EDITORIAL POLICY

Mathematics Magazine aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for the *Magazine*. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Manuscripts on history are especially welcome, as are those showing relationships among various branches of mathematics and between mathematics and other disciplines.

A more detailed statement of author guidelines appears in this *Magazine*, Vol. 83, at pages 73–74, and is available at the *Magazine*'s website www.maa.org/pubs/mathmag.html. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, or published by another journal or publisher.

Please submit new manuscripts by email directly to the editor at mathmag@maa.org. A brief message containing contact information and with an attached PDF file is preferred. Word-processor and DVI files can also be considered. Alternatively, manuscripts may be mailed to Mathematics Magazine, 132 Bodine Rd., Berwyn, PA 19312-1027. If possible, please include an email address for further correspondence.

Cover image by Amy Uyeki

MATHEMATICS MAGAZINE (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Hanover, PA, bimonthly except July/August.
The annual subscription price for *MATHEMATICS MAGAZINE* to an individual member of the Association is $131. Student and unemployed members receive a 66% dues discount; emeritus members receive a 50% discount; and new members receive a 20% dues discount for the first two years of membership.)

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

Advertising correspondence should be addressed to

MAA Advertising
1529 Eighteenth St. NW
Washington DC 20036

Phone: (866) 821-1221
Fax: (202) 387-1208
E-mail: advertising@maa.org

Further advertising information can be found online at www.maa.org

Change of address, missing issue inquiries, and other subscription correspondence:

MAA Service Center, *maahq@maa.org*

All at the address:

The Mathematical Association of America
1529 Eighteenth Street, N.W.
Washington, DC 20036

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

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

Printed in the United States of America

CONTENTS

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

The Mathematical Association of America
1529 Eighteenth Street, NW
Washington, DC 20036

Vol. 83, No. 4, October 2010

MATHEMATICS MAGAZINE

LETTER FROM THE EDITOR

We begin with two articles about applications—to polls and to populations.

Andrew Gelman and Nate Silver tell us how to predict the result of a presidential election during the first hours of the vote count. Their model uses the final polling results for each state, and takes into account the correlations among the separate polling "errors" (which might include last-minute swings). Thus, learning that one state has moved in a particular direction can increase the probability that other states have moved in the same direction. From the earliest reported results, they construct ever-tighter probability distributions for the final electoral vote.

Is there a reader who can apply the same methods to the 2010 congressional elections? Your goal would be to say which party controls each house by, say, 9 or 10 PM on November 2. Post your results ahead of the networks and win fame and fortune!

Population dynamics is the subject of an article by a team from Richard Rebarber's REU group at the University of Nebraska. They describe population models for plant and animal species. Many of us have seen transition matrices used for this purpose, when populations can be partitioned naturally into age or size classes. But what if the distribution is naturally continuous? This team shows how integral models can be used in these cases, and how they compare to their discrete analogs.

The third article, by William Adkins and Mark Davidson, gives us a computational method that might be useful in many applications. How many of us, seeing the linear system $y' = Ay$, have wanted to write the answer $y = \exp(At)$ and be done with it? Alas, computing $\exp(At)$ requires serious attention!

Nicole Oresme reappears on page 327. His work in the 14th century helped us to understand balls rolling down inclined planes, but he might have been alarmed by the device that Stan Wagon describes in the Notes section. Elsewhere in the Notes section you can find the name of an effective middle-school teacher, quotations from Euclid and Gauss, and the first proof that the real numbers are uncountable (not diagonalization!). If you are looking for the answer to the puzzle in our last issue about a Tower-of-Hanoi graph, start on page 257.

The Olympiads This issue contains the problems and solutions for both the USAMO and the IMO. We also have the problems and solutions for the USAJMO—the USA Junior Mathematical Olympiad, for students in 10th grade and below—which was given for the first time this year. The USAMO, USAJMO, and US participation in the IMO are programs of the MAA, carried out by the staff of the MAA's Lincoln, Nebraska office and volunteers led by the MAA's Committee on the American Mathematical Competitions.

These features have appeared in the MAGAZINE every year, but this year we are bringing them to you more quickly. Thanks to the authors, who met tight deadlines.

The Allendoerfer Awards We are also pleased to honor the winners of the 2010 Carl B. Allendoerfer Awards, for articles published in this MAGAZINE during 2009. The winners are Ezra Brown, Keith Mellinger, David Speyer, and Bernd Sturmfels.

Walter Stromquist, Editor

ARTICLES

Structured Population Dynamics: An Introduction to Integral Modeling

JOSEPH BRIGGS
North Carolina State University
Raleigh, NC

KATHRYN DABBS
University of Tennessee
Knoxville, TN

MICHAEL HOLM
JOAN LUBBEN
RICHARD REBARBER
BRIGITTE TENHUMBERG
University of Nebraska
Lincoln, NE

DANIEL RISER-ESPINOZA
Swarthmore College
Swarthmore, PA

Will an exotic species thrive in a new territory? What are the best management options to eradicate a population (pest species) or to facilitate population recovery (endangered species)? Population modeling helps answer these questions by integrating mathematics and biology.

Often, a single species cannot be properly modeled as one population, but instead is best treated as a *structured* population, where the individuals in the population are partitioned into classes, or *stages*. As an example of a stage structured population, it is natural to partition an insect population into egg, larva, pupa, and adult stages. The choice of the stages and the breakdown of the population into stages depend heavily on the type of population, and are informed by biological intuition. For instance, fecundity (number of offspring per capita) in animals often varies with age, while in plants, fecundity typically depends on size. This implies that for mammals, the stages might be best determined by age, so that age is a good *stage variable* for mammals, while size might be a good stage variable for plants. Furthermore, for many animals there are natural classes of ages—the egg/larva/pupa/adult partition of an insect population—while for many plants, the stages can be better described as a continuous function of stem diameter, or another indicator of size. When the stages are discrete, a matrix model is used, and when the stages are continuous, an integral model is used. Both integral and matrix models are commonly used in population viability analysis and are both important tools in guiding population management [**4, 19**]. These models are used to predict long-term and transient behavior of a population, and they inform wildlife managers about which populations are in danger of going extinct or of growing unacceptably large.

Another basic modeling choice is whether time is modeled as a discrete variable or a continuous variable. Field data is often collected at regular time intervals, for

Math. Mag. **83** (2010) 243–257. doi:10.4169/002557010X521778.

instance on a yearly or seasonal basis, so it is often easier and more practical to model time discretely. There is some controversy about the relative merits of discrete-time versus continuous-time modeling [**7**]. Nonetheless, in most of the ecological literature on single-species structured populations, time is modeled as a discrete variable, so in this article we also model time as a discrete variable.

For a population that is partitioned into finitely many stages and modeled at discrete times, the evolution of the population can often be described using a Population Projection Matrix (PPM). The entries in a PPM are determined by the life history parameters of the population, and the properties of the matrix—for instance, its spectrum—determine the behavior of the solutions of the model. In the next section we describe PPMs in detail.

When stages are described by a continuous variable, one can either maintain the continuous stage structure, or partition the continuous range of stages into a finite number of stages. The latter is called a discretization of the population. To do it effectively one must ensure that each stage consists of individuals with comparable growth, survival, and fecundity, because the accuracy of the approximation depends on the similarity of individuals within each stage class. In general, a large number of life history stages increases model accuracy, but at the cost of increasing parameter uncertainty, since each nonzero matrix entry needs to be estimated from data, and the more stages there are, the less data is available per stage. This tradeoff can often be avoided by maintaining the continuous structure, and using an Integral Projection Model (IPM) that uses continuous life history functions that are functions of a continuous range of stages. We discuss IPMs in detail below.

In this article we illuminate the differences and similarities between matrix population models and integral population models for single-species stage structured populations. We illustrate the use of integral models with an application to Platte thistle, following Rose et al. [**22**], showing how the model is determined by basic life history functions. PPMs are ubiquitous in ecology, but for many purposes an IPM might be easier and/or more accurate to use. In TABLE 1 we summarize the similarities between PPMs and IPMs. In order to compare the predictions for PPMs and IPMs, enough data must be available to find the parameters in both models. This is done for models for the plant monkshood in Easterling et al. [**9**]. We should mention that if time is treated as a continuous variable, the analogue of a PPM model is an ordinary differential equation, and the analogue of a IPM is an integro-differential equation.

Matrix models

Matrix models were introduced in the mid 1940s, but did not become the dominant paradigm in ecological population modeling until the 1970s. The modern theory is described in great detail in Caswell [**4**], which also contains a good short history of population projection matrices in its Section 2.6. We summarize some of this history here. The basic theory of describing, predicting, and analyzing population growth by analyzing life history parameters such as survival and fecundity can be traced back to Cannan [**3**] in 1895. Matrix models in particular were developed independently by Bernardelli [**2**], Lewis [**16**], and Leslie [**15**]. The latter is most relevant to the modern theory. P. H. Leslie was a physiologist and self-taught mathematician, who, while working at the Bureau of Animal Population at Oxford between 1935 and 1968, synthesized mortality and fertility data into single models using matrices. We briefly describe his basic models, which are still used for population description, analysis, and prediction.

Although he was highly regarded and well connected in the ecology community, Leslie's work in matrix modeling initially received little attention. One of the few

TABLE 1: Comparison of matrix and integral models

	Population Projection Matrix			**Integral Projection Model**	
vector entry	$n(i,t)$	number of individuals in stage class i at time t	continuous function	$\int_{y_0}^{y_1} n(y,t)\,dy$	number of individuals expected between sizes y_0 and y_1
state vector	$\mathbf{n}(t) = [n(1,t), \cdots, n(m,t)]^T \in \mathbb{R}^m$	stage distribution of population at time t	continuous state function	$n(\cdot,t) \in \mathbb{L}^1(m_s, M_s)$	stage distribution of population at time t
probability	p_{ij}	probability of an individual transitioning from class j to i	probability density function	$\int_{y_0}^{y_1} p(y,x)\,dy$	probability an individual of size x will grow and survive to a size between y_0 and y_1
scalar	f_{ij}	number of newborns size i from parents size j	function	$\int_{y_0}^{y_1} f(y,x)\,dy$	number of newborns between sizes y_0 and y_1 from parents of size x
matrix entry	$k_{ij} = p_{ij} + f_{ij}$	the ijth entry of the transition matrix	function	$k(y,x) = p(y,x) + f(y,x)$	kernel
matrix	$A = [k_{ij}]$		integral operator	$(Av)(y) = \int_{m_s}^{M_s} k(y,x)v(x)\,dx$	
discrete stage variables	$j \sim t$ and $i \sim t+1$	matrix indices associated with time t and time $t+1$	continuous stage variables	$x \sim t$ and $y \sim t+1$	variables associated with time t and time $t+1$
difference equation	$n(j,t+1) = \sum_{i=1}^{m} k_{ji} n(i,t)$		integral equation	$n(y,t+1) = \int_{m_s}^{M_s} k(y,x)n(x,t)\,dx$	
vector form	$\mathbf{n(t+1) = An(t)}$	matrix multiplication	operator form	$\mathbf{n}(t+1) = A\mathbf{n}(t)$	integration

contemporaries who did use the matrix model was Leonard Lefkovitch. He also implemented a matrix model [**14**], but with an innovation: The populations were partitioned into classes based on developmental stage rather than age. This made the method more applicable to plant ecologists, who began defining stage classes by size rather than age—a change that usually resulted in better predictions.

As Caswell points out [**4**], it took some 25 years for the ecology community to adopt matrix projection models after Leslie's influential work. There were two major reasons for this delay. The ecology community at that time thought of matrix algebra as an advanced and esoteric mathematical subject. More importantly, there was a more accessible method, also contributed by Leslie, called life table analysis [**4**, Section 2.3].

Before the widespread use of computers, there was no information that a matrix model could provide that a life table could not. This would change as more sophisticated matrix algebra and computation methods emerged to convince ecologists of the worth of matrix models. For instance, using elementary linear algebra, one can predict asymptotic population growth rates and stable stage distributions from the spectral properties of the matrix. Also, the use of eigenvectors facilitated the development of *sensitivity and elasticity analyses*, giving an easy way to determine how small changes in life history parameters effect the asymptotic population growth rate. This is an especially important question for ecological models, which are typically very uncertain. Sensitivity and elasticity analyses are sometimes used to make recommendations about which stage class conservation managers should focus on in order to increase the population growth rate of an endangered species.

Transition matrices To set up a matrix model we start with a population partitioned into m stage classes. Let $t \in \mathbb{N} = \{0, 1, 2, \dots\}$ be time, measured discretely, and let $\mathbf{n}(t)$ be the population column vector

$$\mathbf{n}(t) = [n(1,t), n(2,t), \dots, n(m,t)]^T,$$

where each entry $n(i,t)$ is the number of individuals belonging to class i at time t. A discrete-time matrix model takes the form

$$\mathbf{n}(t+1) = \mathbf{A}\mathbf{n}(t), \tag{1}$$

where $\mathbf{A} = (k_{ij})$ is the $m \times m$ PPM containing the life-history parameters. It is also called a *transition matrix*, since it dictates the demographic changes occurring over one time step. We can write (1) as

$$n(i, t+1) = \sum_{j=1}^{m} k_{ij} n(j,t), \quad i = 1, \dots n. \tag{2}$$

The entry k_{ij} determines how the number of stage j individuals at time t affects the number of stage i individuals at time $t+1$. This is the form we will generalize when we discuss integral equations.

In their simplest form, the entries of $\mathbf{A}$ are survivorship probabilities and fecundities. What we call a Leslie matrix has the form

$$\mathbf{A} = \begin{pmatrix} f_1 & f_2 & \cdots & f_{m-1} & f_m \\ p_1 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & \cdots & 0 & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & p_{m-1} & 0 \end{pmatrix},$$

where p_i is the probability that an individual survives from age class i to age class $i+1$, and f_i is the fecundity, which is the per capita average number of offspring reaching stage 1 born from mothers of stage class i. The transition matrix has this particular structure when age is the stage class variable and individuals either move into the next class or die. In general, entries for the life-history parameters may appear in any entry of the m x m matrix $\mathbf{A}$.

For any matrix $\mathbf{A}$ and $t \in \mathbb{N}$, let $\mathbf{A}^t$ denote the tth power of $\mathbf{A}$ for any natural number t. It follows from (1) that

$$\mathbf{n}(t) = \mathbf{A}^t \mathbf{n}(0). \tag{3}$$

The long-term behavior of $\mathbf{n}(t)$ is determined by the eigenvalues and eigenvectors of $\mathbf{A}$. We say that $\mathbf{A}$ is nonnegative if all of its entries are nonnegative, and that $\mathbf{A}$ is primitive if for some $t \in \mathbb{N}$, all entries of $\mathbf{A}^t$ are positive. This second condition is equivalent to every stage class having a descendent in every other stage class at some time step in the future. PPMs are generally nonnegative and primitive, thus the following theorem is extremely useful [**23**, Section 1.1]:

PERRON-FROBENIUS THEOREM. *Let* $\mathbf{A}$ *be a square, nonnegative, primitive matrix. Then* $\mathbf{A}$ *has an eigenvalue,* λ*, known as the dominant eigenvalue, that satisfies:*

1. λ *is real and* $\lambda > 0$,
2. λ *has right and left eigenvectors whose components are strictly positive,*
3. $\lambda > |\tilde{\lambda}|$ *for any eigenvalue* $\tilde{\lambda}$ *such that* $\tilde{\lambda} \neq \lambda$,
4. λ *has algebraic and geometric multiplicity* 1.

This theorem is important in the analysis of population models because the dominant eigenvalue is the asymptotic growth rate of the modeled population, and its associated eigenvector is the asymptotic population structure. To see this, assume that $\mathbf{A}$ is primitive. Let $\mathbf{n} = [n_1, n_2, \dots, n_m]$, and $\|\mathbf{n}\|$ denote the ℓ_1 norm:

$$\|\mathbf{n}\| = |n_1| + |n_2| + \dots |n_m|. \tag{4}$$

Denote the unit eigenvector associated with λ by $\mathbf{v}$, so

$$\lim_{t\to\infty} \frac{\|\mathbf{n}(t+1)\|}{\|\mathbf{n}(t)\|} = \lambda \quad \text{and} \quad \lim_{t\to\infty} \frac{\mathbf{n}(t)}{\|\mathbf{n}(t)\|} = \mathbf{v}. \tag{5}$$

Thus as time goes on, the growth rate approaches λ and the stage structure approaches $\mathbf{v}$. In particular, the dynamics of a long-established population is described by λ and $\mathbf{v}$.

Problems with stage discretization To use a population projection matrix model, the population needs to be decomposed into a finite number of discrete stage classes that are not necessarily reflective of the true population structure. As mentioned previously, if stage classes are defined in such a way that there is at least one class in which the life history parameters vary considerably, then it might not be possible to accurately describe individuals in that stage class, which might result in erroneous predictions. Easterling [**8**] and Easterling et al. [**9**] give an example of such a "bad" partition of the population.

Fortunately it is often possible to decompose a particular population in a biologically sensible fashion. Vandermeer [**24**] and Moloney [**18**] have crafted algorithms to minimize errors associated with choosing class boundaries. Such algorithms help to derive more reasonable matrices, but for many populations they cannot altogether eliminate the sampling and distribution errors associated with discretization. For instance, for many plants size is the natural stage variable, and no decomposition of size into discrete stage classes will adequately capture the life history variations. Furthermore, sensitivity and elasticity analyses have both been shown to be affected by changes in stage class division, Easterling, et al. [**9**].

Regardless of how well the population is decomposed into stages, there is also the problem that in a matrix model individuals of a given stage class are treated as though they are identical through every time step. That is, two individuals starting in the same class will always have the same probability of transitioning into a different stage class at every time step in the future, which is not necessarily the case for real populations.

For many populations, these difficulties can be overcome by analyzing a continuum of stages, which is discussed in the next section.

Integral projection models

An alternate approach to discretizing continuous variables such as size is to use Integral Projection Models. These models retain much of the analytical machinery that makes the matrix model appealing, while allowing for a continuous range of stages. Easterling [**8**] and Easterling et al. [**9**] show how to construct such an integral projection model, using continuous stage classes and discrete time, and they provide sensitivity and elasticity formulas analogous to those for matrix models. In Ellner and Rees [**10**] an IPM analogue of the Perron-Frobenius Theorem is given. In particular, there are readily checked conditions under which such a model has an asymptotic growth rate that is the dominant eigenvalue of an operator whose associated eigenvector is the asymptotic stable population distribution.

Just as ecologists were slow to adopt matrix models, they have, so far, not used integral models widely. Stage structured IPMs of the type considered in this paper have appeared in the scientific literature since around ten years ago [**5, 6, 8, 9, 10, 11, 21, 22**]. There is a large literature on integral models for spatial spread of a population [**12, 13**]. The structure of the integral operators describing spatial spread can be very different from those for IPMs. For instance, the integral operators discussed in this paper are compact, while the operators describing spatial spread might not be compact. Compact operators have many properties that are similar to those of matrices [**1**, Chapter 17], and these properties make the spectral analysis, and hence the asymptotic analysis, more analogous to matrix models.

Continuous stage structure and integral operators Let $n(x, t)$ be the population distribution as a function of the stage x at time t. For example, if m_s is the minimum size, and M_s is the maximum size, as determined by field measurements, then $x \in [m_s, M_s]$ would be the size of an individual.

The analogue of the matrix entries $k_{i,j}$ for $i, j \in \{0, 1, \ldots m\}$ is a *projection kernel* $k(y, x)$ for $y, x \in [m_s, M_s]$, and the role of the matrix multiplication operation is analogous to an integral operator. The kernel is time-independent, which is analogous to the time-independent matrix entries. The time unit $t = 1$ represents a time interval in which data is naturally measured; in the example in this paper the unit of time is a year. The analogue of (2) is

$$n(y, t+1) = \int_{m_s}^{M_s} k(y, x) n(x, t)\, dx, \qquad y \in [m_s, M_s]. \tag{6}$$

In particular, the kernel determines how the distribution of stage x individuals at time t contributes to the distribution of stage y individuals at time $t + 1$, in much the same way that in (2) the (i, j)th entry of a projection matrix determines how an individual in stage j at time t contributes to stage i at time $t + 1$.

The kernel is determined by statistically derived functions for life history parameters such as survival, growth, and fecundity. At first the construction of an integral operator model might seem more difficult than the construction of a matrix model. However, the life history functions are assumed to have a particular distributional form, often with only a few parameters to be determined for each function. Hence the total number of parameters to be estimated can be smaller than the number of matrix entries. This of course would not work if the life history functions did not have an appropriate distributional form. Fortunately, ecologists have a toolbox of functional forms for different biological parameters. For instance, size is usually described by a lognormal distribution or truncated normal distribution. TABLE 2 shows all of the life history functions needed to construct the kernel for a particular integral projection model for the Platte thistle [**22**]. An advantage of the integral approach is that data over the entire distribution can be used to estimate the parameters of the life-history functions, thus minimizing parameter uncertainty. In contrast, the transitions between life history stages in matrix models are estimated from subsets of the data.

The stage variable x need not be a scalar, but the range of stage variables should be a compact metric space. In cases where x is not a scalar, the Riemann integration over a subset of $\mathbb{R}$ will be replaced by more general integration over a product space; see [**10**] for such an example.

Integral equations such as (6) can be analyzed in much the same way as matrix-based models of the form (1). Consider the L^1-norm

$$\|f\| := \int_{m_s}^{M_s} |f(x)|\, dx,$$

which is analogous to (4). The space

$$L^1(m_s, M_s) = \{f : (m_s, M_s) \to \mathbb{R} \mid \|f\| < \infty\}$$

is a complete normed linear space (that is, a *Banach space*). For every $t > 0$, the

TABLE 2: Life history functions for the Platte thistle [**22**], where variables x and y are in ln(crown diameter)

Demography	Equation
Survival	$s(x) = \dfrac{e^{-0.62+0.85x}}{(1+e^{-0.62+0.85x})}$
Flowering Probability	$f_p(x) = \dfrac{e^{-10.22+4.25x}}{(1+e^{-10.22+4.25x})}$
Growth Distribution	$g(x, y) =$ Normal Distribution in y with $\sigma^2 = 0.19$ and $\mu(x) = 0.83 + 0.69x$
Individual Seed Set	$S(x) = e^{0.37+2.02x}$
Juvenile Size Distribution	$J(y) =$ Normal Distribution with $\sigma_f^2 = 0.17$ and $\mu_f = 0.75$
Germination Probability	$P_e = .067$ density independent or $P_e = S_T(t)^{-0.33}$ density dependent where $S_T(t)$ is the total seed set

population distribution $n(\cdot, t)$ is in $L^1(m_s, M_s)$, and the total population is $\|\mathbf{n}(t)\|$. Hence $L^1(m_s, M_s)$ plays the same role in an IPM that $\mathbb{R}^m$ (with norm (4)) plays in a PPM.

For a population distribution $n(x, t)$, it is sometimes useful to distinguish between the function $n(x, t)$ of two variables and the $L^1(m_s, M_s)$-valued function of a single variable $\mathbf{n}(t) = n(\cdot, t)$; we refer to $\mathbf{n}(t)$ as a "vector" in $L^1(m_s, M_s)$. Define the operator $\mathbf{A} : L^1(m_s, M_s) \to L^1(m_s, M_s)$ by

$$(\mathbf{Av})(\cdot) := \int_{m_s}^{M_s} k(\cdot, x)\mathbf{v}(x)\, dx.$$

It is not hard to show that $\mathbf{A}$ is bounded on $L^1(m_s, M_s)$. In fact, since

$$\int_{m_s}^{M_s} \int_{m_s}^{M_s} |k(x, y)|^2\, dx\, dy < \infty,$$

it is well known that $\mathbf{A}$ is compact [**1**, p. 403], which implies that $\mathbf{A}$ has nice spectral properties, in a certain sense [**1**, Ch. 21]. Then (6) is equivalent to

$$\mathbf{n}(t + 1) = \mathbf{An}(t), \tag{7}$$

which is analogous to (1).

Ellner and Rees [**10**] show that for a large class of kernels k, the integral operator $\mathbf{A}$ satisfies an analog of the Perron-Frobenius Theorem for matrices. In particular, for a certain class of operators discussed [**10**, Appendix C], $\mathbf{A}$ has a dominant real eigenvalue λ that is the asymptotic growth rate and an associated unit eigenvector $\mathbf{v}$ that is the stable stage distribution. In this case the eigenvectors are functions in $L^1(m_s, M_s)$, rather than vectors in $\mathbb{R}^m$. Additionally

$$\lim_{t\to\infty} \frac{\|\mathbf{n}(t+1)\|}{\|\mathbf{n}(t)\|} = \lambda \quad \text{and} \quad \lim_{t\to\infty} \frac{\mathbf{n}(t)}{\|\mathbf{n}(t)\|} = \mathbf{v},$$

where the convergence of the second equation is interpreted as $L^1(m_s, M_s)$ convergence.

The kernel To construct the kernel, we construct a growth and survival function $p(y, x)$ and a fecundity function $f(y, x)$, and let

$$k(y, x) = p(y, x) + f(y, x).$$

Here $p(y, x)$ is the density of probability that an individual of size x will survive to be an individual of size y in one time step. Therefore, for each $y \in [m_s, M_s]$,

$$\int_{m_s}^{M_s} p(y, x)\, dx \leq 1.$$

The function $f(y, x)$ is the distribution for the number of offspring of size y that an individual of size x will produce in one time step. The fecundity function allows for the possibility of a seedling or newborn moving, in one time step, to a large size, but in practice the probability of this happening is virtually zero.

Estimating the kernel for Platte thistle We now show how a specific model is constructed, using a modification of the model for Platte thistle *(Cirsium canescens)* found in Rose et al. [**22**]. Platte thistle is an indigenous perennial plant in the midgrass sand prairies of central North America. The species is in decline in its native environment, possibly due to a biocontrol agent introduced to manage a different thistle, that is considered invasive. The time unit in this example is one year. It is strictly monocarpic, meaning that plants die after reproducing, so the flowering probability must be incorporated into the kernel. The Platte thistle lives 2–4 years [**17**]. In this model, the continuous class variables x and y are the natural log of the root crown diameter (measured in mm). The maximum and minimum root crown diameter are taken as $m_s = \ln(.5)$ and $M_s = 3.5$, respectively; we found that making M_s larger does not appreciably change the results. To best illustrate the basic concepts, we simplify the model by ignoring the effects of herbivores on fecundity and the possible slight effect of maternal size on offspring size.

We start with some component life-history functions. These are estimated from the data using standard statistical methods. For instance, logistic regression analysis can be used to describe survival as a function of size. Below is a description of these functions, and formulas are given in TABLE 2. All functions are defined for $x \in [m_s, M_s]$.

- $s(x)$ is the probability that a size x individual survives to the next time step. It is statistically fit to the logistic curve

$$s(x) = \frac{e^{ax+b}}{1 + e^{ax+b}},$$

  where $b < 0$.
- $f_p(x)$ is the probability that a size x plant will flower in one time step. This function is chosen to have the same logistic form as $s(x)$.
- $g(y, x)$ is the density of probability that an individual of size x will have size y at the next time step. This can describe both the probability of growing to a larger size and the probability of shrinking to a smaller size. The growth function $g(y, x)$ is a normal distribution in the variable y.
- $S(x)$ is the number of seeds produced on average per plant of size x. It is assumed to be an exponential function.
- $J(y)$ is the distribution of offspring sizes. It is assumed to be a normal distribution.
- P_e is the average probability that a seed will germinate. This is also known as the *recruitment probability*. We first assume that it is constant, but in a more realistic model it will be a function of the number of seeds.

Growth and survival kernel: To construct the growth and survival kernel, note that the probability that a size x individual does not flower is $1 - f_p(x)$. Since the Platte thistle dies after reproduction, the probability that a size x individual survives to the next time step is the survival probability $s(x)$ times the probability of not flowering, or $s(x)(1 - f_p(x))$. Hence the growth and survival kernel is

$$p(y, x) = s(x)(1 - f_p(x))g(y, x).$$

Fecundity kernel: Each plant will produce seeds, and these seeds must germinate for an offspring to be included in the next population count. For a Platte thistle to produce seeds, it must survive through a year and flower. Thus, each plant of root crown diameter size x will produce $s(x)f_p(x)S(x)$ seeds on average, so the total number of seeds resulting from a population distribution of $n(x,t)$ at time t is

$$S_T(t) = \int_{m_s}^{M_s} s(x) f_p(x) S(x) n(x,t) dx \tag{8}$$

and the total number of germinated seeds at time t is $P_e S_T(t)$. Finally, we also need to distribute the offspring into the various sizes by $J(y)$. The distribution of offspring at time $t+1$ resulting from a population distribution of $n(x,t)$ at time t is

$$P_e J(y) S_T(t) = P_e J(y) \int_{m_s}^{M_s} s(x) f_p(x) S(x) n(x,t)\, dx.$$

Therefore the fecundity kernel is

$$f(y,x) = P_e J(y) s(x) f_p(x) S(x). \tag{9}$$

FIGURE 1 shows a graph of the total kernel

$$k(y,x) = p(y,x) + f(y,x) = s(x)(1 - f_p(x)) g(y,x) + P_e J(y) s(x) f_p(x) S(x).$$

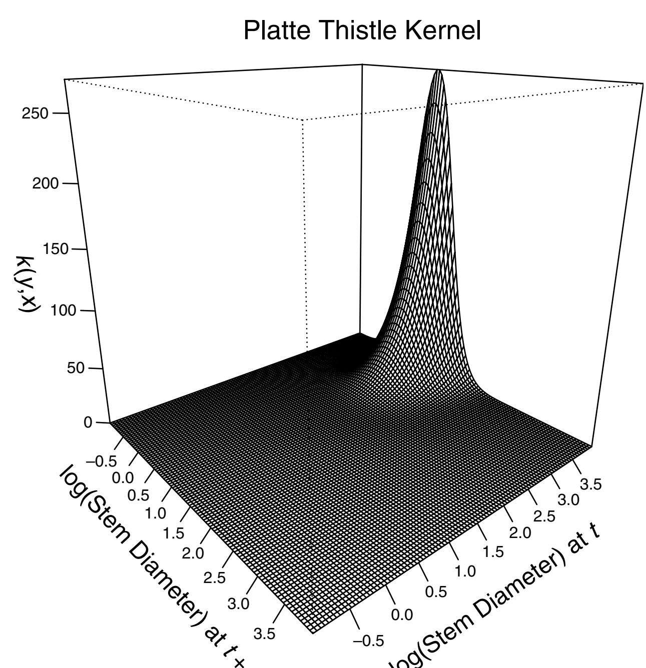


Figure 1 The kernel for the Platte thistle integral projection model

Numerical solution of the integrodifference equation Analytic evaluation of the integral operator is difficult if not impossible to perform. Thus, we use numerical integration to obtain an estimate of the population. A conceptually easy and reasonably accurate method is the midpoint rule. Let N be the number of equally sized intervals, and let $\{x_j\}$ be the midpoints of the intervals. Then

$$(\mathbf{An})(y,t) = \int_{m_s}^{M_s} k(y,x) n(x,t)\, dx \approx \frac{M_s - m_s}{N} \sum_{j=1}^{N} k(y,x_j) n(x_j,t). \tag{10}$$

Let

$$k_{ij} = \frac{M_s - m_s}{N} k(x_i, x_j) \text{ for } i, j = 1, 2, \ldots N, \quad \mathbf{A}_N = (k_{ij})$$

and

$$\mathbf{n}_N(t) = [n(x_1, t), n(x_2, t), \ldots n(x_N, t)]^T.$$

Then $\mathbf{n}_N(t)$ is a discrete approximation of $n(x, t)$, $\mathbf{A}_N$ is a discrete approximation of the integral operator $\mathbf{A}$, and

$$\mathbf{A}_N \mathbf{n}_N = \frac{M_s - m_s}{N} \sum_{j=1}^{N} k(x_i, x_j) n(x_j, t).$$

Since $k(x, y)$ is continuous, the Riemann sum uniformly approximates the integral as $N \to \infty$. Hence the integrodifference equation $\mathbf{n}(t + 1) = \mathbf{A}\mathbf{n}(t)$ can be approximated at the midpoints x_j by $\mathbf{n}_N(t + 1) = \mathbf{A}_N \mathbf{n}_N(t)$.

This matrix model can be analyzed much like a traditional matrix model. Since the dominant eigenvalue λ_N of $\mathbf{A}_N$ converges to the dominant eigenvalue λ of $\mathbf{A}$ as $N \to \infty$ [**10, 8**], the long term growth rate is easily estimated. FIGURE 2 shows this convergence of λ_N to $\lambda = 1.325$ as N increases. The leading eigenvalue of $\mathbf{A}_5$ is 1.332, so we see that fairly small dimensional approximations of $\mathbf{A}$ lead to very good approximations of the long-term growth of the system.

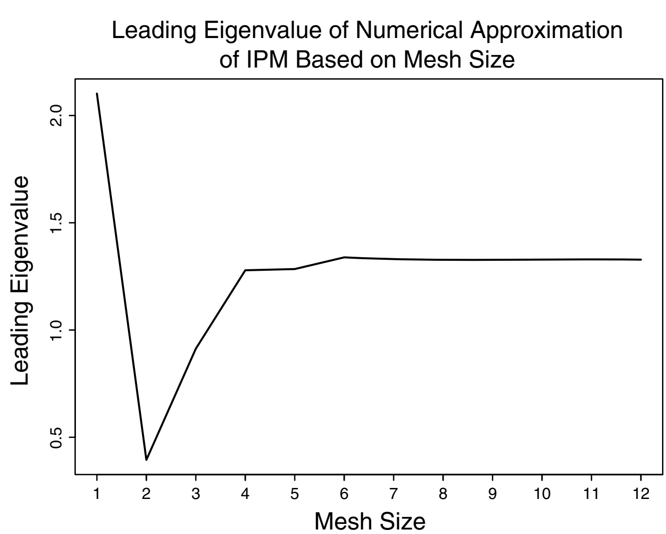


Figure 2 The leading eigenvector of the numerical approximation of the integral projection model as a function of number of subintervals in the Riemann sum

We should emphasize the difference between a PPM and the matrix model obtained from an IPM. In the former every nonzero entry is estimated directly; a large matrix of this type is not intended to approximate an IPM, and is subject to the discretization problems we described above. In the latter, the life history functions are estimated, giving rise to a kernel, and this kernel is used to obtain a matrix that approximates the integral operator for large N. As indicated above, an IPM is often preferable to a PPM, and in these cases the matrix model based on the IPM is also preferable to a PPM.

We now turn to the stable size distribution, that is, the limiting distribution given by the second equation in (5). This can be found by approximating the leading eigenvector of **A**, and normalizing it so that it has $L^1(m_s, M_s)$ norm of 1. This eigenvector is the curve labeled "Density Independent" in FIGURE 3. Note that the x-axis is in mm rather than ln(mm). The curve is obtained by computing the unit leading eigenvector of $\mathbf{A}_N$ for large N, and noting that this is a good approximation of the unit leading eigenvector [**10**].

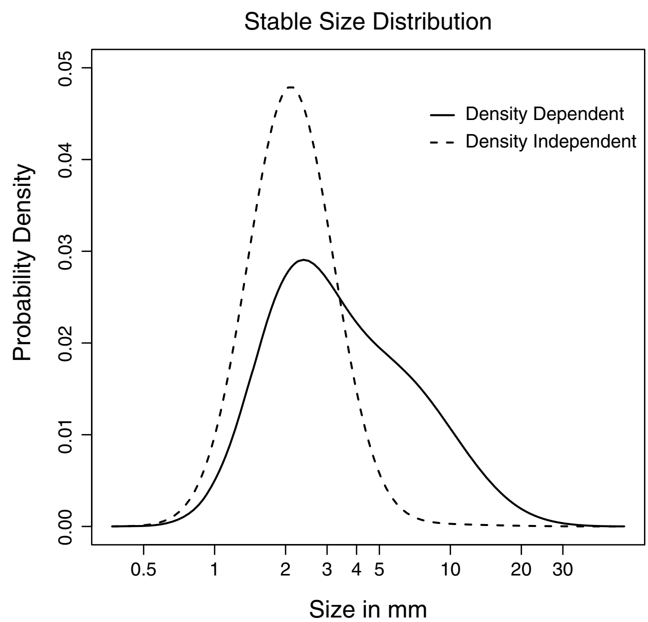


Figure 3 Stable state population densities of Platte thistle

Density dependence In the Platte thistle model above, we made the simplifying assumption that the germination probability, P_e, is constant, and obtained a *density independent* model. By "density independence" we mean that $\mathbf{n}(t+1)$ is a linear function of $\mathbf{n}(t)$, or equivalently, that the operator **A** does not depend upon $\mathbf{n}(t)$. Using the average germination probability, the growth rate of 1.325 we obtain from this model does not match the observed data. In particular, the data in Rose et al. [**22**] does not indicate that there is a constant growth rate, but rather shows a leveling off of the population over time. Furthermore, ecologists consider density dependent recruitment more realistic, since as the total number of seeds increases, the chance that each individual seed will germinate declines. Therefore, the germination probability is taken to be a nonlinear function of $S_T(t)$, the total number of seeds produced at time t, instead of a constant. Since the number of seeds produced depends on $n(x,t)$, the resulting system will be density dependent. In [**22**] the germination probability is modeled by $P_e(t) = (S_T(t))^{-.33}$. The resulting nonlinear system is

$$\begin{aligned} n(y,t+1) &= \int_{m_s}^{M_s} p(y,x)n(x,t)\,dx + J(y)(S_T(t))^{-.33}\int_{m_s}^{M_s} s(x)f_p(x)S(x)n(x,t)\,dx \\ &= \int_{m_s}^{M_s} p(y,x)n(x,t)\,dx + J(y)(S_T(t))^{.67}. \end{aligned}$$

The solutions to the resulting nonlinear system matches the data better than the solutions to the linear system.

This nonlinearity substantially changes the qualitative and quantitative nature of the model. For instance, as discussed above, in the linear model an asymptotic growth rate is determined by the leading eigenvalue and a stable age structure is determined by the eigenvector associated with the leading eigenvalue. We prove in another paper that for this nonlinear model the solutions $n(\cdot, t)$ converge in $L_1(m_s, M_s)$ as $t \to \infty$, and that this limit is independent of the initial population vector (provided that the initial population vector is nonzero) [**20**]. We denote the limit by $\mathbf{w}(\cdot)$, and the normalized limit $\mathbf{v}(\cdot) = \mathbf{w}(\cdot)/\|\mathbf{w}(\cdot)\|$. This latter vector is the stable age distribution for this system, and is shown by FIGURE 3 (the "Density Dependent" curve). It follows from the Dominated Convergence Theorem that the total population $N(t) = \|n(\cdot, t)\|$ converges to $\|\mathbf{w}\|$ as $t \to \infty$, and that the limiting total population is independent of the initial population vector. This is illustrated in FIGURE 4, where the total population as a function of time is shown for five different initial conditions.

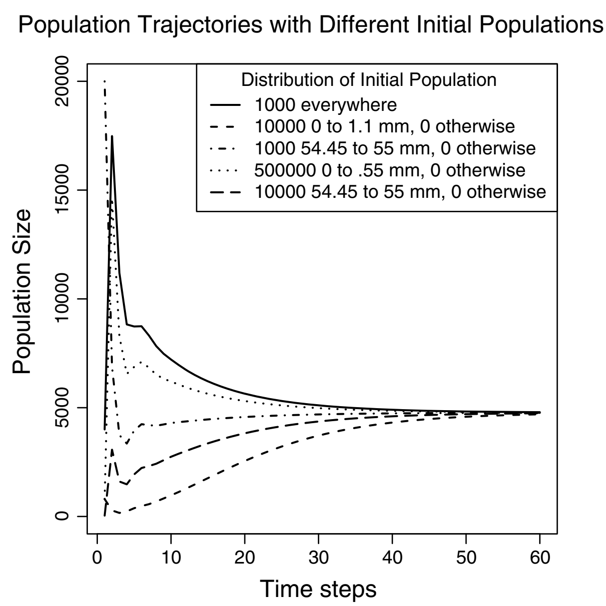


Figure 4 Asymptotic behavior of the total population

Acknowledgment This work was supported by NSF REU Site Grant 0354008, which funded the Applied Mathematics REU Site at the University of Nebraska. RR was supported in part by NSF Grant 0606857. We would like to thank Professor Stephen Ellner (Cornell University, Department of Ecology and Evolution) who suggested that we write this expository description of integral projection models.

REFERENCES

1. G. Bachman and L. Narici, *Functional Analysis*, Academic Press, New York, 1966. (Reprinted by Dover, Mineola, NY, 1998.)
2. H. Bernardelli, Population waves, *J. Burma Research Society* **31** (1941) 1–18.
3. E. Cannan, The probability of a cessation of the growth of population in England and Wales during the next century, *Economic Journal* **5**(20) (1895) 505–515. doi:10.2307/2956626
4. H. Caswell, *Matrix Population Models, Construction, Analysis, and Interpretation*, 2nd ed., Sinauer Associates, Sunderland, MA, 2001.

5. D. Z. Childs, M. Rees, K. E. Rose, P. J. Grubb, and S. P. Ellner, Evolution of complex flowering strategies: an age- and size-structured integral projection model, *Proc. Roy. Soc. London B* **270** (2003) 1829–1838. doi:10.1098/rspb.2003.2399
6. D. Z. Childs, M. Rees, K. E. Rose, P. J. Grubb, and S. P. Ellner, Evolution of size-dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model, *Proc. Roy. Soc. London B* **271** (2004) 425–434. doi:10.1098/rspb.2003.2597
7. B. Deng, The Time Invariance Principle, the absence of ecological chaos, and a fundamental pitfall of discrete modeling, *Ecological Modelling* **215** (2008) 287–292. doi:10.1016/j.ecolmodel.2008.03.013
8. M. R. Easterling, The integral projection model: theory, analysis and application, PhD dissertation, North Carolina State University, Raleigh, NC, (1998).
9. M. R. Easterling, S. P. Ellner, and P. M. Dixon, Size-specific sensitivity: applying a new structured population model, *Ecology* **81** (2000) 694–708. doi:10.1890/0012-9658(2000)081[0694:SSSAAN]2.0.CO;2
10. S. P. Ellner and M. Rees, Integral projection models for species with complex demography, *American Naturalist* **167** (2006) 410–428. doi:10.1086/499438
11. S. P. Ellner and M. Rees, Stochastic stable population growth in integral projection models: theory and application, *J. Math. Biology* **54** (2007) 227–256. doi:10.1007/s00285-006-0044-8
12. M. Kot, Discrete-time traveling waves: ecological examples, *J. Math. Biology* **30** (1992) 413–436. doi:10.1007/BF00173295
13. M. Kot, M. A. Lewis, and P. van den Driessche, Dispersal data and the spread of invading organisms, *Ecology* **77** (1996) 2027–2042. doi:10.2307/2265698
14. L. P. Lefkovitch, The study of population growth in organizms grouped by stages, *Biometrics* **21** (1965) 1–18. doi:10.2307/2528348
15. P. H. Leslie, On the use of matrices in certain population mathematics, *Biometrika* **33** (1945) 183–212. doi:10.1093/biomet/33.3.183
16. E. G. Lewis, On the generation and growth of a population, *Sankhya: The Indian Journal of Statistics* **6** (1942) 93–96.
17. S. M. Louda and M. Potvin, Effect of Inflorescence-Feeding Insects on the Demography and Lifetime of a Native Plant, *Ecology* **76** (1995) 229–245. doi:10.2307/1940645
18. K. A. Moloney, A generalized algorithm for determining category size, *Oecologia* **69** (1986) 176–180. doi:10.1007/BF00377618
19. W. F. Morris and D. F. Doak, *Quantitative Conservation Biology: Theory and Practice in Conservation Biology*, Sinauer Associates, Sunderland, MA (2002).
20. R. Rebarber, B. Tenhumberg, and S. Townley, Global asymptotic stability of a class of density dependent population projection systems (in preparation).
21. M. Rees and K. E. Rose, Evolution of flowering strategies in *Oenothera glazioviana*: an integral projection model approach, *Proc. Roy. Soc. London B* **269** (2002) 1509–1515. doi:10.1098/rspb.2002.2037
22. K. E. Rose, S. M. Louda, and M. Rees, Demographic and evolutionary impacts of native and invasive insect herbivores on *Cirsium canescens*, *Ecology* **86** (2005) 453–465. doi:10.1890/03-0697
23. E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd ed., Springer-Verlag, New York, NY, 2006.
24. J. Vandermeer, Choosing category size in a stage projection matrix, *Oecologia* **32** (1978) 79–84. doi:10.1007/BF00344691

Summary A single species is often modeled as a structured population. In a matrix projection model, individuals in the population are partitioned into a finite number of stage classes. For example, an insect population can be partitioned into egg, larva, pupa and adult stages. For some populations the stages are better described by a continuous variable, such as the stem diameter of a plant. For such populations an integral projection model can be used to describe the population dynamics, and might be easier to use or more accurate than a matrix model. In this article we discuss the similarities and differences between matrix projection models and integral projection models. We illustrate integral projection modeling by a Platte thistle population, showing how the model is determined by basic life history functions.

JOSEPH BRIGGS graduated from North Carolina State University with a BS in mathematics in May 2010. He is currently enrolled at New York University pursuing a PhD in economics with an anticipated focus in microeconomics and game theory. He participated in two NSF-funded Research Experiences for Undergraduates in mathematics at the University of Nebraska–Lincoln and the University of Illinois at Urbana-Champaign during the summers of 2008 and 2009, respectively. Besides a passion for mathematical modeling, Joseph enjoys playing water polo for NCSU's club team.

KATHRYN DABBS graduated from the University of Tennessee in 2010 with a BS in mathematics. She is currently doing an internship with the Center for Transportation Analysis at Oak Ridge National Laboratory and applying to PhD programs in mathematics for fall 2011. She also enjoys traveling, photography, and tutoring math and Latin.

MICHAEL HOLM earned a BA from Northwestern College, IA and a masters degree from the University of Nebraska-Lincoln, both in mathematics. He currently lives in Lincoln, NE with his wife Hannah and expects to complete his PhD work at the University of Nebraska in 2011. Afterward, hopes to teach mathematics at the college level. His current research interests are in fractional calculus and time scales analysis.

JOAN LUBBEN is currently an Assistant Professor of Mathematics at Dakota Wesleyan University where she tries to sneak in as much applied mathematics as possible into the courses she teaches. She did her graduate work at the University of Nebraska-Lincoln with Richard Rebarber and Brigitte Tenhumberg as her thesis advisors. She is interested in the modeling of biological populations, in particular on how models can be used as a guide for population management. When not doing mathematics, she can be found riding her bike around South Dakota.

RICHARD REBARBER is a Professor of Mathematics at the University of Nebraska. He received his BA from Oberlin College and his PhD from the University of Wisconsin. His primary research interests are in population ecology, and in control theory, in particular infinite dimensional systems theory. He is very interested in undergraduate research mentoring, and is currently the Director of the Nebraska REU Site in Applied Mathematics. He also composes, records, and performs music with a large music ensemble.

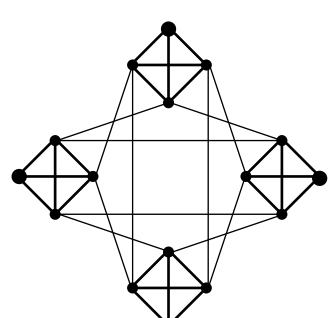
DANIEL RISER-ESPINOZA graduated from Swarthmore College in June, 2009, with a degree in mathematics and statistics. Danny returned home to rural Maine and after suffering a sports injury in July he remained at home for the fall and winter. During the fall he began learning Spanish and started doing statistical analysis for the avian risk assessments of proposed wind farms. In the winter he did some substitute teaching for high school math classes and became a grant writing/research intern for IBParticipation, a non-profit organization dedicated to increasing ethnic diversity in math, science, engineering, and technology. In March Danny went off to the woods of Massachusetts, where he now resides, to join an SCA trail crew. Danny's future plans include doing more trail work for Mass Audubon and applying for grad school.

BRIGITTE TENHUMBERG, PhD, did her undergraduate and graduate degree in Germany (University of Hannover and University of Gttingen, respectively). Then she worked as a postdoctoral fellow in Canada (Simon Fraser University, BC) and Australia (University of Adelaide, South Australia, and University of Queensland, Queensland) before joining the University of Nebraska. She currently is an Assistant Professor at the School of Biological Sciences and the Department of Mathematics. Her research synergistically combines mathematical modeling and empirical work on life history, behavior, and population level dynamics of consumer-resource interactions. Additionally, her models also address applied questions in conservation biology and invasion ecology.

Is This Graph Planar?

In their June article [**1**] Danielle Arett and Suzanne Dorée challenged readers to draw the graph H_4^2 in a plane without crossings. The graph—which describes the Tower-of-Hanoi puzzle with four pegs and two disks—is shown below.

Is it really possible? For a solution, keep turning pages.



REFERENCES

1. D. Arett and S. Dorée, "Coloring and Counting on the Tower of Hanoi Graphs," *Mathematics Magazine* **83** (2010) 200–209.

What Do We Know at 7 PM on Election Night?

ANDREW GELMAN
Columbia University
New York, NY 10027
gelman@stat.columbia.edu

NATE SILVER
FiveThirtyEight.com
nsilver@baseballprospectus.com

On the evening of November 8, 1988, I (AG) was working with my colleague Gary King in his Harvard office. It was election night, and Massachusetts Governor Michael Dukakis was a candidate for President against Vice President George H. W. Bush. Gary somehow had gotten his hands on a pair of tickets to Dukakis's victory party in Boston, and we were trying to decide whether to go. Dukakis was expected to lose, but—who could say, right? We had the TV on, and the first state to report, at 7 PM, was Kentucky, which Bush won by over 10 points. Gary informed me that the election was over: Kentucky, at the time, was near the political center of America, and there was no way that Dukakis would do much better nationally than he did in Kentucky. So we saved ourselves a subway ride and kept on working.

What about the most recent election night—November 4, 2008, when the candidates were Senators Barack Obama and John McCain? Was it possible for a viewer to play along at home with the election and decide at 7 PM, or 8 PM, or 9 PM what the outcome would be? On the night before the election we (AG and NS) analyzed a probabilistic election forecast to make some guesses at what might be known at different times during the evening. This article is a report of that analysis. Most of the results presented were obtained on November 3, the night before the election, and were intended as a guide for interpreting what we would hear the next night, as the votes were counted.

We performed one set of calculations using statewide vote margins; that is, reports like we remembered from Kentucky in 1988; and another set of calculations using only the tally of states won or lost, without the margins of victory.

In the next sections we describe the model we used, and then present our results from November 3. We then discuss what actually happened and ways in which election-night reporting could be improved in the future.

Our model on the night before the election

If one wants to make probabilistic forecasts, one needs a model. We used the model developed by one of us (NS) to make election forecasts at the website fivethirtyeight.com. It is described at that site [**6**] but it is too rich a model for us to describe in detail here. We will describe it in general terms, and discuss the tools one might use to develop a similar model of one's own.

A joint probability distribution The model takes the form of a joint probability distribution on 51 variables, one corresponding to a state's popular vote result. (We are treating the District of Columbia as a state, since it has electoral votes.) For each $j = 1, \ldots, 51$, let Z_j represent the margin of victory, expressed as a percentage of the two-party popular vote, in the j-th state. We take the margin as positive if Obama

Math. Mag. **83** (2010) 258–266. doi:10.4169/002557010X521787. © Mathematical Association of America

wins and negative if McCain wins (an arbitrary choice) and, for convenience, we express the value of Z_j in percentage points. For example, if McCain wins Kentucky (the eighteenth of the 51 states on our list) by a margin of 58.2% to 41.8% (counting only these two candidates' votes) then we say that $Z_{18} = -16.4$. The outcome of the election is summarized by the vector of random variables $\mathbf{Z} = \{Z_1, \ldots, Z_{51}\}$. If we know the values of these variables, then we know the number of electoral votes for each candidate, as well as (with reasonable assumptions about voter turnout in each state) the national popular vote margin.

In the next section we will describe how one might build a probabilistic model of this form, combining the best available mean forecasts of the variables with a realistic error model. In this section we describe how we used the model we had.

We used the model as the basis for a simulation. In the language of statistics, this is called a *Monte Carlo* method. We drew random vectors $z_i = \{z_{i,1}, \ldots, z_{i,51}\}$ for $i = 1, \ldots, 10000$, independently from the joint distribution defined by the model. By "independently" we mean that the 10,000 vectors were independent of each other—not that the variables were independent, since their correlations were dictated by the model. These 10,000 independent simulations amounted to a discrete approximation to the essentially continuous distribution of election outcomes.

All of the results in this paper were derived from this 10000×51 matrix of numbers. Each row represents a possible result of the election. In effect, we act as if these are the only possible results, and that each row of the matrix is equally likely to occur. The full matrix represents the state of our knowledge just before the polls close. (Elsewhere we have analyzed a similar set of simulations to estimate the probability that a single vote in any given state would be decisive in determining the outcome of the election [**4**]).

For our calculations below, we will need to not merely average over our 10,000 simulations but also to use them for conditioning—that is, making probabilistic inferences about future events given information that has already occurred, or has been assumed to occur. We implement the conditional probability formula $P(A \mid B) = P(A \cap B)/P(B)$ in a numerical computation fashion by keeping only the simulations consistent with the conditioned event.

For instance, suppose we were to learn at 7 PM that McCain has won Kentucky by 14 percentage points; in that case, what can we say about the outcome in the remaining states? What we can do is restrict our inferences to those simulations in which $z_{18} = -14\%$. Because of the discreteness of the simulations, we have to allow some slack in the computations; instead of conditioning on the precise event $z_{18} = -14.0\%$, we would keep all simulations in which z_{18} is between -15.0% and -13.0%. The smoothness of the underlying distribution ensures that our results computed in this way are virtually identical to those that would be obtained by exact calculations on the continuous space.

How can one build a model to forecast the election?

There are several different ways that one could construct a probabilistic election forecast (see Campbell [**1**] and Erikson and Wlezien [**2**] for recent reviews). Three natural approaches are: (1) aggregation of pre-election polls, (2) forecasts based on what are sometimes called "the fundamentals" (forecasting the national vote based on the state of the economy and making state-level adjustments based on long-term trends in voting patterns), and (3) judgments of political experts.

The model we used is based on a combination of these sources of information, but primarily the recent pre-election polls. When considering forecasts made months before Election Day, there is a distinct difference between the snapshots obtained by poll aggregation and the predictions obtained by historical forecasting models. In discussing a particular method for incorporating public opinion surveys into forecasting models, Lock and Gelman [**5**] find that early polls provide some information about the relative positions of the states in the final voting but, when it comes to the national vote, they reveal essentially nothing beyond forecasts based on the fundamentals. The day before the election, the balance between polls and external forecasts is slightly different, because the polls tend to converge to the true election outcome (Gelman and King [**3**]; see also FIGURE 3 below). This is why we can pretty much make our election-day forecast based on aggregated polls, with slight adjustments based on a regression model and simulations capturing state-level and national uncertainty. The third part of the forecasting process—judgments of political experts—comes in the form of adjustments to the raw survey numbers.

Suppose you wanted to create your own forecasts of a national election, which can be studied in aggregate (using national polls) or separately (with state polls). How would you do it? To start, you might want to put together these sources of information:

- National polls. Average them how you like, fit a trend line if you want, but in any case use these to predict the aggregate vote.
- A national-level forecast. It makes sense to pull your poll aggregate toward whatever external forecasts you have (for example, based on the historical relation between the economy and election outcomes).
- Historical patterns of state voting relative to the national average. If nothing else, you can take your national forecast and perturb it, adding and subtracting percentages for each state to get an estimated state-by-state vote assuming a uniform national swing from past elections.
- State polls. You can use these to adjust state-level estimates. There are different ways to actually combine all these numbers, but the key is to separate the national forecast from the relative positions of the states.

Now you have a point estimate—which becomes an estimate of the mean of the vector $\{Z_1, \ldots, Z_{51}\}$. The next step is to add variation to capture uncertainty in the forecast. It is natural for this variation to take the form of a covariance matrix for these 51 variables. You want for the correlation matrix to reflect uncertainty at the national, regional, and state levels.

Our particular simulations are based on a poll aggregation method developed for the website fivethirtyeight.com and involving several layers of adjustments (Silver [**6**]). What is relevant for the present article is that our 10,000 simulations represent forecast uncertainty about the election, as expressed in a joint distribution of the candidates' vote shares in the fifty states.

The vote shares in these distributions are correlated—for example, if Barack Obama had outperformed the average forecast in Ohio, we would expect (with some level of uncertainty) that he would've outperformed the forecast in Oklahoma, Oregon, or any other state. This is a sensible property for a set of forecasts, and it implies that our inferences for the outcome of the election in any state is affected by our knowledge of what happened in any other states where the polls closed earlier.

As we demonstrate below, we can use the 10000×51 matrix of state-level forecasts to make conditional inferences, to step through the hours of election night making our best prediction for the future based on the data available at any given time. Whatever the form of the probability distribution you select, the computations can be performed most transparently and generally by operating directly on the simulations.

The prior distribution Let's start by asking what we could say, based on this model, before any election returns were in. In Bayesian statistics this is called the *prior distribution,* and we can describe it by averaging all 10,000 simulations. By this method Obama was expected to win the national popular vote by 4.8 percentage points, with an expected electoral vote total of 340 (compared to 270 needed for victory), and a 96% chance of winning the electoral vote (with a 0.2% chance of a tie in the electoral college). The top row of FIGURE 1 shows the forecast distributions of Obama's popular and electoral vote shares.

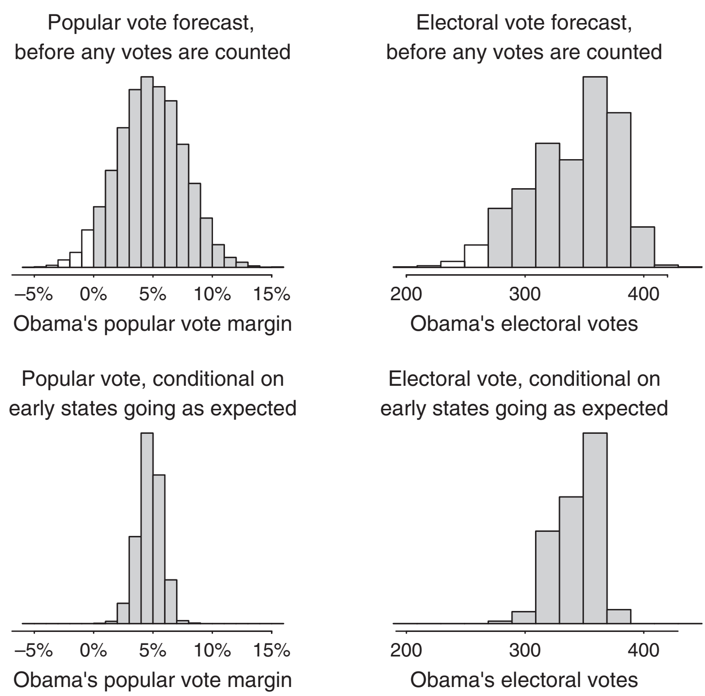


Figure 1 Uncertainty distribution for the presidential election outcome, expressed as Obama's popular vote margin and his electoral vote total. Top row is based on a poll-based forecast the day before the election; bottom row is based on these forecasts, conditional on the states whose polls close at 7 PM going as expected. The averages of the distributions in the top and bottom row are the same, but the distributions on the bottom show less variation: the vote margins of the early states tell us a lot.

Our model on election night

In 2008 the polls closed at 7 PM Eastern Time in six states, which we list in decreasing order of Obama's predicted vote margin: Vermont (Obama predicted to win by 21%), Virginia (+5%), Indiana (−2%), Georgia (−5%), South Carolina (−11%), and Kentucky (−15%). (The map has certainly changed since the days when Kentucky was a swing state!)

Real-time predictions given vote margins in the early states We start by assuming that the viewer would know the vote margin (either as exactly tabulated or as estimated from exit polls) in each of these states, which can be summarized by a simple unweighted average of Virginia, Indiana, Georgia, South Carolina, and Kentucky. (We excluded Vermont because it is the smallest of these six states and farthest from the national median.) Based on the last pre-election poll aggregates, the estimated five-state average vote margin was −5.7%; that is, McCain was expected to beat Obama by an average of 5.7 percentage points on the way to winning Kentucky, South Carolina, Georgia, Indiana, and losing Virginia.

If the expected happens, what have we learned? That is, what if the vote margin in Virginia, Indiana, Georgia, South Carolina, and Kentucky were to equal the expected −5.7%? We pipe this assumption through our model by calculating, for each of our 10,000 simulations, the average vote margin in these five states, and then restricting our analysis to the subset of simulations for which this vote margin is within 1 percentage point of its expected value (that is, between −6.7% and −4.7%). Out of our 10,000 simulations, 2800 fall in this range; that is, we predict there is a 28% chance that McCain's average vote margin in these five states will be between 4.7% and 6.7%. What is of more interest is what happens if this occurs. Considering this subset of simulations, Obama's expected national popular vote margin is +4.7%, his expected electoral vote total is 343, almost the same as the prior distribution. But now the conditional probability of an Obama victory is 100%: he wins the electoral college in all 2800 simulations in this condition. The bottom row of FIGURE 1 shows the forecast distributions of Obama's popular and electoral vote shares, conditional on him doing exactly as expected in the first round of states.

So if the 7 PM states were to go as expected, we would know a lot.

How about other possibilities? We repeat the above calculation under scenarios in which Obama's average vote margin in Virginia, Indiana, Georgia, South Carolina, and Kentucky takes on each possible value between −12% and +1%; based on our simulations, there is a 97% chance that the average 7 PM vote margin (excluding Vermont) will fall in this range.

In the unlikely event that McCain were to get an average vote margin of 12 percentage points in the five 7 PM states, we would have forecast a 1.3% margin for McCain in the national popular vote, an expected 265 electoral votes for McCain, and a 38% chance of him winning the electoral college (with a 60% chance of Obama winning and a 2% chance of a tie).

At the other extreme, if Obama happened to get an average vote margin of 1 percentage point in these early states, we would have predicted his national popular vote margin to be 10% with 394 electoral votes and a 100% chance of winning.

What about the possibilities in between? FIGURE 2 shows Obama's expected popular vote margin, the expected division of electoral votes, and the probability of each

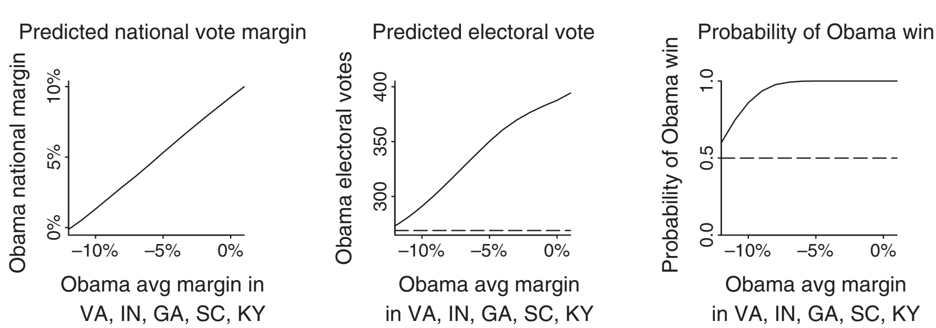


Figure 2 Election predictions at 7 PM Eastern time, after the polls were to close in Virginia, Indiana, Georgia, South Carolina, and Kentucky. Obama's average vote margin in these five states was predicted to be −5.7 percentage points, but it could plausibly have fallen between −12 and +1 percentage point. For each of these possible outcomes, we computed Obama's expected share of the national popular vote, his expected electoral vote total, and the probability he would win in the electoral college. Unless McCain's average vote margin in these five states was at least 9 percentage points, we could confidently call the election for Obama at this point.

candidate winning under the different scenarios. According to these simulations, if McCain hadn't secured an average vote margin of at least 9 percentage points in Virginia, Indiana, Georgia, South Carolina, and Kentucky, he could pretty much have thrown in the towel. And if Obama had lost these states by an average of more than 10 points, we might have been up all night worrying (rather than celebrating or crying).

Real-time predictions given only the state winners What if the TV stations do not give vote margins but just report winners? With six states reporting at 7 PM, there are 64 possible outcomes. Some of these are impossible or uninteresting, however: if McCain had won Vermont, or if Obama had won South Carolina or Kentucky, the election would have been over. What remains are Virginia, Indiana, and Georgia.

TABLE 1 gives the eight possibilities, our forecast probability of each happening, and Obama's expected popular vote margin, electoral vote margin, and win probability under each scenario. The only interesting possibility is if McCain had swept all three states: then it would have been a contest, and he would have had an even chance of pulling it out.

TABLE 1: Scenarios of interest at 7 PM Eastern time, with the first states reporting. In order to have a chance, McCain needed to win Virginia in this first round (in which case he would almost certainly have won Indiana and Georgia as well).

			Prob. of	Pop.vote	Electoral vote		Probability of...		
VA	IN	GA	scenario	margin	Oba	McC	O.win	M.win	Tie
McC	McC	McC	7%	+0.2%	268	270	.47	.49	.04
McC	McC	Oba	0						
McC	Oba	McC	0						
McC	Oba	Oba	0						
Oba	McC	McC	66%	+4.0%	330	208	.99	.01	.00
Oba	McC	Oba	<1%						
Oba	Oba	McC	22%	+7.4%	375	162	1.00	.00	.00
Oba	Oba	Oba	5%	+9.9%	399	139	1.00	.00	.00

We would then have had to wait until 7:30 PM Eastern time, when we would hear from Ohio, North Carolina, and West Virginia. Our pre-election forecasts gave Obama less than a 1% lead in North Carolina and a 2% lead in Ohio, with McCain having a 10% lead in West Virginia. But in the (unlikely) event of McCain sweeping Virginia, Indiana, and Georgia, the story would change. At that point, Obama would be expected to lose Ohio, North Carolina, and West Virginia, by margins of 3, 5, and 15 percentage points, respectively. If McCain's chances were still alive at 7 PM, there was an 87% chance he'd win all three of these must-win states that close the polls at 7:30.

Onward to 8:00, when most of the remaining eastern states closed the polls. The 8 PM states range from Maryland and Massachusetts (where Obama was forecast to win by 23 and 19 points, respectively) to Oklahoma and Alabama (predicted McCain victory margins of 27 and 24 points). Conditional on McCain winning the key early states of Virginia, Indiana, Ohio, and North Carolina, these predictions shift by about 5 percentage points in his favor.

The states to watch at 8 PM—if there was anything worth watching at all—would be New Hampshire, Pennsylvania, and Florida, for which McCain's predicted vote margins, conditional on his previous success, would be −3%, −2%, and +4%. At this point, winning Pennsylvania would pretty much guarantee victory for McCain; his other possibilities are winning New Hampshire and Florida (which would give him an expected 277 electoral votes and a 79% chance of winning, with an amazing 12% chance of a tied electoral college), and winning Florida alone (the most likely possibility, with an 8% chance of happening) which would take him to an expected 264 electoral votes and a 33% chance of winning.

At this point, McCain would have an 80% chance of winning both Missouri and Florida, which would move him up to an expected 281 electoral votes and a 66% chance of winning (compared to 30% for Obama and a 4% chance of a tie, at this point). The next most likely possibility is McCain winning Missouri but losing Florida, in which case his expected electoral vote count drops to 243 and his probability of winning declines to 3%.

If McCain were to win Virginia, Indiana, Ohio, North Carolina, and Florida, while losing Pennslyvania, we would have to keep the TV on. The news at 8:30 wouldn't help much: at this point, McCain would be expected to win handily in Arkansas, the only state to close the polls at that time. At 9 PM we would hear from a bunch of states further west, including Colorado (with Obama expected to win by less than 1% at this point), New Mexico (Obama expected to win by 4%), Minnesota (Obama by 4%), Wisconsin (Obama by 5%), and Michigan (Obama by 6%). If Obama won all five of these, he would have a 97% chance of winning. If McCain won any of them (most likely Colorado), he would have been almost home free, with a 90% chance of an electoral vote win, and if McCain had won two or more, he would have basically won the election.

Again, though, based on the polls the day before the election, we only estimated a 4% chance overall of this happening. The most likely outcome a priori was that the election would be over by 7 PM.

Discussion

With the ubiquity of polling and the rise of internet communication, state-level election forecasts and public opinion estimates are increasingly available to the general public. Attention has generally focused on point predictions and the odds of each candidate winning; here we demonstrate how the correlations in a forecast implicitly allow inferences for the election outcome, conditional on partial information. This is straightforward probability theory, or Bayesian inference; here we are applying standard practice and summarizing inferences using simulations. With 10,000 simulations, we can compute all the relevant conditional probabilities directly. If we wanted to work out some obscure possibilities (for example, Obama winning Virginia and Georgia with McCain winning Indiana), we would need to produce more simulations or else use some analytical approximations, but here we pursued only the more plausible scenarios.

Beyond whatever interest there may be in election-night dynamics, this work is relevant to understanding election forecasts and, more generally, inference about vector outcomes in which there is correlation (so that, for example, the outcome in Virginia is informative, not just about that state, but about national and regional swings). As we have demonstrated, simulation-based calculation allows us to condition on virtually any plausible outcome.

On a practical level, we are only interested here in moderate or large probabilities—we are not trying to pick out 1-in-1000 longshots—and so we only need to condition on events that have a nontrivial forecast probability of occurring. For example, we could safely ignore Vermont in the first set of results, and look only at the aggregate margin in the other five states—and then we only needed to look at one-percent intervals even though (according to the story) we would be observing exact values. These choices are not lazy but rather reflect a realistic understanding of the problem we are studying. If there is any worry on this point, one could always repeat the analysis with 10 or 100 times as many simulations and increase the precision.

What actually happened?

Obama won Vermont and Virginia, and McCain won Georgia, South Carolina, and Kentucky as predicted. As a small surprise, Obama won Indiana. The actual margins for Obama in these six states were Vermont +37.8%, Virginia +6.4%, Indiana +1.0%, Georgia −5.3%, South Carolina −9.1%, Kentucky −16.5%. The average for the last five states was −4.7%. Had we known that average, or just the fact that Obama won Virginia, we could have been nearly certain that Obama would win the election.

Of course, all of our inferences are only as good as the forecasting model. Unsurprisingly, given that the forecasts were based on the latest polls, they did pretty well; see FIGURE 3. And, in a sense, the election really was over by 7 PM—although it was not so easy to learn this from watching the networks' broadcasts.

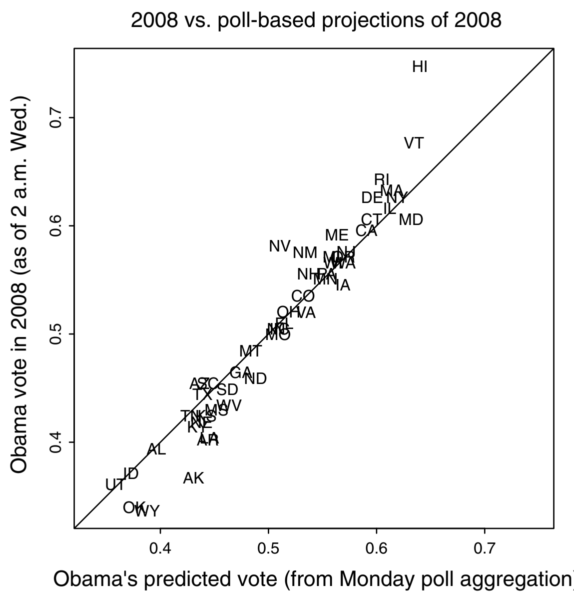


Figure 3 Obama's share of the two-party vote in each state, plotted vs. poll-based predictions made the day before. The poll-based model gave accurate forecasts of the national vote and also most individual states, although with a systematic pattern of underpredicting Obama's vote share where he was expected to do well, and overpredicting in most of the states where he was expected to lose.

Our conditional probability calculations worked reasonably well too—in theory. In practice, though, our advice on what to look for on election night was useless, because the TV networks did not report total vote margins and they did not immediately declare winners. Instead, they reported partial results as they came in from each state (14% of the precincts in Virginia, 12% of the precincts in Indiana, and so forth). As the tallies mounted it became clearer who would win each state, but the broadcasters missed out on the opportunity to provide additional information in the form of conditional probabilities.

In particular, what if the newscasters, after reporting the election results based on 14% of the precincts reporting in Virginia, were to also tell us the results in the previous election in those same precincts? The viewers would then know the swing in vote proportions, which would be much more informative than the raw numbers. (When watching the election returns on TV, we heard occasional comparisons with the 2004 Kerry-Bush numbers, but only sporadically.) Our suggestion would give the information needed to make inferences about the national election without the speculation involved in calling states early based on exit poll results.

For future elections it would be fun to set up an online widget so that users could enter election returns as they are happening, and the relevant probabilities would pop out. It would also be desirable to connect this to election returns by county and even precinct. TV networks aren't supposed to make early calls of the election but maybe there's be some way of doing this informally.

Acknowledgment We thank the Columbia University Applied Statistics Center, National Science Foundation, National Institutes of Health, National Security Agency, Department of Energy, Institute for Educational Sciences, and Yahoo Research for partial support for this research.

REFERENCES

1. J. E. Campbell, Forecasting the 2008 national elections, *PS: Political Science and Politics* **41** (2008) 679–681.
2. R. S. Erikson and C. Wlezien, Leading economic indicators, the polls, and the presidential vote, *PS: Political Science and Politics* **41** (2008) 703–707.
3. A. Gelman and G. King, Why are American Presidential election campaign polls so variable when votes are so predictable? *British Journal of Political Science* **23** (1993) 409–451. doi:10.1017/S0007123400006682
4. A. Gelman, N. Silver, and A. Edlin, What is the probability your vote will make a difference? *Economic Inquiry* (to appear).
5. K. Lock and A. Gelman, Bayesian combination of state polls and election forecasts, *Political Analysis* (to appear).
6. N. Silver, Frequently asked questions, `http://www.fivethirtyeight.com/2008/03/frequently-asked-questions-last-revised.html`.

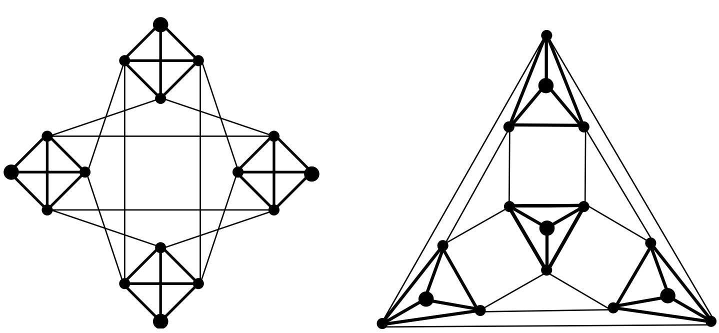
Summary We use a probability forecasting model to estimate the chance of different branches on the tree of state-by-state outcomes on election night. Forecasting models can use data from pre-election surveys as well as extrapolation based on previous election results. We implement conditional probability calculations numerically using a matrix representing 10,000 simulations of the outcomes in the 50 states.

ANDREW GELMAN is a professor of statistics and political science at Columbia University. His books include "Bayesian Data Analysis," "Teaching Statistics: A Bag of Tricks," and "Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do."

NATE SILVER is a statistician at Baseball Prospectus and the founder of FiveThirtyEight.com.

Yes!

Follow the complete 4-graphs shown in bold.



—Suzanne Dorée

Putzer's Algorithm for e^{At} via the Laplace Transform

WILLIAM A. ADKINS
Louisiana State University
Baton Rouge, LA 70803
adkins@math.lsu.edu

MARK G. DAVIDSON
Louisiana State University
Baton Rouge, LA 70803
davidson@math.lsu.edu

The solution of the system of differential equations

$$y' = Ay, \qquad y(0) = y_0, \tag{1}$$

in which A is an $n \times n$ complex matrix while $y(t)$ and y_0 are n-vectors, is

$$y(t) = e^{At} y_0.$$

There are many ways to compute the matrix exponential e^{At}, at least on a theoretical level, using various amounts of linear algebra [**7, 8**]. One traditional method [**4**] uses the transformation of A into its Jordan Canonical Form. Other methods due to Putzer [**2, 5, 9, 10**], Kirchner [**6**], and Fulmer [**3**] compute e^{At} using only the eigenvalues of A and their algebraic multiplicities. Putzer's paper [**9**] actually has two somewhat different algorithms. The purpose of this note is to adapt one of Putzer's arguments to give a very simple proof of a formula for the resolvent matrix $(sI - A)^{-1}$, which we can then combine with the Laplace transform formula

$$e^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\} \tag{2}$$

to get a formula like that of Putzer for e^{At}.

EXAMPLE 1. To illustrate the type of formula that will be obtained, let A be any 3×3 matrix with characteristic polynomial $c_A(s) = (s-1)^2(s-2)$—that is, whose eigenvalues are 1, 1, 2. In this case Putzer's method gives this closed formula for e^{At}:

$$e^{At} = e^t I + te^t(A - I) + (e^{2t} - e^t - te^t)(A - I)^2. \tag{3}$$

We obtain this result by first obtaining a formula for $(sI - A)^{-1}$ as a function of s:

$$(sI - A)^{-1} = \frac{1}{s-1} I + \frac{1}{(s-1)^2}(A - I) + \frac{1}{(s-1)^2(s-2)}(A - I)^2. \tag{4}$$

Equation (3) follows by applying the inverse Laplace transform to (4):

$$\begin{aligned} e^{At} &= \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\} \\ &= \mathcal{L}^{-1}\left\{\frac{1}{s-1}\right\} I + \mathcal{L}^{-1}\left\{\frac{1}{(s-1)^2}\right\}(A - I) + \mathcal{L}^{-1}\left\{\frac{1}{(s-1)^2(s-2)}\right\}(A - I)^2 \\ &= e^t I + te^t(A - I) + (e^{2t} - e^t - te^t)(A - I)^2. \end{aligned}$$

Math. Mag. **83** (2010) 267–275. doi:10.4169/002557010X521796. © Mathematical Association of America

Since the coefficients of (4) are rational functions, the inverse Laplace transforms are easily computed from the standard formula $\mathcal{L}^{-1}\left\{1/(s-\lambda)^k\right\} = t^{k-1}e^{\lambda t}/(k-1)!$ combined with a partial fraction expansion when needed. ■

The method of Putzer that is being adapted uses knowledge of the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the $n \times n$ matrix A, counted according to their algebraic multiplicities, to reduce the system (1) to a recursive system of first order linear differential equations of the form

$$\begin{aligned} r_1' &= \lambda_1 r_1 & r_1(0) &= 1, \\ r_j' &= \lambda_j r_j + r_{j-1} & r_j(0) &= 0, \quad \text{for } j \geq 2. \end{aligned} \tag{5}$$

The solutions $r_j(t)$ then become coefficient functions that are used to write

$$e^{At} = \sum_{j=0}^{n-1} r_{j+1}(t) P_j, \tag{6}$$

in which the matrices P_j are easily determined from A and the eigenvalues as

$$\begin{cases} P_0 = I & \text{if } j = 0, \\ P_j = (A - \lambda_j I) \cdots (A - \lambda_1 I) & \text{if } j \geq 1. \end{cases} \tag{7}$$

Our approach to computing e^{At} makes use of the Laplace transform and replaces the recursive differential equations (5) with the inverse Laplace transform

$$r_j(t) = \mathcal{L}^{-1}\left\{\frac{1}{(s-\lambda_j)\ \cdots\ (s-\lambda_1)}\right\}. \tag{8}$$

In the example above, the expression for e^{At} in (3) is the same as that given by (6), with the $r_j(t)$ being determined by (8), that is as the inverse Laplace transforms of the coefficient functions in (4), rather than by solving the system of differential equations (5). The only Laplace transform properties that we will need are $\mathcal{L}\left\{t^k e^{\lambda t}\right\} = k!/(s-\lambda)^{k+1}$ and the derivative formula $\mathcal{L}\{y'(t)\} = s\mathcal{L}\{y(t)\} - y(0)$.

The theoretical justification of the method we describe, like that of Putzer, relies on the Cayley-Hamilton theorem. That theorem can itself be proved by means of the Laplace transform [**1**]. A simplified version of the proof in [**1**] is given at the end of this note.

Our starting point is Equation (2), which we now verify. Let A be an arbitrary $n \times n$ complex matrix. It follows from the power series description of e^{At} that $(e^{At})' = Ae^{At}$ and $e^{At}|_{t=0} = I$, the $n \times n$ identity matrix. The derivative formula for $\mathcal{L}$ gives

$$s\mathcal{L}\left\{e^{At}\right\} - I = \mathcal{L}\left\{(e^{At})'\right\} = \mathcal{L}\left\{Ae^{At}\right\} = A\mathcal{L}\left\{e^{At}\right\}.$$

Solving for $\mathcal{L}\left\{e^{At}\right\}$ gives $\mathcal{L}\left\{e^{At}\right\} = (sI - A)^{-1}$, which is equivalent to (2).

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A listed in any desired order, each counted according to its algebraic multiplicity. Define a sequence of rational functions $R_j(s)$ and matrices P_j for $0 \leq j \leq n$ by

$$\begin{cases} R_0(s) = 1 & \text{if } j = 0, \\ R_j(s) = (s-\lambda_j)^{-1} \cdots (s-\lambda_1)^{-1} & \text{if } j \geq 1, \end{cases} \tag{9}$$

and

$$\begin{cases} P_0 = I & \text{if } j = 0, \\ P_j = (A - \lambda_j I)\cdots(A - \lambda_1 I) & \text{if } j \geq 1. \end{cases} \tag{10}$$

With these notations, there is the following formula for $(sI - A)^{-1}$.

THEOREM 1. *With $R_j(s)$ and P_j as defined in* (9) *and* (10),

$$(sI - A)^{-1} = \sum_{j=0}^{n-1} R_{j+1}(s)P_j. \tag{11}$$

Proof. Let $B(s)$ denote the right-hand side of Equation (11). It is necessary to show that $(sI - A)B(s) = I$. To this end, note that the functions $R_j(s)$ are related by $(s - \lambda_j)R_j(s) = R_{j-1}(s)$ for $1 \leq j \leq n$, and the matrices P_j are related by $(A - \lambda_{j+1}I)P_j = P_{j+1}$ for $0 \leq j \leq n - 1$. Then

$$\begin{aligned} (sI - A)R_{j+1}(s)P_j &= \big((sI - \lambda_{j+1}I) - (A - \lambda_{j+1}I)\big)\, R_{j+1}(s)P_j \\ &= (s - \lambda_{j+1})R_{j+1}(s)P_j - R_{j+1}(s)(A - \lambda_{j+1}I)P_j \\ &= R_j(s)P_j - R_{j+1}(s)P_{j+1}, \end{aligned}$$

so that

$$\begin{aligned} (sI - A)B(s) &= \sum_{j=0}^{n-1}(sI - A)R_{j+1}(s)P_j \\ &= \sum_{j=0}^{n-1}\big(R_j(s)P_j - R_{j+1}(s)P_{j+1}\big) \\ &= R_0(s)P_0 - R_n(s)P_n \\ &= I - R_n(s)P_n. \end{aligned}$$

But

$$P_n = (A - \lambda_n I)\cdots(A - \lambda_1 I) = c_A(A),$$

in which

$$c_A(s) = \det(sI - A)$$

is the characteristic polynomial of A. By the Cayley-Hamilton Theorem, $c_A(A) = 0$, so we conclude that $(sI - A)B(s) = I$, as required. ■

EXAMPLE 2. Equation (11) produces the standard adjoint formula for $(sI - A)^{-1}$ if $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is a 2×2 matrix. To see this, recall that the characteristic polynomial of A is

$$c_A(s) = \det(sI - A) = s^2 - (a + d)s + (ad - bc) = (s - \lambda_1)(s - \lambda_2)$$

so that $\lambda_1 + \lambda_2 = a + d$. Then the formula in Equation (11) becomes

$$\begin{aligned} (sI - A)^{-1} &= R_1(s)P_0 + R_2(s)P_1 \\ &= (s - \lambda_1)^{-1}I + (s - \lambda_2)^{-1}(s - \lambda_1)^{-1}(A - \lambda_1 I) \end{aligned}$$

$$\begin{aligned}
&= (s-\lambda_2)^{-1}(s-\lambda_1)^{-1}\left((s-\lambda_2)I + (A-\lambda_1 I)\right) \\
&= (c_A(s))^{-1}\left(A + (s-\lambda_1-\lambda_2)I\right) \\
&= \frac{1}{c_A(s)}\begin{bmatrix} s-d & b \\ c & s-a \end{bmatrix}.
\end{aligned}$$

EXAMPLE 3. If A is a 3×3 matrix with eigenvalues $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, then Equation (11) becomes

$$(sI - A)^{-1} = \frac{1}{s-\lambda}I + \frac{1}{(s-\lambda)^2}(A-\lambda I) + \frac{1}{(s-\lambda)^3}(A-\lambda I)^2.$$

Combining Equation (2) and Theorem 1 gives the Laplace transform version of Putzer's formula for e^{At}.

THEOREM 2. *Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then*

$$e^{At} = \sum_{j=0}^{n-1} r_{j+1}(t) P_j, \tag{12}$$

in which

$$\begin{cases} P_0 = I & \text{if } j = 0, \\ P_j = (A-\lambda_j I)\cdots(A-\lambda_1 I) & \text{if } j \geq 1, \end{cases} \tag{13}$$

and

$$r_j(t) = \mathcal{L}^{-1}\left\{R_j(s)\right\} = \mathcal{L}^{-1}\left\{(s-\lambda_j)^{-1}\cdots(s-\lambda_1)^{-1}\right\} \tag{14}$$

for $1 \leq j \leq n$.

In Putzer's original formulation, the functions $r_j(t)$ $(1 \leq j \leq n)$ are determined recursively as solutions of the system

$$\begin{aligned}
&r_1' = \lambda_1 r_1 && r_1(0) = 1 \\
&r_j' = \lambda_j r_j + r_{j-1} && r_j(0) = 0, \quad \text{for } j = 2, \ldots, n.
\end{aligned} \tag{15}$$

To see that this is equivalent to $r_j(t) = \mathcal{L}^{-1}\left\{R_j(s)\right\}$ it is only necessary to observe, using the derivative formula for Laplace transforms, that the solution $r_1(t)$ of (15) satisfies $\mathcal{L}\left\{r_1'(t)\right\} = s\mathcal{L}\left\{r_1(t)\right\} - 1 = \lambda_1\mathcal{L}\left\{r_1(t)\right\}$ so that $\mathcal{L}\left\{r_1(t)\right\} = (s-\lambda_1)^{-1} = R_1(s)$. Apply the Laplace transform to the second equation in (15) and apply induction to conclude that

$$\mathcal{L}\left\{r_j(t)\right\} = (s-\lambda_j)^{-1}\mathcal{L}\left\{r_{j-1}(t)\right\} = (s-\lambda_j)^{-1}\cdots(s-\lambda_1)^{-1} = R_j(s).$$

In Putzer's formulation the computation of the $r_j(t)$ is truly recursive since $r_j(t)$ is determined by means of the first order linear differential equation (15) involving $r_{j-1}(t)$. In our version, $r_j(t) = \mathcal{L}^{-1}\left\{R_j(s)\right\}$, and while it is true that $R_j(s) = (s-\lambda_j)^{-1}R_{j-1}(s)$, the partial fraction method of computing $r_j(t)$ is not facilitated by knowledge of the partial fraction expansion for $R_{j-1}(s)$. However, the recursive nature is still present when one takes into account the relationship between the convolution product of two functions and the product of their Laplace transforms. Recall that if $\mathcal{L}\left\{f(t)\right\} = F(s)$ and $\mathcal{L}\left\{g(t)\right\} = G(s)$, then

$$\mathcal{L}^{-1}\left\{F(s)G(s)\right\}(t) = (f * g)(t) = \int_0^t f(x)g(t-x)\,dx.$$

The integral defines the convolution product of the functions $f(t)$ and $g(t)$. Since $R_j(s) = (s-\lambda_j)^{-1}R_{j-1}(s)$ and $\mathcal{L}^{-1}\{(s-\lambda)^{-1}\} = e^{\lambda t}$ it follows that $r_j(t) = r_{j-1}(t) * e^{\lambda_j t}$; that is, $r_j(t)$ is the convolution product of $r_{j-1}(t)$ and $e^{\lambda_j t}$.

EXAMPLE 4. If A is a 3×3 matrix with eigenvalues $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$, then Equations (14) become

$$\begin{aligned} r_1(t) &= \mathcal{L}^{-1}\left\{(s-\lambda)^{-1}\right\} = e^{\lambda t}, \\ r_2(t) &= \mathcal{L}^{-1}\left\{(s-\lambda)^{-2}\right\} = te^{\lambda t}, \\ r_3(t) &= \mathcal{L}^{-1}\left\{(s-\lambda)^{-3}\right\} = (t^2/2)e^{\lambda t}. \end{aligned}$$

Thus

$$e^{At} = e^{\lambda t}I + te^{\lambda t}(A-\lambda I) + \frac{t^2}{2}e^{\lambda t}(A-\lambda I)^2.$$

As long as one can find the eigenvalues and their multiplicities, Theorem 2 involves only matrix multiplications and the partial fraction expansions that are needed to compute $r_j(t)$. Here is a somewhat larger example.

EXAMPLE 5. Use Theorem 2 to find e^{At} if

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Since A is upper triangular, the eigenvalues are $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = \lambda_4 = 2$. This gives the following for the matrices P_j. $P_0 = I$,

$$P_1 = A - I = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad P_2 = (A-I)P_1 = P_1^2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$P_3 = (A-2I)P_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now compute the coefficient functions $r_j(t)$. Since $R_1(s) = (s-1)^{-1}$, $R_2(s) = (s-1)^{-2}$, $R_3(s) = (s-1)^{-2}(s-2)^{-1}$, and $R_4 = (s-1)^{-2}(s-2)^{-2}$, we see that $r_1(t) = e^t$ and $r_2(t) = te^t$. Use partial fraction expansions of $R_3(s)$ and $R_4(s)$ to compute $r_3(t)$ and $r_4(t)$:

$$R_3(s) = \frac{1}{s-2} - \frac{1}{s-1} - \frac{1}{(s-1)^2} \Longrightarrow r_3(t) = e^{2t} - te^t - e^t$$

$$R_4(s) = \frac{1}{(s-2)^2} - \frac{2}{s-2} + \frac{2}{s-1} + \frac{1}{(s-1)^2} \Longrightarrow r_4(t) = te^{2t} - 2e^{2t} + te^t + 2e^t.$$

Put these together as in Equation (12) to get

$$e^{At} = e^t \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + te^t \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$+ (e^{2t} - te^t - e^t) \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} + (te^{2t} - 2e^{2t} + te^t + 2e^t) \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

After adding the matrices we get

$$e^{At} = \begin{bmatrix} e^t & 0 & e^{2t} - e^t & te^{2t} - e^{2t} + e^t \\ 0 & e^t & 0 & e^{2t} - e^t \\ 0 & 0 & e^{2t} & te^{2t} \\ 0 & 0 & 0 & e^{2t} \end{bmatrix}. \tag{16}$$

In Example 5 it is worth noting that the term te^t appears in some of the coefficient functions $r_j(t)$, but it does not appear in the final expression for e^{At} in Equation (16). This is simply a reflection of the fact that the *same* coefficient functions $r_1(t), \dots, r_4(t)$ are used in the Putzer expansion of e^{At} for *all* 4×4 matrices A with characteristic polynomial $(s-1)^2(s-2)^2$. Some of these may need one or more of the terms te^t and te^{2t}, but this will not be known in advance. If they are not needed in e^{At}, the unneeded terms will cancel each other out.

All of the preceding examples have used matrices with only real eigenvalues. However, the algorithm of Theorem 2 is equally proficient at computing e^{At} for real matrices that have some non-real eigenvalues. We illustrate the requisite calculations in the following example. Observe that the intermediate steps involve complex functions, but the imaginary parts all cancel in the final result to give a real expression for e^{At}. Of course, this is inevitable since the power series definition of e^{At} involves only real matrices.

EXAMPLE 6. Compute e^{At} for the matrix

$$A = \begin{bmatrix} 3 & 5 \\ -1 & -1 \end{bmatrix}$$

using Theorem 2.

The characteristic polynomial $c_A(s) = \det(sI - A) = s^2 - 2s + 2 = (s-1)^2 + 1$ has roots $\lambda_1 = 1 + i$ and $\lambda_2 = 1 - i$. Thus,

$$R_1(s) = (s - (1+i))^{-1} \implies r_1(t) = e^{(1+i)t} = e^t \cos t + ie^t \sin t,$$
$$R_2(s) = ((s-1)^2 + 1)^{-1} \implies r_2(t) = e^t \sin t.$$

Since $P_0 = I$ and

$$P_1 = A - \lambda_1 I = A - (1+i)I = \begin{bmatrix} 2-i & 5 \\ -1 & -2-i \end{bmatrix},$$

Theorem 2 gives

$$\begin{aligned} e^{At} &= e^{(1+i)t}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + e^t \sin t \begin{bmatrix}2-i & 5\\ -1 & -2-i\end{bmatrix} \\ &= \begin{bmatrix} e^t\cos t + ie^t \sin t & 0 \\ 0 & e^t \cos t + ie^t \sin t\end{bmatrix} + \begin{bmatrix}(2-i)e^t \sin t & 5e^t \sin t \\ -e^t \sin t & -(2+i)e^t \sin t\end{bmatrix} \\ &= \begin{bmatrix} e^t \cos t + 2e^t \sin t & 5e^t \sin t \\ -e^t \sin t & e^t \cos t - 2e^t \sin t\end{bmatrix}. \end{aligned}$$

The Cayley-Hamilton Theorem While it is not needed for understanding the main results of the paper (Theorems 1 and 2), it is interesting to note that the Cayley-Hamilton theorem itself can be proven by means of Laplace transform techniques. We will conclude with such a proof. While this is not the most natural of the many proofs of the Cayley-Hamilton theorem, it is very much within the spirit of the present paper.

The key ingredient is the following calculation, which simplifies the original proof in [**1**].

LEMMA 1. *Let j be a nonnegative integer and let $a \in \mathbb{C}$. Then*

$$\boldsymbol{D}^l\left(t^j e^{at}\right)\Big|_{t=0} = \boldsymbol{D}^j t^l \Big|_{t=a},$$

for all nonnegative integers l.

Proof. The derivative formula $\boldsymbol{D}^l(ye^{at}) = ((\boldsymbol{D}+a)^l y)e^{at}$ implies

$$\begin{aligned} \boldsymbol{D}^l(t^j e^{at})\Big|_{t=0} &= ((\boldsymbol{D}+a)^l t^j)\Big|_{t=0} \\ &= \sum_{k=0}^{l}\binom{l}{k} a^{l-k}(\boldsymbol{D}^k t^j)\Big|_{t=0} \\ &= \begin{cases} 0 & \text{if } l < j \\ \dfrac{a^{l-j} l!}{(l-j)!} & \text{if } l \geq j \end{cases} \\ &= \boldsymbol{D}^j t^l\Big|_{t=a}. \end{aligned}$$

■

We can now give the Laplace transform proof of the Cayley-Hamilton theorem.

THEOREM 3. (CAYLEY-HAMILTON) *Let A be an $n \times n$ complex matrix and let $c_A(s)$ be its characteristic polynomial. Then*

$$c_A(A) = 0.$$

Proof. Let $c_A(s)$ be the characteristic polynomial of A. If $\lambda_1, \ldots, \lambda_m$ are the roots of $c_A(s)$, i.e., the eigenvalues of A, with corresponding multiplicities $r_1, \ldots, r_m$, then

$$c_A(s) = (s-\lambda_1)^{r_1} \cdots (s-\lambda_m)^{r_m}. \tag{17}$$

The adjoint formula for the inverse of a matrix gives

$$(sI - A)^{-1} = \frac{1}{c_A(s)}\operatorname{Adj}(sI - A) = \left[\frac{b_{\mu\nu}(s)}{c_A(s)}\right],$$

in which the μ, ν cofactor $b_{\mu\nu}(s)$ of $\det(sI - A)$ is a polynomial of degree at most $n-1$. Thus each entry of $(sI-A)^{-1}$ is, by means of partial fractions, a linear combination of terms of the form $(s-\lambda_k)^{-j}$ with $1 \le j \le r_k$. Note that

$$\mathcal{L}^{-1}\left\{\frac{1}{(s-\lambda_k)^j}\right\} = \frac{1}{(j-1)!}t^{j-1}e^{\lambda_k t}.$$

By collecting the coefficients of $t^{j-1}e^{\lambda_k t}$ in $\mathcal{L}^{-1}\left\{b_{\mu\nu}(s)/c_A(s)\right\}$ for each index μ, ν, we may assume that there are $n \times n$ matrices $M_{j,k}$, $j = 0, \ldots, r_k - 1$, $k = 1, \ldots, m$, so that

$$e^{At} = \mathcal{L}^{-1}\left\{(sI-A)^{-1}\right\} = \sum_{k=1}^{m}\sum_{j=0}^{r_k-1} t^j e^{\lambda_k t} M_{j,k}.$$

Differentiating both sides l times and evaluating at $t = 0$ gives

$$A^l = \sum_{k=1}^{m}\sum_{j=0}^{r_k-1} \boldsymbol{D}^l(t^j e^{\lambda_k t})\Big|_{t=0} M_{j,k} = \sum_{k=1}^{m}\sum_{j=0}^{r_k-1} \boldsymbol{D}^j t^l\Big|_{t=\lambda_k} M_{j,k},$$

with the second equality coming from the Lemma. Now let $p(t) = c_0 + c_1 t + \cdots + c_N t^N$ be any polynomial. Then

$$\begin{aligned} p(A) &= \sum_{l=0}^{N} c_l A^l \\ &= \sum_{l=0}^{N}\sum_{k=1}^{m}\sum_{j=0}^{r_k-1} c_l \boldsymbol{D}^j t^l\Big|_{t=\lambda_k} M_{j,k} \\ &= \sum_{k=1}^{m}\sum_{j=0}^{r_k-1} \boldsymbol{D}^j\left(\sum_{l=0}^{N} c_l t^l\right)\Big|_{t=\lambda_k} M_{j,k} \\ &= \textstyle\sum_{k=1}^{m}\sum_{j=0}^{r_k-1} p^{(j)}(\lambda_k) M_{j,k}. \end{aligned} \tag{18}$$

For the characteristic polynomial, Equation (17) shows that $c_A^{(j)}(\lambda_k) = 0$ for all $j = 0, \ldots, r_k - 1$ and $k = 1, \ldots, m$. Now let $p(t) = c_A(t)$ in Equation (18) to get $c_A(A) = \sum_{k=1}^{m}\sum_{j=0}^{r_k-1} c_A^{(j)}(\lambda_k) M_{j,k} = 0$. ■

REFERENCES

1. William A. Adkins and Mark G. Davidson, The Cayley-Hamilton and Frobenius theorems via the Laplace transform, *Linear Algebra Appl.* **371** (2003) 147–152. doi:10.1016/S0024-3795(03)00427-0
2. Tom M. Apostol, *Calculus*, Vol. II, 2nd ed., John Wiley, New York, 1969 (Section 7.13).
3. Edward P. Fulmer, Computation of the matrix exponential, *Amer. Math. Monthly* **82** (1975) 156–159. doi: 10.2307/2319662
4. Morris W. Hirsch and Stephen Smale, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, 1974.
5. Roger A. Horn and Charles R. Johnson, *Topic in Matrix Analysis*, Cambridge University Press, New York, 1991 (Section 6.5).
6. R. B. Kirchner, An explicit formula for e^{At}, *Amer. Math. Monthly* **74** (1967) 1200–1204. doi10.2307/2315661
7. Cleve Moler and Charles Van Loan, Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.* **20** (1978) 801–836. doi:10.1137/1020098
8. Cleve Moler and Charles Van Loan, Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, *SIAM Rev.* **45** (2003) 3–49. doi:10.1137/S00361445024180

9. E. J. Putzer, Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients, *Amer. Math. Monthly* **73** (1966) 2–5. doi:10.2307/2313914
10. Paul Waltman, *A Second Course in Elementary Differential Equations*, Dover, Mineola, NY, 2004 (Section 1.7).

Summary A method due to E. J. Putzer computes the matrix exponential e^{At} for an $n \times n$ matrix A without transforming A to Jordan canonical form. A variation of Putzer's algorithm is presented. This approach is based on an algorithmically produced formula for the resolvent matrix $(sI - A)^{-1}$ that is combined with simple Laplace transform formulas to give a formula, similar to Putzer's, for e^{At}.

NOTES

The Geometry of the Snail Ball

STAN WAGON
Macalester College
St. Paul, Minnesota 55105
wagon@macalester.edu

An amusing device called the snail ball can be found at some puzzle shops (e.g., [**1**], which has a video of the device). The ball rolls down an inclined ramp, but very, very slowly. It rolls just a few millimeters at a time, with long periods during which the ball just stands motionless. On a six-inch ramp tilted at an angle of 4°, the ball can take six minutes or more to descend. This is indeed a snail-like pace of an inch a minute: a typical snail race has fastest times of about 6 inches per minute. The device is shown in FIGURE 1.

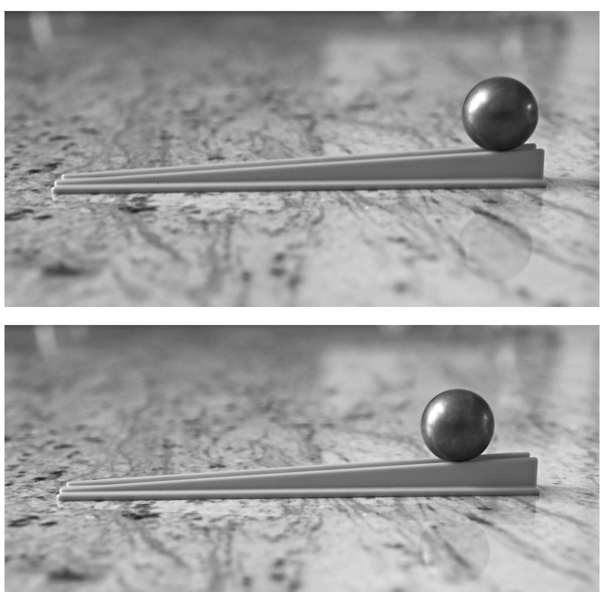


Figure 1 Two views of the snail ball; the second was taken 30 seconds after the first.

The mystery can be explained by some familiar concepts: the center of gravity and the curtate cycloid curve. Recall that a point on a rolling wheel of radius r traces out the curve $r(t - a\sin t, 1 - a\cos t)$ where a is the proportion of the radius that the tracing point is away from the rim. When $a < r$ this is called a *curtate cycloid*, but we will just use the term *cycloid* here for brevity.

The snail ball consists of a spherical metal shell (diameter 19 mm) hiding a smaller core (a solid ball of diameter 12 mm) in its interior. There is also a viscous liquid in

Math. Mag. **83** (2010) 276–279. doi:10.4169/002557010X521804.

the interior (perhaps glycerol), which delays the ability of gravity to lower the core to the bottom position. By making the simplifying assumption that the viscosity acts very strongly to delay motion, we can use some simple geometry to explain the effect. (The unit of viscosity is the *poise*, for Poiseuille; water has poise 0.894 at 25°C while glycerol has 1500 poise at the same temperature—roughly the same as corn syrup.) As a bonus we can figure out the center of gravity of the snail ball without breaking it open and weighing the components.

If a wheel has a fixed center of gravity (call it C) located below its center, then it cannot be rolled very far. The center of gravity cannot, in the ideal case where its initial speed is infinitesimally small, rise above its starting point (the sum of potential and kinetic energy cannot rise and the kinetic energy is 0 at the start). Of course a small initial speed generates some momentum, and this raises C a little before the slope of the curtate cycloid (FIGURE 2) forces it back down to its original level; the effect of friction thus returns it to its starting state via a damped oscillation.

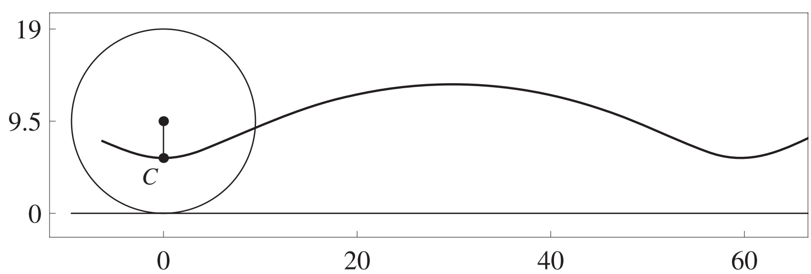


Figure 2 When the center of gravity is not at the geometrical center of a wheel, rolling is inhibited because of the local minimum of the curtate cycloid.

But suppose the ball is on a ramp, tilted downward at angle θ (FIGURE 3; where the angle is an artificially steep 15°). The cycloidal curve is tilted as well and its local minimum is shifted a bit to the right. Thus the ball wants to roll down a short distance. If the center of gravity were fixed, then the ball would simply settle into this new position. But the viscous fluid allows the core to slowly return to the bottom of the shell. When that happens the device is back to its initial state and again moves forward a very little bit.

This model also explains the surprising back-and-forth motion of the snail ball. For as it falls into the new minimum its momentum carries it up the other side before it falls backward and settles into its new location.

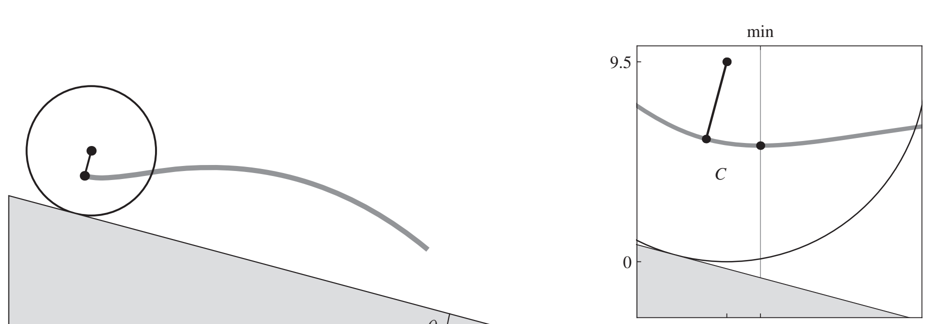


Figure 3 The tilted cycloid has its minimum at a level slightly lower than the initial center of gravity, as shown in the magnified view.

This geometrical view allows us to determine the center of gravity without knowing the exact details of what is hidden inside the shell. For if the ramp is very steep, the ball will just roll down without ever stopping: the center of gravity will rise and fall over the bumps of the cycloid but the viewer will just see this is a variation in the speed, which never becomes 0. Thus we can, by tilting the ramp, determine the steepest angle that causes the ball to rest during its journey; this resting time is what allows the core to return to the bottom and the motion to restart.

To start the experiment we place the ball on a flat surface so that the core can settle (as in FIGURE 2) and then we tilt the ramp and ball as shown in FIGURE 3 and immediately release the ball. From this we learn that, for the particular snail ball shown in FIGURE 1, the critical angle is about 9.5°. What this means is that the first local maximum of the tilted cycloid is at the same vertical level as the center of gravity C. We can use calculus to find this maximum. The y-coordinate of the cycloid after clockwise rotation through θ is $r(1 - a\cos(t+\theta) - t\sin\theta)$, where r is the radius of the ball, θ is the tilt angle, and a is the proportion of the total radius that locates the center of gravity below the center of the ball. The derivative with respect to t is $r(a\sin(t+\theta) - \sin\theta)$, which we can set to 0 and solve. Making the correct choice of solution, one finds that the maximum corresponds to $t = \pi - \arcsin((\sin\theta)/a) - \theta$ and the corresponding local maximum of the height is

$$r\left(1 - (\pi - \theta)\sin\theta + \arcsin[(\sin\theta)/a]\sin\theta + \sqrt{a^2 - \sin^2\theta}\right).$$

So now we can determine the center of gravity for the snail ball by setting $r = 9.5$ and θ to be the critical angle of 9.5°. Numerical root-finding then gives the final answer: the value of a so that the maximum just given coincides with $r(1 - a\cos\theta)$, the height of the center of gravity in its initial position. This occurs at $a \approx 0.2113$, meaning that C lies at height $r(1 - a\cos 9.5°) \approx 7.5$. A visual check (see FIGURE 4, which shows the tilted cycloid's tangency to the horizontal from C) shows that we have successfully located it. If the center of gravity were any higher, the ball would roll over the invisible bump and continue down the ramp without pausing.

As typically happens, the real world device is a little more complicated than the simplest mathematical model. The friction between the ball and the ramp plays a role: as the core returns to its low point, the motion will not restart until the gravity force is enough to overcome the frictional force. This helps explain why the core must end up quite near the bottom for the restart to occur, and perhaps explains why the device does not just move slowly down in some equilibrium state.

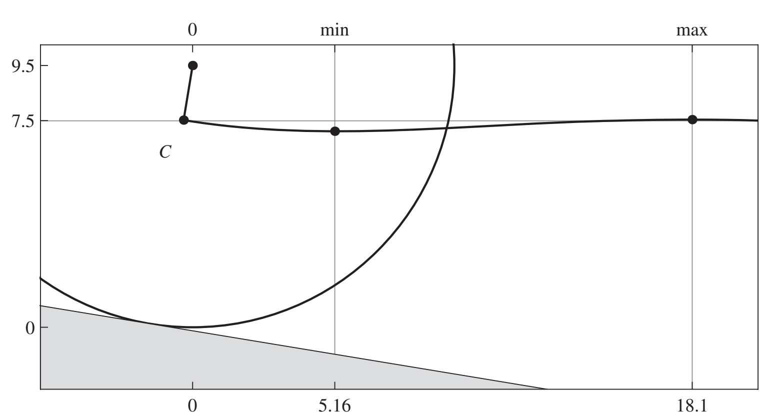


Figure 4 When the center of gravity is at the exact same level as the next local maximum of the tilted cycloid, the ramp is at its critical steepness.

Figure 5 One can make a transparent snail ball using a jar, some corn syrup, and a moderately heavy metal cylinder.

One can easily build a simple transparent version of the device. Just partially fill a glass jar with corn syrup and then insert a heavy metal cylinder. I used a 3.5-inch diameter jar with a 1.5-inch diameter cylindrical piece of aluminum to act as the weight (FIGURE 5). One can then see the effect of friction by wrapping elastic bands around the jar and also varying the ramp surface. A more detailed study using advanced techniques from fluid dynamics is available in [**2**].

REFERENCES

1. Grand Illusions, `http://www.grand-illusions.com/acatalog/Snail_Ball.html`.
2. N. J. Balmforth, J. W. M. Bush, D. Vener, and W. R. Young, Dissipative descent: Rocking and rolling down an incline, *Journal of Fluid Mechanics* 590 (2007) 295–318. doi:10.1017/S0022112007008051

Summary The snail ball is a device that rolls down an inclined plane, but very slowly, repeatedly coming to a stop and staying motionless for several seconds. The interior of the ball is hollow, with a smaller solid ball inside it, surrounded by a very viscous fluid. We show how to model the stop-and-start motion by analyzing the cycloidal curve that would correspond to the motion of the center of gravity as the ball rolls down an inclined plane.

Another Morsel of Honsberger

MOWAFFAQ HAJJA
Department of Mathematics
Yarmouk University
Irbid–Jordan
mhajja@yu.edu.jo, mowhajja@yahoo.com

In both of his interesting books [**5**] and [**6**], Ross Honsberger presents and proves the following property:

> In FIGURE 1, as A moves along the circular arc $\widehat{BC}$, $AB + AC$ attains its maximum when A is the midpoint M. (1)

In [**5**, Problem 9, pp. 16–17], he describes (1) as *intuitively obvious* and gives a proof. Actually, the last paragraph there can be thought of as another proof. In [**6**, pp. 21–24], he gives two more proofs. He attributes the first to I. van Yzeren and describes it as *full*

Math. Mag. **83** (2010) 279–283. doi:10.4169/002557010X521813.

of ingenuity, and attributes the second to K. A. Post and describes it as *most elegant*. He describes the problem as *unusually rich in interesting approaches*. In this note, we confirm this last phrase by giving three more proofs. We also examine how certain steps in our proofs and in the proofs in Honsberger's books are related to propositions in Euclid's *Elements* and to other problems in geometry.

Our proofs are simple, short, and transparent, and they prove the following statement that is slightly stronger than (1):

> In FIGURE 1, as A moves along the circular arc $\widehat{BC}$ from B to the midpoint M, $AB + AC$ increases. (2)

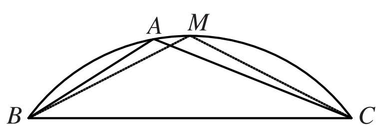


Figure 1

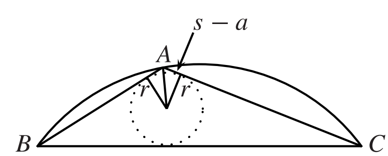


Figure 2

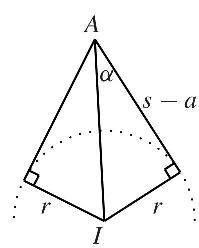


Figure 3

In these proofs, we follow common practice and we denote the side-lengths and angles of triangle ABC by a, b, c, A, B, C in the standard order, the semi-perimeter by s, the area by $[ABC]$, and the inradius by r. We also let $A = 2\alpha$, $B = 2\beta$, $C = 2\gamma$. The angle A can have any positive measure less than 180°. This means that the arc $\widehat{BC}$ can have any size, and is not necessarily a minor arc as shown in FIGURE 1.

Proof #1. Referring to FIGURE 1 and using the law of sines and simple trigonometric identities, we obtain

$$\frac{b+c}{a} = \frac{\sin 2\beta + \sin 2\gamma}{\sin 2\alpha} = \frac{2\sin(\beta+\gamma)\cos(\beta-\gamma)}{2\sin\alpha\cos\alpha} = \frac{\cos(\beta-\gamma)}{\sin\alpha}.$$

Since a and α are fixed, $b + c$ is proportional to $\cos(\beta - \gamma)$, and thus increases as A moves from B to M. ■

Proof #2. We again refer to FIGURE 1. Using the law of cosines $b^2 + c^2 - a^2 = 2bc\cos A$, the area formula $2[ABC] = bc\sin A$, and the double-angle formulas $\sin A = 2\sin\alpha\cos\alpha$, $1 + \cos A = 2\cos^2\alpha$, we obtain

$$(b+c)^2 - a^2 = (b^2 + c^2 - a^2) + 2bc = 2bc\cos A + 2bc = 2bc(1+\cos A)$$
$$= \frac{bc\sin A}{2}\,\frac{4(1+\cos A)}{\sin A} = [ABC](4\cot\alpha).$$

Since a is fixed and cot α is fixed and positive, it follows that $(b+c)^2$ (and hence $b+c$) increases with $[ABC]$. Since $[ABC]$ increases as A moves from B to M along $\widehat{BMC}$, so does $b+c$, as desired. ■

Proof #3. We refer to FIGURES 2 and 3. Multiplying the obvious relation $s - a = r\cot\alpha$ by 2 and then by $a + b + c$, and using $2[ABC] = r(a + b + c)$, we obtain

$$b + c - a = 2r\cot\alpha \tag{3}$$

$$(b + c)^2 - a^2 = 4[ABC]\cot\alpha. \tag{4}$$

This is the last relation in the previous proof. ∎

Beside being very short, Proof #3 has the advantage of showing that (2) still holds if $b + c$ is replaced by the inradius r. In fact, (4) implies that if [*ABC*] increases, then $b + c$ increases, and (3) implies that if $b + c$ increases, then r increases. Of course, a is fixed and hence one can also replace $b + c$ by the perimeter $2s = a + b + c$ of *ABC*. We combine this with (2) in the following theorem:

THEOREM 1. *If A moves along the circular arc $\widehat{BC}$ from B to the midpoint M, then*

(i) *the area* [*ABC*], (ii) *the perimeter* $2s$, *and* (iii) *the inradius r*

of triangle ABC increase.

So far, we have taken Part (i) of Theorem 1 for granted. Its weaker form that *the maximum of* [*ABC*] *occurs at* $A = M$ has also been used as *obvious* in Post's proof in [**6**, p. 23]. Looking for a proof, we discovered that Theorem 1(i) appears, in disguise, within Proposition EE.III.15—meaning Book III, Proposition 15—in Euclid's *Elements*. To see this, complete the circle in FIGURE 1, draw a diameter *UV* parallel to *BC*, and drop a perpendicular *AX* on *BC* that meets *UV* at *Y* and the circle at *Z*; see FIGURE 4. Since a is fixed, [*ABC*] is proportional to, and hence increases with, *AX*. Since *XY* is fixed and $AZ = 2AY$, *AX* increases with *AZ*. Thus Theorem 1(i) can be restated as follows:

In FIGURE 4, as A moves along the circular arc $\widehat{BC}$ from B to its midpoint M, AZ increases. (5)

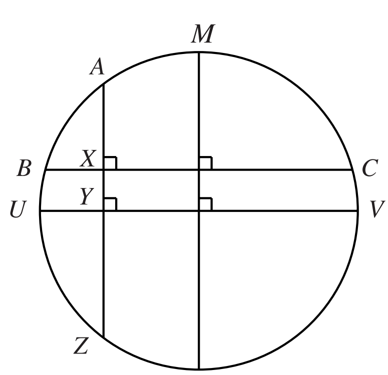


Figure 4

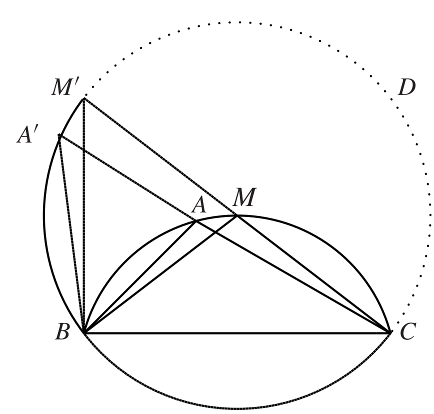


Figure 5

This statement is a special case, and the first step in the proof, of EE.III.15 which reads as follows:

PROPOSITION EE.III.15. *For two chords in a circle, the one that is nearer to the center is longer.*

Although this follows immediately from Pythagoras' theorem, Euclid's longer proof has the advantage of showing that this is a theorem in neutral geometry.

It is interesting that another obscure proposition of Book III is needed if the proof of (1) that appears in [**5**] is to be modified so that it yields (2). That proof has two components. The first consists in drawing an auxiliary circular arc centered at M and passing through B and C and lying on the same side of BC as M; see FIGURE 5. For each A on the arc $\widehat{BMC}$, one lets A' be the point where the ray CA meets the new arc. From

$$\angle AA'B + \angle ABA' = \angle BAC = \angle BMC = 2\angle BM'C = 2\angle BA'C,$$

it follows that $\angle ABA' = \angle AA'B$ and $AA' = AB$ and hence $AB + AC = AA' + AC = CA'$. In view of this, the second component of the proof (of (2)) would need the following statement:

> In FIGURE 5, as a point A' moves along a semi-circle $\widehat{CBM'}$ from B to M', the length of CA' increases. (6)

But this is the special case $P = U$ of EE.III.7.

PROPOSITION EE.III.7. *If $\widehat{UV}$ is a semi-circle with center O and diameter UV and if P is between O and U but not equal to O, then as X moves from U to V on $\widehat{UV}$, the length of PX increases.*

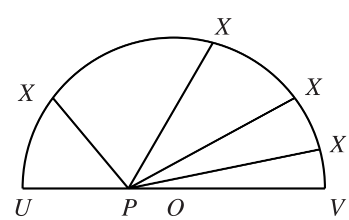


Figure 6

For the sake of completeness, we mention that Proposition EE.III.8 deals with the case when P is on the extension of OU. We also mention that it may be debatable whether Proposition EE.III.7 was meant to cover the extreme case $P = U$. This may be why this special case is added as a separate theorem in Heath's book [**4**, p. 20, last paragraph]. We also add that for proving (1) one does not need Proposition EE.III.7 but rather the simpler fact that *the length of UX attains its maximum when UX is a diameter*.

We end this note with two remarks. First, the first component of the proof in [**5**] described above is interesting on its own since it uses exactly the same configuration used in the proof of the celebrated *Broken Chord Theorem of Archimedes*; see [**7**, pp. 1–2] and compare with FIGURE 5. Secondly, it should be mentioned that our Proof #1 is inspired by a lemma that Robert Breusch had designed to solve a *Monthly* problem; see [**8**]. That lemma, together with its proof, is reproduced in [**1**] and [**2**], where it is used by the present author to give short proofs of Urquhart's theorem and of a stronger form of the Steiner-Lehmus theorem. Hyperbolic versions of Breusch's Lemma, Urquhart's Theorem, and the Steiner-Lehmus Theorem can be found in [**9**, 4.19–4.21, pp. 151–158].

(The title of this note alludes to the author's paper [**3**] on a different Honsberger topic.)

REFERENCES

1. M. Hajja, A very short and simple proof of the most elementary theorem in Euclidean geometry, *Forum Geom.* **6** (2006) 167–169.
2. M. Hajja, Stronger forms of the Steiner-Lehmus theorem, *Forum Geom.* **8** (2008) 157–161.
3. M. Hajja, On a morsel of Ross Honsberger, *Math. Gaz.* **93** (2009) 309–312.
4. Sir Thomas L. Heath, *Euclid—The Thirteen Books of the Elements, Vol. 2 (Books III–IX)*, 2nd ed., Dover Publications, New York, 1956.
5. R. Honsberger, *Mathematical Morsels*, Mathematical Association of America, Washington, DC, 1978.
6. R. Honsberger, *Mathematical Gems III*, Mathematical Association of America, Washington, DC, 1985.
7. R. Honsberger, *Episodes in Nineteenth and Twentieth Century Euclidean Geometry*, Mathematical Association of America, Washington, DC, 1995.
8. E. Trost, Problem 4964, *Amer. Math. Monthly* **68** (1961) 384; solution by R. Breusch, *Amer. Math. Monthly* **69** (1962) 672–674. doi:10.2307/2311604
9. A. A. Ungar, *A Gyrovector Space Approach to Hyperbolic Geometry*, Synthesis Lectures on Mathematics and Statistics, Morgan & Claypool, San Rafael, CA, 2009.

Summary In two of his books, Ross Honsberger presented several proofs of the fact that the point A on the circular arc $\widehat{BC}$ for which $AB + AC$ is maximum is the midpoint of the arc. In this note, we give three more proofs and examine how these proofs and those of Honsberger are related to propositions in Euclid's *Elements* and, less strongly, to other problems in geometry such as the broken chord theorem, Breusch's lemma, Urquhart's theorem, and the Steiner-Lehmus theorem.

Cantor's Other Proofs that $\mathbb{R}$ Is Uncountable

JOHN FRANKS
Northwestern University
Evanston, IL 60208-2730
j-franks@math.northwestern.edu

There are many *theorems* that are widely known among serious students of mathematics, but there are far fewer *proofs* that are part of our common culture. One of the best known proofs is Georg Cantor's diagonalization argument showing the uncountability of the real numbers $\mathbb{R}$. Few people know, however, that this elegant argument was not Cantor's first proof of this theorem, or, indeed, even his second! More than a decade and a half before the diagonalization argument appeared Cantor published a different proof of the uncountability of $\mathbb{R}$. The result was given, almost as an aside, in a paper [**1**] whose most prominent result was the countability of the algebraic numbers. Historian of mathematics Joseph Dauben has suggested that Cantor was deliberately downplaying the most important result of the paper in order to circumvent expected opposition from Leopold Kronecker, an important mathematician of the era who was an editor of the journal in which the result appeared [**4**, pp. 67–69]. A fascinating account of the conflict between Cantor and Kronecker can be found in Hal Hellman's book [**6**]. A decade later Cantor published a different proof [**2**] generalizing this result to perfect subsets of $\mathbb{R}^k$. This still preceded the famous diagonalization argument by six years.

Mathematical culture today is very different from what it was in Cantor's era. It is hard for us to understand how revolutionary his ideas were at the time. Many mathematicians of the day rejected the idea that infinite sets could have different cardinalities. Through much of Cantor's career many of his most important ideas were treated with skepticism by some of his contemporaries (see [**6**] for an interesting account of some of the disputes).

Math. Mag. **83** (2010) 283–289. doi:10.4169/002557010X521822.

As mentioned above, Cantor's first proof was in a paper [**1**] whose main result was the countability of the algebraic numbers—those real numbers which are roots of polynomials with integer coefficients. Since the real numbers are uncountable and the algebraic numbers are only countable there must be infinitely many (in fact, uncountably many) real numbers which are not algebraic. Such numbers are called *transcendental*. The fact that transcendental numbers exist had been established by Joseph Liouville, only about thirty years earlier and was itself still the subject of controversy.

Cantor's early proofs of uncountability are nearly as simple as his more famous diagonalization proof and deserve to be better known. In this expository note we present all three of these proofs and explore the relationships between them. Understanding multiple proofs of an important result almost always leads to a deeper understanding of the concepts involved.

Cantor's first proof

Recall that a set X is *countably infinite* if there is a bijection (or one-to-one correspondence) between the elements of X and the natural numbers $\mathbb{N} = \{1, 2, 3 \dots\}$. Equivalently, X is *countably infinite* if there is a sequence $\{x_k\}_{k=1}^{\infty}$ of distinct elements in which every element of X occurs precisely once. An infinite set that is not countable is called *uncountable*. So to prove that a set X is uncountable we must show that for every sequence $\{x_k\}_{k=1}^{\infty}$ of distinct elements of X there must be an element of X which is omitted by that sequence. Different sequences will omit different elements, of course, but there is no one sequence which contains every element of X.

Cantor's first proof of the uncountability of $\mathbb{R}$ was published in 1874 and is based on the fact that bounded monotonic sequences of real numbers converge.

THEOREM 1. (CANTOR [**1**]) *If $\{x_k\}_{k=1}^{\infty}$ is a sequence of distinct real numbers there is at least one $z \in \mathbb{R}$ which does not occur in this sequence.*

Proof. Let $\{x_k\}_{k=1}^{\infty} = x_1, x_2, \dots$ be a sequence of distinct real numbers. Define a sequence of closed intervals $I_n = [a_n, b_n]$ as follows. Let a_1 be the smaller of x_1 and x_2 and b_1 be the larger. Define I_1 to be $[a_1, b_1]$. We define I_n recursively. Given the non-trivial interval $I_{n-1} = [a_{n-1}, b_{n-1}]$ let y and y' be the first two elements of the sequence $\{x_k\}_{k=1}^{\infty}$ which lie in the open interval (a_{n-1}, b_{n-1}). (Clearly such y and y' must exist or there are infinitely many choices of elements of the interior of I_{n-1} which are not in the sequence $\{x_k\}_{k=1}^{\infty}$ and our proof is done.) Define a_n to be the smaller of y and y' and b_n to be the larger and let $I_n = [a_n, b_n]$.

From their construction it is clear that these closed intervals are non-trivial and nested. That is, for each index n,

$$a_{n-1} < a_n < b_n < b_{n-1},$$

and hence $I_n \subset I_{n-1}$. So the sequence $\{a_k\}_{k=1}^{\infty}$ is strictly increasing and bounded above (for example any b_n is an upper bound) and the sequence $\{b_k\}_{k=1}^{\infty}$ is strictly decreasing and bounded below.

Cantor then appealed to the fact that bounded monotonic sequences always have limits. He defined:

$$a_\infty = \lim_{n\to\infty} a_n \quad \text{and}$$

$$b_\infty = \lim_{n\to\infty} b_n$$

He observed that since $a_n < b_n$ for all n, we have $a_\infty \le b_\infty$ and the interval $[a_\infty, b_\infty]$ contains at least one point.

If $z \in [a_\infty, b_\infty]$ then

$$a_n < z < b_n \quad \text{for all } n \in \mathbb{N}, \tag{1}$$

and in particular $z \neq a_n$ and $z \neq b_n$.

We will prove by contradiction that z cannot occur in the sequence $\{x_k\}_{k=1}^{\infty}$. To do this we assume z is in the sequence and show this assumption leads to a contradiction. If z does occur in this sequence then there are only finitely many points preceding it in the sequence and hence only finitely many elements of the subsequence $\{a_n\}_{n=1}^{\infty}$ preceding it. Let a_m be the last element of the subsequence $\{a_n\}_{n=1}^{\infty}$ which precedes z in the sequence $\{x_k\}_{k=1}^{\infty}$.

We defined a_{m+1} and b_{m+1} to be the first two elements of the sequence $\{x_k\}_{k=1}^{\infty}$ which lie in the interior of I_m. Since z is in the interior of I_m, by Equation (1), and is not equal to either a_{m+1} or b_{m+1}, it must be that a_{m+1} and b_{m+1} precede z in the sequence $\{x_k\}_{k=1}^{\infty}$. This contradicts the definition of a_m as the last element of the subsequence $\{a_n\}_{n=1}^{\infty}$ preceding z in this sequence. This contradiction implies that the assumption that z is in the sequence $\{x_k\}_{k=1}^{\infty}$ is false and hence proves the result. ■

Cantor also remarked that, in fact, the sequence $\{x_k\}_{k=1}^{\infty}$ omits at least one point in any non-empty open interval (a, b), because we could choose a_1 and b_1 to be the first two points of the sequence which lie in this interval. According to historian Joseph Dauben, this published proof benefited from some simplifications due to the German mathematician Richard Dedekind who had seen a more complicated early draft [**4**, pp. 50–52].

Indeed, the heart of this proof is the fact that bounded monotonic sequences have limits. Mathematicians in 1874 would have accepted this as a fact, but it is worth remembering that the rigorous foundations for results such as this were still being established. It was only two years earlier, in 1872, that Dedekind had published his monograph, *Stetigkeit und irrationale Zahlen,* or *Continuity and irrational numbers* [**5**]. It was in this monograph that he introduced what we now call "Dedekind cuts" as a foundation for the construction of the real numbers. This construction provided the basis for what in modern terminology is called the *completeness* of the real numbers and in particular the existence of limits for bounded monotonic sequences.

Perfect sets

In 1884 Cantor published a generalization of Theorem 1 which asserts that any perfect subset of $\mathbb{R}^k$ is uncountable. Recall that a subset X of $\mathbb{R}^k$ is said to be *perfect* if X is closed and every point x of X is a limit of a sequence of points in X which are distinct from x.

THEOREM 2. (CANTOR [**2**]) *Suppose X is a perfect subset of $\mathbb{R}^k$. Then X is uncountable.*

Proof. We will again show that if $\{x_n\}_{n=1}^{\infty}$ is a sequence in X, then there is a $z \in X$ which is not a term in this sequence.

Since X is perfect, for every x in X, a ball of any positive radius centered at x contains infinitely many points of X. From this it is easy to see that if B is a closed ball in $\mathbb{R}^k$ centered at a point of X, and y is any point of X, then there is another closed ball B' which is contained in B, is centered at a point of X, and does not contain the point y.

This property is used to construct recursively a sequence $\{z_n\}_{n=1}^{\infty}$ which has a limit z which is not an element of our original sequence. At the same time we construct a nested sequence of closed balls $\{B_n\}_{n=1}^{\infty}$ with each B_n centered at z_n. Let B_0 be any closed ball of positive diameter D centered at a point z_0 of X. Given B_{n-1} choose a closed ball B_n such that

- The ball B_n is a subset of B_{n-1};
- The center of the ball B_n, which we denote z_n, is a point of X;
- The ball B_n does not contain the point x_n; and
- The diameter of B_n is at most half the diameter of B_{n-1}.

Notice that, for $1 \leq m \leq n$, the point x_m is not in B_n.

From the fact that each successive diameter is at most half of the previous one, it is easy to see by induction that the diameter of $B_n \leq D/2^n$. Cantor observed that the sequence $\{z_n\}_{n=1}^{\infty}$ is what we now call a Cauchy sequence. This is because if $n, m > N$, then z_n and z_m are in B_N so

$$\|z_n - z_m\| < \frac{D}{2^N}.$$

Since the sequence $\{z_n\}_{n=1}^{\infty}$ is a Cauchy sequence it has a limit in $\mathbb{R}^k$ which we will denote z. Since X is a closed set and $z_n \in X$, the limit point z is also in X.

For any $n > 0$, all of the points z_m with $m \geq n$ are in the closed ball B_n so their limit z must also be in B_n. But recall that by construction the point x_n is *not* in B_n. Hence for every n it must be that $z \neq x_n$. ■

Using what we now know about compactness the proof above can be significantly simplified. Having constructed a nested family of balls B_n each of which contains some point of X and with $x_n \notin B_n$, we don't need to worry about centers or diameters or Cauchy sequences. Instead we let $Z_n = X \cap B_n$. Then Z_n is closed and bounded and hence compact. It is also non-empty since each B_n contains at least one point of X. And, of course, $Z_n \subset Z_{n-1}$. These properties imply that the nested intersection $\bigcap_{n=1}^{\infty} Z_n$ is non-empty. If z is a point of this intersection then for each $n \in \mathbb{N}$, $z \in B_n$ and hence $z \neq x_n$. So the point z is not in the sequence $\{x_n\}_{n=1}^{\infty}$.

Of course this line of proof was not available to Cantor. He could not have known that a nested intersection of non-empty compact sets is non-empty—indeed the concept of compactness was unknown at the time he wrote this paper. It was not until 1894 that Émile Borel proved that an open cover of a closed interval has a finite subcover. See [**7**] for a history of the concept of compactness. What we consider the standard properties of compactness were not developed until the 20th century.

Cantor published this result in §16 of [**2**]. It is interesting that it appeared a decade after his first proof (Theorem 1) and still well prior to the so-called diagonalization proof which we discuss in the next section. It certainly bears a resemblance to his first proof but, as we will see, it also strongly foreshadows the more famous diagonalization argument.

The diagonalization proof

More than a decade and a half after his first proof Cantor published the much more famous proof of the uncountability of $\mathbb{R}$ which has become associated with his name. This was the introduction of what is now called the *Cantor diagonalization argument*.

THEOREM 3. (CANTOR [**3**]) *The unit interval* [0, 1] *is not countable.*

Proof. Let X denote the subset of [0, 1] consisting of those numbers which have decimal representations containing only the digits 4 and 9. We choose 4 and 9 for concreteness; other choices would work as well. We know in general that two different decimal expansions can represent the same real number. For example,

$$0.4999\cdots = 0.5000\ldots,$$

where the first decimal ends in an infinitely repeating sequence of 9's and the second in an infinitely repeating sequence of 0's. But if we allow ourselves only to use the digits 4 and 9 there is only one way to write this number.

Indeed, the representation for any number in the set X using only the digits 4 and 9 is unique. To see this suppose u and v are elements of X, so they have decimal representations using only 4 and 9; or more formally, suppose

$$u = \sum_{i=1}^{\infty} \frac{u_i}{10^i}, \quad \text{and} \quad v = \sum_{i=1}^{\infty} \frac{v_i}{10^i},$$

where each u_i and v_i is either 4 or 9. Suppose these decimal representations differ first in the nth place, so $u_i = v_i$ for $1 \le i < n$ and $u_n \ne v_n$. Let w denote the number with decimal representation equal to the decimal representation of u and v in places 1 to $n-1$ (where they agree) and with 0 in all other places so

$$w = \sum_{i=1}^{n-1} \frac{u_i}{10^i} = \sum_{i=1}^{n-1} \frac{v_i}{10^i}$$

Since u and v disagree in the nth place the larger of them has a 9 in this place and must be greater than $w + 9 \times 10^{-n}$. Similarly, the smaller of u and v has a 4 in the nth place and must be at most $w + 5 \times 10^{-n}$. Hence $|u - v| > 4 \times 10^{-n} > 0$ so $u \ne v$. This shows that two different decimal representations, which use only the digits 4 and 9, must actually represent different numbers.

Now given any sequence $\{x_k\}_{k=1}^{\infty}$ in X we define an element z by specifying its decimal expansion using a process called *diagonalization*. Specifically let

$$z = \sum_{k=1}^{\infty} \frac{z_k}{10^i},$$

where

$$z_k = \begin{cases} 4, & \text{if the } k\text{th decimal digit of } x_k \text{ is 9;} \\ 9, & \text{if the } k\text{th decimal digit of } x_k \text{ is 4.} \end{cases}$$

We conclude that z is in X, since its decimal expansion contains only the digits 4 and 9. But it is not an element of the sequence $\{x_k\}_{k=1}^{\infty}$ since z differs from x_k in the kth decimal place. It follows that it is not possible to enumerate the elements of the set X. In other words, there is no sequence $\{x_k\}_{k=1}^{\infty}$ of elements of X which contains all the elements of X. This proves X is uncountable.

There is a subtle point here. We have *not* found one z which is omitted from every sequence $\{x_k\}_{k=1}^{\infty}$. Instead we have shown that for each sequence $\{x_k\}_{k=1}^{\infty}$ there is an omitted z—different sequences will omit different elements of X.

Since [0, 1] contains the uncountable set X, it must also be uncountable. (We are using the fact that a subset of a countable set is also countable.) ■

There are two parts to this proof. In the first part we showed that there is a subset X whose elements can be uniquely specified by a decimal expansion containing only the digits 4 and 9, i.e., an infinite sequence of 4's and 9's. In fact, Cantor did not include this part of the proof in his original paper. It is not difficult to show and he probably considered it obvious. He also did not use 4 and 9 but instead used the letters m and w to represent arbitrary distinct digits. Essentially the same argument given above will show that two decimal representations of a single number must be identical if they both use only the same two digits.

The second part of the proof uses what has come to be called a *diagonalization argument* to show that the collection of all such infinite sequences is not countable. The term diagonalization is used because one way to view the construction of z given in the proof is to use the sequence $\{x_n\}$ in X to make an infinite matrix M. The first row of the matrix M consists of the decimal digits of x_1, the second row the decimal digits of x_2, and the nth row the decimal digits of x_n. So M_{ij} is the jth decimal digit of x_i. Then the element z which does not occur in the sequence is obtained from the *diagonal* of M. More precisely z_n, the nth decimal digit of z, is 4 if $M_{nn} = 9$ and 9 if $M_{nn} = 4$. Then z does not correspond to any row of the matrix M because the nth decimal digit of z is different from the diagonal entry M_{nn}. So z does not correspond to any row of the matrix M and hence z is not in the sequence $\{x_n\}$.

As mentioned above the proof for perfect subsets of $\mathbb{R}^k$ (Theorem 2 above) strongly foreshadows the diagonalization argument. To see this, let X be the subset of $[0, 1]$ consisting of those numbers with decimal representations containing only the digits 4 and 9. It is an easy exercise to show that X is perfect, though we will not need this fact. Let $X_0 = X$ and let X_n be the subset of X_{n-1} consisting of all of those points whose nth decimal digit (4 or 9) is different from the nth decimal digit of x_n. Then $\{X_n\}_{n=0}^{\infty}$ is a nested family of compact sets and $\bigcap_{n=1}^{\infty} X_n$ consists of the single point produced by the diagonalization in the proof of Theorem 3.

There is a slightly different and very clever way to make the diagonalization part of Cantor's argument. Recall that $\mathcal{P}(\mathbb{N})$, the power set of the natural numbers $\mathbb{N}$, is the set of all subsets of $\mathbb{N}$. We first observe that there is a bijection from X, the set of infinite sequences of 4's and 9's, to $\mathcal{P}(\mathbb{N})$. This bijective correspondence is given by

$$A \longleftrightarrow \{x_n\}_{n=1}^{\infty}$$

where A is a subset of $\mathbb{N}$ and $x_i = 9$ if $i \in A$ and $x_i = 4$ otherwise. Thus, it suffices to show that the set $\mathcal{P}(\mathbb{N})$ is uncountable. This can be done as a special case of a more general argument.

PROPOSITION 4. *Suppose S is a non-empty set and $f : S \to \mathcal{P}(S)$ is a function from S to its power set. Then f is not surjective.*

Proof. For each $x \in S$ either $x \in f(x)$ or $x \notin f(x)$. Let $Y = \{y \in S \mid y \notin f(y)\}$. Let x be any element of S. From the definition of Y we observe that x is in Y if and only if x is not in the set $f(x)$. Hence the sets Y and $f(x)$ can never be equal since one of them contains x and the other does not. Therefore, there is no x with $f(x) = Y$, so f is not surjective. ■

This proposition implies that any set S has a cardinality which is less than that of its power set $\mathcal{P}(S)$ and, in particular, $\mathcal{P}(\mathbb{N})$ is uncountable. The proof of Proposition 4 is really just a disguised version of the diagonalization argument and consequently this proposition is also sometimes referred to as Cantor's diagonalization theorem.

Acknowledgment Supported in part by NSF grant DMS0901122.

REFERENCES

1. Georg Cantor, Ueber eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen, *Journal für die reine und angewandte Mathematik* (1874) 258–261. Reprinted and translated in [**8**]. doi:10.1515/crll.1874.77.258

2. Georg Cantor, Ueber unendliche lineare Punktmannichfaltigkeiten, Nr.6, *Math. Annalen* **23** (1884) 453–488. doi:10.1007/BF01446598
3. Georg Cantor, Über eine elementare Frage de Mannigfaltigketslehre, *Jahresber. Deutsch. Math.-Verein.* **1** (1890–1891) 75–78.
4. Joseph Dauben, *Georg Cantor: His Mathematics and Philosophy of the Infinite*, Princeton University Press, Princeton, NJ, 1990.
5. Richard Dedekind, *Essays on the Theory of Numbers. I: Continuity and Irrational Numbers; II: The Nature and Meaning of Numbers*, authorized translation by Wooster Woodruff Beman, Dover, New York, 1963.
6. Hal Hellman, *Great Feuds in Mathematics: Ten of the Liveliest Disputes Ever*, John Wiley, Hoboken, NJ, 2006.
7. Manya Janaky Raman, *Understanding Compactness: A Historical Perspective*, M.A. thesis, University of California Berkeley, 1997.
8. Jacqueline Stedall, *Mathematics Emerging: A Sourcebook 1540–1900,* Oxford University Press, New York, 2008.

Summary This expository note describes some of the history behind Georg Cantor's proof that the real numbers are uncountable. In fact, Cantor gave three different proofs of this important but initially controversial result. The first was published in 1874 and the famous diagonalization argument was not published until nearly two decades later. We explore the different ideas used in each of his three proofs.

Nothing Lucky about 13

B. SURY
Stat-Math Unit, Indian Statistical Institute
8th Mile Mysore Road
Bangalore 560 059 India
sury@isibang.ac.in

Recently, a high school teacher came across the following problem which he passed on to a forum for mathematics teachers:

$$\text{Evaluate } \cos\left(\frac{2\pi}{13}\right) + \cos\left(\frac{6\pi}{13}\right) + \cos\left(\frac{8\pi}{13}\right).$$

One could solve this in a number of elementary ways, and as we will show below, the value turns out to be $\frac{-1+\sqrt{13}}{4}$. The point here is to find what is special about 13 and about 2, 6, 8.

Without further ado, let us break the illusion that 13 might be particularly "lucky" to admit such a simple expression: We show a corresponding result for every prime number congruent to 1 modulo 4 and, indeed, for every prime.

Here we will explain briefly how to prove for any prime number $p \equiv 1$ modulo 4 the identity

$$\sum_{a \in Q} \cos\left(\frac{2a\pi}{p}\right) = \frac{-1+\sqrt{p}}{2}, \tag{1}$$

where the sum is over the set Q of quadratic residues mod p; that is, $a \in Q$ if $1 \leq a \leq p-1$ and for some integer b, $a \equiv b^2 \bmod p$. When $p \equiv 1 \bmod 4$, then -1 is a square mod p; indeed, for those who know it, we mention that Wilson's congruence $(p-1)! \equiv -1 \bmod p$ simplifies to $((\frac{p-1}{2})!)^2 \equiv -1 \bmod p$ in the case $p \equiv 1 \bmod 4$. Thus the squares mod p (as well as the nonsquares mod p) come in pairs $a, -a$ with

Math. Mag. **83** (2010) 289–293. doi:10.4169/002557010X521840.

exactly one of these less than or equal to $(p-1)/2$. As $\cos(t) = \cos(-t)$, the identity (1) could be rewritten as

$$\sum_{\substack{a\in Q \\ a\le (p-1)/2}} \cos\left(\frac{2a\pi}{p}\right) = \frac{-1+\sqrt{p}}{4},$$

and this explains the opening result, as the squares mod 13 are $\pm 1, \pm 3, \pm 4$.

The identity mentioned for primes congruent to 1 mod 4 has an analog for primes congruent to -1 mod 4.

The secret is the so-called Gauss sum, which Gauss used to prove the quadratic reciprocity law. Let p be an odd prime. The Gauss sum is the expression $\sum_{a=1}^{p-1} \pm z^a$, with $z = e^{2i\pi/p}$, where we use a plus sign if a is a square mod p and put a minus sign if a is not. The Legendre symbol $\left(\frac{a}{p}\right)$, which denotes 1 or -1 depending on whether a is a square mod p or not, allows us to express this more clearly: Write

$$G := \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) z^a.$$

It is a remarkable fact that $G^2 = \pm p$ with the sign determined by whether $p \equiv \pm 1$ mod 4. With some care for the signs, we can show that G is $\sqrt{p}$ or $i\sqrt{p}$ as p is 1 or -1 mod 4 [**1**, pp. 70–76]. We have, therefore,

$$G = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) z^a = \begin{cases} \sqrt{p} & \text{if } p \equiv +1 \text{ mod } p \text{, and} \\ i\sqrt{p} & \text{if } p \equiv -1 \text{ mod } p. \end{cases} \tag{2}$$

(A different choice of primitive pth root of unity as z can lead to a different sign for G.)

In this note we will first show how to use (2) to prove identities like (1) and its analogy for primes congruent to -1 mod 4. After this main result, we make a brief tour of the history of Gauss sums and, specifically, of the determination of their signs. Finally, we give a proof of (2).

The main result is :

THEOREM. *Let p be an odd prime and let Q be the subset of squares in $\mathbb{Z}_p^*$. Then,*

(i) *if $p \equiv 1$ mod 4, so that $Q = T \cup -T$ with $T \subseteq \{1, \dots, \frac{p-1}{2}\}$, then*

$$\sum_{a\in Q} \cos\left(\frac{2a\pi}{p}\right) = 2\sum_{b\in T} \cos\left(\frac{2b\pi}{p}\right) = \frac{-1+\sqrt{p}}{2}.$$

(ii) *If $p \equiv -1$ mod 4, so that $\mathbb{Z}_p^* = Q \cup -Q$, then*

$$\sum_{a\in Q} \sin\left(\frac{2a\pi}{p}\right) = \frac{\sqrt{p}}{2} \quad \textit{and} \quad \sum_{a\in Q} \cos\left(\frac{2a\pi}{p}\right) = \frac{-1}{2}.$$

Remark The prime 2 is special and, when we apply the proof below to it, we get the trivial identity $\cos(\pi) = -1$.

Proving the theorem We consider first the case when $p \equiv 1$ mod 4, when $Q = T \cup -T$ with $T \subseteq \{1, \dots, \frac{p-1}{2}\}$. If N denotes the nonsquares mod p,

$$G = \sqrt{p} = \sum_{a\in Q} z^a - \sum_{b\in N} z^b.$$

On the other hand, since it is well known that the sum of the roots of unity sum to zero, we have

$$-1 = \sum_{c\in\mathbb{Z}_p^*} z^c = \sum_{a\in Q} z^a + \sum_{b\in N} z^b. \tag{3}$$

Adding the two equations and substituting $z = e^{2ip/p}$, we have

$$-1 + \sqrt{p} = 2\sum_{a\in Q} z^a = \sum_{a\in Q}(z^a + z^{-a}) = \sum_{a\in Q} 2\cos\left(\frac{2a\pi}{p}\right)$$

which is our first claim.

If $p \equiv -1 \bmod 4$, so that $\mathbb{Z}_p^* = Q \cup -Q$, then $G = i\sqrt{p}$ gives

$$\sum_{a\in Q}(z^a - z^{-a}) = i\sqrt{p};$$

which quickly simplifies to $\sum_{a\in Q} 2\sin(2a\pi/p) = \sqrt{p}$. The second identity in (ii) follows immediately from (3).

Some history of Gauss sums Gauss sums were introduced by Gauss in 1801, when he stated some of their properties and used them to prove the quadratic reciprocity law in different ways. Gauss wrote that he had studied since 1805 the theory of cubic and biquadratic residues and, since results for these proved elusive, that he was motivated to find more proofs of the quadratic reciprocity law, hoping that one of them would yield a generalization for higher reciprocity laws. Gauss's fourth and sixth proofs of the quadratic reciprocity law used Gauss sums and, indeed, proved successful in investigating higher reciprocity laws.

The sign of the Gauss sum was a notoriously difficult question; he recorded the correct assertion in his mathematical diary in May 1801, but could find a proof only in 1805. He says in a letter to Olbers written in September 1805 that he was annoyed by this inability to determine the sign and that hardly a week went by for those 4 years when he did not make one or more unsuccessful attempt. He says that finally the mystery was solved "the way lightning strikes" [**1**].

As mentioned earlier, if z is a primitive pth root of unity, the sign of the sum $\sum_{a=1}^{p-1}\left(\frac{a}{p}\right)z^a$ depends on the choice of z. However, the key observation seems to be that the equality

$$\sum_{a=0}^{p-1}\left(\frac{a}{p}\right)z^a = \prod_{b=1}^{(p-1)/2}(z^{-b/2} - z^{b/2})$$

holds for any choice of primitive pth root of unity z. This can be deduced from a result on polynomials and, it is in this context that Gauss introduced the so-called Gaussian polynomials which generalize the binomial coefficients. Proofs to determine the sign of the Gauss sum were found later by Kronecker, Schur, Mertens, etc. A beautiful proof by Schur appears in Landau's classic German text [**2**, pp. 162–166]. Although Ireland and Rosen is a convenient modern reference for the Gauss sum computation [**1**, pp. 70–76], we take the liberty of recalling Schur's proof briefly for the sake of English-speaking readers.

THEOREM. *Let $n > 0$ be odd. Then $S := \sum_{s=0}^{n-1} e^{2i\pi s^2/n} = \sqrt{n}$ or $i\sqrt{n}$ depending on whether $n \equiv \pm 1$ mod 4.*

Before proving this result, we mention that when n is prime, $S = G$, the Gauss sum. This is again due to the observation that $\sum_{a=0}^{p-1} e^{2i\pi a/p} = 0$ mentioned earlier.

Proof (Schur). Put $z = e^{2i\pi/n}$ and consider the $n \times n$ matrix $A = (z^{kl})_{0\le k,l<n}$. Our sum is $S = \sum_k z^{k^2} = \operatorname{tr} A = \sum_{r=1}^{n} \lambda_r$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Viewing S as the trace of a matrix involving roots of unity proves advantageous because sums involving roots of unity often admit lots of cancellations.

The u, v entry of A^2 is $(A^2)_{u,v} = \sum_w z^{(u+v)w} = b_{u+v}$, where $b_m = \sum_w z^{mw}$. If $n \mid m$, then evidently $b_m = \sum_w z^{mw} = \sum_w 1 = n$. On the other hand, if $n \nmid m$, we have $z^m b_m = \sum_w z^{m(w+1)} = b_m$, which gives $b_m = 0$ since $z^m \neq 1$. Note that $\sum_r \lambda_r^2 = \operatorname{tr} A^2 = \sum_u b_{2u} = n$. Also, $(A^4)_{uv} = \sum_w b_{u+w} b_{w+v} = n^2$ or 0, depending on whether $u = v$ or not. Thus $A^4 = n^2 I$ where I is the $n \times n$ identity matrix.

The characteristic polynomial $\chi_{A^4}(\lambda)$ of A^4 is $(\lambda - n^2)^n$, which means that the eigenvalues $\lambda_1^4, \ldots, \lambda_n^4$ are all equal to n^2. In particular, $\lambda_r = i^{a_r}\sqrt{n}$ where $a_r = 0$, 1, 2, or 3. For each $k = 0, 1, 2, 3$, we count the number of eigenvalues with that power of i, by setting $m_k = |\{a_r : a_r = k\}|$. Note that $m_0 + m_1 + m_2 + m_3 = n$, because there are n eigenvalues.

We first show that $|S|^2 = n$. We start with

$$|S|^2 = S\bar{S} = \sum_{s=0}^{n-1} z^{s^2} \sum_{t=0}^{n-1} z^{-t^2} = \sum_{s,t} z^{s^2-t^2} = \sum_{s,t} z^{(s+t)^2-t^2}$$
$$= \sum_{s,t} z^{s^2+2st} = \sum_s \left(z^{s^2} \sum_t z^{2st} \right).$$

As $z = e^{2i\pi/n}$, we have $\sum_t z^{2st} = \sum_t e^{4i\pi st/n} = n$ or 0 depending on whether $n \mid s$ or not. Therefore, $|S|^2 = n$ and it remains to establish which square root gives the correct value of S.

Since S is determined in terms of the eigenvalues λ_rs which, in turn, depend on the m_is, we try to obtain linear equations satisfied by the m_is as a consequence of the equality $|S|^2 = n$. Continuing with the proof, since

$$S = \sum_r \lambda_r = \sum_r i^{a_r}\sqrt{n} = \sqrt{n}(m_0 + im_1 - m_2 - im_3)$$

and $|S|^2 = n$, we have $(m_0 - m_2)^2 + (m_1 - m_3)^2 = 1$. In other words, either $m_0 - m_2 = \pm 1$ and $m_1 = m_3$ or $m_0 = m_2$ and $m_1 - m_3 = \pm 1$. Hence $S = v\eta\sqrt{n}$ where $v = \pm 1$ and $\eta = 1$ or i. Thus, we have in terms of the m_is, the equation

$$m_0 + im_1 - m_2 - im_3 = v\eta$$

and its conjugate

$$m_0 - im_1 - m_2 + im_3 = v\eta^{-1}.$$

Also, the equality $\operatorname{tr} A^2 = \sum_r \lambda_r^2 = n$ observed earlier gives the equation

$$m_0 - m_1 + m_2 - m_3 = 1.$$

Thus, the system of four linear equations can be written as a matrix equation $Bx = y$, where

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}, \quad x = \begin{pmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{pmatrix}, \quad \text{and} \quad y = \begin{pmatrix} n \\ v\eta \\ 1 \\ v\eta^{-1} \end{pmatrix}.$$

Inverting this matrix, we get $x = B^{-1}y$ with

$$B^{-1} = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}.$$

In particular, $m_2 = \frac{n+1-v(\eta+\eta^{-1})}{4}$ being an integer implies that $\eta = 1$ or i, depending on whether $n \equiv 1 \bmod 4$ or $n \equiv 3 \bmod 4$. Further, $\det A = \prod_r \lambda_r = n^{n/2} i^{m_1+2m_2-m_3}$ $= n^{n/2} i^{3(n-1)/2} v = n^{n/2} i^{n(n-1)/2} v$; to obtain this, we have used the fact, obtained from $x = B^{-1}y$, that $m_1 + 2m_2 - m_3$ is $\frac{n+1}{2} - v$ or $\frac{n+1}{2} + v$ depending on whether $n \equiv 1$ or $3 \bmod 4$. We have also used $i^v = iv$ to simplify.

Finally, we show that $v = 1$: This will be a consequence of evaluating—in two different ways—the determinant of the matrix A:

$$\det A = \prod_{0 \le l < k < n} \left(e^{2i\pi k/n} - e^{2i\pi l/n}\right) = \prod_{l<k} e^{i\pi(k+l)/n}\left(e^{i\pi(k-l)/n} - e^{i\pi(l-k)/n}\right).$$

From $\sum_{0 \le l < k < n}(k+l) = n(n-1)^2/2$, we have

$$\prod_{l<k} e^{i\pi(k+l)/n} = e^{i\pi(n-1)^2/2} = i^{(n-1)^2} = 1.$$

Hence

$$\det A = \prod_{l<k}(e^{i\pi(k-l)/n} - e^{i\pi(l-k)/n}) = \prod_{l<k}\left(2i \sin\frac{\pi(k-l)}{n}\right)$$
$$= i^{n(n-1)/2}\prod_{l<k}\left(2 \sin\frac{\pi(k-l)}{n}\right).$$

As the last mentioned product is positive, the two expressions $\det A = n^{n/2} i^{n(n-1)/2} v =$ $i^{n(n-1)/2} \prod_{l<k}\left(2 \sin\frac{\pi(k-l)}{n}\right)$ imply that $v > 0$ and is, therefore, equal to 1. ■

Acknowledgment Initially, I was tempted to discuss a bit of the fascinating history surrounding Gauss sums but avoided it as this note was a short one. However, one of the referees raised this matter and mentioned that this is a "missed opportunity." This encouraged and emboldened me to add historical as well as mathematical material, which perhaps gives a better understanding of the background. It is a great pleasure to acknowledge the constructive comments from both the referees.

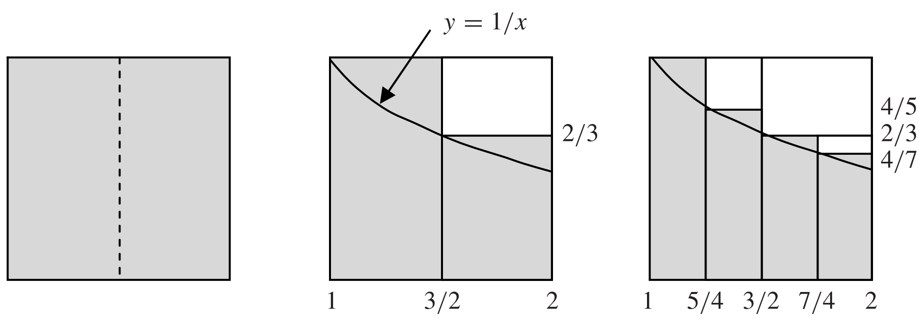
REFERENCES

1. K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed., Springer-Verlag, New York, 1990.
2. E. Landau, *Vorlesungen Uber Zahlentheorie aus der elementaren zahlentheorie*, Chelsea Publishing, New York, 1950.

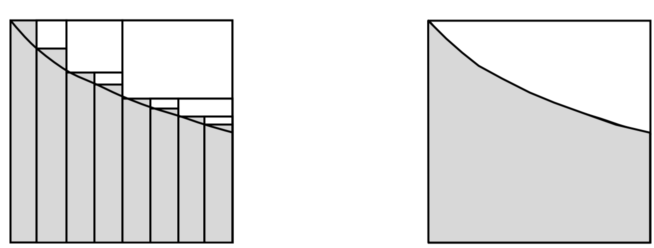
Summary Gauss sums were introduced by Gauss in 1801, when he stated some of their properties and used them to prove the quadratic reciprocity law in different ways. The determination of the sign of the Gauss sum was a notoriously difficult question; Gauss recorded the correct assertion in his mathematical diary in May 1801, but could find a proof only in 1805. This note uses the Gauss sums to evaluate certain sums of trigonometric functions.

Proof Without Words: The Alternating Harmonic Series Sums to ln 2

CLAIM. $\displaystyle\sum_{n=0}^{\infty}(-1)^n\frac{1}{n+1} = \ln 2.$



$1\cdots \qquad -\frac{1}{2}+\frac{1}{3}-\cdots \qquad -\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\frac{1}{7}-\cdots$



$-\frac{1}{8}+\frac{1}{9}-\frac{1}{10}+\frac{1}{11}-\frac{1}{12}+\frac{1}{13}-\frac{1}{14}+\frac{1}{15}-\cdots \qquad = \int_1^2 \frac{1}{x}\,dx = \ln 2$

—Matt Hudelson
Washington State University
Pullman WA 99164

Summary We demonstrate graphically the result that the alternating harmonic series sums to the natural logarithm of two. This is accomplished through a sequence of strategic replacements of rectangles with others of lesser area. In the limit, we obtain the region beneath the curve $y = 1/x$ and above the x-axis between the values of one and two.

Math. Mag. **83** (2010) 294. doi:10.4169/002557010X521831.

Period Three Begins

CHENG ZHANG
Department of Bioengineering
Rice University
Houston, Texas 77005
cz1@rice.edu

The logistic map

$$x_{n+1} = f(x_n) \equiv r x_n (1 - x_n), \tag{1}$$

where $x_n \in [0, 1]$ and $r \in [0, 4]$, manifests many trademark features in nonlinear dynamics, such as period-doubling and chaos [**1, 2**]. Surprisingly, all the complex behaviors can be explored by setting the parameter r to different values. We focus on the particular value $r^* = 1 + \sqrt{8} \approx 3.8284$, which gives birth to the period-three cycle, where the system repeats itself every three iterations, meaning that $x_{n+3} = x_n$. One way to visualize the period-three cycle is to draw the iterating process on the cobweb plot, see FIGURE 1.

There are several different ways to derive the value r^* [**3, 4, 5, 6**]. The first demonstration, given by Saha and Strogatz [**3**], unfortunately involves heavy algebraic manipulation. Later on, Bechhoefer [**4**] gives a simpler proof, where $f(f(f(x)))$ is expanded directly as a polynomial of x, and the coefficients are compared with their expected values. Gordon [**5**] approaches the problem by writing down the Fourier transformed version of (1) and then comparing the coefficients of different components on both sides of the equation. A more recent derivation is provided by Burm and Fishback [**6**] using Sylvester's theorem.

Here we present a new elementary derivation based on the geometry of the cobweb plot FIGURE 1.

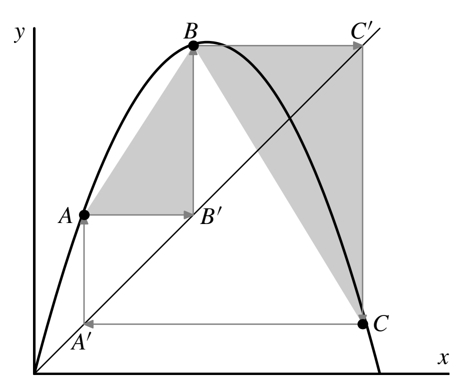


Figure 1 The cobweb plot of the logistic map for $r = 3.84$

The proof Let us first denote the x-coordinates of A, B, and C in FIGURE 1 as a, b, and c, respectively. The map defines a cyclic relation of a, b, and c: $b = f(a)$, $c = f(b)$, and $a = f(c)$. Our derivation is based on

$$\overline{A'A} + \overline{B'B} + \overline{C'C} = 0, \tag{2}$$

Math. Mag. **83** (2010) 295–297. doi:10.4169/002557010X521859.

where we have used overbars to denote signed distance. Since $\overline{AB'} = \overline{A'A}$, we can express $\overline{B'B}$ as $(\overline{B'B}/\overline{AB'})\,\overline{A'A}$. The ratio $\overline{B'B}/\overline{AB'}$ can be directly calculated as $[f(b) - f(a)]/(b-a) = r(1-a-b)$. Similarly, we calculate the ratio $\overline{C'C}/\overline{BC'} = r(1-b-c)$ and replace $\overline{C'C}$ by $(\overline{C'C}/\overline{BC'})\,(\overline{B'B}/\overline{AB'})\,\overline{A'A}$. With the common factor $\overline{A'A}$ eliminated, (2) becomes

$$1 + r(1-a-b) + r^2(1-a-b)(1-b-c) = 0. \tag{3}$$

The other two symmetrical versions of (3) can be reached by cycling the symbols $a \to b$, $b \to c$, $c \to a$ twice. We now sum over the three resulting equations

$$\begin{aligned} &3 + r[3 - 2(a+b+c)] \\ &\quad + r^2[3 - 4(a+b+c) + (a+b+c)^2 + ab + bc + ca] = 0. \end{aligned} \tag{4}$$

We can further reduce (4) to an equation of a single variable $X = a + b + c$ by using the following two identities $2(ab + bc + ca) = (a+b+c)^2 - (a^2 + b^2 + c^2)$, and $r(a^2 + b^2 + c^2) = (r-1)(a+b+c)$; the latter follows from the fact $b + c + a = f(a) + f(b) + f(c)$. Thus, we have

$$r^2X^2 - (3r+1)rX + 2(1 + r + r^2) = 0. \tag{5}$$

The discriminant of the quadratic is $\Delta = r^2(r^2 - 2r - 7)$. If it is positive, the equation has two roots; correspondingly, the system has two period-three cycles: a stable one and an unstable one. If it is negative, there is no root, which means that there is no period-three cycle. At the onset of period-three, the discriminant is zero, and the solution of $\Delta = 0$ gives the desired result $r^* = 1 + \sqrt{8}$. Note, the other root $r^* = 1 - \sqrt{8} \approx -1.8284$ also gives the onset of period three for the negative r case, see FIGURE 2.

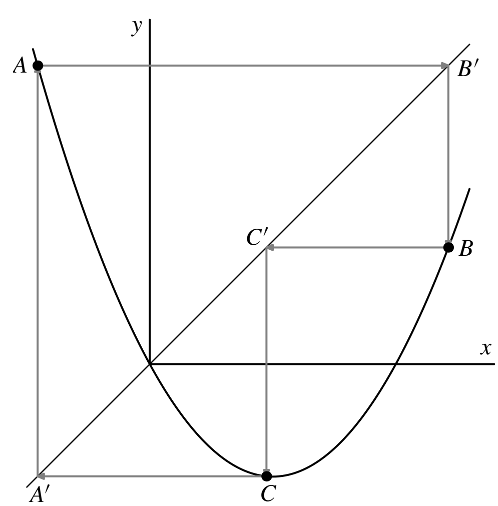


Figure 2 The cobweb plot for $r = -1.8285$

The same derivation applies to any quadratic function $f(x)$. Actually, the algebra is much simpler if we first transform the original logistic map to $y_{n+1} = R - y_n^2$ through a change of variables, $y_n = r(x_n - 1/2)$, $R = (r^2 - 2r)/4$ [**2**]. In this case, the counterpart of (5) is $X^2 - X + 2 - R = 0$, with the discriminant being $4R - 7$. The zero-discriminant condition is readily translated to $R^* = 7/4$, or equivalently $r^* = 1 \pm \sqrt{8}$, which agrees with the previous result.

Acknowledgment I thank the referees for reading the manuscript carefully and for pointing out the meaning of the negative r^*.

REFERENCES

1. R. M. May, Simple mathematical models with very complicated dynamics, *Nature* **261** (1976) 459–467. doi: 10.1038/261459a0
2. S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA, 1994
3. P. Saha and S. H. Strogatz, The birth of period three, *Mathematics Magazine* **68** (1995) 42–47.
4. J. Bechhoefer, The birth of period 3, revisited, *Mathematics Magazine* **69** (1996) 115–118.
5. W. B. Gordon, Period three trajectories of the logistic map, *Mathematics Magazine* **69** (1996) 118–120.
6. J. Burm and P. Fishback, Period-3 orbits via Sylvester's theorem and resultants, *Mathematics Magazine* **74** (2001) 47–51.

Summary By exploiting the geometry of the cobweb plot, we provide a simple and elementary derivation of the parameter for the period-three cycle of the logistic map.

Stacking Blocks and Counting Permutations

LARA K. PUDWELL
Valparaiso University
Valparaiso, Indiana 46383
Lara.Pudwell@valpo.edu

In this paper we will explore two seemingly unrelated counting questions, both of which are answered by the same formula. In the first section, we find the surface areas of certain solids formed from unit cubes. In second section, we enumerate permutations with a specified set of restrictions. Next, we give a bijection between the faces of the solids and the set of permutations. We conclude with suggestions for further reading. First, however, it is worth explaining how this paper came about.

The author received an email from David Harris while he was helping his 12-year-old daughter Julia complete a project for her math class. Together the Harrises constructed triangular piles of cubes. After creating an increasing sequence of these piles, they computed the surface area of each pile, and hoped to find a formula for the surface area of their nth pile. This project and its solution are described in the next section. At the time of their correspondence, David and Julia had deduced several facts about the construction but had not yet found a formula for the surface area in general. When they searched for the first few terms in their sequence, Google returned only one hit: a Maple data file on the author's website.

The sequence that the Harrises discovered online was originally generated in the context of pattern-avoiding words and permutations. Their web search produced a conjecture that gives a nice geometric interpretation of a question about permutation patterns. This serendipitous discovery of the surprising and beautiful connection between a geometry problem and an enumeration problem illustrates how attractive new results may sometimes appear in such a surprising place as a middle-school homework exercise.

Math. Mag. **83** (2010) 297–302. doi:10.4169/002557010X521868.

The surface area of cubes We begin with the Harrises' original geometry question. We first describe a recursive construction involving unit cubes, and then compute the surface area of the nth solid in this construction.

The first solid is a unit cube, which has surface area 6. To construct the nth solid, first form a row of $2n-1$ cubes. Then, center the $(n-1)$st construction on top of this row. For example, the second solid is shown in FIGURE 1. It has surface area 18. The third solid is also shown. It has surface area 34.

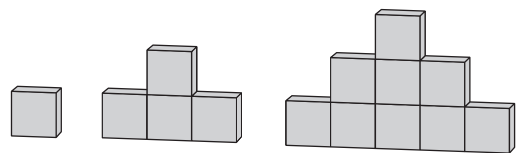


Figure 1 The first, second, and third solids

Now, we wish to compute the surface area $SA(n)$ of the nth solid. We have already computed $SA(1)$, $SA(2)$, and $SA(3)$ above.

To construct $SA(n)$, we glue together a solid of surface area $SA(n-1)$ together with a rectangular prism of surface area $4 \cdot (2n-1) + 2 = 8n-2$. However, there are $2n-3$ pairs of squares that overlap, and that become part of the interior of the shape. Thus, the surface area only increases by $(8n-2) - 2(2n-3) = 4n+4$ units; that is, $SA(n) - SA(n-1) = 4n+4$. Since we know that $SA(1) = 6$, it is easy to prove by induction that $SA(n) = 2n^2 + 6n - 2$ for all $n \geq 1$.

Permutation patterns We have proved that the surface area of the Harrises' nth solid is $2n^2 + 6n - 2$. We now give the necessary definitions to produce a set of permutations with $2n^2 + 6n - 2$ elements.

For this paper, a *permutation* is just a string of digits, such as 112 or 2671165. (Really, a permutation is a string of integers, but fortunately, we will never need integers with more than one digit in our examples.) In particular, we are interested in strings in which each digit appears exactly twice. We write $S_n^{(2)}$ to denote the set of permutations of two 1's, two 2's, and so on up to two n's. For example $S_1^{(2)} = \{11\}$ and $S_2^{(2)} = \{1122, 1212, 1221, 2112, 2121, 2211\}$. Typically, a permutation refers to an ordering of n *distinct* letters. Since we are considering permutations with more than one copy of each letter we may refer to our permutations as *multiset permutations*.

Now, we will say what it means for a permutation to *contain* a certain pattern, or to *avoid* a pattern. Given a string of numbers s, the *reduction* of s is the string obtained in the following way: find the smallest number in the string and replace all occurrences of that number with 1, then find the second smallest number in the string and replace all occurrences of that number with 2, and so forth, replacing the occurrences of the ith smallest number with the number i. For example, the reduction of 2671165 is 2451143. Now, given strings of numbers $p = p_1 \cdots p_n$ and $q = q_1 \cdots q_m$, we say that p *contains* q as a pattern if there exist indices $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ such that $p_{i_1} \cdots p_{i_m}$ reduces to q. Otherwise, we say that p *avoids* q. For example, 2671165 contains the pattern 2321 because it contains the subsequence 6765, which reduces to 2321. However, 2671165 avoids the pattern 1234 because it has no strictly increasing subsequence of length 4.

Finally, we expand on our permutation set notation. Let $S_n^{(2)}(Q)$ denote the set of permutations of $S_n^{(2)}$ avoiding all patterns in the list Q. For example

$$S_2^{(2)}(112) = \{1221, 2121, 2211\}$$

and

$$S_2^{(2)}(132, 231, 2134) = \{1122, 1212, 1221, 2112, 2121, 2211\}.$$

The first of these equalities can be checked by looking at all 6 members of $S_2^{(2)}$ and noting which do not contain a 112 pattern. The second is equal to the set $S_2^{(2)}$ because permutations with only the digits 1 and 2 cannot contain a pattern with 3 or 4 distinct digits. A more interesting example is that

$$\begin{aligned} S_3^{(2)}(132, 231, 2134) = \{&112233, 121233, 122133, 211233, 212133, 221133, \\ &311223, 312123, 312213, 321123, 321213, 322113, \\ &331122, 331212, 331221, 332112, 332121, 332211\}. \end{aligned}$$

From our examples, $\left|S_2^{(2)}(132, 231, 2134)\right| = 6$ and $\left|S_3^{(2)}(132, 231, 2134)\right| = 18$. These are precisely the surface areas of the Harrises' first and second solids. This is the coincidence that we explain in the remainder of this note.

The author's website contains data about $S_n^{(2)}(Q)$ for many different lists of permutations Q. It turns out the Harrises' sequence corresponds to $\left|S_n^{(2)}(132, 231, 2134)\right|$ for $n = 2, 3, 4, \ldots$. Notice that this sequence begins with $n = 2$ instead of $n = 1$; that is, the Harrises' nth solid is related to the set $S_{n+1}^{(2)}(132, 231, 2134)$. In the following sections we will provide a bijection between the faces of the Harrises' nth solid and the members of this set of permutations to show that $\left|S_{n+1}^{(2)}(132, 231, 2134)\right| = 2n^2 + 6n - 2$, but first we need a lemma.

A permutation lemma The lemma presented here is a special case of a result of Burstein [**3**].

LEMMA 1. $\left|S_n^{(2)}(132, 231, 213)\right| = 2n + 2$ *for* $n \geq 2$.

Proof. Since we will only consider permutations that avoid the set of patterns $\{132, 231, 213\}$ in this proof, we will write A_n instead of $S_n^{(2)}(132, 231, 213)$.

First consider the case of $n = 2$. Because no string of 1's and 2's can contain a pattern in $\{123, 231, 213\}$, we have that $A_2 = \{1122, 1212, 1221, 2112, 2121, 2211\}$, and $|A_2| = 6$, which is equal to $2n + 2$ as desired.

We proceed by induction. Consider $p \in A_n$. Let p' be the multiset permutation formed by deleting the two copies of n in p. For example if $p = 312123$, then $p' = 1212$. Notice that since $p \in A_n$, we have that $p' \in A_{n-1}$.

Now, given $p' \in A_{n-1}$, we consider all the ways to insert two copies of n into p' to obtain a multiset permutation in A_n. Notice that if n is inserted *between* two letters of p', we have necessarily created either a 132 pattern or a 231 pattern. Thus, the n's can be inserted in one of only 3 ways: (i) both n's are prepended to the beginning of p', (ii) both n's are appended to the end of p', or (iii) one n is prepended to the beginning of p' and the other n is appended to the end of p'. Clearly, (i) always produces a member of $S_n^{(2)}$, however, (ii) and (iii) must be considered more carefully. In particular, appending an n to the end of p' only produces a 213-avoiding multiset permutation if p' avoids the pattern 21, i.e. if p' is weakly increasing. Thus, $|A_n| = |A_{n-1}| + 2$, since we may prepend two n's to the beginning of any member of A_{n-1}, but we may also append two n's to the end of the unique increasing permutation of A_{n-1}, or we may prepend an n to the beginning of it and append an n to the end of it.

Finally, since $|A_n| - |A_{n-1}| = 2$, we know that $|A_n|$ grows linearly, and use the fact that $A_2 = 6$ to compute the formula $|A_n| = 2n + 2$. ■

This lemma is key to our main theorem, which is given at the end of the next section.

The bijection We now give a bijection between the faces of the nth solid of the Harrises' construction and the multiset permutations of $S_{n+1}^{(2)}(132, 231, 2134)$. While we could count the permutations of $S_{n+1}^{(2)}(132, 231, 2134)$ directly, a bijection not only will show that the two quantities in question are equal, but a bijection will also illuminate some parallels between the cube construction and the structure of the members of $S_{n+1}^{(2)}(132, 231, 2134)$. To find such a bijection, it suffices to associate each permutation in $S_{n+1}^{(2)}(132, 231, 2134)$ with a unique unit square on the surface of the Harrises' nth solid.

To this end, we consider another description of the Harrises' construction. To construct the nth solid from the $(n-1)$st solid, we first remove the bottom face of the solid and move it one unit lower as in FIGURE 2 (i). Next, we form a rectangular ring of $4n$ squares. This ring should be constructed so that it has two opposing sides of length 1 and two opposing sides of length $2n-1$, as shown in FIGURE 2 (ii). Now, attach a new square to the top and bottom of each end of the ring, as shown in FIGURE 2 (iii), for a total of $4n+4$ new squares. We may glue the modified version of the $(n-1)$st solid together with this new modified ring of $4n+4$ squares to form the nth solid. Two views of this gluing are shown in FIGURE 2 (iv).

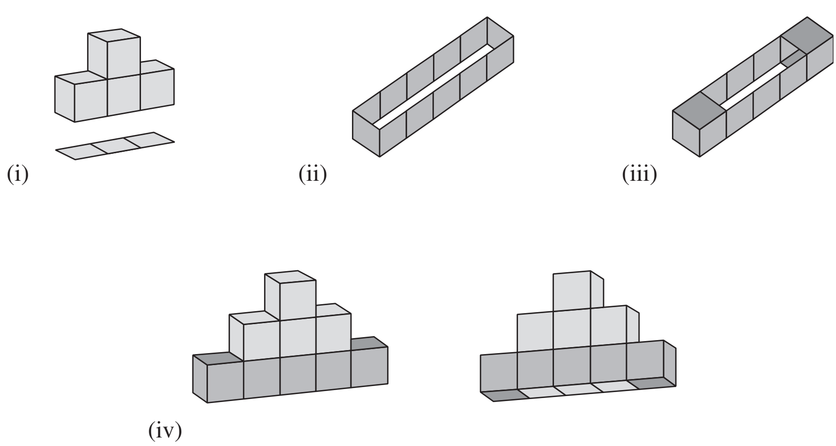


Figure 2 Constructing the $n=3$ solid from the $n=2$ solid

This alternate construction has a clear advantage. Although it is more complicated to explain, this revised description allows us to associate each square on the surface of the $(n-1)$st solid with squares on the nth solid, rather than "gluing" some squares into the interior.

The permutations of $S_{n+1}^{(2)}(132, 231, 2134)$ also have a nice recursive structure. Given $p' \in S_{n}^{(2)}(132, 231, 2134)$, there are three ways to insert two copies of $(n+1)$ into p' to obtain a multiset permutation in $S_{n+1}^{(2)}(132, 231, 2134)$: (i) both $(n+1)$'s are prepended to the beginning of p', (ii) both $(n+1)$'s are appended to the end of p', or (iii) one $(n+1)$ is prepended to the beginning of p' and the other $(n+1)$ is appended to the end of p'. As with the permutations of Lemma 1, (i) always produces a member of $S_{n+1}^{(2)}(132, 231, 2134)$, but (ii) and (iii) must be considered in more detail. In particular, appending $(n+1)$ to the end of p' may induce a copy of a forbidden 2134 pattern if p' contains a 213 pattern.

Now, we may recursively define a bijection between the squares of the nth solid and the permutations of $S_{n+1}^{(2)}(132, 231, 2134)$.

To begin, since there are 6 elements of $S_2^{(2)}(132, 231, 2134)$, and 6 faces in a unit cube, we may assign each one of these permutations to a unique face of the cube.

Now, consider the nth solid, constructed as described in this section. In the $(n-1)$st solid, each of the light gray squares was associated with some permutation $p \in S_n^{(2)}(132, 231, 2134)$. Let each such square now be associated with the permutation $(n+1)(n+1)p \in S_{n+1}^{(2)}(132, 231, 2134)$.

We must now account for the four dark gray squares (the tops and bottoms of the left and right cubes in the bottom row of the solid) and the $4n$ medium gray squares (the side faces of all cubes in the bottom row of the solid). Clearly, these must correspond to the permutations of $S_{n+1}^{(2)}(132, 231, 2134)$ that either begin and end with $(n+1)$ or that end with two copies of $(n+1)$. Notice that each of *these* permutations was formed by taking one of the $2n+2$ permutations in $S_n^{(2)}(132, 231, 213)$ and inserting two $(n+1)$'s in one of the two ways just described. Thus the $4n+4$ permutations of the form $p(n+1)(n+1)$ or $(n+1)p(n+1)$ where $p \in S_n^{(2)}(132, 231, 213)$ are precisely the members of $S_{n+1}^{(2)}(132, 231, 213)$ that correspond to the $4n+4$ dark gray and medium gray squares. We now have established a recursive bijection between the exterior faces of the Harrises' piles of cubes and the members of $S_{n+1}^{(2)}(132, 231, 2134)$. This correspondence gives a combinatorial proof of the following theorem, which was first observed using the method of enumeration schemes found in [**6**].

THEOREM 1. $\left|S_{n+1}^{(2)}(132, 231, 2134)\right| = 2n^2 + 6n - 2$ *for* $n \geq 1$.

For further reading In this paper we found a bijection between the squares on the faces of the Harrises' nth construction, and certain pattern-avoiding permutations. This bijection illustrates the nice and unexpected connection between a question of middle-school geometry and enumerative combinatorics.

Interested readers may wish to learn more about other enumeration problems related to this paper. Permutations which avoid other permutations have been actively studied since the seminal paper of Simion and Schmidt [**7**]. Applications have been made in other areas of combinatorics. A friendly introduction to permutation patterns can be found in [**2**]. The permutations in this paper, with precisely two copies of each letter, are a special case of multiset permutations in which there may be an arbitrary numbers of copies of each letter. More detailed work with pattern avoidance involving multiset permutations can be found in [**1**], [**3**], [**5**], and [**6**].

The bijection demonstrated in this paper illustrates one of several connections between the Harrises' cube constructions and pattern-avoiding permutations. To see another bijection that relies on different geometric and combinatorial properties, visit the author's website at `http://faculty.valpo.edu/lpudwell/papers.html`.

Acknowledgment Thank you to David and Julia Harris for inspiring this paper, and to Andrew Baxter for many valuable presentation suggestions. Thanks, also, to Julia's teacher: Julie McDaniel at Ottoson Middle School in Arlington, MA.

REFERENCES

1. M. H. Albert, R. E. L. Aldred, M. D. Atkinson, C. Handley, and D. Holton, Permutations of a multiset avoiding permutations of length 3, *European. J. Combin.* **22** (2001) 1021–1031. doi:10.1006/eujc.2001.0538
2. M. Bóna, *Combinatorics of Permutations*, Chapman & Hall, New York, 2004, Chapter 4, 129–173.
3. A. Burstein, *Enumeration of Words with Forbidden Patterns*, Ph.D. thesis, University of Pennsylvania, 1998.
4. D. Harris, Unpublished correspondence, June 4, 2009.
5. S. Heubach and T. Mansour, Avoiding patterns of length three in compositions and multiset permutations, *Advances in Applied Mathematics* **36** (2006) 156–174. doi:10.1016/j.aam.2005.06.005

6. L. Pudwell, *Enumeration Schemes for Pattern-Avoiding Words and Permutations*, Ph.D. thesis, Rutgers University, 2008.
7. R. Simion and F. W. Schmidt, Restricted permutations, *European J. Combin.* **6** (1985) 383–406.

Summary This paper explores a surprising connection between a geometry problem and a result in enumerative combinatorics. First, we find the surface areas of certain solids formed from unit cubes. Next, we enumerate multiset permutations which avoid the patterns {132, 231, 2134}. Finally, we give a bijection between the faces of the solids and the set of permutations.

Counting Ordered Pairs

DAVID M. BRADLEY
University of Maine
Orono, ME 04469-5752
dbradley@member.ams.org

After Cantor [**3**, p. 107] (cf. also [**2**]), the standard method of enumerating the set $\mathbf{Z}^+ \times \mathbf{Z}^+$ of ordered pairs of positive integers is to list the entries by traversing successive diagonals, beginning with (1, 1), (1, 2), (2, 1), (1, 3), (2, 2), (3, 1), and so on. An explicit bijection that accomplishes this is $\varphi : \mathbf{Z}^+ \times \mathbf{Z}^+ \to \mathbf{Z}^+$ defined by

$$\varphi(m, n) = m + \frac{(m+n-1)(m+n-2)}{2}.$$

Providing an algebraic proof that φ is indeed a bijection is an instructive exercise.

By exploiting the multiplicative structure of the codomain, we can construct a map $\psi : \mathbf{Z}^+ \times \mathbf{Z}^+ \to \mathbf{Z}^+$ which is immediately recognized as a bijection. (No need to resort to algebraic calculation or a pictorial argument with diagonals.) For each pair of positive integers m and n, let $\psi(m, n) = 2^{m-1}(2n-1)$. Bijectivity of ψ is equivalent to the fact that every positive integer has a unique representation as the product of an odd positive integer and a non-negative integer power of 2. As one referee noted, this fact is also key to Glaisher's bijection between partitions of a positive integer into odd parts and partitions with distinct parts [**1**, Ex. 2.2.6; **4**, p. 12].

REFERENCES

1. David Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, MAA and Cambridge University Press, 1999.
2. Georg Cantor, Ueber eine Eigenschaft des Inbegriffes aller reelen algebraischen Zahlen, *J. Reine Angew. Math.* **77** (1874) 258–261. doi:10.1515/crll.1874.77.258
3. Georg Cantor, *Contributions to the Founding of the Theory of Transfinite Numbers*, trans. by Philip E. B. Jourdain, Open Court Publishing, La Salle, Illinois, 1941.
4. Percy A. MacMahon, *Combinatory Analysis*, Vol. II, Chelsea Publishing, New York, 1984.

Math. Mag. **83** (2010) 302. doi:10.4169/002557010X528032.

PROBLEMS

BERNARDO M. ÁBREGO, *Editor*
California State University, Northridge

Assistant Editors: SILVIA FERNÁNDEZ-MERCHANT, California State University, Northridge; JOSÉ A. GÓMEZ, Facultad de Ciencias, UNAM, México; ROGELIO VALDEZ, Facultad de Ciencias, UAEM, México; WILLIAM WATKINS, California State University, Northridge

PROPOSALS

To be considered for publication, solutions should be received by March 1, 2011.

1851. *Proposed by Éric Pité, Paris, France.*

Let a be an arbitrary integer. Consider the recursive sequence of integers defined by $u_0 = 4$, $u_1 = 0$, $u_2 = 2$, $u_3 = 3$, and $u_{n+4} = u_{n+2} + u_{n+1} + a \cdot u_n$ for every integer $n \geq 0$. Prove that p divides u_p for every prime p.

1852. *Proposed by Radu Gologan, Institute of Mathematics "Simion Stoilow" of the Romanian Academy, Bucharest, Romania; and Cezar Lupu (student), University of Bucharest, Bucharest, Romania.*

Let $f : [0, 1] \to \mathbb{R}$ be a differentiable function with a continuous derivative such that $f(0) = f(1) = -\frac{1}{6}$. Prove that

$$\int_0^1 (f'(x))^2\,dx \geq 2\int_0^1 f(x)\,dx + \frac{1}{4}.$$

1853. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

A function f is continuous nearly everywhere if it is continuous on its domain except for a countable set. Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function.

(a) Prove that if f is continuous nearly everywhere, then for every open set $G \subseteq \mathbb{R}$ there are an open set O and a countable set C such that $f^{-1}(G) = O \cup C$.

(b) Is the converse of part (a) true? Prove or disprove.

Math. Mag. **83** (2010) 303–310. doi:10.4169/002557010X521877.

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. Submitted problems should not be under consideration for publication elsewhere.

Solutions should be written in a style appropriate for this MAGAZINE.

Solutions and new proposals should be mailed to Bernardo M. Ábrego, Problems Editor, Department of Mathematics, California State University, Northridge, 18111 Nordhoff St, Northridge, CA 91330-8313, or mailed electronically (ideally as a LaTeX or pdf file) to mathmagproblems@csun.edu. All communications, written or electronic, should include **on each page** the reader's name, full address, and an e-mail address and/or FAX number.

1854. *Proposed by Marian Tetiva, National College "Gheorghe Roşca Codreanu", Bârlad, Romania.*

Let n and d be nonnegative integers. Find the number of all subsets of $\{1, 2, \dots, n\}$ which do not contain two numbers whose difference is d. (Subsets with at most one element satisfy the condition by vacuity.)

1855. *Proposed by Michael Goldenberg and Mark Kaplan, The Ingenuity Project, Baltimore Polytechnic Institute, Baltimore, MD.*

Prove that the Euler line of a triangle is perpendicular to one of the medians, if and only if, the Brocard line is perpendicular to the symmedian line from the same vertex as the median.

(The Brocard line passes through the circumcenter and the symmedian point of a triangle. The symmedian point is the point of concurrency of the symmedians and a symmedian through a vertex is the line symmetric to the median with respect to the angle bisector from the same vertex.)

Quickies

Answers to the Quickies are on page 310.

Q1003. *Proposed by Rick Mabry, Louisiana State University in Shreveport, Shreveport, LA.*

What are the zeros of the nth derivative of $f(x) = x^2 e^x$?

Q1004. *Proposed by Daniel Edelman (student), Mason–Rice Elementary School, Newton Centre, MA; and Alan Edelman, MIT, Cambridge, MA.*

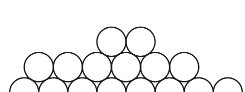
If we concatenate (run together) a finite sequence of nonzero numbers in base 10, can this number equal the product? For example, given 6, 54, and 321, we are comparing 654 321 with $6 \cdot 54 \cdot 321 = 104\,004$ which is less than 654 321.

Solutions

Counting block fountains of coins **October 2009**

1826. *Proposed by Michael Woltermann, Washington & Jefferson College, Washington, PA.*

A block fountain of coins is an arrangement of identical coins in rows such that the coins in the first row form a contiguous block, each row above that forms a contiguous block, and each coin in a higher row is supported by two adjacent coins in the row below. As an example,



If a_n denotes the number of block fountains with exactly n coins in the base, then $a_n = F_{2n-1}$, where F_k denotes the kth Fibonacci number. (Wilf, *generatingfunctionology*, 1994.) How many block fountains are there if two fountains that are mirror images of each other are considered to be the same? That is, if two fountains such as

are the same, while two fountains such as

are different?

Solution by Michael Cap Khoury, University of Michigan, Ann Arbor, MI.

For the purposes of this problem, index the Fibonacci sequence so that $F_1 = F_2 = 1$; this is the convention that makes $a_n = F_{2n-1}$ as claimed in the problem statement.

We begin by counting those block fountains with exactly n coins in the base that have mirror symmetry. Let B_n be the set of such fountains and $b_n = |B_n|$. (Note that by identifying a symmetric fountain with the length of its rows, b_n is also the number of decreasing sequences of positive integers that start with n and alternate in parity.) We claim that $b_n = F_{n+1}$. It is easy to see that $b_1 = 1$ and $b_2 = 2$. To prove the result in general, note that the fountains in B_n are of two types: those that contain a row of $n-1$ coins directly above the base and those that do not. In the former case, remove the base to obtain a fountain in B_{n-1}; in the latter case, remove one coin from each side of the base to obtain a fountain in B_{n-2}. This gives a bijection between B_n and $B_{n-1} \cup B_{n-2}$, so $b_n = b_{n-1} + b_{n-2}$, and our claim follows.

Now, when we identify a fountain with its mirror image, a_n double-counts the asymmetric fountains but not the symmetric ones. So $a_n + b_n$ double-counts all fountains, and the answer to our problem is $(a_n + b_n)/2 = (F_{2n-1} + F_{n+1})/2$. (Note that as a consequence F_{2n-1} and F_{n+1} have the same parity for all n.)

Also solved by Elizabeth Bentley and Todd Lee; Elton Bojaxhiu (Germany) and Enkel Hysnelaj (Australia); Shane M. Bryan; Robert Calcaterra; Chip Curtis; Daniele Degiorgi (Switzerland); Emeric Deutsch; Sergio Falcón (Spain), José M. Pacheco (Spain), and Ángel Plaza (Spain); David Getling (Germany); Ruth A. Koelle; Kathleen E. Lewis; Jeff Lutgen; Joseph McKenna; Ray Rosentrater; Ossama A. Saleh and Terry J. Walters; Nicholas C. Singer; Skidmore College Problem Group; John Sumner and Aida Kadic-Galeb; James Swenson; Taylor Problem Solving Group; Yen Yanosko; and the proposer. There was one incorrect submission.

The row sum divides the determinant **October 2009**

1827. *Proposed by Christopher Hilliar, Texas A & M University, College Station, TX.*

Let A be an $n \times n$ matrix with integer entries and such that each column of A is a permutation of the first column. Prove that if the entries in the first column do not sum to 0, then this sum divides $\det(A)$.

Solution by Reiner Martin, Bad Soden-Neuenhain, Germany.

More generally, only assume that the integer entries of each row of A sum to the same number s. Clearly, the n-vector of 1s is an eigenvector with eigenvalue s. Thus, s is a zero of the integer characteristic polynomial $p(x) = \sum_{i=0}^{n} a_i x^i$ of A. Consequently,

$$a_0 = s \cdot \left(-\sum_{i=1}^{n} a_i s^{i-1} \right),$$

so s divides $a_0 = \det(A)$. Note that the condition $s \neq 0$ is not needed.

Editor's Note. As was noted by Francisco Vial and Mark Ashbaugh this problem appears as Problem 7.2.9 in P. N. de Souza and J. Silva, *Berkeley Problems in Mathematics*, Springer, New York, 2004.

Also solved by Michael Andreoli; George Apostolopoulos (Greece); Armstrong Problem Solvers; Michel Bataille (France); Jany C. Binz (Switzerland); Elton Bojaxhiu (Germany) and Enkel Hysnelaj (Australia); Shane M. Bryan; Nicholas Buck (Canada); Robert Calcaterra; Minh Can; Hongwei Chen; John Christopher; Daniele Degiorgi (Switzerland); "Fejéntaláltuka Szeged" problem solving group (Hungary); Rod Hardy, David McFarland, and Alin A. Stancu; Eugene A. Herman; Tom Hoffman; Bianca-Teodora Iordache (Romania); Michael Cap Khoury; Ruth A. Koelle; Omran Kouba (Syria); Victor Y. Kutsenok; Elias Lampakis (Greece); Kathleen E. Lewis; Kim McInturff; Missouri State University Problem Solving Group; Northwestern University Math Problem Solving Group; Occidental College Problem Solving Group; Éric Pité (France); Ángel Plaza (Spain) and José M. Pacheco (Spain); Gabriel Prăjitură; Henry Ricardo; Raúl A. Simón (Chile); Nicholas C. Singer; John H. Smith; John Sumner and Aida Kadic-Galeb; Taylor Problem Solving Group; Marian Tetiva (Romania); Francisco Vial (Chile); Michael Vowe (Switzerland); Stanley Y. Xiao (Canada); Ken Yanosko; John T. Zerger; and the proposer. There were three incorrect submissions.

A Stirling product **October 2009**

1828. *Proposed by Stephen J. Herschkorn, Department of Statistics, Rutgers University, New Brunswick, NJ.*

Let α_0 be the smallest value of α for which there exists a positive constant C such that

$$\prod_{k=1}^{n} \frac{2k}{2k-1} \le Cn^{\alpha}$$

for all positive integers n.

a. Find the value of α_0.
b. Prove that the sequence

$$\left\{ \frac{1}{n^{\alpha_0}} \prod_{k=1}^{n} \frac{2k}{2k-1} \right\}_{n=1}^{\infty}$$

is decreasing and find its limit.

Solution by Ossama A. Saleh and Stan Byrd, Department of Mathematics, University of Tennessee at Chattanooga, Chattanooga, Tennessee.

Let $p_n = \prod_{k=1}^{n}(2k)/(2k-1)$, then after a straight forward calculation, we see that $p_n = 2^{2n}(n!)^2/(2n)!$. We prove by induction that $p_n \le 2\sqrt{n}$ for $n \ge 1$. For $n = 1$, $p_1 = 2$. Assume that $p_n \le 2\sqrt{n}$. Then

$$p_{n+1} = \left(\frac{2(n+1)}{2n+1}\right) p_n \le \frac{4(n+1)\sqrt{n}}{2n+1} = \frac{4\sqrt{n+1}\sqrt{n(n+1)}}{2n+1}.$$

By the Arithmetic Mean–Geometric Mean Inequality,

$$\sqrt{n(n+1)} < \frac{n+n+1}{2} = \frac{2n+1}{2},$$

so $p_{n+1} \le 2\sqrt{n+1}$. Therefore $\alpha_0 \le \frac{1}{2}$ with $C = 2$.

Next, we prove that if $\alpha < \frac{1}{2}$, then there is no positive constant C such that

$$\prod_{k=1}^{n} \frac{2k}{2k-1} \le Cn^{\alpha}$$

for all positive integers n. Assume the contrary, that is, for some $\alpha < \frac{1}{2}$ and some positive constant C, $p_n \leq Cn^{\alpha}$ for all positive integers n. We employ the well-known Stirling's Formula:

$$\lim_{n\to\infty} \frac{n!e^n}{n^n\sqrt{2\pi n}} = 1.$$

For $n \geq 1$ let $g(n) = n!e^n/n^n\sqrt{2\pi n}$. Therefore,

$$1 = \lim_{n\to\infty} \frac{(g(n))^2}{g(2n)} = \lim_{n\to\infty} \frac{2^{2n}(n!)^2}{(2n!)\sqrt{\pi n}} = \lim_{n\to\infty} \frac{p_n}{\sqrt{\pi n}} \leq \lim_{n\to\infty} \frac{C}{\sqrt{\pi} n^{\frac{1}{2}-\alpha}} = 0,$$

which is a contradiction. Therefore $\alpha_0 = 1/2$.

For part b, we define $h_n = n^{-\alpha_0} p_n$. Then, $h_n = 2^{2n}(n!)^2/(2n)!\sqrt{n}$ and hence, for $n \geq 1$,

$$\frac{h_{n+1}}{h_n} = \frac{2\sqrt{n(n+1)}}{2n+1} < 1,$$

once again by the Arithmetic Mean–Geometric Mean Inequality.

Therefore, the sequence $\{h_n\}$ is decreasing. Because $h_n = \sqrt{\pi}(g(n))^2/g(2n)$ and $\lim_{n\to\infty}(g(n))^2/g(2n) = 1$, it follows that $\lim_{n\to\infty} h_n = \sqrt{\pi}$. This completes the proof.

Editor's Note. Most of the solutions submitted used either Stirling's Formula or Wallis Formula. Michael Andreoli points out that essentially the same problem appears on page 328 of M. Spivak, *Calculus*, W. A. Benjamin Inc., New York, 1967. Mark Ashbaugh and Francisco Vial notice that the same problem can be seen in H. Hochstadt, *The Functions of Mathematical Physics*, Dover, 1986 (originally published by Wiley in 1971). Moreover, they generalized in the following way; they asked to find the largest real number β such that $(1/\sqrt{n+\beta})\prod_{k=1}^{n} \frac{2k}{2k-1}$ is decreasing and they claimed that $\beta = \frac{1}{4}$.

Also solved by Michael Andreoli, Mark Ashbaugh and Francisco Vial (Chile), Michel Bataille (France), Paul Bracken, Bruce S. Burdick, Robert Calcaterra, Hongwei Chen, Daniele Degiorgi (Switzerland), Robert L. Doucette, Dmitry Fleischman, Michael Goldenberg and Mark Kaplan, Eugene A. Herman, Enkel Hysnelaj (Australia) and Elton Bojaxhiu (Germany), Bianca-Teodora Iordache (Romania), Michael Cap Khoury, Santiago de Luxán (Spain) and Ángel Plaza (Spain), Occidental College Problem Solving Group, Paolo Perfetti (Italy), Robert C. Rhoades, John M. Sayer, Joel Schlosberg, Nicholas C. Singer, John Sumner and Aida Kadic-Galeb, Marian Tetiva (Romania), Michael Vowe (Switzerland), Stan Wagon, Haohao Wang and Jerzy Wojdyło, John B. Zacharias, and the proposer. There were two incorrect submissions and one incomplete solution.

An inequality for the excircles of a triangle **October 2009**

1829. *Proposed by Oleh Faynshteyn, Leipzig, Germany.*

Let ABC be a triangle with $BC = a$, $CA = b$, and $AB = c$. Let r_a denote the radius of the excircle tangent to BC, r_b the radius of the excircle tangent to CA, and r_c the radius of the excircle tangent to AB. Prove that

$$\frac{r_a r_b}{(a+b)^2} + \frac{r_b r_c}{(b+c)^2} + \frac{r_c r_a}{(c+a)^2} \leq \frac{9}{16}.$$

I. *Solution by Ercole Suppa, Teramo, Italy.*

Let $\Delta = \text{Area}(\mathbf{ABC})$ and $s = (a+b+c)/2$. Taking into account the identities $r_a = \Delta/(s-a)$, $r_b = \Delta/(s-b)$, $r_c = \Delta/(s-c)$, as well as Heron's formula $\Delta = \sqrt{s(s-a)(s-b)(s-c)}$, the inequality in the problem becomes

$$\frac{s(s-c)}{(a+b)^2} + \frac{s(s-a)}{(b+c)^2} + \frac{s(s-b)}{(c+a)^2} \le \frac{9}{16} \qquad \Leftrightarrow$$

$$\frac{(a+b)^2-c^2}{(a+b)^2} + \frac{(b+c)^2-a^2}{(b+c)^2} + \frac{(c+a)^2-b^2}{(c+a)^2} \le \frac{9}{4} \qquad \Leftrightarrow$$

$$\frac{a^2}{(b+c)^2} + \frac{b^2}{(a+c)^2} + \frac{c^2}{(a+b)^2} \ge \frac{3}{4}. \tag{1}$$

Now, by using Nesbitt's inequality

$$\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} \ge \frac{3}{2}$$

and the well-known inequality $3\left(x^2+y^2+z^2\right) \ge (x+y+z)^2$ we get

$$\left(\frac{a}{b+c}\right)^2 + \left(\frac{b}{c+a}\right)^2 + \left(\frac{c}{a+b}\right)^2 \ge \frac{1}{3}\left(\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b}\right)^2$$

$$\ge \frac{1}{3}\left(\frac{3}{2}\right)^2 = \frac{3}{4}$$

and (1) is proved. The equality holds if and only if $a = b = c$.

II. *Solution by Robert L. Doucette, Department of Mathematics, Computer Science and Statistics, McNeese State University, Lake Charles, LA.*

As in the first solution, the inequality to be proved is equivalent to

$$\frac{a^2}{(b+c)^2} + \frac{b^2}{(a+c)^2} + \frac{c^2}{(a+b)^2} \ge \frac{3}{4}.$$

Letting $f(x) := x^2/(1-x)^2$ for $x \in (0,1)$, this may be written as

$$\frac{1}{3}\left(f\left(\frac{a}{a+b+c}\right) + f\left(\frac{b}{a+b+c}\right) + f\left(\frac{c}{a+b+c}\right)\right) \ge f\left(\frac{1}{3}\right).$$

Since $f''(x) = 2x/(1-x)^3 > 0$ for $x \in (0,1)$, the function f is strictly convex on the interval $(0,1)$. Our inequality follows from Jensen's inequality. It also follows that equality occurs if and only if $\triangle ABC$ is equilateral.

Editor's Note. The statement of the problem was incorrectly published with the reversed inequality. We thank all the readers who noticed the mistake and still managed to solve the correct problem.

Also solved by Arkady Alt, George Apostolopoulos (Greece), Herb Bailey, Michel Bataille (France), Robert Calcaterra, Minh Can, Chip Curtis, Daniele Degiorgi (Switzerland), Marian Dinca (Romania), Sebastián García Saenz (Chile), John G. Heuver (Canada), Enkel Hysnelaj (Australia) and Elton Bojaxhiu (Germany), Bianca-Teodora Iordache (Romania), Omran Kouba (Syria), Victor Y. Kutsenok, Elias Lampakis (Greece), Kee-Wai Lau (China), Peter Nüesch (Switzerland), Jennifer Pajda, Ángel Plaza (Spain) and Sergio Falcón (Spain), Gabriel Prăjitură, Fary Sami, Achilleas Sinefakopoulos (Greece), John Sumner and Aida Kadic-Galeb, Marian Tetiva (Romania), Philip Todd, Michael Vowe (Switzerland), Haohao Wong and Jerzy Wojdyło, and the proposer. There were two incorrect submissions.

A rational quotient of floors of irrationals **October 2009**

1830. *Proposed by H. A. ShahAli, Tehran, Iran.*

Let α and β be positive real numbers and let r be a positive rational number. Find necessary and sufficient conditions to ensure that there exist infinitely many positive integers m such that

$$\frac{\lfloor m\alpha \rfloor}{\lfloor m\beta \rfloor} = r.$$

I. *Solution by Omran Kouba, Higher Institute for Applied Sciences ad Technology, Damascus, Syria.*

The desired condition is $\alpha = r\beta$. Denote by $\{x\}$ the fractional part of x. Note that

$$m(\alpha - r\beta) = \lfloor m\alpha \rfloor - r\lfloor m\beta \rfloor + \{m\alpha\} - r\{m\beta\},$$

so if $\lfloor m\alpha \rfloor = r\lfloor m\beta \rfloor$, then

$$m|\alpha - r\beta| = |\{m\alpha\} - r\{m\beta\}| \leq 1 + r.$$

Hence, if we suppose that there exists a sequence $(m_k)_{k\geq 1}$ of positive integers satisfying $\lim_{k\to\infty} m_k = \infty$ and $\lfloor m_k\alpha \rfloor = r\lfloor m_k\beta \rfloor$, then for every $k \geq 1$, $|\alpha - r\beta| \leq (1 + r)/m_k$. Letting k tend to infinity we conclude that $\alpha = r\beta$ is a necessary condition.

Conversely, let us suppose that $\alpha = r\beta$. By hypothesis $r = p/q$ for some positive integers p and q. If $\ell > 1/\beta$ is a positive integer satisfying $\{\ell\beta\} < 1/(p + q)$, then from $\ell\beta = \lfloor \ell\beta \rfloor + \{\ell\beta\}$ we conclude that $p\ell\beta = p\lfloor \ell\beta \rfloor + p\{\ell\beta\}$, $q\ell\beta = q\lfloor \ell\beta \rfloor + q\{\ell\beta\}$, and $0 \leq p\{\ell\beta\}, q\{\ell\beta\} < 1$. Hence $\lfloor p\ell\beta \rfloor = p\lfloor \ell\beta \rfloor$ and $\lfloor q\ell\beta \rfloor = q\lfloor \ell\beta \rfloor$. Finally, if $m = q\ell$, then

$$\frac{\lfloor m\alpha \rfloor}{\lfloor m\beta \rfloor} = \frac{\lfloor q\ell\alpha \rfloor}{\lfloor q\ell\beta \rfloor} = \frac{\lfloor p\ell\beta \rfloor}{\lfloor q\ell\beta \rfloor} = \frac{p}{q} = r.$$

Thus the result would follow if we prove that the set of positive integers ℓ satisfying $\{\ell\beta\} < 1/(p + q)$ is infinite. This is true if β is rational by taking multiples of the denominator of β. If β is irrational, then by Kronecker's Theorem the set $\{\{\ell\beta\} : \ell > 1\}$ is dense in the interval [0, 1] and consequently infinitely many of its members belong to the interval $[0, 1/(p + q))$. This achieves the proof of sufficiency.

II. *Solution by Nicholas C. Singer, Annandale, VA.*

The necessity is obtained as in the last solution. For the sufficiency suppose $\alpha/\beta = r$. Then α and β are rational or irrational together. If they are rational, say $\beta = t/u$ for positive integers t and u, then $qu\alpha = qu\beta r = pu\beta = pt$ is an integer. Thus for all positive integers k,

$$\frac{\lfloor kqu\alpha \rfloor}{\lfloor kqu\beta \rfloor} = \frac{kqu\alpha}{kqu\beta} = \frac{kpt}{kqt} = r.$$

If α and β are irrational, use continued fractions to obtain strictly increasing sequences of positive integers (k_j) and (m_j) such that

$$\frac{k_j}{m_j} < \frac{\alpha}{p} = \frac{\beta}{q} < \frac{k_j}{m_j} + \frac{1}{m_j^2}.$$

Then $pk_j < m_j\alpha < pk_j + p/m_j$ and $qk_j < m_j\beta < qk_j + q/m_j$. As soon as j is large enough such that $m_j > \max\{p, q\}$, we have $\lfloor m_j\alpha \rfloor = pk_j$ and $\lfloor m_j\beta \rfloor = qk_j$. Thus

$$\frac{\lfloor m_j\alpha \rfloor}{\lfloor m_j\beta \rfloor} = \frac{pk_j}{qk_j} = r.$$

Also solved by George Apostolopoulos (Greece), Michel Bataille (France), Elton Bojaxhiu (Germany) and Enkel Hysnelaj (Australia), Robert Calcaterra, John Christopher, Andrés Fielbaum (Chile), Dmitry Fleischman, Bianca-Teodora Iordache (Romania), Michael Cap Khoury, Northwestern University Math Problem Solving Group, Occidental College Problem Solving Group, Brad Pearson, John Sumner and Aida Kadic-Galeb, Marian Tetiva (Romania), and the proposer.

Answers

Solutions to the Quickies from page 304.

A1003. It is clear that $f^{(n)}(x) = (a_n x^2 + b_n x + c_n)e^x$ for constants a_n, b_n, and c_n. Differentiating once gives the following simple recursive formulas: $(a_0, b_0, c_0) = (1, 0, 0)$, and $(a_{n+1}, b_{n+1}, c_{n+1}) = (a_n, 2a_n + b_n, c_n + b_n)$. Thus $a_n = 1$ for all n, $b_n = 2n$ for all n, and $c_j - c_{j-1} = 2(j-1)$ for all $j \geq 1$. Summing the last equality over $1 \leq j \leq n$ gives $c_n = n(n-1)$. Thus $f^{(n)}(x) = (x^2 + 2nx + n(n-1))e^x$, whose zeros are $-n \pm \sqrt{n}$.

A1004. The concatenation is never equal to the product. The concatenation is larger than the product of the first number and every other number rounded up to the next power of 10. This is in turn larger than the product itself. (Example: 54 and 321 round up to 100 and 1 000. Therefore, $654\,321 > 600\,000 = 6 \cdot 100 \cdot 1000 > 6 \cdot 54 \cdot 321$.)

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

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

Deolalikar, Vinay, $P \neq NP$: Synopsis of proof, `http://www.hpl.hp.com/personal/Vinay_Deolalikar/Papers/pnp_synopsis.pdf`.

Lipton, R. J., Gödel's lost letter and $P = NP$, blog at `http://rjlipton.wordpress.com/`.

AN HP researcher has circulated a 120+-page paper claiming to resolve the most famous problem in theoretical computer science by establishing that $P \neq NP$, that is, that the complexity class NP of problems is strictly larger than the class P. The class P ("polynomial time") is composed of problems for which a problem of size n has solutions whose algorithmic cost is bounded by a polynomial in n; such problems are considered "easy," and examples include primality testing. The class NP ("nondeterministic polynomial time") consists of problems for which a solution can be verified in polynomial time, but for which finding a solution in may take exponential time. Examples include factoring integers and the traveling sales problem. The Clay Mathematics Institute offers a $1 million prize for resolving the question of whether $P = NP$. As of this writing, Deolalikar's paper has not yet been refereed, verified, or refuted. The site by Lipton (and his associated blog) contains links to further developments.

Origami crease pattern design proved NP-hard, `http://www.technologyreview.com/blog/arxiv/25591/`.

Demaine, Erik D., Sandor P. Fekete, and Robert J. Lang, Circle packing for origami design is hard, `http://arxiv.org/abs/1008.1224`.

A problem is NP-hard if it is at least as "hard" as the hardest problem in NP, meaning that the latter can be reduced to the former by a polynomial-time algorithm. If $P \neq NP$, then there is no polynomial-time solution to any NP-hard problem. The traveling sales problem is NP-hard, and the paper by Demaine et al. establishes that determining a crease pattern to fold a square of paper into a three-dimensional shape is NP-hard. They prove that fact by reducing the origami problem into a circle-packing problem known to be NP-hard.

Davidson, Morley, John Dethridge, Herbert Kociemba, and Tomas Rokicki, God's number is 20, `http://www.cube20.org`.

The authors establish that every position of Rubik's Cube can be solved in 20 moves or fewer. Using 35 CPU-years of computer time, Davidson et al. analyzed every position of the puzzle. A position that requires 20 moves has been known since 1995 (they suggest that there are probably at least 100 million of them!), and the upper bound has slowly drifted down from 52 in 1981 to the now-established 20. The authors broke the positions down into sets of positions via cosets and solved each set in about 20 seconds on a PC. They did not find an optimal solution for every position, just one in 20 moves or fewer. Meanwhile, contestants commonly solve Cube positions blindfolded (after visually examining the cube) in under 10 seconds.

Math. Mag. **83** (2010) 311–312. doi:10.4169/002557010X521886. © Mathematical Association of America

Gallian, Joseph A. (ed.), *Mathematics and Sports*, MAA, 2010; xi + 329 pp, $39.95(P) (member price: $29.95). ISBN 978-0-88385-349-8.

This book contains 25 all-new essays, solicited for the 2010 Mathematics Awareness Month. There are several each on baseball, basketball, football, golf, and track and field, and one or two each on NASCAR, tournament scheduling, soccer, and tennis. The mathematics used ranges from probability (including cumulative distribution functions and moments), hypothesis testing, recursion, matrix multiplication, ballistics, combinatorics, to graph theory—with only two or three passing uses of calculus. I found most interesting "Down by 4 with a minute to go," by G. Edgar Parker, about last-minute strategy in basketball, and (after watching Tiger Woods miss many putts the day before) "Tigermetrics," by Roland Minton, which notes that at more than 8 ft, pros make less than half their putts.

Epstein, Richard A., *The Theory of Gambling and Statistical Logic*, 2nd ed., Academic Press, 2009; xiii + 442 pp, $42.95. ISBN 978-0-12-374940-6.

This new edition of a 1967 classic (revised in 1995) gives a wealth of information about strategies in a diverse collection of games of chance and strategy, together with preliminaries about probability, statistics, game theory, and gambling in general. Much of this distilled information cannot be found conveniently elsewhere. This is not a book for a general audience; summation signs, random variables, and exponentials appear early and often. New in this edition is a chapter on Parrondo's paradox (that alternating between two losing games can produce positive expectation). The chapter on betting systems does not consider a house limit on bets, and not all criticisms by reviewers of previous editions have been addressed. [Mathematical variables are rendered inconsistently in math italic or roman, the index didn't help me find what I sought (e.g., double factorial), and there are a few misprints: p. 58, Thm. IV, line 4, α should be a; p. 323, last line should be *American Mathematical Monthly* 107.]

Stewart, Ian, *Cows in the Maze and Other Mathematical Explorations*, Oxford University Press, 2010; xi + 296 pp, $17.95(P). ISBN 978-0-19-956207-7.

This is the third collection of Ian Stewart's Mathematical Recreations columns (with the addition of feedback from readers) from *Scientific American* and *Pour la Science*. The 20 columns feature mainly geometry: opaque fences, quadruped locomotion, knotted tiles, wormholes, sphericons, teardrop shapes, knight's tours, string figures, Klein bottles, knot energies, dodecahedra, and the fiendish maze of the title. There are also non-transitive dice, Hex, primes in progression, the interrogator's fallacy, magic squares, and more. There are (almost) no equations except in the column on probability in jurisprudence.

Ehrenberg, Rachel, Elusive symmetry appears in nature: Complex E_8 patterns detected in supercold physical system, *Science News* (30 January 2010) 15.

The exceptional Lie group E_8, whose admissible representations were computed in 2007 (`http://aimath.org/E8/`), has been found to occur in nature. When cobalt niobate is chilled toward absolute zero in a magnetic field, the electron spins form quasiparticles that resonate. "Two of the frequencies are in the ratio of the golden mean.... Ratios of the five frequences found correspond to the complex E_8 Lie group symmetry."

Nickerson, Raymond S., *Mathematical Reasoning: Patterns, Problems, Conjecture, and Proofs*, Psychology Press, 2010; xi + 583 pp, $69.95. ISBN 978-1-84872-837-1.

Few authors who are not mathematicians have the understanding and sympathy for the subject and its practitioners that is shown in this book, whose author is an experimental psychologist and former VP of Bolt Beranek and Newman Inc. He surveys the nature, joys, usefulness, foundations, and teaching of mathematics, in its manifestations of pattern-finding, problem-solving, conjecture-making, and proof-devising. He concludes, "the supreme reason for acquiring some competence in mathematics is the door it opens to an immensely attractive and rewarding workspace—or playground—for the mind." This is a great book for nonmathematicians, one that a mathematician can praise.

NEWS AND LETTERS

39th USA Mathematical Olympiad
1st USA Junior Mathematical Olympiad

JACEK FABRYKOWSKI
Youngstown State University / Central Michigan University
Mount Pleasant, MI 48859
jfabrykowski@ysu.edu

STEVEN R. DUNBAR
MAA American Mathematics Competitions
University of Nebraska-Lincoln
Lincoln, NE 68588-0658
sdunbar@maa.org

In 2010 the Committee on the American Mathematics Competitions introduced the USA Junior Mathematical Olympiad (USAJMO) for students in 10th grade and below. Offered for the first time in April 2010, the USAJMO bridges the computational solutions of the AIME and the proofs required on the USAMO.

The USAJMO sets 6 problems over 2 days, the same as the USAMO. Problems J1, J2 on Day 1, and Problems J4, J5 do not appear on the USAMO. Problem J3 is the same as Problem 1 on the USAMO, Problem J6 is the same as Problem 4 on the USAMO.

USAMO Problems

1. Let $AXYZB$ be a convex pentagon inscribed in a semicircle of diameter AB. Denote by P, Q, R, S the feet of the perpendiculars from Y onto lines AX, BX, AZ, BZ, respectively. Prove that the acute angle formed by lines PQ and RS is half the size of $\angle XOZ$, where O is the midpoint of segment AB.
2. There are n students standing in a circle, one behind the other. The students have heights $h_1 < h_2 < \cdots < h_n$. If a student with height h_k is standing directly behind a student with height h_{k-2} or less, the two students are permitted to switch places. Prove that it is not possible to make more than $\binom{n}{3}$ such switches before reaching a position in which no further switches are possible.
3. The 2010 positive numbers $a_1, a_2, \ldots, a_{2010}$ satisfy the inequality $a_i a_j \le i + j$ for all distinct indices i, j. Determine, with proof, the largest possible value of the product $a_1 a_2 \cdots a_{2010}$.
4. Let ABC be a triangle with $\angle A = 90°$. Points D and E lie on sides AC and AB, respectively, such that $\angle ABD = \angle DBC$ and $\angle ACE = \angle ECB$. Segments BD and CE meet at I. Determine whether or not it is possible for segments AB, AC, BI, ID, CI, IE to all have integer lengths.
5. Let $q = \frac{3p-5}{2}$ where p is an odd prime, and let

$$S_q = \frac{1}{2 \cdot 3 \cdot 4} + \frac{1}{5 \cdot 6 \cdot 7} + \cdots + \frac{1}{q(q+1)(q+2)}.$$

Prove that if $\frac{1}{p} - 2S_q = \frac{m}{n}$ for integers m and n, then $m - n$ is divisible by p.

Math. Mag. **83** (2010) 313–319. doi:10.4169/002557010X521895. © Mathematical Association of America

6. A blackboard contains 68 pairs of nonzero integers. Suppose that for each positive integer k at most one of the pairs (k, k) and $(-k, -k)$ is written on the blackboard. A student erases some of the 136 integers, subject to the condition that no two erased integers may add to 0. The student then scores one point for each of the 68 pairs in which at least one integer is erased. Determine, with proof, the largest number N of points that the student can guarantee to score regardless of which 68 pairs have been written on the board.

USAMO Solutions

1. Let T be the foot of the perpendicular from Y to line AB. We note the P, Q, T are the feet of the perpendiculars from Y to the sides of triangle ABX. Because Y lies on the circumcircle of triangle ABX, points P, Q, T are collinear, by Simson's theorem. Likewise, points S, R, T are collinear. We need to show that $\angle XOZ = 2\angle PTS$ or

$$\begin{aligned}\angle PTS = \frac{\angle XOZ}{2} = \frac{\overset{\frown}{XZ}}{2} &= \frac{\overset{\frown}{XY}}{2} + \frac{\overset{\frown}{YZ}}{2}\\ &= \angle XAY + \angle ZBY = \angle PAY + \angle SBY.\end{aligned}$$

Because $\angle PTS = \angle PTY + \angle STY$, it suffices to prove that

$$\angle PTY = \angle PAY \quad \text{and} \quad \angle STY = \angle SBY;$$

that is, to show that quadrilaterals $APYT$ and $BSYT$ are cyclic, which is evident, because $\angle APY = \angle ATY = 90^\circ$ and $\angle BTY = \angle BSY = 90^\circ$.

Titu Andreescu suggested this problem.

2. Let h_i also denote the student with height h_i. We prove that for $1 \le i < j \le n$, h_j can switch with h_i at most $j - i - 1$ times. We proceed by induction on $j - i$, the base case $j - i = 1$ being evident because h_i is not allowed to switch with h_{i-1}.

For the inductive step, note that h_i, h_{j-1}, h_j can be positioned on the circle either in this order or in the order h_i, h_j, h_{j-1}. Since h_{j-1} and h_j cannot switch, the only way to change the relative order of these three students is for h_i to switch with either h_{j-1} or h_j. Consequently, any two switches of h_i with h_j must be separated by a switch of h_i with h_{j-1}. Since there are at most $j - i - 2$ of the latter, there are at most $j - i - 1$ of the former.

The total number of switches is thus at most

$$\begin{aligned}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-i-1) &= \sum_{i=1}^{n-1}\sum_{j=0}^{n-i-1} j = \sum_{i=1}^{n-1}\binom{n-i}{2}\\ &= \sum_{i=1}^{n-1}\left(\binom{n-i+1}{3} - \binom{n-i}{3}\right) = \binom{n}{3}.\end{aligned}$$

Kiran Kedlaya suggested this problem.

3. Multiplying together the inequalities $a_{2i-1}a_{2i} \le 4i - 1$ for $i = 1, 2, \dots, 1005$, we get

$$a_1 a_2 \cdots a_{2010} \le 3 \cdot 7 \cdot 11 \cdots 4019. \tag{1}$$

The tricky part is to show that this bound can be attained.

Let

$$a_{2008} = \sqrt{\frac{4017 \cdot 4018}{4019}}, \quad a_{2009} = \sqrt{\frac{4019 \cdot 4017}{4018}}, \quad a_{2010} = \sqrt{\frac{4018 \cdot 4019}{4017}},$$

and define a_i for $i < 2008$ by downward induction using the recursion $a_i = (2i+1)/a_{i+1}$. We then have

$$a_i a_j = i + j \qquad \text{whenever } j = i+1 \qquad \text{or} \qquad i = 2008,\ j = 2010. \tag{2}$$

We will show that (2) implies $a_i a_j \le i + j$ for all $i < j$, so that this sequence satisfies the hypotheses of the problem. Since $a_{2i-1}a_{2i} = 4i - 1$ for $i = 1, \dots, 1005$, the inequality (1) is an equality, so the bound is attained.

We show that $a_i a_j \le i + j$ for $i < j$ by downward induction on $i + j$. There are several cases:

- If $j = i + 1$, or $i = 2008$, $j = 2010$, then $a_i a_j = i + j$, from (2).
- If $i = 2007$, $j = 2009$, then

$$a_i a_{i+2} = \frac{(a_i a_{i+1})(a_{i+2} a_{i+3})}{(a_{i+1} a_{i+3})} = \frac{(2i+1)(2i+5)}{2i+4} < 2i + 2.$$

  Here the second equality comes from (2), and the inequality is checked by multiplying out: $(2i+1)(2i+5) = 4i^2 + 12i + 5 < 4i^2 + 12i + 8 = (2i+2)(2i+4)$.
- If $i < 2007$ and $j = i + 2$, then we have

$$a_i a_{i+2} = \frac{(a_i a_{i+1})(a_{i+2} a_{i+3})(a_{i+2} a_{i+4})}{(a_{i+1} a_{i+2})(a_{i+3} a_{i+4})} \le \frac{(2i+1)(2i+5)(2i+6)}{(2i+3)(2i+7)} < 2i + 2.$$

  The first inequality holds by applying the induction hypothesis for $(i+2, i+4)$, and (2) for the other pairs. The second inequality can again be checked by multiplying out: $(2i+1)(2i+5)(2i+6) = 8i^3 + 48i^2 + 82i + 30 < 8i^3 + 48i^2 + 82i + 42 = (2i+2)(2i+3)(2i+7)$.
- If $j - i > 2$, then

$$a_i a_j = \frac{(a_i a_{i+1})(a_{i+2} a_j)}{a_{i+1} a_{i+2}} \le \frac{(2i+1)(i+2+j)}{2i+3} < i + j.$$

  Here we have used the induction hypothesis for $(i+2, j)$, and again we check the last inequality by multiplying out: $(2i+1)(i+2+j) = 2i^2 + 5i + 2 + 2ij + j < 2i^2 + 3i + 2ij + 3j = (2i+3)(i+j)$.

This covers all the cases and shows that $a_i a_j \le i + j$ for all $i < j$, as required.

Gabriel Carroll suggested this problem.

4. Let $BD = m$, $AD = x$, $DC = y$, $AB = c$, $BC = a$, and $AC = b$. The Bisector Theorem implies $\frac{x}{b-x} = \frac{c}{a}$ and the Pythagorean Theorem yields $m^2 = x^2 + c^2$. Both equations imply that

$$2ac = \frac{(bc)^2}{m^2 - c^2} - a^2 - c^2$$

and therefore a is rational. Therefore, $x = \frac{bc}{a+c}$ is also rational, and so is y. Let now $\angle ABD = \alpha$ and $\angle ACE = \beta$ where $\alpha + \beta = \pi/4$. It is obvious that $\cos\alpha$ and $\cos\beta$ are both rational and the above shows that $\sin\alpha = x/m$ is rational. On the other hand, $\cos\beta = \cos(\pi/4 - \alpha) = (\sqrt{2}/2)(\sin\alpha + \sin\beta)$, which is a contradiction. The solution shows that a stronger statement is true: There is no right triangle with both legs and bisectors of acute angles all with integer lengths.

Zuming Feng suggested this problem. Jacek Fabrykowski suggested the given solution.

5. We have

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

Hence

$$\begin{aligned} 2S_q &= \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{q} + \frac{1}{q+1} + \frac{1}{q+2}\right) - 3\left(\frac{1}{3} + \frac{1}{6} + \cdots + \frac{1}{q+1}\right) \\ &= \left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{\frac{3p-1}{2}}\right) - \left(1 + \frac{1}{2} + \cdots + \frac{1}{\frac{p-1}{2}}\right), \end{aligned}$$

and so

$$\begin{aligned} 1 - \frac{m}{n} = 1 + 2S_q - \frac{1}{p} &= \frac{1}{\frac{p+1}{2}} + \cdots + \frac{1}{p-1} + \frac{1}{p+1} + \cdots + \frac{1}{\frac{3p-1}{2}} \\ &= \left(\frac{1}{\frac{p+1}{2}} + \frac{1}{\frac{3p-1}{2}}\right) + \cdots + \left(\frac{1}{p-1} + \frac{1}{p+1}\right) \\ &= \frac{p}{\left(\frac{p+1}{2}\right)\left(\frac{3p-1}{2}\right)} + \cdots + \frac{p}{(p-1)(p+1)}. \end{aligned}$$

Because all denominators are relatively prime with p, it follows that $n - m$ is divisible by p and we are done.

Titu Andreescu suggested this problem.

6. The answer is 43.

We first show that we can always get 43 points. Without loss of generality, we assume that the value of x is positive for every pair of the form (x, x) (otherwise, replace every occurrence of x on the blackboard by $-x$, and every occurrence of $-x$ by x). Consider the ordered n-tuple $(a_1, a_2, \ldots, a_n)$ where $a_1, |a_2|, \ldots, a_n$ denote all the distinct absolute values of the integers written on the board.

Let $\phi = \frac{\sqrt{5}-1}{2}$, which is the positive root of $\phi^2 + \phi = 1$. We consider 2^n possible underlining strategies: Every strategy corresponds to an ordered n-tuple $s = (s_1, \ldots, s_n)$ with $s_i = \phi$ or $s_i = 1 - \phi$ $(1 \le i \le n)$. If $s_i = \phi$, then we underline all occurrences of a_i on the blackboard. If $s_i = 1 - \phi$, then we underline all occurrences of $-a_i$ on the blackboard. The weight $w(s)$ of strategy s equals the product $\prod_{i=1}^{n} s_i$. It is easy to see that the sum of weights of all 2^n strategies is equal to $\sum_s w(s) = \prod_{i=1}^{n}[\phi + (1-\phi)] = 1$.

For every pair p on the blackboard and every strategy s, we define a corresponding cost coefficient $c(p, s)$: If s scores a point on p, then $c(p, s)$ equals the weight $w(s)$. If s does not score on p, then $c(p, s)$ equals 0. Let $c(p)$ denote the sum of coefficients $c(p, s)$ taken over all s. Now consider a fixed pair $p = (x, y)$:

(a) In this case, we assume that $x = y = a_j$. Then every strategy that underlines a_j scores a point on this pair. Then $c(p) = \phi \prod_{i \ne j}^{n}[\phi + (1-\phi)] = \phi$.

(b) In this case, we assume that $x \ne y$. We have

$$c(p) = \begin{cases} \phi^2 + \phi(1-\phi) + (1-\phi)\phi = 3\phi - 1, & (x, y) = (a_k, a_\ell); \\ \phi(1-\phi) + (1-\phi)\phi + (1-\phi)^2 = \phi, & (x, y) = (-a_k, -a_\ell); \\ \phi^2 + \phi(1-\phi) + (1-\phi)^2 = 2 - 2\phi, & (x, y) = (\pm a_k, \mp a_\ell). \end{cases}$$

By noting that $\phi \approx 0.618$, we can easily conclude that $c(p) \ge \phi$.

We let C denote the sum of the coefficients $c(p, s)$ taken over all p and s. These observations yield that

$$C = \sum_{p,s} c(p, s) = \sum_{p} c(p) \geq \sum_{p} \phi = 68\phi > 42.$$

Suppose for the sake of contradiction that every strategy s scores at most 42 points. Then every s contributes at most $42w(s)$ to C, and we get $C \leq 42 \sum_s w(s) = 42$, which contradicts $C > 42$.

To complete our proof, we now show that we cannot always get 44 points. Consider the blackboard contains the following 68 pairs: for each of $m = 1, \ldots, 8$, five pairs of (m, m) (for a total of 40 pairs of type (a)); for every $1 \leq m < n \leq 8$, one pair of $(-m, -n)$ (for a total of $\binom{8}{2} = 28$ pairs of type (b)). We claim that we cannot get 44 points from this initial stage. Indeed, assume that exactly k of the integers $1, 2, \ldots, 8$ are underlined. Then we get at most $5k$ points on the pairs of type (a), and at most $28 - \binom{k}{2}$ points on the pairs of type (b). We can get at most $5k + 28 - \binom{k}{2}$ points. Note that the quadratic function $5k + 28 - \binom{k}{2} = -\frac{k^2}{2} + \frac{11k}{2} + 28$ obtains its maximum 43 (for integers k) at $k = 5$ or $k = 6$. Thus, we can get at most 43 points with this initial distribution, establishing our claim and completing our solution.

Gerhard Woeginger suggested this problem.

USAJMO Problems

J1. A *permutation* of the set of positive integers $[n] = \{1, 2, \ldots, n\}$ is a sequence $(a_1, a_2, \ldots, a_n)$ such that each element of $[n]$ appears precisely one time as a term of the sequence. For example, $(3, 5, 1, 2, 4)$ is a permutation of $[5]$. Let $P(n)$ be the number of permutations of $[n]$ for which ka_k is a perfect square for all $1 \leq k \leq n$. Find with proof the smallest n such that $P(n)$ is a multiple of 2010.

J2. Let $n > 1$ be an integer. Find, with proof, all sequences $x_1, x_2, \ldots, x_{n-1}$ of positive integers with the following three properties:

(a) $x_1 < x_2 < \cdots < x_{n-1}$;

(b) $x_i + x_{n-i} = 2n$ for all $i = 1, 2, \ldots, n - 1$;

(c) given any two indices i and j (not necessarily distinct) for which $x_i + x_j < 2n$, there is an index k such that $x_i + x_j = x_k$.

J4. A triangle is called a *parabolic triangle* if its vertices lie on a parabola $y = x^2$. Prove that for every nonnegative integer n, there is an odd number m and a parabolic triangle with vertices at three distinct points with integer coordinates with area $(2^n m)^2$.

J5. Two permutations $a_1, a_2, \ldots, a_{2010}$ and $b_1, b_2, \ldots, b_{2010}$ of the numbers $1, 2, \ldots,$ 2010 are said to *intersect* if $a_k = b_k$ for some value of k in the range $1 \leq k \leq 2010$. Show that there exist 1006 permutations of the numbers $1, 2, \ldots, 2010$ such that any other such permutation is guaranteed to intersect at least one of these 1006 permutations.

USAJMO Solutions

J1. Every integer in S_n can be uniquely written in the form $x^2 \cdot q$, where q is either 1 or *square free*. Let $\langle q \rangle$ denote the set $\{1^2 \cdot q, 2^2 \cdot q, 3^2 \cdot q, \ldots\}$.

Note that for f to satisfy the square-ness property, it must permute $\langle q \rangle$ for every q. To see this, notice that given an arbitrary square-free q, in order for $q \cdot f(q)$ to be a square, $f(q)$ needs to contribute one of every prime factor in q, after which it can take only even powers of primes. Thus, $f(q)$ is equal to the product of q and some perfect square.

The number of f that permute the $\langle q\rangle$ is equal to

$$\prod_{\substack{q\le n\\ q \text{ is square-free}}} \left\lfloor \sqrt{\frac{n}{q}} \right\rfloor !$$

For $2010 = 2\cdot 3\cdot 5\cdot 67$ to divide $P(n)$, we simply need 67! to appear in this product, which will first happen in $\langle 1\rangle$ so long as $\sqrt{n/q} \ge 67$ for some n and q. The smallest such n is $67^2 = 4489$.

Andy Niedermaier suggested this problem.

J2. Assume $x_1, x_2, \ldots, x_{n-1}$ satisfies the conditions. By condition (a)

$$x_1,\ 2x_1,\ x_1+x_2,\ x_1+x_3,\ x_1+x_4,\ldots,\ x_1+x_{n-2}$$

is an increasing sequence. By condition (c), this new sequence is a subsequence of the original sequence. Because both sequences have exactly $n-1$ terms, these two sequences are identical; that is, $2x_1 = x_2$ and $x_1 + x_j = x_{j+1}$ for $2 \le j \le n-2$. It follows that $x_j = jx_1$ for $1 \le j \le n-1$. By condition (b), $(x_1, x_2, \ldots, x_{n-1}) = (2, 4, \ldots, 2n-2)$.

Razvan Gelca suggested this problem. Richard Stong suggested the given solution.

J4. Let $A = (a, a^2)$, $B = (b, b^2)$, and $C = (c, c^2)$, with $a < b < c$. Then the area of triangle ABC is

$$[ABC] = (2^n m)^2 = \frac{(b-a)(c-a)(c-b)}{2}.$$

Set $b - a = x$ and $c - b = y$ (where both x and y are positive integers), the above equation becomes

$$(2^n m)^2 = \frac{xy(x+y)}{2}. \tag{3}$$

If $n = 0$, we take $m = x = y = 1$. If $n = 1$, we take $m = 3$, $x = 1$, $y = 8$. Assume that $n \ge 2$. Let a, b, c be a primitive Pythagorean triple with b even. Let $b = 2^r d$ where d is odd and $r \ge 2$. Let $x = 2^{2k}$, $y = 2^{2k} b$, and $z = 2^{2k} c$ where $k \ge 0$. We let $m = adc$ and $r = 2$ if $n = 3k+2$, $r = 3$ if $n = 3k+3$, and $r = 4$ if $n = 3k+4$.

Assuming that $x = a\cdot 2^s$, $y = b\cdot 2^t$, other triples are possible:

(a) If $n = 3k$, then let $m = 1$ and $x = y = 2^{2k}$.

(b) If $n = 3k+1$, then take $m = 3$, $x = 2^{2k}$, $y = 2^{2k+3}$.

(c) If $n = 3k+2$, then take $m = 63$, $x = 49\cdot 2^{2k}$, and $y = 2^{2k+5}$.

Zuming Feng suggested this problem. Jacek Fabrykowski suggested the given solution.

J5. Create 1006 permutations $X_1, X_2, \ldots, X_{1006}$, the first 1006 positions of which are all possible cyclic rotations of the sequence $1, 2, 3, 4, \ldots, 1005, 1006$, and the remaining 1004 positions are filled arbitrarily with the remaining numbers $1007, \ldots, 2009, 2010$:

We claim that at least one of these 1006 sequences has the same integer at the same position as the initial (unknown) permutation. Suppose not. Then the set of the first (the left) 1006 integers of every sequence X_i, $i = 1, \ldots, 1006$, contains no integers from 1007 to 2010, but there are only 1004 such integers, therefore any other permutation must have two of the integers $1, 2, 3, 4, \ldots, 1005, 1006$ within the first 1006 places. Consequently, at least two sequences X_i satisfy the conclusion of the problem.

Gregory Galperin suggested this problem.

2010 Olympiad Results. The top twelve students on the 2010 USAMO were (in alphabetical order):

Timothy Chu	12	Lynbrook High School	San Jose	CA
Calvin Deng	9	William G. Enloe High School	Raleigh	NC
Michael Druggan	11	Tates Creek High School	Lexington	KY
Brian Hamrick	12	Thomas Jefferson High School	Alexandria	VA
Travis Hance	12	Lakota West High School	West Chester	OH
Xiaoyu He	10	Acton-Boxborough High School	Acton	MA
Mitchell Lee	10	Thomas Jefferson High School	Alexandria	VA
In Sung Na	11	Northern Valley High School	Old Tappan	NJ
Evan O'Dorney	11	Berkeley Math Circle	Berkeley	CA
Toan Phan	12	Taft School	Watertown	CT
Hunter Spink	11	Western Canada High School	Calgary	AB
Allen Yuan	11	Detroit Country Day School	Beverly Hills	MI

The top thirteen students on the 2010 USAJMO were (in alphabetical order):

Yury Aglyamov	9	Liberal Arts and Science Academy HS	Austin	TX
Ravi Bajaj	10	Phillips Exeter Academy	Exeter	NH
Evan Chen	8	Horner Junior High School	Fremont	CA
Zijing Gao	9	North Carolina State University	Raleigh	NC
Gill Goldshlager	9	Walton High School	Marietta	GA
Youkow Homma	10	Carmel High School	Carmel	IN
Jesse Kim	9	Henry M. Gunn High School	Palo Alto	CA
Sadik Shahidain	10	Princeton High School	Princeton	NJ
Alexander Smith	9	La Plata High School	La Plata	MD
Susan Di Yun Sun	10	West Vancouver SS	Vancouver	BC
Jiaqi Xie	10	Cypress Bay High School	Weston	FL
Jeffrey Yan	9	Palo Alto High School	Palo Alto	CA
Kevin Zhou	10	Woburn CI	North York	ON

Allen Yuan received a $20,000 scholarship from the Akamai Foundation for first place and Xiaoyu He and Toan Phan received a $12,500 scholarship for tying for second place on the IMO. Each of the 12 USAMO winners received a $500 bond from Robert Balles.

51st International Mathematical Olympiad

ZUMING FENG
Phillips Exeter Academy
Exeter, NH 03833
zfeng@exeter.edu

PO-SHEN LOH
Carnegie Mellon University
Pittsburgh, PA 15213
ploh@cmu.edu

YI SUN
952 Windsor Hills Circle
San Jose, CA 95123
yi.sun@post.harvard.edu

Problems (Day 1)

1. Determine all functions $f : \mathbb{R} \to \mathbb{R}$ such that the equality

$$f(\lfloor x \rfloor y) = f(x)\lfloor f(y) \rfloor$$

holds for all $x, y \in \mathbb{R}$. (Here $\lfloor z \rfloor$ denotes the greatest integer less than or equal to z.)

2. Let I be the incentre of triangle ABC and let Γ be its circumcircle. Let the line AI intersect Γ again at D. Let E be a point on the arc $\widehat{BDC}$ and F a point on the side BC such that

$$\angle BAF = \angle CAE < \tfrac{1}{2}\angle BAC.$$

Finally, let G be the midpoint of the segment IF. Prove that the lines DG and EI intersect on Γ.

3. Let $\mathbb{N}$ be the set of positive integers. Determine all functions $g : \mathbb{N} \to \mathbb{N}$ such that

$$\big(g(m) + n\big)\big(m + g(n)\big)$$

is a perfect square for all $m, n \in \mathbb{N}$.

Problems (Day 2)

4. Let P be a point inside triangle ABC. The lines AP, BP, and CP intersect the circumcircle Γ of triangle ABC again at the points K, L, and M, respectively. The tangent to Γ at C intersects the line AB at S. Suppose that $SC = SP$. Prove that $MK = ML$.

5. In each of six boxes B_1, B_2, B_3, B_4, B_5, B_6 there is initially one coin. There are two types of operation allowed:

   Type 1: Choose a nonempty box B_j with $1 \le j \le 5$. Remove one coin from B_j and add two coins to B_{j+1}.

   Type 2: Choose a nonempty box B_k with $1 \le k \le 4$. Remove one coin from B_k and exchange the contents of (possibly empty) boxes B_{k+1} and B_{k+2}.

   Determine whether there is a finite sequence of such operations that results in boxes B_1, B_2, B_3, B_4, B_5 being empty and box B_6 containing exactly $2010^{2010^{2010}}$ coins. (Note that $a^{b^c} = a^{(b^c)}$.)

Math. Mag. **83** (2010) 320–323. doi:10.4169/002557010X528041.

6. Let $a_1, a_2, a_3, \ldots$ be a sequence of positive real numbers. Suppose that for some positive integer s, we have
$$a_n = \max\{a_k + a_{n-k} \mid 1 \le k \le n-1\}$$
for all $n > s$. Prove that there exist positive integers ℓ and N, with $\ell \le s$ and such that $a_n = a_\ell + a_{n-\ell}$ for all $n \ge N$.

Solutions

1. The answer is $f(x) = c$ for all x, where $c = 0$ or $1 \le c < 2$. To prove that these are the only possible solutions, consider two cases. First suppose that $\lfloor f(y) \rfloor = 0$ whenever $0 \le y < 1$. Then $f(\lfloor x \rfloor y) = f(x)\lfloor f(y) \rfloor = 0$ whenever $0 \le y < 1$. Since every real number can be represented as a product of the form $\lfloor x \rfloor y$ with $x \in \mathbb{R}$ and $0 \le y < 1$, in this case f is identically zero.

   Otherwise, suppose $\lfloor f(y_0) \rfloor \ne 0$ for some $0 \le y_0 < 1$. For any x_n satisfying $n \le x_n < n+1$, set $y = y_0$ and $x = x_n$ in the given equality to obtain $f(ny_0) = f(x_n)\lfloor f(y_0) \rfloor$. Letting $c_n = \frac{f(ny_0)}{\lfloor f(y_0) \rfloor}$, it follows that $f(x_n) = c_n$ for all $x_n \in [n, n+1)$. In particular, we have $\lfloor c_0 \rfloor = \lfloor f(y_0) \rfloor \ne 0$, hence $c_0 \ne 0$. Now set $x = y = 0$ in the given equality to obtain $c_0 = f(0) = f(0)\lfloor f(0) \rfloor = c_0 \lfloor c_0 \rfloor$, hence $\lfloor c_0 \rfloor = 1$. Finally, setting $y = 0$ and $x = n$ in the given equality, we find $c_n = f(n) = \frac{f(0)}{\lfloor f(0) \rfloor} = \frac{c_0}{\lfloor c_0 \rfloor} = c_0$. Therefore, in this case we have $f(x) = c_0$ for all x, and $\lfloor c_0 \rfloor = 1$.

   This problem was proposed by Pierre Bornsztein of France.

2. Let P be the second intersection of ray EI and Γ, and let segments PD and FI meet at M. We wish to show that $M = G$, or, equivalently, $FM = MI$. Let Q be the intersection of segments PD and AF. Applying Menelaus's theorem to triangle AFI and line QMD gives $\frac{FQ \cdot AD \cdot IM}{QA \cdot DI \cdot MF} = 1$. Hence it suffices to show that $\frac{FQ \cdot AD}{QA \cdot DI} = 1$ or equivalently that $AD/AQ = (DI + DA)/FA$.

   Triangles QAD and IAE are similar, so $AD/AQ = EA/AI$. Also, triangles ABF and AEC are similar, so we have $AF/AB = AC/AE$. Together these imply that $\frac{AD}{AQ} = \frac{AB \cdot AC}{AF \cdot AI}$. Now, let H be the intersection of BC and AD; notice that triangles DHC and DCA are similar, hence $DC^2 = DH \cdot DA$. Now because $\angle DCI = \angle CID$, we have $DC = DI$, hence $DA^2 - DI^2 = DA^2 - DC^2 = DA^2 - DH \cdot DA = DA \cdot HA$. On the other hand, notice that triangles ABH and ADC are similar, so $DA \cdot HA = AB \cdot AC$. Putting these together, we see that $\frac{AD}{AQ} = \frac{AB \cdot AC}{AF \cdot AI} = \frac{DA \cdot HA}{AF \cdot AI} = \frac{DI + DA}{FA}$, as needed.

   This problem was proposed by Wai Ming Tai of Hong Kong and Chongli Wang of China.

3. All functions of the form $g(n) = n + c$ for a constant nonnegative integer c satisfy the problem conditions. We claim that these are the only such functions.

   We first show that g must be injective. Suppose instead that $g(a) = g(b)$ for some $a \ne b$. Choose n so that $n + g(a) = p$ is prime and greater than $|a - b|$. From the hypothesis both $p(g(n) + a)$ and $p(g(n) + b)$ must be perfect squares, meaning that $g(n) + a$ and $g(n) + b$ are both divisible by p. But this is impossible, as $p > |a - b|$. Therefore, g is injective as claimed.

   We now show that $|g(k+1) - g(k)| = 1$ for all k. Suppose instead that some prime p divides $g(k+1) - g(k)$. Now, choose an integer n as follows. If $p^2 \mid g(k+1) - g(k)$, then take n so that $n + g(k+1)$ is divisible by p but not p^2. Otherwise, take n so that $n + g(k+1)$ is divisible by p^3 but not p^4. Note that the maximum power of p dividing $n + g(k+1)$ and $n + g(k)$ is odd. Now, the hypothesis implies that $(n + g(k+1))(g(n) + k + 1)$ and $(n + g(k))(g(n) + k)$ are both squares, meaning that $g(n) + k + 1$ and $g(n) + k$ are both divisible by p, a contradiction.

   For each k, we now have either $g(k+1) = g(k) + 1$ or $g(k+1) = g(k) - 1$. But g is injective, so if the latter occurs for some k, then it occurs for all $k' > k$, an impossibility because g takes positive values. Therefore, we have $g(k+1) = g(k) + 1$ for all k, hence $g(k) = k + g(1) - 1$.

This problem was proposed by Gabriel Carroll of the USA.

4. Without loss of generality, we may assume that S is on ray BA. Set $x_1 = \angle PAB$, $y_1 = \angle PBC$, $z_1 = \angle PCA$, $x_2 = \angle PAC$, $y_2 = \angle PBA$, and $z_2 = \angle PCB$. Because SC is tangent to Γ, we have $SC^2 = SA \cdot SB$ by the Power of a Point Theorem, and $\angle SCP = \angle SCM = \angle ACM + \angle ACS = z_1 + \angle ABC = z_1 + y_1 + y_2$. Because $SP = SC$, we have $SP^2 = SC^2 = SA \cdot SB$, so triangles SAP and SPB are similar. It follows that $\angle SPA = \angle SBP = y_2$ and that $\angle ASP = \angle BAP - \angle SPA = x_1 - y_2$. Now, $SP = SC$ implies $\angle SPC = \angle SCP = z_1 + y_1 + y_2$, so $\angle PSC = 180^\circ - 2(z_1 + y_1 + y_2) = x_1 + x_2 + z_2 - (z_1 + y_1 + y_2)$. Notice that $\angle ASC = \angle BAC - \angle ACS = (x_1 + x_2) - (y_1 + y_2)$, so we have $\angle ASP = \angle ASC - \angle PSC = z_1 - z_2$. Combining our two computations of $\angle ASP$ yields $x_1 - y_2 = z_1 - z_2$ or $x_1 + z_2 = y_2 + z_1$. That is, we have $(\widehat{KB} + \widehat{BM})/2 = (\widehat{LA} + \widehat{AM})/2$, hence $\widehat{KM}/2 = \widehat{LM}/2$ and $MK = ML$.

   This problem was proposed by Marcin E. Kuczma of Poland.

5. The answer is yes. Although the problem specifies that the number of boxes is $n = 6$, the operations extend in the obvious way to general values of n. Our proof will consider several different values of n on the way to the final result. For this, it is convenient to let $(b_1, \dots, b_n)$ denote the n-box configuration where b_1 balls are in box B_1, b_2 balls are in box B_2, etc. Write $(b_1, \dots, b_n) \to (b'_1, \dots, b'_n)$ if we can obtain the configuration $(b'_1, \dots, b'_n)$ from $(b_1, \dots, b_n)$ following the rules in the n-box setting. We begin with two lemmas.

LEMMA 1. *Let a be a positive integer. Then* $(a, 0, 0) \to (0, 2^a, 0)$.

Proof. We will show that $(a, 0, 0) \to (a-k, 2^k, 0)$ for every $1 \le k \le a$, by inducting on k. For $k = 1$, applying a Type 1 operation to the first number gives $(a, 0, 0) \to (a-1, 2, 0) = (a-1, 2^1, 0)$. Now assume the statement holds for some $k < a$. Starting from $(a-k, 2^k, 0)$, repeatedly applying 2^k many Type 1 operations at the middle box yields $(a-k, 2^k, 0) \to \cdots \to (a-k, 0, 2^{k+1})$. A final Type 2 operation applied at the first box produces $(a-k, 0, 2^{k+1}) \to (a-k-1, 2^{k+1}, 0)$, completing the induction. ∎

LEMMA 2. *Define* $P_n = \underbrace{2^{2^{\cdot^{\cdot^{2}}}}}_{n}$. *Then* $(a, 0, 0, 0) \to (0, P_a, 0, 0)$ *for every positive integer a.*

Proof. We will show that $(a, 0, 0, 0) \to (a-k, P_k, 0, 0)$ for $1 \le k \le a$, by inducting on k. For $k = 1$, a Type 1 operation applied at the first box gives $(a, 0, 0, 0) \to (a-1, P_1, 0, 0)$. Now assume that $(a, 0, 0, 0) \to (a-k, P_k, 0, 0)$ for some $a < k$. Applying Lemma 1 to the last three boxes, we obtain $(a-k, P_k, 0, 0) \to (a-k, 0, P_{k+1}, 0)$. A final Type 2 operation applied at the first box gives $(a-k, 0, P_{k+1}, 0) \to (a-k-1, P_{k+1}, 0, 0)$, completing our induction. ∎

We now describe the construction for the original 6-box setting. Write $A = 2010^{2010^{2010}}$. First, apply a Type 1 operation to B_5, giving $(1, 1, 1, 1, 1, 1) \to (1, 1, 1, 1, 0, 3)$. Second, apply Type 2 operations to B_4, B_3, B_2, and B_1 in this order, obtaining $(1, 1, 1, 1, 0, 3) \to (1, 1, 1, 0, 3, 0) \to (1, 1, 0, 3, 0, 0) \to (1, 0, 3, 0, 0, 0) \to (0, 3, 0, 0, 0, 0)$. Third, apply Lemma 2 twice, giving the sequence $(0, 3, 0, 0, 0, 0) \to (0, 0, P_3, 0, 0, 0) \to (0, 0, 0, P_{16}, 0, 0)$. It is easy to check that $P_{16} > A$, so there are more than $A = 2010^{2010^{2010}}$ coins in B_4 at this point. Fourth, decrease the number of coins in B_4 by applying Type 2 operations repeatedly to B_4 until its size decreases to $\frac{A}{4}$. This gives $(0, 0, 0, P_{16}, 0, 0) \to \cdots \to \left(0, 0, 0, \frac{A}{4}, 0, 0\right)$. Finally, apply Type 1 operations repeatedly to first empty B_4 and then B_5, obtaining $\left(0,0,0,\frac{A}{4},0,0\right) \to \cdots \to \left(0,0,0,0,\frac{A}{2},0\right) \to \cdots \to (0,0,0,0,0,A)$, as desired.

This problem was proposed by Hans Zantema of Netherlands.

Note. Following a practice established last year, Fields Medalist (and IMO gold medalist) Terence Tao hosted an online project for others to collaborate in solving this problem, which he identified as the most challenging problem on the exam (`http://polymathprojects.org/2010/07/08/minipolymath2-project-imo-2010-q5/`).

6. We generalize to the setting where the a_n may assume negative values. For any $r \in \mathbb{R}$, note that the transformation $a_n \mapsto a_n + rn$ does not change the problem conditions or the result to be proved. Picking $\ell \le s$ such that a_ℓ/ℓ is maximal, we can thus assume without loss of generality that $a_\ell = 0$. This means all of $a_1, \ldots, a_s$ are non-positive, hence all a_n are non-positive. Let $b_n = -a_n \ge 0$. For $n > s$, we have $b_n = \min\{b_k + b_{n-k} \mid 1 \le k \le n-1\}$ and in particular $b_n \le b_{n-\ell} + b_\ell = b_{n-\ell}$.

   From this, we draw two conclusions. First, all b_n must be bounded above by $M = \max\{b_1, \ldots, b_s\}$. Second, if we let S be the set of all linear combinations of the form $c_1b_1 + c_2b_2 + \cdots c_sb_s$, where the c_i are nonnegative integers, and let $T = \{x \le M : x \in S\}$, then since $b_n = \min\{b_k + b_{n-k} \mid 1 \le k \le n-1\}$, it is clear that every b_n must be in T. Crucially, T is a finite set.

   Now, for each integer i satisfying $\ell i + 1 > s$, let β_i denote the ℓ-tuple $(b_{\ell i+1}, b_{\ell i+2}, \ldots, b_{\ell i+\ell})$. By the previous paragraph, the number of such ℓ-tuples is at most $|T|^\ell$, a finite number. Further, because $b_n \le b_{n-\ell}$ for $n > s$, the individual indices of these β_i are non-increasing functions of i. Thus, there can only be finitely many i for which $\beta_i \ne \beta_{i+1}$. Let i_0 be greater than the largest such value; then, all ℓ-tuples β_i with $i \ge i_0$ are identical. Choosing $N = \ell(i_0+1)$ finishes the problem, since any $n \ge N$ gives $b_n = b_{n-\ell} = b_\ell + b_{n-\ell}$.

   This problem was proposed by Morteza Saghafian of Iran. This solution is by Evan O'Dorney.

Results. The IMO was held in Astana, Kazakhstan, on July 7–8, 2010. There were 517 competitors from 96 countries and regions. On each day contestants were given four and a half hours for three problems.

On this challenging exam, a perfect score was achieved by only one student, Zipei Nie (China). The USA team ranked third, behind China and Russia. Although the American team has consistently finished in the top ten at the IMO, this year's performance was particularly impressive because none of the team members were in their final year of high school. The students' individual results were as follows.

- Calvin Deng, who finished 9th grade at William G. Enloe High School in Raleigh, NC, won a silver medal.
- Ben Gunby, who finished 10th grade at Georgetown Day School in Washington, DC, won a gold medal.
- Xiaoyu He, who finished 10th grade at Acton-Boxborough Regional High School in Acton, MA, won a gold medal.
- In-Sung Na, who finished 11th grade at Northern Valley Regional High School in Old Tappan, NJ, won a silver medal.
- Evan O'Dorney from Danville, CA, who finished 11th grade (homeschooled through Venture School), won a gold medal. Furthermore, he placed 2nd overall with a score of 39/42. For his spectacular performance, he received a private congratulatory telephone call from the President of the United States, Barack Obama.
- Allen Yuan, who finished 11th grade at Detroit Country Day School in Beverly Hills, MI, won a silver medal.

2010 Carl B. Allendoerfer Awards

The Carl B. Allendoerfer Awards, established in 1976, are made to authors of expository articles published in *Mathematics Magazine.* The Awards are named for Carl B. Allendoerfer, a distinguished mathematician at the University of Washington and President of the Mathematical Association of America, 1959–1960.

Ezra Brown and Keith Mellinger, "Kirkman's Schoolgirls Wearing Hats and Walking Through Fields of Numbers," *Mathematics Magazine,* 82:1 (2009), pp. 3–15.

The historical basis for this interesting article is a problem in recreational mathematics posed by T. P. Kirkman in 1850. Kirkman's problem states:

> "Fifteen young ladies of a school walk out three abreast for seven days in succession: it is required to arrange them daily so that no two shall walk abreast more than once."

Using this problem as a springboard, the authors treat the reader to a captivating exploration of the theory and applications of block designs. In the process, solutions to the schoolgirls problem are uncovered in such seemingly unrelated areas as the subfield structure of algebraic number fields and the configuration of "spreads" and "packings" in finite projective geometry.

Additional connections to the schoolgirls problem are revealed by the authors' extension of Todd Ebert's "Three Hats" problem to an analogous problem involving fifteen hats. Their solution to this extension is based upon an interesting relationship between Kirkman's problem and the theory of error correcting codes. In particular, a solution to the schoolgirls problem leads to a Hamming code, which then can be used to solve the fifteen hats problem.

This well-written and accessible article invites the reader to join the authors on a fascinating journey into the modern theory of block designs and the surprising connections of these designs to diverse areas of mathematics. Readers who accept the invitation will be left with both a deeper understanding of Kirkman's problem and an appreciation of the ubiquitous nature of its solution.

Biographical Notes

Ezra (Bud) Brown grew up in New Orleans, has degrees from Rice and Louisiana State University, and has been at Virginia Tech since 1969, where he is currently Alumni Distinguished Professor of Mathematics. His research interests include number theory and combinatorics. He particularly enjoys discovering connections between apparently unrelated areas of mathematics and working with students who are engaged in research. He has been a frequent contributor to the Mathematical Association of America journals, and he just finished a term as the Maryland, District of Columbia, and Virginia Section Governor. He and Art Benjamin edited *Biscuits of Number Theory*, a collection of number theory articles published in 2009 by the Mathematical Association.

In his spare time, Bud enjoys singing (from opera to rock and roll), playing jazz piano, and talking about his granddaughter Phoebe Rose. Under the direction of his wife Jo, he has

Math. Mag. **83** (2010) 324–326. doi:10.4169/002557010X528050.

become a fairly tolerable gardener, and the two of them enjoy kayaking. He occasionally bakes biscuits for his students, and he once won a karaoke contest.

Originally from Lancaster County, Pennsylvania, **Keith E. Mellinger** graduated with his Ph.D. in mathematics from the University of Delaware and was a post-doc at the University of Illinois at Chicago before moving to the University of Mary Washington in 2003. He is currently an Associate Professor and Chair of the Department of Mathematics at the University of Mary Washington. His research interests are in discrete mathematics, usually connected to finite geometry, and he regularly mentors undergraduate researchers. In 2005 he earned a research grant from the National Security Agency, and in 2008 he won the Outstanding Young Faculty Award presented by the University of Mary Washington.

Outside of mathematics, Keith is a decent cook, an average tennis player, and an accomplished musician. He has performed in different bands over the years, usually on acoustic guitar. However, most of his time outside the office these days is spent with his two beautiful children, Gabriel and Cecilia.

Response from Ezra Brown and Keith Mellinger. It was in graduate school that each of us first learned Thomas P. Kirkman's elegant solution to his Fifteen Schoolgirls problem, Bud in Dave Roselle and Brooks Reid's graduate course in Combinatorics, and Keith in Gary Ebert's similar course. In both of us, there began a fascination with combinatorial designs, and both of us have written on the topic for *Mathematics Magazine* in the past. It happens that Kirkman's solution turns up in a variety of different settings, including algebraic number fields, finite geometries, coding theory, and the so-called fifteen hats problem in recreational mathematics. We had this wild idea of presenting our discoveries to a general audience, and so we wrote this paper together.

Actually, truth be told, we wrote the paper for each other. We will be forever grateful to Frank Farris, who pointed out that we need to stop writing for each other and start writing for others. Frank, they just don't pay you enough. We want to thank the paper's two referees, who suggested numerous improvements that greatly improved the paper's readability—and how! We also thank our families for their support and encouragement over the years. It is a great honor to receive the Allendoerfer Award: we are truly grateful for all that the Mathematical Association of America does for our community!

David Speyer and Bernd Sturmfels, "Tropical Mathematics," *Mathematics Magazine,* 92:3 (2009), pp. 163–173.

Contrary to popular opinion, mathematics is not static. New branches of mathematics are born and develop in response to impulses from within the subject and from areas of application. But it is hard to find good expository articles which capture the excitement of a new field. This article by David Speyer and Bernd Sturmfels succeeds brilliantly in making the tropical approach to mathematics attractive and accessible to a wide readership.

The adjective "tropical" was chosen by French mathematicians to honor Imre Simon, the Brazilian originator of min-plus algebra, which grew into this field. The basic object of study is the tropical semiring consisting of the real numbers $\mathbf{R}$ with a point at infinity under the operations of $\oplus$ equals minimum and $\odot$ equals plus. The arithmetic and algebra of this simple number system yield surprising connections with well-studied branches of classical mathematics. For example, tropical polynomials in n variables are precisely the piecewise-linear concave functions on $\mathbf{R}^n$ with integer coefficients.

Further developments and generalizations connect with combinatorics, algebraic geometry, and computational biology. This article offers an introduction to diverse aspects of this new subject. Each section begins with very elementary material and ends with research problems. In a short amount of space the authors manage to convey the depth of this attractive new field and its broad reach.

Biographical Notes

David Speyer has just begun serving as an associate professor of mathematics at the University of Michigan. Before that, he was a research fellow funded by the Clay Mathematics Institute. He has an A.B. and a Ph.D. in mathematics, from Harvard University and the University of California Berkeley, respectively. Speyer's research focuses on problems which combine questions of algebraic geometry and combinatorics; both the many such questions which arise in tropical geometry and those which occur in the study of classical algebraic varieties and representation theory. He blogs at `http://sbseminar.wordpress.com`.

Bernd Sturmfels received doctoral degrees in mathematics in 1987 from the University of Washington, Seattle, and the Technical University Darmstadt, Germany. After two postdoctoral years in Minneapolis and Linz, Austria, he taught at Cornell University, before joining the University of California Berkeley in 1995, where he is Professor of Mathematics, Statistics, and Computer Science. His honors include a National Young Investigator Fellowship, a Sloan Fellowship, a David and Lucile Packard Fellowship, a Clay Senior Scholarship, and an Alexander von Humboldt Senior Research Prize. Presently, he serves as Vice President of the American Mathematical Society. A leading experimentalist among mathematicians, Sturmfels has authored or edited fifteen books and 180 research articles, in the areas of combinatorics, algebraic geometry, polyhedral geometry, symbolic computation, and their applications. His current research focuses on algebraic methods in optimization, statistics, and computational biology.

Response from David Speyer. I am highly honored to receive the Carl B. Allendoerfer Award. Tropical mathematics is an exciting but frustrating field to introduce people to: the motivating computations are extremely explicit and elegant, but learning what questions to ask often requires absorbing a great deal of background and sophisticated technology. We hope that our article has helped explain what this field is about and excited our readers to consider entering it.

If it is odd to say so, it is also true that my first thanks should go to my co-author and advisor Bernd, who has driven me both to discover and to expound tropical mathematics. I also want to thank the Clay Institute for their support of my research. Finally, I must thank my fellow graduate students at the University of California Berkeley where I began writing this article and my colleagues at the Massachusetts Institute of Technology where I finished it, for listening to my many attempts to explain the tropical perspective.

Response from Bernd Sturmfels: It is a great honor for me receive the Carl B. Allendoerfer Award for the article with David Speyer on tropical mathematics. This prize means a lot to me, especially since I find myself following in the footsteps of my late advisor, Victor Klee, who received this award in 1999. Vic's passion for the combinatorics of convex polyhedra has been a great inspiration for me over the years.

Tropical mathematics is a delightful subject that challenges our basic assumptions about arithmetic and geometry and thus leads us to a deeper understanding of familiar structures we are so accustomed to. It has been a great pleasure for me to embark on the tropical journey with numerous students and postdocs. I wish to thank them for being so patient with their impatient mentor.

Please allow me also to use this opportunity to share my view that the Mathematical Association of America is doing a marvelous job in its various programs. They have always reminded me of the unity of mathematics, and the fact that the benefits of integrating research with teaching at all levels cannot be overestimated.

Pronouncing "Oresme"

(From the Editor)

In an article in this MAGAZINE (February 2010) Olympia Nicodemi gave the pronunciation hint "Orezzmay" for the name of Nicole Oresme, who lived in the Normandy region of France in the 14th Century. Was it right?

Fernando Gouvea wrote to express doubt. "Most folks I know pronounce his name something like oh-RHAYM, though of course you should ask someone from France to be sure." For such questions he suggested this useful website: `http://www.waukesha.uwc.edu/mat/kkromare/up.html`.

"Well, it's quite complicated," writes Professor Nicodemi. "The pronunciation clue 'Orezzmay' was a quick insertion of my own at the Editor's request, unresearched but not entirely unfounded. The French say 'Oreme' and spell it that way. At one time, of course, all the esses were pronounced, as in cotes-coast. One of the times of change was the 14th century, just when he lived. And whether it is French or Latin is a question. I have asked a colleague in our French Department about it. One of the ways they guess (her phrase) at pronunciation is through rhyme schemes in poetry. In Oresme's time, 'quaresme' was found rhymed with 'escame' indicating that the ess in 'esme' was not pronounced. But it was quite fluid."

Here is another relevant website: `http://hsci.cas.ou.edu/exhibits/exhibit.php?exbid=45&exbpg=1`. It is a wonderful website, rich in information about early science. At this site you will find Oresme himself telling students how to pronounce his name, if you believe him. I wrote to Kerry Magruder, Curator of the History of Sciences Collections at Oklahoma University, whose site it is. His reply:

> Professor Nicodemi's account seems exactly on target to me. It agrees with what I have always heard among historians of science; namely, that although pronunciation was going through a transition, the best guess (based on rhyming patterns, as noted) is that even though the "s" was still included, without the circumflex accent, it was not pronounced. So Oresme himself most likely would have pronounced his name, so far as can now be determined, as "Or–em."
>
> But the more I think about it the less certain I am. Maybe the question of pronunciation is unsolvable. Perhaps it is a modern question, based on the assumption that one would always spell or pronounce one's name the same way. In the transmission of Oresme's works there appears to have been some variation, at least in spelling, so that the quest for an original spelling (let alone pronunciation) may be unsolvable. Edward Grant, for instance, lists the following early versions of Oresme's spelling in French: Oresmius, Oréme, Oresmes, d'Oresme, d'Oresmieux, Orem, and Orême [**1**, p. 3, citing Meunier]. So I'm not sure where this leaves us.

He also passes along this from Steven J. Livesey, a professor of medieval science at Oklahoma:

> I suspect that the pronunciation—to the extent that it is even recoverable—is a regionalism and dependent on oral transmission of the text. So when copies of *De proportionibus* were circulating in Italy or Poland, for example, librarians may have written their own contents lists on flyleaves based on the local pronunciation, not the

Math. Mag. **83** (2010) 327–328. doi:10.4169/002557010X528069.

way the name was pronounced in Paris or Lisieux, to say nothing of how Oresme himself pronounced his name. (And my own experience with a confusing surname suggests that some people give up correcting others' pronunciation and allow for variation.)

Edward Grant comments, "I have always pronounced the name as 'orayme' (or 'oh-Rhaym') as have all other Oresme scholars that I know, but I cannot say how it was actually pronounced in the 14th century."

The "ay" part of the original hint is probably wrong. The final vowel was either silent or barely pronounced, since there is no accent (not "Oresmé"). The weight of opinion seems to favor not pronouncing the "s" either, although that is much less certain. But just as we converge on a single view, Gouvea passes along another website, `http://www.forvo.com/word/nicolas_oresme/`, where a modern speaker clearly pronounces the "s."

Can anyone provide a better answer?

REFERENCE

1. Edward Grant, ed. and trans., *Nicole Oresme: De proportionibus proportionum and Ad pauca respicientes,* Univ. of Wisconsin Press, Madison, WI, 1966.