It is not necessary to submit duplicate copies.

Send all communications to the problems department to Leroy F. Meyers, Mathematics Department, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.

1205. Show that for $x \in [0,1]$,

(a) $$\left|\frac{\pi}{4} - \operatorname{Arcsin} x\right| \leqq \frac{\pi}{4}\sqrt{1 - 2x\sqrt{1-x^2}} \quad \text{and}$$

(b) $$\operatorname{Arcsin} x \leqq \frac{x}{\frac{2}{\pi} + \frac{\pi}{12}(1-x^2)}.$$

When does equality hold? [*Vania D. Mascioni, student, ETH Zürich, Switzerland.*]

Quickie

Answer to the Quickie appears on p. 305.

Q695. Triangles ABC and $A'B'C'$ are said to be skew-congruent iff $\angle A = \angle A'$, $AB = A'B'$, and $BC = B'C'$. Suppose we have triangles $A_0B_0C_0, A_1B_1C_1, \ldots, A_nB_nC_n$ such that $A_{k-1}B_{k-1}C_{k-1}$ is skew-congruent to $A_kB_kC_k$ for $1 \leqq k \leqq n$. Is it possible for $A_0B_0C_0$ and $A_nB_nC_n$ to be similar but not congruent? [*James Propp, University of California at Berkeley.*]

Solutions

A Tricky Infinite Nested Radical **May 1983**

1174. The sixth morning problem of the 1953 Putnam Competition was to determine the limit of the sequence

$$\left(\sqrt{k}, \sqrt{k-\sqrt{k}}, \sqrt{k-\sqrt{k+\sqrt{k}}}, \sqrt{k-\sqrt{k+\sqrt{k-\sqrt{k}}}}, \ldots\right)$$

in the specific instance when $k = 7$. (The limit is 2.) For which real numbers A does there exist a k such that the sequence converges to A? For such A, write k explicitly in terms of A. [*Thomas P. Dence, California State University, Los Angeles.*]

Solution (adapted by the editor): If $k = 0$, the sequence converges to 0. Otherwise, for all $k \geqq k_0 \approx 1.7548777$ (where k_0 is the unique positive solution to $k^3 - 2k^2 + k - 1 = 0$), the sequence converges to $A = \sqrt{k - \frac{3}{4}} - \frac{1}{2}$, whence $k = A^2 + A + 1$. Moreover, for each $A \geqq A_0 = \sqrt{k_0 - \frac{3}{4}} - \frac{1}{2} \approx 0.5024359$, as well as for $A = 0$, but for no other value of A, there is a (unique) value of k for which the sequence has limit A.

For fixed $k > 0$ let the given sequence $(x_n)_{n=1}^{\infty}$ and an auxiliary sequence $(y_n)_{n=1}^{\infty}$ be defined by

$$x_1 = y_1 = \sqrt{k}; \quad x_{n+1} = \sqrt{k - y_n} \quad \text{and} \quad y_{n+1} = \sqrt{k + x_n} \quad \text{for } n \geqq 1.$$

If $A = \lim_{n\to\infty} x_n$ exists, then so does $B = \lim_{n\to\infty} y_n$, and we have $A = \sqrt{k-B}$ and $B = \sqrt{k+A}$. Elimination of B yields the equation $k^2 - (2A^2+1)k + A^4 - A = 0$ with solution $k = A^2 - A$ or $k = A^2 + A + 1$. But $0 \leqq x_n \leqq \sqrt{k}$ for all n, so that $0 \leqq A \leqq \sqrt{k}$ and $A^2 - A \leqq \max\{0, k - \sqrt{k}\} < k$; hence the first possibility is eliminated. From the second possibility we obtain the unique positive solution $A = \sqrt{k - \frac{3}{4}} - \frac{1}{2}$; then $B = k - A^2 = A + 1$.

In order to ensure that all terms x_n are real, it suffices to require that $x_3 \geqq 0$, since if this holds, then we can prove by induction that $x_3 \leqq x_n \leqq x_1 = \sqrt{k}$ for all n. Now if $k > 0$, then the condition $x_3 \geqq 0$ is equivalent to $k^3 - 2k^2 + k - 1 \geqq 0$, and elementary analysis suffices to show that this holds if and only if $k \geqq k_0 \approx 1.7548777$. We now show that if $k \geqq k_0$, then $\lim_{n\to\infty} x_n$ exists. Since from $x_m \leqq x_n$ it follows that $x_{m+2} \geqq x_{n+2}$, an easy induction yields

$$0 \leqq x_3 \leqq x_4 \leqq x_7 \leqq x_8 \leqq x_{11} \leqq x_{12} \leqq \cdots \leqq x_{10} \leqq x_9 \leqq x_6 \leqq x_5 \leqq x_2 \leqq x_1;$$

hence

$$y_1 \leqq y_4 \leqq y_5 \leqq y_8 \leqq y_9 \leqq y_{12} \leqq y_{13} \leqq \cdots \leqq y_{11} \leqq y_{10} \leqq y_7 \leqq y_6 \leqq y_3 \leqq y_2.$$

Convergence is guaranteed by an estimate of the form

$$|x_{n+1} - x_n| \leqq \alpha |x_{n-1} - x_{n-2}| \quad \text{for } n \geqq 5 \quad \text{and} \quad k \geqq k_0,$$

where α is a constant in $(0,1)$, since for even n the nested interval theorem can be applied. Indeed, for $n \geqq 5$ and $k \geqq k_0$ we have

$$|x_{n+1} - x_n| = \frac{|y_{n-1} - y_n|}{x_{n+1} + x_n} = \frac{|x_{n-1} - x_{n-2}|}{(x_{n+1} + x_n)(y_{n-1} + y_n)}$$

and $(x_{n+1} + x_n)(y_{n-1} + y_n) \geqq 2x_4 \cdot 2y_4$. Now $y_3 \geqq y_1 = \sqrt{k} \geqq \sqrt{k_0} > \frac{1}{2}$ implies that

$$\frac{dx_2}{dk} = \frac{1}{2x_2}\left(1 - \frac{1}{2y_1}\right) > 0$$

and

$$\frac{dx_4}{dk} = \frac{1}{2x_4}\left(1 - \frac{1}{2y_3}\left(1 + \frac{1}{2x_2}\left(1 - \frac{1}{2y_1}\right)\right)\right) > \frac{1}{2x_4}\left(1 - \frac{1}{2y_3}\right)\left(1 - \frac{1}{2x_2 \cdot 2y_3}\right).$$

Hence x_2 increases as k increases from k_0, so that $2x_2 \cdot 2y_3 \geqq 2x_2^* \cdot 2y_1 > 1$, where the star denotes the value for $k = k_0$. Thus $dx_4/dk > 0$ for $k \geqq k_0$, and so $2x_4 \cdot 2y_4 \geqq 4x_4^* y_1 > 7/3$, and we may choose α to be $3/7$.

CHICO PROBLEM GROUP
California State University

Also solved partially by G. A. Heuer, L. Kuipers (Switzerland), David Lindsay, Richard Parris, and Daniel A. Rawsthorne. In addition, there were twelve seriously incomplete or incorrect solutions.

None of the solutions submitted was completely correct. In most of the incomplete solutions it was shown only that *if* the sequence converges to A, *then* $k = A^2 + A + 1$, as the proposer had intended, before the editors strengthened the problem. Richard Parris, as well as David F. Paget (Australia), used Cardan's formula to solve the cubic equation $k^3 - 2k^2 + k - 1 = 0$ for k, obtaining

$$k_0 = \left[(25 + 3\sqrt{69})/54\right]^{1/3} + \left[(25 - 3\sqrt{69})/54\right]^{1/3} + \tfrac{2}{3}$$

and the corresponding value for A_0. Related, but simpler, problems can be found in *National Mathematics Magazine*, v. 9 (1934–5), pp. 208–210, 252, problems 75, 78.

Divisibility of a Combinatorial Sum — September 1983

1175. Suppose that $m = nq$, where n and q are positive integers. Prove that the sum of binomial coefficients

$$\sum_{k=0}^{n-1} \binom{(n,k)q}{(n,k)}$$

is divisible by m, where (x, y) denotes the greatest common divisor of x and y. [*Anon, Erewhon-upon-Spanish River.*]

Solution II: Since

$$\binom{(n,k)q}{(n,k)} = q\binom{(n,k)q-1}{(n,k)-1},$$

q divides each term of the sum and thus divides the sum. Hence it needs to be shown only that

$$n \text{ divides } \sum_{k=0}^{n-1}\binom{(n,k)q-1}{(n,k)-1}.$$

For that, it is sufficient to show that if p is a prime and $p\nmid a$, then

$$p^e \text{ divides } \sum_{k=0}^{p^e a-1}\binom{(p^e a,k)q-1}{(p^e a,k)-1} \quad \text{for } e \geqq 1.$$

Now the only possible values of $(p^e a, k)$ are divisors of $p^e a$; indeed, each divisor d of $p^e a$ occurs as the value of $(p^e a, k)$ exactly $\phi(p^e a/d)$ times, where ϕ is Euler's totient function. Thus

$$\sum_{k=0}^{p^e a-1}\binom{(p^e a,k)q-1}{(p^e a,k)-1} = \sum_{d\mid p^e a}\binom{dq-1}{d-1}\phi\left(\frac{p^e a}{d}\right) = \sum_{b\mid a}\left(\sum_{j=0}^{e}\binom{p^j bq-1}{p^j b-1}\phi(p^{e-j})\right)\phi\left(\frac{a}{b}\right).$$

The solution will be complete provided it can be shown that if $(p,b)=1$, then

$$p^e \text{ divides } \sum_{j=0}^{e}\binom{p^j bq-1}{p^j b-1}\phi(p^{e-j}).$$

Now

$$\sum_{j=0}^{e}\binom{p^j bq-1}{p^j b-1}\phi(p^{e-j}) = \sum_{j=0}^{e-1}\binom{p^j bq-1}{p^j b-1}(p^{e-j}-p^{e-j-1}) + \binom{p^e bq-1}{p^e b-1}$$

$$= p^e\binom{bq-1}{b-1} + \sum_{j=1}^{e} p^{e-j}\left\{\binom{p^j bq-1}{p^j b-1} - \binom{p^{j-1}bq-1}{p^{j-1}b-1}\right\}.$$

Hence it is sufficient to show that

$$\binom{p^j bq-1}{p^j b-1} \equiv \binom{p^{j-1}bq-1}{p^{j-1}b-1} \pmod{p^j} \quad \text{for } j=1,2,\ldots,e.$$

To see this, note that

$$\binom{p^j bq-1}{p^j b-1} = \prod_{t=1}^{p^j b-1}\frac{p^j bq-t}{p^j b-t} = \left(\prod_{\substack{p\nmid t\\ 1\leqq t\leqq p^j b-1}}\frac{p^j bq-t}{p^j b-t}\right)\prod_{s=1}^{p^{j-1}b-1}\frac{p^j bq-sp}{p^j b-sp}.$$

The second product on the right, after p is canceled from each factor, is seen to be

$$x = \binom{p^{j-1}bq-1}{p^{j-1}b-1}.$$

Since $p\nmid b$, the numerator y and the denominator z of the first product on the right are each the product of the elements of b reduced residue systems modulo p^j, and consequently $y \equiv z \pmod{p^j}$, $yx \equiv zx \pmod{p^j}$, and, since yx/z is an integer and z is relatively prime to p^j, $yx/z \equiv x \pmod{p^j}$. This establishes the required congruence, and so the solution is complete.

J. M. METZGER
University of North Dakota

Solution I appeared in the September 1984 issue.

A Euler Product **November 1983**

1178. Evaluate

$$\lim_{n\to\infty}\left[\prod_{i=1}^{n}\left(a+\frac{i-1}{n}\right)\right]^{1/n},$$

where a is any positive constant. [*Russell Euler, Northwest Missouri State University.*]

Solution I: Let $g_n(a)=[\prod_{i=1}^{n}(a+(i-1)/n)]^{1/n}$. Then

$$\log g_n(a)=\frac{1}{n}\sum_{i=1}^{n}\log\left(a+\frac{i-1}{n}\right).$$

Thus, $\log g_n(a)$ is a Riemann sum for the logarithm function on the interval $[a,a+1]$. Hence we have

$$\log g_n(a)\to\int_a^{a+1}(\log x)\,dx=(a+1)\log(a+1)-a\log a-1$$

$$=\log\frac{(a+1)^{a+1}}{a^a}-1\text{ as }n\to\infty.$$

Since the exponential function is continuous, we have, finally,

$$g_n(a)\to\frac{(a+1)^{a+1}}{a^ae}\quad\text{as }n\to\infty.$$

BENJAMIN G. KLEIN
Davidson College

Solution II: If Γ denotes the gamma function, then

$$\prod_{i=1}^{n}\left(a+\frac{i-1}{n}\right)=n^{-n}\prod_{i=1}^{n}(na+i-1)=n^{-n}\frac{\Gamma(n(a+1))}{\Gamma(na)}.$$

From Stirling's formula (see E. Artin, *The Gamma Function*, ch. 3), we obtain

$$\Gamma(n(a+1))=\sqrt{2\pi}\,[n(a+1)]^{n(a+1)-1/2}\exp[-n(a+1)]\exp\frac{\theta_n}{12n(a+1)}$$

and

$$\Gamma(na)=\sqrt{2\pi}\,(na)^{na-1/2}\exp(-na)\exp\frac{\tilde{\theta}_n}{12na}$$

for some θ_n and $\tilde{\theta}_n$ in $(0,1)$. Therefore we have

$$\left[\prod_{i=1}^{n}\left(a+\frac{i-1}{n}\right)\right]^{1/n}=\frac{a+1}{e}\left(\frac{a+1}{a}\right)^a\left(\frac{a}{a+1}\right)^{1/(2n)}\exp\left(\frac{\theta_n}{12n^2(a+1)}-\frac{\tilde{\theta}_n}{12n^2a}\right),$$

which yields

$$\lim_{n\to\infty}\left[\prod_{i=1}^{n}\left(a+\frac{i-1}{n}\right)\right]^{1/n}=\frac{a+1}{e}\left(\frac{a+1}{a}\right)^a.$$

HEINZ-JÜRGEN SEIFFERT, student
Freie Universität Berlin

Also solved as in Solution I by Mangho Ahuja, P. J. Anderson (Canada), Jacques Bélair (Canada), David Boduch (student), F. Lee Cook, Charles R. Diminnie, Thomas E. Elsner, G. A. Heuer, Edwin T. Hoefer, Hans Kappus (Switzerland), M. S. Klamkin (Canada), L. Kuipers (Switzerland), J. C. Linders (The Netherlands), Vania D. Mascioni (student, Switzerland), Armel Mercier (Canada), J. M. Metzger, Roger B. Nelsen, William A. Newcomb, *David Paget (Australia), Richard Parris, Daniel Plotnick, Daniel M. Rosenblum, Vincent P. Schielack, Jr., Jens Schwaiger (Austria), Harry Sedinger, M. Selby (Canada), Dennis Spellman, Murray R. Spiegel, Anders Szepessy (Sweden), Charles H. Toll, Michael Vowe (Switzerland), William P. Wardlaw, H. G. Williams, and the proposer; and solved as in Solution II by Ralph Garfield, Chico Problem Group, Douglas Henkin (student), Jens Schwaiger (Austria, second solution), and Jan Söderkvist (student, Sweden). There was one incorrect solution.*

Mercier and Rosenblum generalized the problem. Not one solver mentioned the appropriateness of this problem for the issue in which it was proposed.

Minimum Polynomial and Rank — November 1983

1179. Let A be a square matrix of rank r. Show that the minimum polynomial of A has degree at most $r+1$. [*William P. Wardlaw, U.S. Naval Academy.*]

Solution I: Let the $n\times n$ matrix A have rank r. Then $A=BC$ for some matrices B and C, where B is $n\times r$ and C is $r\times n$. Thus the $r\times r$ matrix CB has minimum polynomial $m(x)$ of degree $d\leqq r$, and

$$Am(A)=BCm(BC)=Bm(CB)C=0.$$

Since A is a zero of the polynomial $xm(x)$ of degree $d+1$, the minimum polynomial of A has degree $\leqq d+1\leqq r+1$.

WILLIAM P. WARDLAW
U.S. Naval Academy

Solution II: Let A be an $n\times n$ matrix of rank r. Then A is similar to a matrix B of the form

$$B=\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix},$$

where A_1 is an $r\times r$ matrix and A_2 is an $r\times(n-r)$ matrix. Hence B has the same rank and minimum polynomial as A. It is easy to verify by induction that

$$B^n=\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}^n=\begin{bmatrix} A_1^n & A_1^{n-1}A_2 \\ 0 & 0 \end{bmatrix}$$

for every positive integer n. Let

$$p(x)=x^r+a_1x^{r-1}+\cdots+a_r$$

be the characteristic polynomial of A_1. We then have

$$\begin{aligned} Bp(B)&=\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}p\left(\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}\right)\\ &=\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} p(A_1) & \left(A_1^{r-1}+a_1A_1^{r-2}+\cdots+a_{r-1}I_r\right)A_2 \\ 0 & a_rI_{n-r} \end{bmatrix}\\ &=\begin{bmatrix} 0 & p(A_1)A_2 \\ 0 & 0 \end{bmatrix}=0. \end{aligned}$$

Therefore B satisfies the polynomial equation $xp(x)=0$, which is of degree $r+1$, and A does likewise. Hence the degree of the minimum polynomial of A is at most $r+1$.

YAN-LOI WONG, student
University of California, Berkeley

Also solved by Irl C. Bivens, David Boduch (student), Oscar A. Campoli (Argentina), Robert Gilmer, G. A. Heuer, L. R. King, Richard Parris, Daniel A. Rawsthorne, and Jens Schwaiger (Austria).

A Reduced Ring Problem November 1983

1180. Let R be an associative ring with no nonzero nilpotent elements (i.e., R is reduced). Show that if $x \in R$ has only finitely many distinct powers, then there exists an integer $n > 1$ such that $x = x^n$. [*Gary F. Birkenmeier, Southeast Missouri State University.*]

Solution I: Since x has only finitely many distinct powers, there are positive integers j and k with $j < k$ such that $x^j = x^k$. If $j = 1$, we are done; if not, we can show that $x^{j-1} = x^{k-1}$, and repetition of the argument yields $x = x^{k-j+1}$. Indeed, $(x^{j-1} - x^{k-1})^2 = x^{j+j-2} - 2x^{j+k-2} + x^{k+k-2} = x^{j+k-2} - 2x^{j+k-2} + x^{j+k-2} = 0$, since $x^{j+t} = x^{k+t}$ for all $t \geqq 0$. Since R has no nonzero nilpotents, we conclude that $x^{j-1} - x^{k-1} = 0$, or $x^{j-1} = x^{k-1}$. Note that it is enough to assume that no nonzero element of R has square equal to zero.

J. M. Metzger
University of North Dakota

Solution II: Let A denote the subring of R generated by x. Then A is a commutative ring with no nonzero nilpotent elements. Hence A is isomorphic to a subring of a direct sum of fields. (See Theorem 31, p. 123, of *Rings and Ideals* by Neal R. McCoy, MAA, 1948.) Suppose that $x^m = x^n$ for some positive integers m and n with $m > n$. Then $f_i^m = f_i^n$, where f_i is the projection of x onto the i-coordinate in a direct sum representation of A. Hence $f_i^{m-n+1} = f_i$ for each i, and so $x^{m-n+1} = x$.

Gregory P. Wene
The University of Texas at San Antonio

Also solved by Efraim P. Armendariz, David Boduch (student), B. J. Gardner (Australia), Robert Gilmer, Geoffrey A. Kandall, Paul Peck, Dennis Spellman, A. E. Spencer, and Fernand R. Tessier (Canada); partial solutions (under the assumption that R has an identity element) by Irl C. Bivens, Eric Crane Brody (student), Joel K. Haack, G. A. Heuer & Karl W. Heuer, Benjamin G. Klein, Richard Parris, Charles H. Toll, William P. Wardlaw, Gordon Williams, and the proposer. There was one incorrect solution.

Wardlaw cited the related *Monthly* problem 6284, v. 89 (1982), pp. 135–136. Most of the partial solutions used the binomial theorem in such a way that x^i appeared with $i = 0$.

Equal Areas and Centroid of a Triangle November 1983

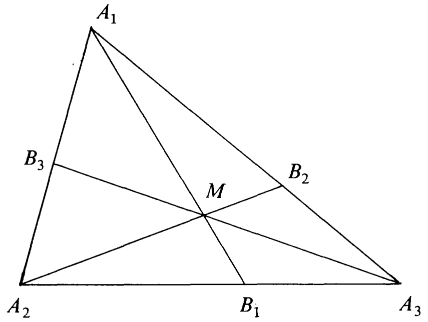
1181. Let $A_1A_2A_3$ be a triangle and M an interior point. The straight lines MA_1, MA_2, MA_3 intersect the opposite sides at the points B_1, B_2, B_3, respectively. Show that if the areas of triangles A_2B_1M, A_3B_2M, and A_1B_3M are equal, then M coincides with the centroid of triangle $A_1A_2A_3$. [*George Tsintsifas, Thessaloniki, Greece.*]

Solution I: Set

$$A_1B_3 : B_3A_2 = p:1, \quad A_2B_1 : B_1A_3 = q:1, \quad \text{and} \quad A_3B_2 : B_2A_1 = r:1.$$

Then p, q, and r are positive, and

$$\frac{A_1A_2B_1}{A_1B_1A_3} = \frac{q}{1} = \frac{MA_2B_1}{MB_1A_3}.$$



Hence

$$\frac{q}{1} = \frac{A_1A_2M}{A_1MA_3} = \frac{\frac{p+1}{p}A_1B_3M}{\frac{r+1}{r}A_3B_2M} = \frac{r(p+1)}{p(r+1)},$$

since $A_1B_3M = A_3B_2M$. Similarly,

$$\frac{r}{1}=\frac{p(q+1)}{q(p+1)} \quad \text{and} \quad \frac{p}{1}=\frac{q(r+1)}{r(q+1)}.$$

Multiplying the three equations together and simplifying yields

$$pqr=1. \tag{1}$$

(This may also be obtained from Ceva's theorem.) On the other hand, multiplying out each equation and using (1) yields

$$1+pq=rp+r, \quad 1+qr=pq+p, \quad \text{and} \quad 1+rp=qr+q,$$

and adding these now yields

$$p+q+r=3.$$

Comparison with (1) now shows that the arithmetic mean of p, q, r is equal to their geometric mean. Hence $p=q=r=1$, showing that M is the centroid of triangle $A_1A_2A_3$.

J. C. LINDERS
Eindhoven, The Netherlands

Solution II: Using barycentric coordinates, let $\mathbf{M}=x_1\mathbf{A}_1+x_2\mathbf{A}_2+x_3\mathbf{A}_3$, where x_1, x_2, and x_3 are positive and $x_1+x_2+x_3=1$. Then $x_1=[MA_2A_3]/[A_1A_2A_3]$, etc., where $[PQR]$ denotes the area of PQR. Also,

$$\mathbf{B}_1=\frac{x_2\mathbf{A}_2+x_3\mathbf{A}_3}{x_2+x_3}, \text{ etc.}$$

Since $\overline{A_2B_1}/\overline{A_2A_3}=x_3/(x_2+x_3)$, we have

$$\frac{[A_2MB_1]}{[A_1A_2A_3]}=\frac{x_3x_1}{x_2+x_3}, \text{ etc.}$$

By hypothesis we have

$$\frac{x_3x_1}{x_2+x_3}=\frac{x_1x_2}{x_3+x_1}=\frac{x_2x_3}{x_1+x_2},$$

or

$$x_2(1-x_1)=x_3(1-x_2)=x_1(1-x_3).$$

Without loss of generality, we assume that $x_1 \geqq x_2 \geqq x_3$. Then $x_1(1-x_3)=x_3(1-x_2) \leqq x_3(1-x_3)$, and so $x_1 \leqq x_3$. Thus $x_1=x_2=x_3$ and M is the centroid.

MURRAY S. KLAMKIN
University of Alberta

Also solved by Mangho Ahuja, J. C. Binz (Switzerland), W. J. Blundon (Canada), Kenneth Fogarty, Geoffrey A. Kandall, L. Kuipers (Switzerland), Richard Parris, Jan Söderkvist (student, Sweden), J. M. Stark, Michael Vowe (Switzerland), and the proposer. There was one interesting incorrect solution.

Answer

Solution to the Quickie on p. 299.

Q695. No, because skew-congruent triangles have the same circumradius $BC/(2\sin\angle A)$. [*John Horton Conway, Cambridge University.*]

REVIEWS

Paul J. Campbell, Editor
Beloit College

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

David, Edward E., Jr., *Renewing U. S. mathematics*, Science 224 (15 June 1984).

The president of Exxon Research and Engineering summarizes and supports the conclusions of the National Research Council's report of the same title. "In 1983 there were as many mathematicians as physicists or chemists in American universities, but only some 60 postdoctoral students in mathematics were receiving federal support compared with about 1200 in physics and 2500 in chemistry" (not to mention the resulting differential effect in starting salaries in academia). Despite two years of upturns, federal support of mathematics is only about two-thirds of its 1968 level (in constant dollars). One can ask how this defunding came about; more important is whether the mathematics constituencies in the sciences, engineering, and government will continue to relegate mathematics, as a low national priority, to chronic underfunding.

Thurston, William P., and Weeks, Jeffrey R., *The mathematics of three-dimensional manifolds*, Scientific American 251:1 (July 1984) 108-120, 138.

In the 19th century mathematicians "classified" all two-manifolds by showing that each can be represented as a polygon with certain edges identified. Thurston now seems close to classifying all three-manifolds; naturally, the latter classification is a bit less simple. Perforce, much of this beautiful article is devoted to examples of two-manifolds in an attempt to develop intuition about three-manifolds; but an ingenious analogy with linkage systems proves quite worthwhile.

Kac, Mark, *Marginalia: more on randomness*, American Scientist 72:3 (May-June 1984) 282-283.

Cites Bertrand's paradox, which concerns chords drawn "at random" in a circle, to show that the phrase "at random" does not have an absolute meaning. Moreover, the question "What is random?" has no answer, as Kac illustrates by means of the results of experiment in physics.

Mathis, Frank H., and Turner, Danny W., *Estimating the length of material wrapped around a cylindrical core*, SIAM Review 26 (1984) 263-266.

Presents three methods ("models") for estimating the amount of carpet remaining on a roll, together with conditions under which the estimates agree to within one per cent. Only calculus is used ("hard" integration, Taylor series). This classroom note resulted from a request by an inventory manager at the authors' campus, where "many" rolls of carpet are in storage!

Kolata, Gina, *Esoteric math has practical result*, Science 225 (3 August 1984) 494-495.

An efficient new method to generate realistic fractal computer pictures of natural objects has been developed by M. Shahshahani (Boeing). His method relies on the "zeta averaging" of P. Diaconis (Stanford), which among other facts accounts for the "first-digit phenomenon" (that the density of numbers with first digit k is $\log(k+1)$). Shahshahani's method uses a Markov chain of affine transformations, and the analysis of the corresponding stationary distribution involves the random infinite sums investigated by Diaconis in 1973. Diaconis had felt that his results were so esoteric that he did not publish them. Here, then, is another story of mathematics pursued for its own sake that has found an unexpected and significant use.

Kolata, Gina, *The art of learning from experience*, Science 225 (13 July 1984) 156-158.

Profile of Bradley Efron (Stanford) and the "new statistics," particularly his technique of the "bootstrap." The bootstrap involves multiple resampling from a universe made up of replicates of the original sample, in an effort to assess variability of that original sample. Current work centers on learning the limitations of this new technique.

Fishburn, Peter C., *Discrete mathematics in voting and group choice*, SIAM Journal of Algebraic and Discrete Methods 5 (1984) 263-275.

Illustrates the diversity of mathematics that has been brought to bear on the analysis and design of voting procedures and group decision processes: discrete ranking structures, nested hierarchies of sign functions, finite topologies, combinatorial impossibility theorems, integer optimization problems, linear separation lemmas, and matters of probability--not to mention paradoxes, axiom systems, and recurrences. A good guide to applications in political science for a course in discrete mathematics.

Campbell, Douglas M., and Higgins, John C., Mathematics: People, Problems, Results, 3 vols. Wadsworth, 1984; xvi + 871 pp, $39.95, $27.95 (P).

In the 30 years since James R. Newman's four-volume colossal collection, The World of Mathematics appeared, much good writing about mathematics has been done. An excellent selection of it is compiled in this new collection, as "an introduction to the spirit of mathematics... to give the nonmathematician some insight into the nature of mathematics and those who create it." The editors: "The truth is that we have no particular message to get across... . We would not presume to tell the readers what mathematics is. Let them read the articles, listen to the voices, and form their own opinions." Regrettably, there is no index.

Schiffer, M. M., and Bowden, L., The Role of Mathematics in Science, MAA, 1984; xi + 207 pp, $14 (P) (discount to members).

This volume begins with an absolute gem concerning the laws of the lever and the inclined plane. Further chapters illuminate growth, optics, time, relativistic motion, and energy. Calculus is presumed, and transformation matrices are used freely after an explanatory chapter.

Guillen, Michael, Bridges to Infinity: The Human Side of Mathematics, Houghton Mifflin, 1983; 204 pp, $12.95.

Seventeen breezy and interesting essays on important themes and areas in modern mathematics. There are no equations and--surprisingly--no figures. Nonmathematicians should be inspired without being overtaxed. (There is a mistake on p. 50 regarding the Continuum Hypothesis.)

Herr, Albert, and Staib, Jack, Mathematical Experiments, (A. Herr, Dept. of Math., Drexel University, 32nd & Chestnut Sts., Philadelphia, PA 19104), 1983; 62 pp + unlocked diskette (for Apple II+, IIe), free (send blank diskette + $2 postage).

Pedagogically ingenious laboratory experiences to enhance the learning of algebra, trigonometry, and analytic geometry. The user must repeat each experiment to gather data, draw graphs, estimate from the graphs, do computations to a specified accuracy, and answer questions on the worksheet. The experiments, actually simulations, include: object falling down well, lines tangent to circles, hyperbola (via sound waves), dipstick (estimating volume of liquid stored in a cylindrical tank), sliding box (inclined plane with friction), and hockey pucks (elastic collisions). A program is provided to calibrate the other programs for the aspect ratio of the user's screen. Extremely well designed, the package would be of great value in physics classes as well.

Wattenberg, Frank, and Wattenberg, Marvin, Interactive Experiments in Calculus, Prentice-Hall, 1984; vii + 143 pp + locked diskette (for Apple II+, IIe), (P).

Calculus supplement containing most of the standard sorts of programs. One unusual program allows successive zooming in on the behavior of a function at a point. The diskette is protected by extra cardboard inside the front cover, but the software is strictly "black box" (no algorithms or source code in the text). Although a printer-customizing program is included, there is no program to correct for the aspect ratio of the monitor used.

Williams, Gareth, and Williams, Donna, Linear Algebra Computer Companion: Student Manual, Allyn and Bacon, 1984; ii + 212 pp + locked program diskette + data diskette (for Apple IIe or Apple II with Applesoft in ROM), $35 (P).

One solution to the logistic, ethical, and financial problems of software use in a course is for each student to buy a "low-priced" copy. The publisher profits from volume and repeat sales to new classes, the students get to keep the tool they have used, and the institution is relieved of having to purchase and protect a small number of expensive copies. Allyn and Bacon is trying out this solution with its companion to G. Williams' Linear Algebra with Applications, which can be used as a supplement in any linear algebra course. However, tucking disks in an envelope inside the back cover of a paperback is inviting disaster (the data diskette of the review copy was unusable). Unfortunately, the software is strictly "black box": no algorithms or source code in the manual. How can you trust it?

Brown, Stephen I., and Walter, Marion I., The Art of Problem Posing, Franklin Institute Pr, 1983; 147 pp, $11.95 (P).

Formulates strategies for generating conjectures, asserting that problem-posing and problem-solving are intimately intertwined. The examples are taken from high-school mathematics, and the authors hope that an emphasis on problem-posing will enhance creativity and reduce anxiety in the classroom.

Gastel, Barbara, Presenting Science to the Public, ISI Pr, 1983; x + 146 pp, $17.95, $11.95 (P).

"In particular, beware of presenting ideas in mathematical terms." The author's sound advice may seem paradoxical and self-defeating when applied to communicating mathematical ideas and discoveries. Conveying mathematics to the public is difficult but immensely important; this book offers advice on all aspects of that task, from interviews with reporters, to working with a public information office, to writing for a popular science magazine. Concerning the last, there has never been so vast an opportunity (at up to a million readers per issue and 50¢ or more per word) for promoting understanding of mathematics.

Patternson, Elizabeth Chambers, Mary Somerville and the Cultivation of Science, 1815-1840, Kluwer Boston, 1983; xiv + 264 pp, $39.50.

Detailed biography of the woman who made an important contribution to the modernization of British mathematics with her translation of Laplace's Mècanique cèleste. Despite her mathematical talent, the support of her husband for her studies and the honors bestowed on her for her pioneering accomplishments, she remained convinced that women had no creative intellectual powers. Belying that conviction, she was friend and tutor to Ada Lovelace, stimulating her creativity.

Hodges, Andrew, Alan Turing: the Enigma, Simon & Schuster, 1983; 587 pp, $24.95.

Frank and sensitive biography of a man who admitted "a lack of reverence for everything except the truth." Turing's discoveries in mathematics and computing, and the outward circumstances of his life, have been related before. Only now has it become possible to discuss his role in deciphering the German Enigma messages, or to plumb the other passion of his life, homosexuality. This portrait reveals a brilliant mind, a courageous man, and a tragic life.

Barnette, David, Map Coloring, Polyhedra, and the Four-Color Problem, MAA, 1983; x + 168 pp, $26 (discount available to members).

This latest in the Dolciani series of Mathematical Expositions leads up to the Four Color Theorem by way of the necessary graph theory. The exposition is at an easier level than Don Crowe's chapter on polyhedra in Beck, Bleicher, and Crowe's Excursions into Mathematics. There are exercises, along with solutions.

Thompson, Thomas M., From Error-Correcting Codes Through Sphere Packings to Simple Groups, MAA 1983; xii + 228 pp, $21 (discount available to members).

Both exposition and history, this newest volume in the Carus Monograph series traces connections: from Hamming's invention of error-correcting codes, to sphere-packing and the Leech lattice, and finally to the last sporadic simple groups being discovered inside the Leech lattice. The book benefits greatly from the author's efforts at interviewing the participants of their recollections of the circumstances of their contributions. The subject and level make the book an ideal text for a seminar for senior undergraduates, who would appreciate seeing contemporary mathematics and its inherent interconnections.

van der Waerden, B. L., Geometry and Algebra in Ancient Civilizations, Springer-Verlag, 1983; xii + 223 pp, $29.50.

The author, the dean of historians of mathematics and astronomy, argues the startling thesis: "The similarity [of mathematics of ancient civilizations] is so close that a pre-Babylonian common origin must be assumed... which is most faithfully reflected in the Chinese texts of the Han-period." Key similarities involve the Pythagorean theorem and Pythagorean triples. Historians and anthropologists, however, object to diffusionism and prefer to assume independent invention; they will criticize the lack of relevant extant pre-Babylonian sources and identified routes of transmission. Even mathematicians will argue with some of van der Waerden's conclusions, but his more than 40 years of research in history of mathematics and astronomy give authority to his judgment.

Weil, André, Number Theory: An Approach through History, from Hammurapi to Legendre, Birkhäuser, 1984; xxi + 375 pp.

Critical examination of pre-Gauss number-theoretic works, necessarily focusing mainly on Fermat, Euler, Lagrange and Legendre. The reader experiences the rare feeling of being taken into the workshops of those masters and discovering both what they knew and what eluded their grasp.

NEWS

BIEBERBACH CONJECTURE PROVED

Another famous conjecture, unsolved for 68 years, has fallen. Last spring, Louis de Branges (Purdue University) announced that he had proved the Bieberbach conjecture, considered to be one of the most important in classical analysis. In the ensuing months, several mathematicians helped de Branges to correct, clean up, and shorten the proof from 300 to 10 pages.

Like Fermat's Theorem, the Bieberbach conjecture is disarmingly simple to state:

If a power series with complex coefficients

$$f(x) = z + a_2z^2 + a_3z^3 + \ldots + a_nz^3 + \ldots$$

converges for all z in the unit disk in the complex plane, and f(z) is one-to-one, then $|a_n| \leq n$.

More details can be found in a report by Paul Zorn (St. Olaf College) in the October 1984 issue of FOCUS, the MAA newsletter.

NOTES ON THE 1984 OLYMPIAD

(Much of the information for this report was supplied by Murray S. Klamkin, a leader of the USA team for the IMO. His full "Reflections on the International Mathematical Olympiads" can be found in his column "The Olympiad Corner", which appears in the Canadian mathematics journal CRUX MATHEMATICORUM. Dedicated to problem-solving, this journal is an excellent addition to any math department library. For information, write to the editor, Léo Sauvé, Algonquin College, 140 Main Street, Ottawa, Ontario, Canada K1S 1C2. New subscriptions ($22 Canadian, $20 U.S., 10 issues per year) are welcomed.

The 25th International Mathematical Olympiad was held June 29 - July 10, 1984 in Prague, Czechoslovakia, with a record 34 countries competing. Team size was six students from each country (except Luxemburg and Norway, with just one student each), for a total of 192 students participating. The six problems of the competition (which follow this report) were given on two consecutive days, three each day, and were assigned 7 points each, for a maximum score of 42. There were eight students who scored a perfect 42:

D. B. Mihov,	Bulgaria
K. Groger,	East Germany
D. Tataru,	Romania
A. Astrelin,	Soviet Union
K. Ignatiev,	Soviet Union
L. Orydoroga,	Soviet Union
D. Moews,	United States

Eleven of the contestants were girls, including K. Groger listed above. Although the official results of the competition are announced only for individuals, unofficial total scores for each country's team are compiled, and teams are ranked accordingly. (The top five were announced in the last issue of this *Magazine*.)

Members of the USA team, D. Davidson, D. Grabiner, J. Kahn, D. Moews, S. Newman, and M. Reid, were led by M.S. Klamkin and Andy Liu, both of the University of Alberta. Members of the Canadian team, M. Bradley, F. D'Ippolito, M. Molloy, M. Piotte, T. Vo Minh, and L. Yen were led by Leon Bowden, University of Victoria, and Edward Barbeau, University of Toronto.

Students did not spend all their time in Prague proving their mathematical mettle. They showed ingenuity in other ways, as evidenced by the following anecdote (supplied by Andy Liu). When teammate David Moews got stuck in a bathroom, Steve Newman used his compass to pick the lock. There was a plank which blocked the space beneath the door, and Mike Reid tried to pry it loose with a metal ruler. In came Jeremy Kahn who quipped "What is this? Euclidean destruction?"

Solutions to the 1984 Canadian and USA Olympiads are in this *Magazine*; solutions to the 1984 IMO will appear in our next issue.

25th INTERNATIONAL MATH OLYMPIAD PROBLEMS

1. Prove that

$$0 \leq yz + zx + xy - 2xyz \leq \frac{7}{27},$$

where x, y, z are non-negative real numbers for which $x + y + z = 1$.

2. Find one pair of positive integers a, b such that:

(1) $ab(a+b)$ is not divisible by 7,

(2) $(a+b)^7 - a^7 - b^7$ is divisible by 7^7.

Justify your answer.

3. In the plane two different points O,A are given. For each point X of the plane, other than O, denote by $a(X)$ the measure of the angle between OA and OX in radians, counterclockwise from OA ($0 \leq a(X) < 2\pi$). Let $C(X)$ be the circle with centre O and radius of length $OX + \frac{a(X)}{OX}$. Each point of the plane is colored by one of a finite number of colors. Prove that there exists a point Y for which $a(Y) > 0$ such that its color appears on the circumference of the circle $C(Y)$.

4. Let $ABCD$ be a convex quadrilateral such that the line CD is a tangent to the circle on AB as diameter. Prove that the line AB is a tangent to the circle on CD as diameter if and only if the lines BC and AD are parallel.

5. Let d be the sum of the lengths of all the diagonals of a plane convex polygon with n vertices ($n > 3$), and let p be its perimeter. Prove that

$$n - 3 < \frac{2d}{p} < \left[\frac{n}{2}\right]\left[\frac{n+1}{2}\right] - 2.$$

($[x]$ denotes the greatest integer not exceeding x.)

6. Let a, b, c, d be odd integers such that $0 < a < b < c < d$ and $ad = bc$. Prove that if $a + d = 2^k$, $b + c = 2^m$ for some integers k and m, then $a = 1$.

SOLUTIONS TO CANADIAN AND USA MATH OLYMPIADS

The solutions to the 1984 Canadian and USA Math Olympiads which follow were especially prepared for publication in this Magazine *by Loren Larson and Bruce Hanson, St. Olaf College. A report on the 1984 USAMO and IMO, containing problems and solutions is available from W.E. Mientka, 917 Oldfather Hall, Univ. of Nebraska, Lincoln, NE 68588.*

16th CANADIAN OLYMPIAD

1. Prove that the sum of the squares of 1984 consecutive positive integers cannot be the square of an integer.

Sol. For arbitrary integers $m \geq 0$ and $n > 0$,

$$\sum_{i=1}^{n} (m+i)^2 = \sum_{i=1}^{n} (m^2 + 2mi + i^2)$$

$$= nm^2 + mn(n+1) + \frac{n(n+1)(2n+1)}{6}.$$

When $n = 1984 = 64 \times 31$, the bottom sum is congruent to 32 modulo 64. But $x^2 \equiv 32 \bmod 64$ has no solutions, which completes the proof.

2. Alice and Bob are in a hardware store. The store sells coloured sleeves that fit over keys to distinguish them. The following conversation takes place.

Alice: Are you going to cover your keys?

Bob: I would like to, but there are only 7 colours and I have 8 keys.

Alice: Yes, but you could always distinguish a key by noticing that the red key next to the green key was different from the red key next to the blue key.

Bob: You must be careful what you mean by "next to" or "three keys over from" since you can turn the key ring over and the keys are arranged in a circle.

Alice: Even so, you don't need 8 colours.

Problem: What is the smallest number of colours needed to distinguish n keys if all the keys are to be covered?

Sol. Let $C(n)$ denote the smallest number of colors needed to distinguish n keys. It is easy to argue that $C(1) = 1$, $C(2) = 2$, $C(3) = 3$, $C(4) = 3$, $C(5) = 3$. For $n \geq 6$, $C(n) = 2$, since the coloring sequence (reading clockwise) $B\ R\ B\ B\ R\ R\ \ldots\ R$ will suffice (R = red, B = blue).

3. An integer is *digitally divisible* if (a) none of its digits is zero and (b) it is divisible by the sum of its digits (e.g., 322 is digitally divisible).

Show that there are infinitely many digitally divisible integers.

Sol. Let A_n be the number whose decimal representation consists of a sequence of 3^n consecutive 1's. For $n \geq 1$, $A_{n+1} = A_n(1 + 10^{3^n} + 10^{2\cdot 3^n})$.

Note that $1 + 10^{3^n} + 10^{2\cdot 3^n}$ is divisible by 3 (the sum of its digits is 3). Using these facts it is easy to show, by induction, that $A_n \equiv 0 \bmod 3^n$ for $n \geq 1$.

4. An acute-angled triangle has unit area. Show that there is a point inside the triangle whose distance from each of the vertices is at least

$$\frac{2}{\sqrt[4]{27}}.$$

Sol. Let C be a circle. We will show that of all triangles inscribed in C the equilateral triangle has maximum area. Let ΔPQR be inscribed in C with $PQ \neq QR$. It is easy to show that there is an inscribed isoceles triangle, with base PR whose area is greater than the area of ΔPQR. Therefore the inscribed triangle with maximum area, (which must exist by continuity considerations), must be the equilateral triangle.

Now let T be an acute triangle of unit area, let D be its circumcircle, and let E be an equilateral triangle inscribed in D. Then

$1 = \text{Area}(T) \leq \text{Area}(E)$, and $\dfrac{\text{Area}(D)}{\text{Area}(E)} = \dfrac{4\pi}{\sqrt{27}}$. Thus, $\text{Area}(D) \geq \dfrac{4\pi}{\sqrt{27}}$, and therefore the center of D is $\geq 2/\sqrt[4]{27}$ from each vertex of T. Since the center of D lies inside T, the solution is complete.

5. Given any 7 real numbers, prove that there are two of them, say x and y, such that

$$0 \leq \frac{x - y}{1 + xy} \leq \frac{1}{\sqrt{3}}.$$

Sol. Denote the seven given real numbers by $y_1, \ldots, y_7$, and let

$x_i = \text{Arctan}\, y_i \in (-\pi/2, \pi/2)$

for $i = 1, \ldots, 7$. Divide the interval $[-\pi/2, \pi/2]$ into six equal subintervals, each of length $\pi/6$. By the pigeonhole principle, two of these, x_i and x_j, $x_i \geq x_j$, lie in the same subinterval; that is

$$0 \leq x_i - x_j \leq \pi/6.$$

Since the tangent function is increasing on $(-\pi/2, \pi/2)$, it follows that

$$\tan 0 \leq \tan(x_i - x_j) \leq \tan(\pi/6),$$

or equivalently,

$$0 \leq \frac{y_i - y_j}{1 + y_i y_j} \leq \frac{1}{\sqrt{3}}.$$

13th USA MATH OLYMPIAD

1. The product of two of the four roots of the quartic equation

$$x^4 - 18x^3 + kx^2 + 200x - 1984 = 0$$

is -32. Determine the value of k.

Sol. Let r_1, r_2, r_3, r_4 be the four roots of the given quartic equation. We know that

$$r_1 + r_2 + r_3 + r_4 = 18, \qquad (1)$$

$$r_1r_2 + r_1r_3 + r_1r_4 + r_2r_3 + r_2r_4 + r_3r_4 = k \qquad (2)$$

$r_1r_2r_3 + r_1r_2r_4$

$$+ r_1r_3r_4 + r_2r_3r_4 = -200 \quad (3)$$

$$r_1r_2r_3r_4 = -1984, \quad (4)$$

and from the hypothesis of the problem, we may assume that $r_1r_2 = -32$ and from (4), $r_3r_4 = 62$. Using this, we can write equations (1), (2), and (3) in the form

$$(r_3 + r_4) + (r_1 + r_2) = 18, \quad (5)$$

$$(r_1 + r_2)(r_3 + r_4) + 30 = k, \quad (6)$$

$$-32(r_3 + r_4) + 62(r_1 + r_2) = -200. \quad (7)$$

Equations (5) and (7) imply that $r_1 + r_2 = 4$ and $r_3 + r_4 = 14$. Thus, from equation (6), $k = 86$.

2. The geometric mean of any set of m non-negative numbers is the m-th root of their product.

(i) For which positive integers n is there a finite set S_n of n distinct positive integers such that the geometric mean of any subset of S_n is an integer?

(ii) Is there an infinite set S of distinct positive integers such that the geometric mean of any finite subset of S is an integer?

Sol.

(i) Let n be any positive integer, and take $S_n = \{1^{n!}, 2^{n!}, \ldots, n^{n!}\}$. Then S_n has the required property.

(ii) There is no infinite set with the stated property. For suppose S is such a set, and let a and b be distinct elements of S. Pick a positive integer m so that $(a/b)^{1/m}$ is an irrational number, and choose m-1 additional elements, $r_2, \ldots, r_m$ from S. Then $(ar_2 \ldots r_m)^{1/m}$ and $(br_2 \ldots r_m)^{1/m}$ are integers, so that

$$\left(\frac{a}{b}\right)^{1/m} = \frac{(ar_2 \ldots r_m)^{1/m}}{(br_2 \ldots r_m)^{1/m}}$$

is rational. This contradiction shows that such an infinite set S cannot exist.

3. P, A, B, C and D are five distinct points in space such that

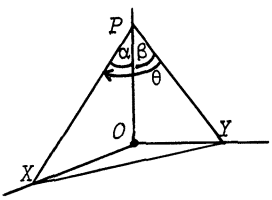
$$\angle APB = \angle BPC = \angle CPD = \angle DPA = \theta,$$

where θ is a given acute angle. Determine the greatest and least values of $\angle APC + \angle BPD$.

Sol. Suppose A, B, C, D are situated so that lines PA and PC coincide and PB and PD coincide. Then $\angle APC + \angle BPD = 0$ and this is the least value of the sum.

$$\text{Let } \alpha = \tfrac{1}{2}\angle APC \text{ and } \beta = \tfrac{1}{2}\angle BPD.$$

If α (or β) is zero, the maximum value of $\alpha + \beta$ is 2θ. So suppose $\alpha > 0$ and $\beta > 0$. The planes determined by points A, P, C and B, P, D are orthogonal, and intersect in a line which we can identify with the z-axis. Identify a point O on this axis as the origin, with $|OP| = 1$, and points X on the x-axis and Y on the y-axis, so that $\angle XPO = \angle APO = \alpha$, $\angle YPO = \angle BPO = \beta$.



By the law of cosines, using $\Delta\, XPY$, $|XY|^2 = |XP|^2 + |YP|^2 - 2|XP||YP|\cos\theta$ $= (1+|OX|^2) + (1+|OY|^2) - |XP||YP|\cos\theta$ $= 2 + |XY|^2 - 2|XP||YP|\cos\theta$. From this equation it follows that

$$1 = |XP||YP|\cos\theta,$$

or equivalently,

$$\cos\alpha \cos\beta = \cos\theta. \quad (1)$$

Because cosine is decreasing on $[0,\pi]$, the sum $\alpha + \beta$ is a maximum when $\cos(\alpha + \beta)$ is a minimum. By expanding $\cos(\alpha + \beta)$ and using (1) twice, one can show that $\cos(\alpha + \beta)$ is a minimum if and only if $\cos^2\alpha + \cos^2\beta$ is a minimum. By the arithmetic mean - geometric mean inequality,

$$\cos^2\alpha + \cos^2\beta \geq 2 \cos^2\alpha \cos^2\beta = 2\cos\theta.$$

It follows that the minimum value of $\cos^2\alpha + \cos^2\beta$ is $2\cos\theta$, and this minimum occurs when $\alpha = \beta$. Thus, the largest value of $\angle APC + \angle BPD$ occurs when $\cos\alpha = \cos\theta$, and is equal to $4 \arccos\sqrt{\cos\theta}$.

4. A difficult mathematical competition consisted of a Part I and a Part II with a combined total of 28 problems. Each contestant solved exactly 7 problems altogether. For each pair of problems, there were exactly two constestants who solved both of them. Prove that there was a contestant who in Part I solved either no problems or at least 4 problems.

Sol. Let n denote the number of contestants. Each contestant solved $\binom{7}{2}$ pairs of problems, and each of the $\binom{28}{2}$ pairs of problems were solved by exactly two contestants. It follows that $n\binom{7}{2} = 2\binom{28}{2}$, and from this, we find that $n = 36$.

Let m_1, m_2, m_3 denote the number of contestants who solved exactly 1, 2, and 3 problems from Part 1 respectively.

Suppose, to the contrary, that

$$m_1 + m_2 + m_3 = 36. \qquad (1)$$

Let k denote the number of problems in Part I. By counting first the number of pairs of problems that consist of two problems from Part I, and then those pairs that consist of one problem from Part I, we find that

$$m_2 + 3m_3 = 2\binom{k}{2} \qquad (2)$$

and

$$6m_1 + 10m_2 + 12m_3 = 2k(28 - k). \qquad (3)$$

Substituting $m_1 = 36 - m_2 - m_3$ into (3) yields

$$2m_2 + 3m_3 = -108 + 28k - k^2. \qquad (4)$$

Subtracting (2) from (4) yields

$$m_2 = -108 + 29k - 2k^2. \qquad (5)$$

It is straightforward to check that $-108 + 29k - 2k^2$ is negative for all k, and therefore, from (5), $m_2 < 0$. But this is contrary to our interpretation of m_2. This contradiction means that $m_1 + m_2 + m_3 < 36$ and the solution is complete.

5. $P(x)$ is a polynomial of degree $3n$ such that

$$P(0) = P(3) = \dots = P(3n) = 2,$$
$$P(1) = P(4) = \dots = P(3n-2) = 1,$$
$$P(2) = P(5) = \dots = P(3n-1) = 0,$$

and $P(3n+1) = 730$.

Determine n.

Sol. Define the polynomial

$$L_i(x) = \frac{x(x-1)\dots\widehat{(x-i)}\dots(x-(3n+1))}{(i-0)(i-1)\dots\widehat{(i-i)}\dots(i-(3n+1))},$$

where the factors with the "hats" are omitted. Then $L_i(j) = \delta_{ij}$, so that

$$P(x) = \sum_{i=0}^{3n+1} L_i(x)\, P(i) \quad .$$

Our problem is to determine n so that the degree of $P(x)$ is $3n$. The coefficient of x^{3n+1} is

$$\sum_{i=0}^{3n+1} \frac{(-1)^{3n+1-i}}{i!(3n+1-i)!} P(i) \quad ,$$

so we need to find n such that

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

For each positive integer n, and for $k = 0, 1, 2$, let

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

Equation (1) can be manipulated into the form

$$S_1^{3n+1} = -243, \text{ if } n \text{ is odd}, \qquad (2)$$

or

$$S_1^{3n+2} = -1458, \text{ if } n \text{ is even}. \qquad (3)$$

By using the recursions

$$S_0^{n+1} = S_0^n - S_2^n ,$$
$$S_1^{n+1} = S_1^n - S_0^n ,$$
$$S_2^{n+1} = S_2^n - S_1^n ,$$

it is easy to prove, by induction, that

$$S_1^{n+6} = (-3)^3\, S_1^n,\ n \geq 1.$$

Using this recurrence we find that equation (2) holds only if $3n+1 = 12$. But this yields $n = 11/3$ which is impossible. Also, equation (3) holds only when $3n+2 = 14$, or equivalently, when $n = 4$. Thus, the n we seek is $n = 4$.

ACKNOWLEDGEMENTS

In addition to our associate editors, the following have assisted the Magazine *by refereeing papers during the past year. We appreciate the time and care they have given.*

Andrews, George E., *The Pennsylvania State University.*
Aspray, William, *The University of Minnesota.*
Barksdale, James B., Jr., *Western Kentucky University.*
Baxley, John V., *Wake Forest University.*
Beaver, Olga R., *Williams College.*
Beineke, Lowell, *The University of Indiana - Purdue at Ft. Wayne.*
Blundon, W. J., *Memorial University.*
Boas, Ralph P., *Northwestern University.*
Braden, Bart, *Northern Kentucky University.*
Brown, Walter, *Moravian College.*
Burr, Stefan A., *City College of N.Y.*
Callahan, James, *Smith College.*
Chakerian, Don, *University of California, Davis.*
Coleman, Roger, *Ursinus College.*
Connelly, Robert, *Cornell University.*
Cundy, H. Martyn, *Kendal, England.*
Cunningham, R. S., *California State College - Stanislaus.*
Davison, T. M. K., *McMaster University.*
Dershem, Herbert L., *Hope College.*
Dixon, Crist, *York College.*
Drucker, Thomas, *University of Wisconsin - Madison.*
Dunham, Douglas, *University of Minnesota - Duluth.*
Edelman, Paul H., *University of Pennsylvania.*
Eisenberg, Bennett, *Lehigh University.*
Emerson, John D., *Middlebury College.*
Epp, Susanna, *DePaul University.*
Feldman, Arnold, *Franklin & Marshall College.*
Finney, Ross L., *Newton, Massachusetts.*
Firey, William J., *Oregon State University.*
Fishburn, Peter C., *Bell Laboratories, Murray Hill.*
Fisher, J. Chris, *University of Regina.*
Flanders, Harley, *Florida Atlantic University.*
Foster, Lorraine L., *California State University, Northridge.*
Garner, Cyril, *Carleton University.*
Getchell, Charles L., *Lycoming College.*
Gilbert, William J., *University of Waterloo.*
Gingrich, Ross, *Moravian College.*
Goldstein, Jerome A., *Tulane University.*
Green, Euline, *Abilene Christian University.*
Groemer, Helmut, *University of Arizona.*
Grosshans, Frank D., *West Chester University.*
Guinand, Andrew, *Trent University.*
Guy, Richard, *University of Calgary.*
Hailperin, Theodore, *Lehigh University.*
Hanson, Denis, *University of Regina.*
Harrell, Ronald E., *Allegheny College.*
Hawkins, William, *University of the District of Columbia.*
Henson, C. Ward, *University of Illinois at Urbana - Champaign.*
Herron, Isom H., *Howard University.*
Hill, Victor E., *Williams College.*
Hilton, Peter, *State University of New York - Binghamton.*
Honsberger, Ross A., *University of Waterloo.*
Jackson, Bradley, *San Jose State University.*
Kalman, Dan, *Augustana College.*
Kelly, Leroy, *Michigan State University.*
Klamkin, Murray S., *University of Alberta.*
Klarner, David, *University of Nebraska.*
Koppelman, Elaine, *Goucher College.*
Lagarias, Jeffrey, *Bell Laboratories, Murray Hill.*
Landau, Martin D., *Lafayette College.*
Larson, Loren C., *St. Olaf College.*
Leinbach, Carl, *Gettysburg College.*
Lenz, Gerald E., *St. John's University.*
L'Heureux, James E., *West Chester State College.*
Liu, A., *University of Alberta.*
Logothetti, David, *University of Santa Clara.*
Loud, Warren S., *University of Minnesota.*
Luker, Mark, *University of Minnesota - Duluth.*
Martens, Susan, *York College.*
Martin, George E., *State University of New York at Albany.*
McCleary, John, *Vassar College.*
Monson, Barry, *University of New Brunswick.*
Moore, David S., *Purdue University.*
Moore, L. C., Jr., *Duke University.*
Murphy, Michael G., *University of Houston - Downtown College.*
Niven, Ivan, *University of Oregon.*
Notz, William I., *Purdue University.*
Paine, Dwight, *Messiah College.*
Parker, Jeffrey D., *Amherst College.*

Rebman, Kenneth R., *Redwood City, California.*
Ross, Peter, *University of Santa Clara.*
Rosenstein, George M., Jr., *Franklin & Marshall College.*
Saaty, Thomas L., *University of Pittsburgh.*
Sallee, G. T., *University of California.*
Salwach, Chester J., *Lafayette College.*
Sandefur, James T., *Georgetown University.*
Sawka, John, *University of Santa Clara.*
Schaumberger, Norman, *Bronx Community College.*
Schrag, Gerald C., *Central Missouri State University.*
Scully, Dan, *St. John's University.*
Senechal, Marjorie, *Smith College.*
Shuchat, Alan, *Wellesley College.*
Sibley, Thomas Q., *St. John's University.*
Siegel, Martha J., *Towson State University.*
Small, Lynne B., *University of San Diego - Alcala Park.*
Somers, Kay, *Moravian College.*
Spiegel, Murray, *East Hartford, Connecticut.*
Stapleton, James H., *Michigan State University.*
Stephens, Stanley L., *Anderson College.*
Straffin, Philip D., Jr., *Beloit College.*
Sved, Marta, *University of Adelaide.*
Vanden Eynden, Charles L., *Illinois State University.*
Vardeman, Stephen, *Iowa State University.*
Vargas, Luis G., *University of Pittsburgh.*
Wagner, Clifford, *Pennsylvania State University.*
Warner, Hoyt D., *Western New England College.*
Wilansky, Albert, *Lehigh University.*
Wilf, Herbert S., *The University of Pennsylvania.*
Wilker, John B., *Cambridge University.*
Yuster, Thomas, *Lafayette College.*

INDEX

AUTHORS

TITLES

PROBLEMS

P, S, and Q *refer to Proposals, Solutions and Quickies, respectively. Page numbers are given in parentheses. Thus,* S1175,I (239) *refers to solution number* 1175 *which appears on page* 239.

NEWS

U.S. Postal Service

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

Required by 39 U.S.C. 3685)

1A. TITLE OF PUBLICATION: MATHEMATICS MAGAZINE	1B. PUBLICATION NO.: M A M G A 8	2. DATE OF FILING: Sept. 12, 1984
3. FREQUENCY OF ISSUE: bi-monthly except July and August	3A. NO. OF ISSUES PUBLISHED ANNUALLY: Five	3B. ANNUAL SUBSCRIPTION PRICE: Library $28.00; Indiv. Member $11.00

4. COMPLETE MAILING ADDRESS OF KNOWN OFFICE OF PUBLICATION *(Street, City, County, State and ZIP+4 Code) (Not printers)*
1529 Eighteenth St., N.W., Washington, D.C. 20036

5. COMPLETE MAILING ADDRESS OF THE HEADQUARTERS OF GENERAL BUSINESS OFFICES OF THE PUBLISHER *(Not printer)*
1529 Eighteenth St., N.W., Washington, D.C. 20036

6. FULL NAMES AND COMPLETE MAILING ADDRESS OF PUBLISHER, EDITOR, AND MANAGING EDITOR *(This item MUST NOT be blank)*

PUBLISHER *(Name and Complete Mailing Address)*
The Mathematical Association of America, 1529 Eighteenth St., N.W., Washington, D.C. 20036

EDITOR *(Name and Complete Mailing Address)*
Doris Schattschneider, Moravian College, Bethlehem, PA 18018

MANAGING EDITOR *(Name and Complete Mailing Address)*
Harry Waldman, MAA, 1529 Eighteenth St., N.W., Washington, D.C. 20036

7. OWNER *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given. If the publication is published by a nonprofit organization, its name and address must be stated.) (Item must be completed.)*

FULL NAME	COMPLETE MAILING ADDRESS
The Mathematical Association of America, Inc.	1529 Eighteenth St., N.W. Washington, D.C. 20036

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES *(If there are none, so state)*

FULL NAME	COMPLETE MAILING ADDRESS
None	

9. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES *(Section 423.12 DMM only)*
The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes *(Check one)*

(1) [X] HAS NOT CHANGED DURING PRECEDING 12 MONTHS
(2) [ ] HAS CHANGED DURING PRECEDING 12 MONTHS
(If changed, publisher must submit explanation of change with this statement.)

10. EXTENT AND NATURE OF CIRCULATION *(See instructions on reverse side)*	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NO. COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A. TOTAL NO. COPIES *(Net Press Run)*	11,064	10,858
B. PAID AND/OR REQUESTED CIRCULATION 1. Sales through dealers and carriers, street vendors and counter sales	100	100
2. Mail Subscription *(Paid and/or requested)*	8,968	8,455
C. TOTAL PAID AND/OR REQUESTED CIRCULATION *(Sum of 10B1 and 10B2)*	9,068	8,555
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	131	131
E. TOTAL DISTRIBUTION *(Sum of C and D)*	9,199	8,686
F. COPIES NOT DISTRIBUTED 1. Office use, left over, unaccounted, spoiled after printing	1,865	2,172
2. Return from News Agents	none	none
G. TOTAL *(Sum of E, F1 and 2 - should equal net press run shown in A)*	11,064	10,858

11. **I certify that the statements made by me above are correct and complete**
SIGNATURE AND TITLE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER: Alfred B. Willcox, Executive Director

PS Form 3526, July 1984 *(See instruction on reverse)*

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

MATHEMATICS MAGAZINE VOL. 57, NO. 5, November 1984