MATHEMATICS MAGAZINE

Vol. 55 No. 2
March, 1982

BOOKS AND MATHEMATICS • POLYGONS OF VECTORS
COMPUTER QUEUES • CONTRIVED SYMMETRIC MATRICES

MATHEMATICS MAGAZINE

Volume 55, Number 2, March 1982

(ISSN—0025-570X)

COVER: "The Bookworm," from Leary's Old Book Store, Philadelphia.

ARTICLES

67 Queueing Network Models in Computer System Design, *by Robert Geist and Kishor Trivedi.*

81 Mathematics and Literature, *by D. O. Koehler.*

NOTES

96 Tristram's Mathematics Problem, *by Richard Moore.*

97 Proofs Without Words: Count the Dots, *by Warren Page.*

98 On the Value of Mathematics (Books), *by G. L. Alexanderson and L. F. Klosinski.*

104 On a Family of Polygons, *by Paul R. Scott.*

106 Symmetric Matrices with Prescribed Eigenvalues and Eigenvectors, *by Konrad J. Heuvers.*

111 A Number-Theoretic Sum, *by John D. Baum.*

PROBLEMS

114 Proposals Number 1140-1143.

115 Quickie Number 671.

115 Solutions to Problems 1114, 1115.

116 Answer to Quickie Number 671.

REVIEWS

117 Reviews of recent books and expository articles.

NEWS AND LETTERS

120 Comments on recent issues; news; 1981 William Lowell Putnam Examination.

EDITORIAL POLICY

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

New manuscripts should be sent to: Doris Schattschneider, Editor, Mathematics Magazine, Moravian College, Bethlehem, Pennsylvania 18018. Manuscripts should be prepared in a style consistent with the format of *Mathematics Magazine*. They should be typewritten and double spaced on $8\frac{1}{2}$ by 11 paper. Authors should submit the original and one copy and keep one copy as protection against possible loss. Illustrations should be carefully prepared on separate sheets of paper in black ink, the original without lettering and two copies with lettering added.

Authors planning to submit manuscripts should read the full statement of editorial policy which appears in the News and Letters section of this *Magazine*, Vol. 54, pp. 44-45. Additional copies of the policy are available from the Editor.

AUTHORS

Robert Geist and **Kishor Trivedi** ("Queueing Network Models in Computer System Design") are both on the faculty of computer science at Duke University. Geist received an M.A. in computer science from Duke and a Ph.D. in mathematics from the University of Notre Dame, and taught mathematics at Pembroke State University prior to his appointment at Duke. He has published in both computer science (performance evaluation) and algebraic topology. Trivedi, who is an associate professor of computer science and electrical engineering, received a Ph.D. in computer science from the University of Illinois. He has published in the areas of computer arithmetic, computer architecture, memory management, and performance evaluation.

Don Koehler ("Mathematics and Literature") is a professor of mathematics at Miami University in Oxford, Ohio. He received his Ph.D. from Indiana University and has taught at Hanover College, DePauw University, and the University of Newcastle in Australia. He has also served as a faculty intern at the J. M. Smucker Company in Orrville, Ohio, and as a Research Associate at Wright Patterson Air Force Base. To help enable students to relate mathematics to other disciplines, he has taught courses relating mathematics to biology, to music, and to literature. His article is an outgrowth of his experiences in the mathematics and literature course.

ILLUSTRATIONS

Candy Baker sketched the scene on p. 88.
The editor provided the illustrations on pp. 84, 91.
All other illustrations were provided by the authors.

BUSINESS INFORMATION. Mathematics Magazine is published by the Mathematical Association of America at Washington, D.C., five times a year in January, March, May, September, and November. The annual subscription price for Mathematics Magazine to an individual member of the Association is $8, included as part of the annual dues of $40. Students receive a 50% discount. Bulk subscriptions (5 or more copies to a single address) are available to colleges and universities for distribution to undergraduate students at a 35% discount. The library subscription price is $25.

Subscription correspondence and notice of change of address should be sent to A. B. Willcox, Executive Director, Mathematical Association of America, 1529 Eighteenth Street, N. W., Washington, D.C. 20036. Back issues may be purchased, when in print, from P. and H. Bliss Co., Middletown, Connecticut 06457.

Advertising correspondence should be addressed to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, New York 14214.

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

ARTICLES

Queueing Network Models in Computer System Design

Computers also serve who only stand and wait.

ROBERT GEIST
KISHOR TRIVEDI
Duke University
Durham, NC 27706

A professional typist at a computer terminal can type information into the machine at a rate of 120 words/minute. A student using the hunt-and-peck technique on the keyboard might manage 25 words/minute. Punched cards of information can be read by a machine at a rate of 300/minute, yet even this card reader is a very slow device. In the time it takes this high-priced unit to read a card, a modestly-priced central processing unit (CPU) can execute 100,000 addition operations or, more to the point, completely finish all computations for several average computer programs. Similar remarks apply to line printers and other input/output (I/O) devices such as magnetic drums, disks, and tape drives.

For this reason, modern computer systems incorporate both *autonomous peripheral devices*, that is, input/output devices capable of operating concurrently with the CPU, and *multiprogramming*, a mode of operation in which several programs are allowed to reside in the main memory of the computer at once, so that when one program requests an I/O operation, the CPU can be switched immediately to another waiting program. Although this concurrent operation of CPU and I/O devices would seem to make the most efficient use of the equipment, serious problems can and do arise. As more jobs compete for the available resources, congestion and job queues develop, and performance (in the eyes of the individual user) suffers. Thus those who design and evaluate such systems ask:

1. *What limit should we place on the number of jobs allowed to reside in the computer?*
2. *To improve performance, should we replace an existing device with a larger one, replace it with a faster one, add a device, remove a device?*

Some answers to such questions have been provided through application of **queueing network models**. We will present some of the basic results and techniques in this area of current research.

§1. Parameters and performance measures

At the most basic level, we can regard a component of a computing system (CPU, drum, disk, card reader, line printer, etc.) as a service center and the jobs in the system as customers which arrive at the component, eventually receive service, and depart. If, upon arrival, a job finds the

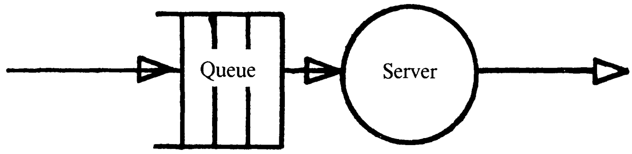


FIGURE 1

server busy, then it must join a queue associated with this component and await selection for service; thus our systems will be made up of nodes of the form shown in FIGURE 1.

The queueing network of FIGURE 2 may be taken as a model for a typical interactive computing system. In this network, requests for computation are generated from M user terminals and must immediately queue to await a slice of main memory. The maximum number of such slices available (the *maximum degree of multiprogramming*) is set by the system operator and can be a key factor in system performance. Once a slice of memory has been allocated to it, the job must await a turn at the central processor, after which it either returns to the user, with some probability q_0, or makes an I/O request of device i, with some probability q_i, and returns to the CPU queue to await another burst of computation.

Such a network generates several important questions. Given values for system parameters such as the number of user terminals, the maximum degree of multiprogramming, the device speeds, and the user demands, what **throughput** (rate of flow along the return path to the user terminals) should we expect? How long should a user expect to wait for the system to respond? How many requests might we expect to find queued at each device? To answer these questions we will need to be more specific about the parameters and performance measures. We focus, for the moment, on a single node (FIGURE 1) in isolation.

The nature of activity at a node is specified by three model parameters: the *arrival pattern*, the *service pattern*, and the *scheduling discipline*.

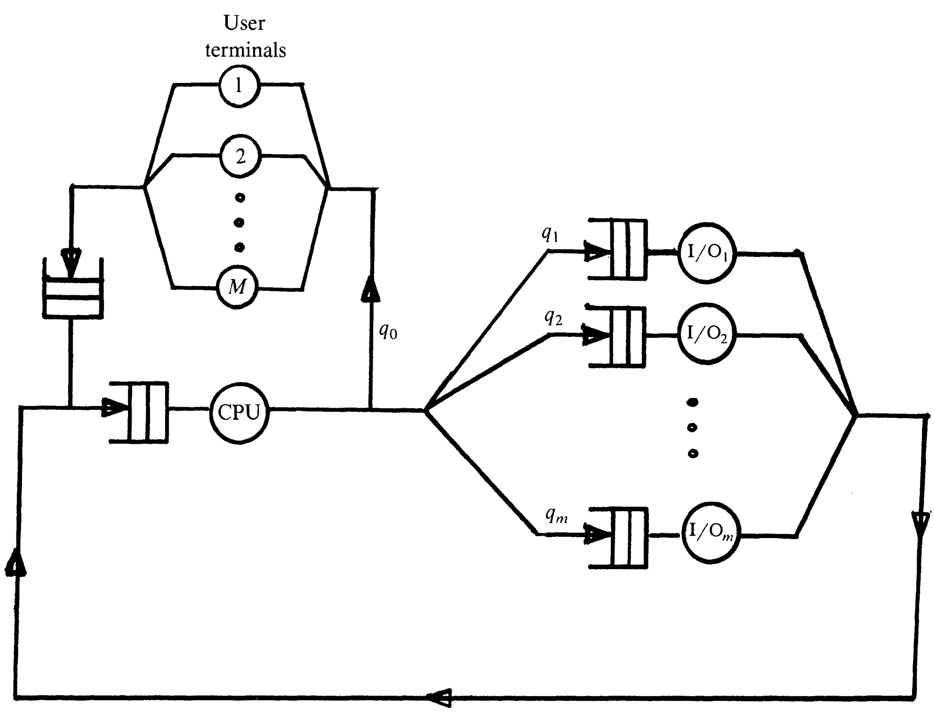


FIGURE 2

A stochastic model for an external arrival pattern is the most difficult specification of the three, because the short-term nature of workload demands upon most large systems is highly unpredictable. Thus we simply postulate an independent random arrival process. Suppose we observe a system for a large time interval T and find an average arrival rate of λ jobs/second over this interval. We then postulate that in any sufficiently small subinterval of length $\Delta = T/n$ we will either have one arrival, with probability $\lambda\Delta$, or no arrivals, and that, in either event, what happens here is independent of what happens in any other nonoverlapping interval of length Δ.

These assumptions lead us to the classical Poisson process [**19**], [**27**]; that is, if $N(T)$ denotes the number of arrivals during the time interval $[0, T]$ then

$$P(N(T)=i)=\frac{(\lambda T)^i}{i!}e^{-\lambda T}.$$

Since the mean and variance of the random variable $N(T)$ are both λT, the random variable $N(T)/T$ has mean λ and variance λ/T, so $\lim_{T\to\infty} N(T)/T=\lambda$, which justifies the notion of "long-term arrival rate λ."

The Poisson arrival assumption is standard in queueing network analysis, and empirical studies [**16**] have shown this model to agree well with observed data. Nevertheless, extensions to general renewal processes have been studied [**12**], [**33**].

Consider next the specification of a distribution for the service time, X, required by a user job at our single node (device). Although techniques exist which allow consideration of any differentiable distribution (see [**21**], [**28**], [**37**]), we shall restrict ourselves to the exponential distribution, that is

$$P(X \leq t) = 1 - e^{-\mu t} \qquad (t \geq 0). \tag{1}$$

Since the mean of this distribution is $1/\mu$, the parameter μ is often termed the **service rate**. This choice is made for several reasons. First, it agrees fairly well with observed data [**38**]. Second, of all nontrivial distributions, (1) is analytically the most tractable, due to the "memoryless" property. This means that if we know that a job has received s time units of service, then the distribution of remaining service time, $X-s$, is

$$P(X-s \leq t \mid X > s) = \frac{P(0 < X-s < t)}{P(X \geq s)} = 1 - e^{-\mu t},$$

which is identical to the distribution of X. Thus, in classifying the state of a node, the amount of service received by the job in process is irrelevant, and we need only know the *number* of customers at that node. Finally, the most common method for extending the single-node case to general service time distributions involves decomposing the general distribution into a series-parallel cascade of exponential stages and then applying the results obtained from the exponential case [**37**].

We shall not attempt to discuss the variety of scheduling disciplines, a topic deserving separate consideration [**22**]. Most of our results will apply to a large class of disciplines known as work-conserving, of which first-come, first-served is the most well known. A scheduling discipline is said to be **work-conserving** [**23**] if:

(i) the service demand of each customer is independent of the discipline;
(ii) the discipline does not allow customers to wait when the server is idle.

In addition to **first-come, first-served** (FCFS), there are other well-known work-conserving disciplines. In a **round-robin** (RR) schedule, each job in turn receives service for a small time slice. If the job is not completed by the end of this time slice, it is placed at the rear of the queue, eventually to be allotted another time slice. This is perhaps the most common CPU scheduling discipline in large time-sharing systems, but the use of such a discipline for an I/O device, such as a line printer, would seem to be sheer folly. One could easily imagine a group of users arguing about which sentences, or even words, on a particular page of output belonged to whom!

Nevertheless, it has been tried [3], because preemptive disciplines such as this alleviate certain other difficulties, in particular that known as operating system deadlock (see [**10**]). A **processor-sharing** (PS) schedule is the limiting case of RR as the specified time slice approaches 0. All jobs at a node are considered to be moving through the server at the same rate, and hence there is no queue. Though PS does not represent an actual scheduling discipline, it can be extremely useful as an analytically tractable approximation to RR. A **last-come, first-served, preemptive resume** (LCFSPR) is a push-down-stack procedure in which a job receives service immediately upon arrival. If one job happens to be in progress when another arrives, the job in progress is preempted and pushed onto a stack of waiting jobs, much like a stack of plates in a cafeteria. Upon completion of a job, the top job (plate) on the waiting stack is returned to service.

The whole point of the analysis of a system is to measure '*performance*,' and for a given installation one or several different measures may be used, depending on the user. We shall concentrate on those most commonly used in system design and tuning, namely, the equilibrium values for utilization, throughput, waiting time, and queue length. By **utilization**, we mean the fraction of time that a device is busy, and by **system throughput**, we mean the average number of jobs completed per unit of time.

If we consider a single node (device) with infinite customer population, then computations of utilization and throughput are easy. Given an arrival rate of λ and a service rate of μ, we see that for any sufficiently long time interval T, the utilization of that device is given by

$$\rho=\min\left\{\frac{\lambda T/\mu}{T},1\right\}=\min\{\lambda/\mu,1\}.$$

For this case, the throughput is $\rho\mu=\min\{\lambda,\mu\}$. For infinite customer population systems in general, eventually flow in equals flow out (λ), unless the system is swamped ($\lambda>\mu$). For the more realistic finite customer population systems (FIGURE 2), computation of node and system utilization and throughput can be exceedingly difficult.

From a customer's point of view, the most important measure is the **average waiting time** $\overline{W}$ for each job, which is related to the average queue length, $\overline{Q}$. The actual relationship, which holds under completely general conditions, is surprisingly simple:

$$\overline{Q}=\lambda\overline{W}, \tag{2}$$

where λ is the mean arrival rate [**26**]. The proof, on the other hand, is nontrivial, and the result is extremely useful.

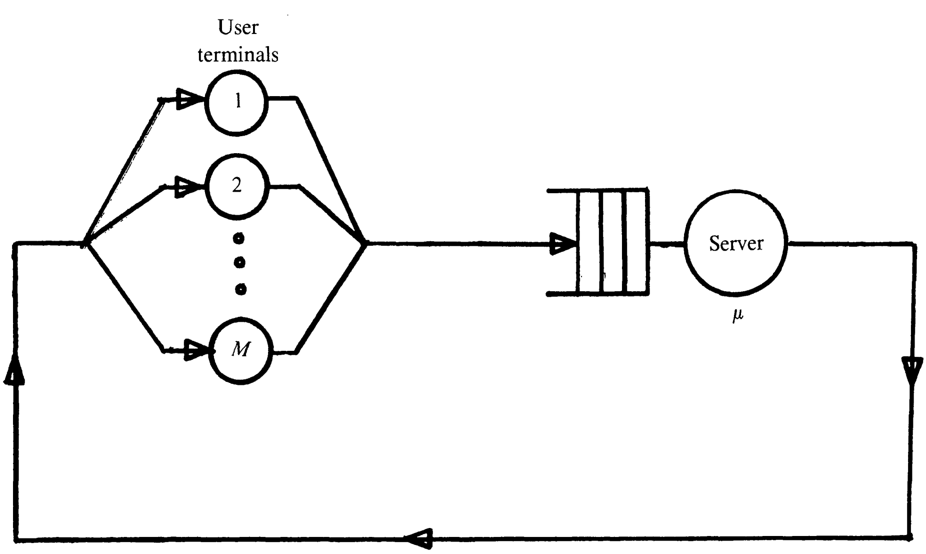


FIGURE 3

§2. Single-node systems

Most queueing network models (e.g., FIGURE 2) do not yield elegant answers to the questions we have posed, and we must either face a formidable analytic problem or resort to approximation techniques. Most work has been directed toward the latter approach, and we shall follow this course. (For partial results on the former, see [**24**].)

Consider the single-node system of FIGURE 3, which might be termed an "outer-level" approximation to that of FIGURE 2. At any given time t, there are $N(t)$ customers waiting to be served by the single device. It is usually the case that at time $t+h$, the number of customers $N(t+h)$ in line is directly influenced by both $N(t)$ and h. Often such a single-node system can be analyzed as a birth-death process. A stochastic process $N(t)$ with values in the nonnegative integers is called a **birth-death process** (see [**29**], [**27**]) if the transition probabilities $\lambda_{ij}(t,t+h)=P[N(t+h)=j|N(t)=i]$ satisfy:

$$\text{(i)}\quad \lambda_{i,i+1}(t,t+h)=\lambda_i h+o(h),$$

$$\text{(ii)}\quad \lambda_{i,i-1}(t,t+h)=\mu_i h+o(h),$$

$$\text{(iii)}\quad \lambda_{i,i}(t,t+h)=1-(\lambda_i h+\mu_i h)+o(h),$$

$$\text{(iv)}\quad \lambda_{ij}(t,t+h)=o(h)\qquad |i-j|>1.$$

(Recall that a function $f(h)$ is called $o(h)$ if $\lim_{h\to 0}f(h)/h=0$.) In particular, transitions are "nearest neighbor only" and independent of time t. If we let $P_i(t)=P[N(t)=i]$, then

$$P_i(t+h)=P_i(t)[1-(\lambda_i+\mu_i)h+o(h)]+P_{i-1}(t)[\lambda_{i-1}h+o(h)]+P_{i+1}(t)[\mu_{i+1}h+o(h)],$$

and hence

$$P_i'(t)=-(\lambda_i+\mu_i)P_i(t)+\lambda_{i-1}P_{i-1}(t)+\mu_{i+1}P_{i+1}(t),\qquad i\geqslant 1;$$

$$P_0'(t)=-\lambda_0P_0(t)+\mu_1P_1(t).$$

We postulate a constant "equilibrium" solution to these equations given by $p_i=\lim_{t\to\infty}P_i(t)$ and $\lim_{t\to\infty}P_i'(t)=0$, so

$$\lambda_i p_i-\mu_{i+1}p_{i+1}=\lambda_{i-1}p_{i-1}-\mu_i p_i,\qquad i\geqslant 1\quad\text{and}\quad \lambda_0p_0-\mu_1p_1=0.$$

From these equations we find recursively $p_{i+1}=(\lambda_i/\mu_{i+1})p_i$, $i\geqslant 0$, so

$$p_{i+1}=\frac{\prod_{k=0}^{i}\lambda_k}{\prod_{k=1}^{i+1}\mu_k}p_0,\qquad i\geqslant 0. \tag{3}$$

Finally, since $\Sigma_{i=0}^{\infty}p_i=1$, we must have

$$p_0=\left(1+\sum_{i=0}^{\infty}\prod_{k=0}^{i}\lambda_k/\mu_{k+1}\right)^{-1}. \tag{4}$$

We can apply these results to analyze the system of FIGURE 3. Using the generating function transform for discrete random variables [**23**], one can easily show that the superposition of n independent Poisson arrival processes with respective arrival rates $\lambda_1,\lambda_2,\ldots,\lambda_n$ yields a Poisson arrival process with rate

$$\lambda=\sum_{i=1}^{n}\lambda_i$$

(see FIGURE 4). Suppose $N(t)$ is the number of customers at the service node at time t. If we assume independent requests for service at rate λ and an exponential distribution of service time

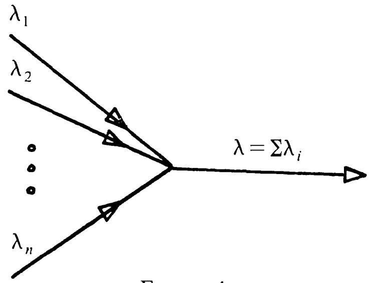


FIGURE 4

at service rate μ, then

$$P(\text{remaining service time} \leq h) = 1 - e^{-\mu h} = \mu h + o(h).$$

This implies that we have a birth-death process with

$$\lambda_i = \begin{cases} (M-i)\lambda, & 0 \leq i \leq M, \\ 0, & i > M, \end{cases}$$

and $\mu_i = \mu$, $i \geq 1$, where M denotes the number of terminals. But then equations (3) and (4) show that

$$\begin{aligned} p_i &= \rho^i \frac{M!}{(M-i)!} p_0, \qquad 1 \leq i \leq M, \\ p_0 &= 1 \Big/ \sum_{i=0}^{M} \frac{M!}{(M-i)!} \rho^i, \end{aligned} \tag{5}$$

where $\rho = \lambda/\mu$. In this model with finite customer population, system throughput is no longer trivially determined by the external arrival rate. Instead, we now measure CPU utilization (probability that the processor is not idle) as $1 - p_0$. From (5), we have

$$1 - p_0 = \frac{\sum_{i=1}^{M} \frac{M!}{(M-i)!} \rho^i}{\sum_{i=0}^{M} \frac{M!}{(M-i)!} \rho^i} = \frac{F_X(M-1; \rho^{-1})}{F_X(M; \rho^{-1})},$$

where $F_X(x; \rho^{-1})$ is the cumulative distribution function of a Poisson random variable with parameter ρ^{-1} [**19**]. Thus the system throughput is

$$\frac{\mu F_X(M-1; \rho^{-1})}{F_X(M; \rho^{-1})}. \tag{6}$$

We now consider the terminals and CPU together as a single node. In this context, equation (2), which gives Little's formula, can be reinterpreted. The queue length $\bar{Q}$ is replaced by M, the number of terminals to be served, the arrival rate λ is replaced by the system throughput given in (6), and the waiting time $\bar{W}$ is replaced by the sum of the average request service time, $\bar{R}$, and the average terminal user's "think time," $1/\lambda$. Thus (2) becomes

$$M = \frac{\mu F_X(M-1; \rho^{-1})}{F_X(M; \rho^{-1})} (\bar{R} + 1/\lambda). \tag{7}$$

The average request service time $\bar{R}$ is an important performance measure, and using (7) we can solve for $\bar{R}$:

$$\bar{R} = \frac{M F_X(M; \rho^{-1})}{\mu F_X(M-1; \rho^{-1})} - \frac{1}{\lambda}. \tag{8}$$

§3. Closed networks

The network of FIGURE 3 was an "outer-level" approximation to that of FIGURE 2 in that the CPU-I/O subsystem was assumed to be a *single* server with constant service rate μ. We now consider the other extreme, an "inner-level" approximation to the system, for which we postulate a constant outer-level system load. We assume that when one job leaves the system another immediately enters, so the number n of jobs circulating in the system (the degree of multiprogramming) *remains fixed*. An appropriate model is the closed system shown in FIGURE 5, in which a completed job immediately reenters along the "new program path." Each of the n jobs circulating in the system is active and has been allocated a partition of main memory.

Consider then an arbitrary closed system of $m+1$ service nodes, each having an exponential distribution of service time, with service rates μ_i, $i=0,1,\dots,m$. In this system, a job, upon completion of service at node i, requests service at node j with probability q_{ij}. Then the matrix $Q=[q_{ij}]$ is the transition matrix of a Markov chain, and thus, under reasonable requirements (that all states are ergodic [**15**]), Q will have 1 as an eigenvalue, associated to a one-dimensional eigenspace. We take $\lambda=(\lambda_0,\lambda_1,\dots,\lambda_m)$ as an arbitrary nonzero vector in this eigenspace. A state of the system is denoted by a vector $(n_0,n_1,\dots,n_m)$ where n_i equals the number of customers at node i, and $p(n_0,n_1,\dots,n_m)$ denotes the probability of the system being in that state. The steady-state balance equation (flow out = flow in) is then

$$\sum_{\{j|n_j>0\}} \mu_j p(n_0,n_1,\dots,n_m)=\sum_{\{j|n_j>0\}} \sum_i \mu_i q_{ij} p(n_0,\dots,n_i+1,\dots,n_j-1,\dots,n_m). \tag{9}$$

If $\rho_i=\lambda_i/\mu_i$ is the "relative utilization" of node i, and

$$C(n)=\sum \prod_{i=0}^{m} \rho_i^{n_i}, \tag{10}$$

the sum taken over all $(n_0,\dots,n_m)$ with $\sum n_i=n$, ($C(n)$ is called the **normalization constant**), then it can be shown that the solution to (9) can be written in the product form

$$p(n_0,n_1,\dots,n_m)=\frac{1}{C(n)}\prod_{i=0}^{m}\rho_i^{n_i}. \tag{11}$$

This result is due to Williams and Bhandiwad [**39**], who extended the work of Gordon and Newell [**18**].

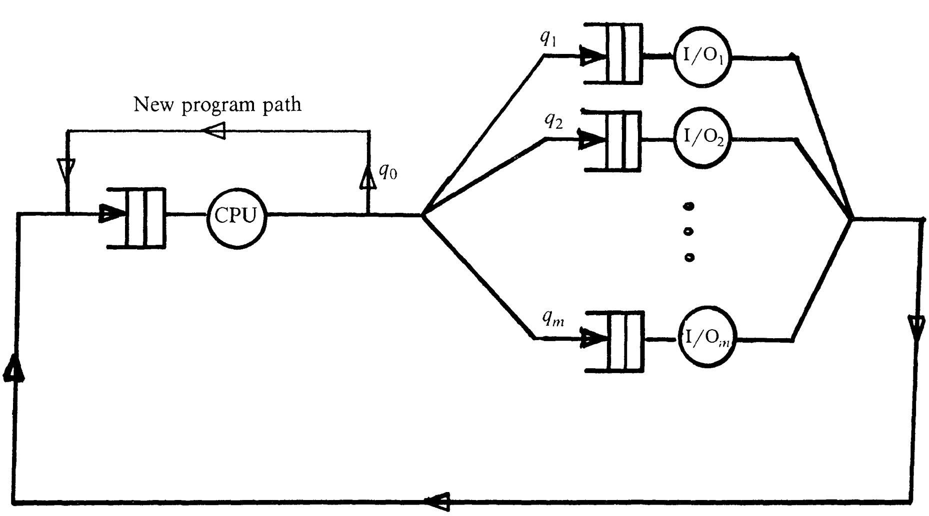


FIGURE 5

Most of the useful system measurements depend heavily on the normalization constant $C(n)$. As an example, we compute the utilization of the jth node, which we denote $U_j(n)$; $U_j(n)$ is just $P(n_j \geq 1)$. Following Williams and Bhandiwad [**39**], we let

$$X_i(z) = 1 + \rho_i z + \rho_i^2 z^2 + \cdots = 1/(1-\rho_i z),$$

and let $G(z) = \prod_{i=0}^{m} X_i(z)$, so $G(z)$ is the generating function for the normalization constants, i.e., $G(z) = 1 + C(1)z + C(2)z^2 + \cdots$. Now let

$$h_j(z) = G(z)\left[X_j(z) - 1\right]/X_j(z) = G(z)\left[1 - 1/X_j(z)\right],$$

and observe that the coefficient $H_j(n)$ of z^n in $h_j(z)$ is the sum of the terms $\prod_{i=0}^{m} \rho_i^{n_i}$ for which $n_j \geq 1$. Thus, $H_j(n)/C(n) = U_j(n)$. But

$$h_j(z) = G(z)\left[1 - (1-\rho_j z)\right] = G(z)\rho_j z, \quad \text{so } H_j(n) = \rho_j C(n-1)$$

and

$$U_j(n) = \rho_j C(n-1)/C(n). \tag{12}$$

By a similar argument,

$$P(n_j \geq k) = \rho_j^k C(n-k)/C(n). \tag{13}$$

By definition, the throughput of the ith node is $\mu_i U_i(n)$, which by (12) is $\lambda_i C(n-1)/C(n)$. Using (13), the average number $\overline{N}_i(n)$ of customers at node i can be calculated:

$$\overline{N}_i(n) = \sum_{j=0}^{n} jP(n_i = j) = \sum_{j=1}^{n} P(n_i \geq j) = \sum_{j=1}^{n} \frac{\rho_i^j C(n-j)}{C(n)}. \tag{14}$$

Equations (12), (13), and (14) show that most system performance measures are obtained quickly from the normalization constants $C(n)$. However, for systems of any size, computation of the $C(n)$'s from definition (10) is always too expensive and often numerically unstable. Fortunately, there is an efficient recursive algorithm for this computation. Using (14), the total number n of customers in the system can be written

$$n = \sum_{i=0}^{m} \overline{N}_i(n) = \sum_{i=0}^{m} \sum_{j=1}^{n} \frac{\rho_i^j C(n-j)}{C(n)},$$

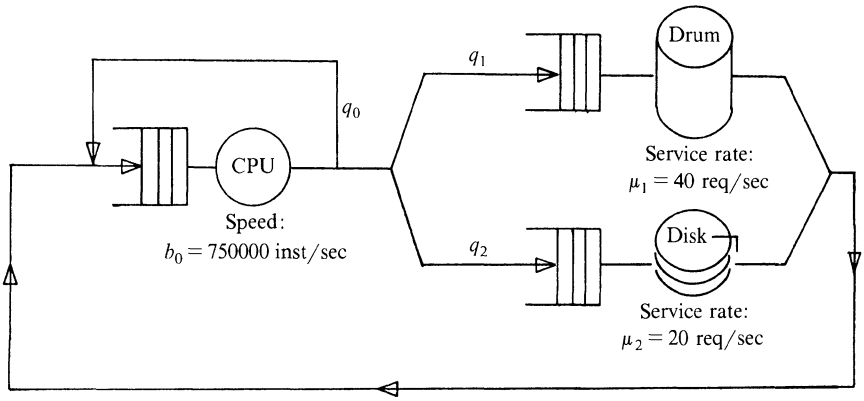
and thus

$$C(n) = \frac{1}{n} \sum_{i=0}^{m} \sum_{j=1}^{n} \rho_i^j C(n-j), \tag{15}$$

where $C(0) = 1$. Other methods of computing the normalization constants can be found in [**9**]. Another approach to computation of system measures which avoids the $C(n)$'s can be found in [**31**].

EXAMPLE. We illustrate the application of some of our results. Suppose we have two I/O devices ($m = 2$), a drum and a disk, with mean request service times of $1/\mu_1 = .025$ and $1/\mu_2 = .05$ seconds, respectively. Assume that the CPU can execute $b_0 = 750{,}000$ instructions/second, and that the average program (job) requires $J_0 = 50{,}000$ instructions for its completion (so the job service time J_0/b_0 would be $1/15$ second/job, if there were no interruptions for I/O). However, we assume there are $I_1 = 5{,}000$ instructions between successive drum requests and $I_2 = 10{,}000$ instructions between successive disk requests, so the average job will be frequently interrupted. The situation is illustrated in FIGURE 6A.

We can evaluate system throughput (i.e., flow along the new program path) as follows. The



Instructions Executed:	
Per Job	$J_0 = 50000$
Between Successive Drum Requests	$I_1 = 5000$
Between Successive Disk Requests	$I_2 = 10000$

FIGURE 6A

transition matrix Q is of the form

$$Q = \begin{bmatrix} q_0 & q_1 & q_2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

so an eigenvector corresponding to eigenvalue 1 has the form $\lambda = (\lambda_0, \lambda_0 q_1, \lambda_0 q_2)$. The average CPU service time is $1/\mu_0 = J_0/(b_0 \times \text{number of trips to CPU}) = J_0 q_0 / b_0$, so our relative utilizations are

$$\rho_0 = \frac{\lambda_0}{\mu_0} = \frac{\lambda_0 J_0 q_0}{b_0},$$

$$\rho_1 = \frac{\lambda_1}{\mu_1} = \frac{\lambda_0 q_1}{\mu_1},$$

$$\rho_2 = \frac{\lambda_2}{\mu_2} = \frac{\lambda_0 q_2}{\mu_2}.$$

Further, for $i > 0$, q_i is given by the quotient (number of trips to device i)/(number of trips to CPU) $= (J_0/I_i)/(1/q_0) = q_0 J_0 / I_i$. For convenience, we can choose $\lambda_0 = 1/q_0$. Then

$$\rho_0 = J_0/b_0$$
$$\rho_1 = J_0/(\mu_1 I_1)$$
$$\rho_2 = J_0/(\mu_2 I_2)$$

and the system throughput is

$$T(n) = U_0(n)\mu_0 q_0 = \rho_0 \frac{C(n-1)}{C(n)} \mu_0 q_0 = \frac{C(n-1)}{C(n)}.$$

Note that the choice of λ_0 as $1/q_0$ eliminates the variables μ_0 and q_0 from the equation for $T(n)$. Thus by (15), $T(n)$ depends only on the ρ_i's, and can be calculated when branching probabilities and service rates are not available. In FIGURE 6B, we plot $T(n)$ for $n = 1$ to 10. Note that the system throughput increases monotonically with the degree of multiprogramming, a result which is realistic only if the increase in multiprogramming is effected through the purchase of additional main memory.

More often multiprogramming is implemented through a paged *virtual memory* system: both programs and main memory are partitioned into fixed equal size segments called *pages* and *page frames*, respectively. Programs are then allowed to execute with only a fraction of their total pages

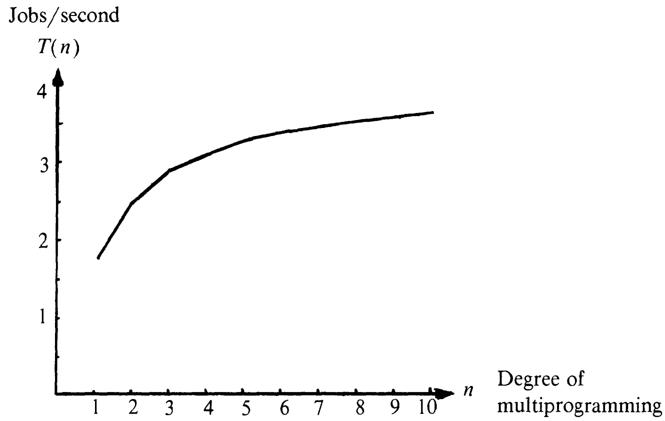


FIGURE 6B

resident in main memory frames. During execution, if a reference is made to a page which is not resident (*page fault*), then a transfer request must be made to the paging device (usually a drum) where the desired page does reside.

In such systems, an increase in multiprogramming implies a reduced allocation of main memory frames per program and hence an increase in the number of page faults and subsequent paging (drum) activity. This increased activity tends to degrade system throughput, in direct conflict with the beneficial effects of multiprogramming already observed.

We can illustrate this with a slight modification of our previous example. Assume the drum is a paging device so that I_1 is no longer a constant .005 million, but rather an exponentially decreasing function of n, say $I_1(n)=0.1e^{-.415n}$. Note that $I_1(8)<.005<I_1(7)$. We plot $T(n)$ for $n=1$ to 10 in FIGURE 7. After reaching a peak throughput of 3.95 jobs/second at degree of multiprogramming $n=5$, the system experiences a severe degradation in performance (due to excessive paging) called *thrashing*.

Since thrashing is a real problem in virtual memory systems currently in use, it has been studied extensively, with both hardware and software solutions proposed [**14**]. One such solution, an increase in the speed of the paging device, can be illustrated with our present example by varying μ_1 from 10 to 100. The result, throughput as a function of degree of multiprogramming n and drum service rate μ_1, is shown in FIGURE 8. Note that thrashing is not eliminated by increasing μ_1, but its onset is forestalled, and its severity is considerably lessened.

The Gordon-Newell result has been extended to closed networks involving multiple servers, multiple job classes, and general service time distributions (see [**28**], [**2**], and [**8**]). For analogous results on open networks, see [**4**] and [**20**].

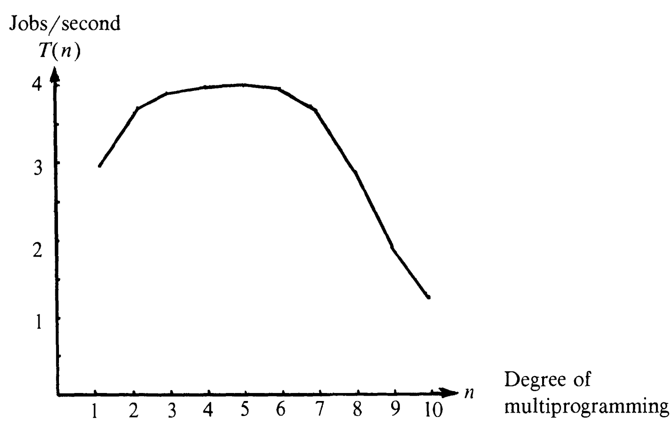


FIGURE 7

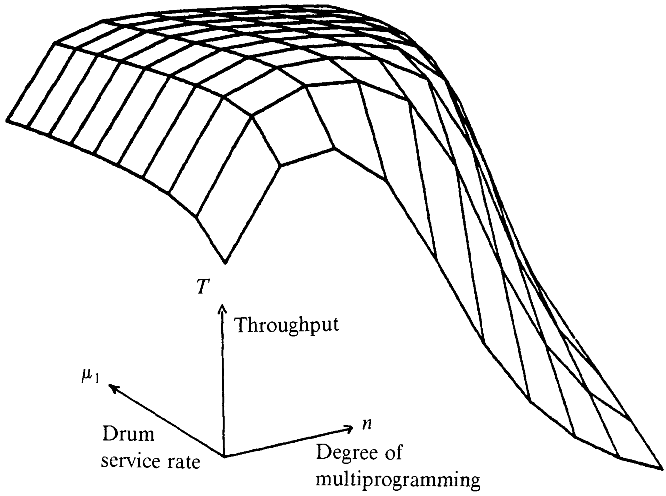


FIGURE 8

§4. Norton's Theorem Reduction

Often the rate of interaction within the CPU-I/O subsystem is orders of magnitude larger than any external arrival rate, and therefore the CPU-I/O subsystem will reach equilibrium much faster than any external job queue. Thus it is reasonable to represent the entire CPU-I/O subsystem as a *single equivalent server* (FIGURE 3) whose load dependent service rate is the system throughput of the closed, inner-level, CPU-I/O model.

Applying this to FIGURE 2, we again obtain a birth-death process with

$$\lambda_i = \begin{cases} (M-i)\lambda, & 0 \leqslant i \leqslant M, \\ 0, & i > M, \end{cases}$$

where M is the number of user terminals, but now

$$\mu_i = \begin{cases} T(i), & 1 \leqslant i \leqslant n, \\ T(n), & i > n, \end{cases}$$

where $T(i)$ is the throughput of the closed CPU-I/O subsystem of FIGURE 5 under degree of multiprogramming i, and n is the maximal degree of multiprogramming supported.

We leave it as an exercise for the reader to derive a computable expression for p_i, the probability of i customers waiting at the equivalent server. The expected throughput $\bar{T}$ of this server is then

$$\bar{T} = \sum_{i=1}^{n} T(i) p_i + \sum_{i=n+1}^{M} T(n) p_i,$$

and the expected system response time is

$$\bar{R} = (M/\bar{T}) - (1/\lambda),$$

as in section 2.

We illustrate with an example. Assume a system has three I/O devices with parameters as follows:

device	CPU	drum	fast disk	slow disk	
service rate μ_i	89.3	44.6	26.8	13.4	jobs/sec
branching probability q_i	.05	.5	.3	.15	

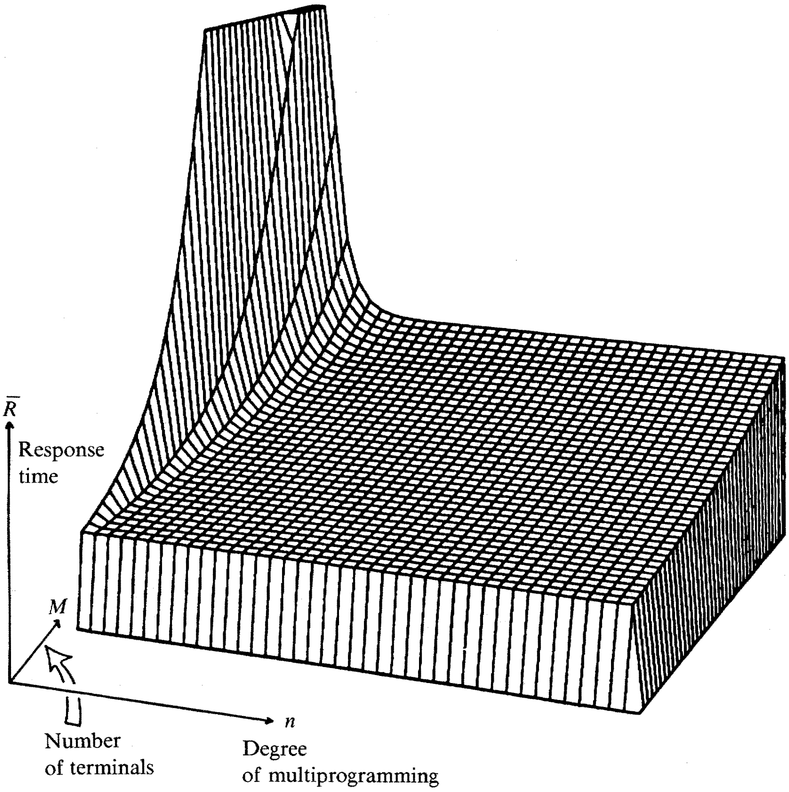


FIGURE 9

Additionally, assume an arrival rate of $\lambda = 1/15$ requests/second. In FIGURE 9 we plot $\bar{R}$ as a function of the number M of terminals (1–40) and maximal degree n of multiprogramming (1–40). Note that even for as many as 40 terminals, an increase in multiprogramming beyond about 6 or 7 is probably wasteful.

The decomposition technique presented here by way of example was developed by Chandy, Herzog, and Woo [7], and is referred to in the literature as *Norton's Theorem Reduction* (from the analogous Norton's theorem of electrical circuit theory).

There are several assumptions implicit in the application of this technique, which, if violated in the system being modeled, can introduce substantial errors. These potential sources of error are identified in [**32**], along with the methods for adjusting Norton's Theorem Reduction to account for each. One such technique is the adjustment of the coefficient of variation (σ_s/μ_s) of the equivalent server away from its default value (which equals 1 for an exponential distribution) to more accurately reflect the characteristics of either the service time distribution or the departure process.

Other decomposition and approximation techniques have been proposed, e.g., [**11**]. An excellent survey and a unified approach to improving these approximations can be found in [**34**].

§5. Decision models

Queueing network models form the foundation for the decision models used by systems analysts and designers. To make decisions effectively, one needs clear, quantitative reformulations of phrases such as, "...an increase in multiprogramming beyond about 6 or 7 is probably wasteful." To illustrate, we will discuss one such decision model [**35**].

Consider an arbitrary closed network (such as in FIGURE 5) consisting of $m+1$ service centers, each of which has an exponential distribution of service time using a first-come, first-served discipline. Service center i has speed b_i work units/time unit, and each of the n jobs in the system will, upon completion of service at center i, request service from center j with probability q_{ij}. The workload description consists of the average total number J_i of work units processed per job at the ith service facility, and the average number t_i of visits per job to the ith facility ($i=0,1,\ldots,m$).

System throughput is formulated as in section 3: the mean service time per visit to the ith node is $1/\mu_i = J_i/b_i t_i$. If $\lambda=(\lambda_0,\lambda_1,\ldots,\lambda_m)$ is a left eigenvector of the transition matrix $Q=[q_{ij}]$ and $\rho_i=\lambda_i J_i/b_i t_i$, then throughput

$$T(n)=\frac{\lambda_0}{t_0}\frac{C(n-1)}{C(n)},$$

where $C(n)=\Sigma\Pi_{i=0}^{m}\rho_i^{n_i}$, the sum over all $(m+1)$-tuples $(n_0,\ldots,n_m)$ with $\Sigma n_i=n$. Again we can simplify by choosing our eigenvector λ to be the vector of trips $(t_0,t_1,\ldots,t_m)$, since this reduces ρ_i to J_i/b_i, and we can calculate $T(n)$ without knowing either Q or $(t_0,t_1,\ldots,t_m)$.

One important decision model is the "optimal" selection of device speeds $b_0,b_1,\ldots,b_m$. Total cost is important in almost any definition of "optimal," so let us incorporate it by assuming that the cost K_i of device i is a polynomial in the device speed b_i: thus, $K_i=\Sigma_{j=1}^{s}a_{ij}b_i^j$.

Suppose we wish to define the "optimal" vector $(b_0,b_1,\ldots,b_m)$ as that which maximizes system throughput subject to a fixed cost constraint. Assuming the parameters J_i are fixed, we can regard the relative utilizations ρ_i as the decision variables and rewrite total cost K as

$$K=\sum_{i=0}^{m}\sum_{j=1}^{s}a'_{ij}\rho_i^{-j}$$

where $a'_{ij}=a_{ij}J_i^j$. Thus our problem can be expressed as a constrained, nonlinear optimization: *minimize* $C(n)/C(n-1)$ *subject to* $K\leq$ *cost constraint and* $\rho_i\geq 0$, $i=0,1,\ldots,m$.

This problem is less formidable than one might suspect since $C(n)/C(n-1)$ is a convex function of the ρ_i's [**30**], and hence we have to minimize a convex function over a convex feasible region. It is well known [**1**] that for any such function a local minimum is also a global minimum, and thus, at least in theory, we need only crank up our machines and tell them to march downhill to the solution! The real world isn't so nice. Although this problem (or its counterpart: minimize cost subject to a throughput constraint) may well be an easy mark for standard minimization routines, practical variations on this model can prove extremely elusive at the bottom (optimization) line.

Consider the file assignment problem in which a collection of often-accessed files are to reside permanently in the various storage devices of a network. We wish to decide how much X_{ji} of file j to place on device i, to optimize throughput. In this case the b_i's are fixed (as is some activity measure of each file), and the decision variables X_{ji} determine the other factors of the ρ_i's. Although it is still possible to formulate the problem as a convex function of the X_{ji}'s to be minimized over a convex feasible region, we now have a population explosion in variables and constraints!

For example, in [**17**] we consider an airline reservation system in which 42 files are to be allocated across five devices. We thus have a function of 210 variables with five capacity constraints, 42 file size constraints, and 210 positivity constraints. The estimated run time for a standard optimization routine attacking this problem on a fairly fast machine was ten hours. Should we then wish to implement a device capacity selection model by iterating our file assignment model over those capacity choices which lie within a fixed budget, we could easily exceed 10,000 hours run time! New techniques are needed to solve such problems, and the reader is encouraged to view those few which have recently appeared (such as [**17**] and [**36**]) as continuations of this paper.

We would like to thank Stephen Daniel for the three-dimensional plots used in our examples, and David Smith and the editor for many useful suggestions.

This work was supported in part under the National Science Foundation Grant Number MCS 78-22327 and the National Library of Medicine Program Project Grant Number LM-03373.

References

[1] M. Avriel, Nonlinear Programming, Prentice-Hall, 1976.

[2] F. Baskett, K. Chandy, R. Muntz, and F. Palacios, Open, closed, and mixed networks of queues with different classes of customers, J. Assoc. Comput. Mach., 22 (1975).

[3] C. Bron, Allocation of virtual store in THE multiprogrammed system, International Seminar on Operating System Techniques, Belfast, Northern Ireland, 1971.

[4] P. Burke, The output of a queueing system, Oper. Res., 4 (1956).

[5] J. Buzen, Computational algorithms for closed queueing networks with exponential servers, Comm. ACM, 16 (1973).

[6] ______, Analysis of system bottlenecks using a queueing network model, Proc. ACM-SIGOPS Workshop on Sys. Perf. Eval., Harvard, 1971.

[7] K. Chandy, U. Herzog, and L. Woo, Approximate analysis of general queueing networks, IBM J. Res. Dev., 19 (1975).

[8] K. Chandy, J. Howard, and D. Towsley, Product form and local balance in queueing networks, J. Assoc. Comput. Mach., 24 (1977).

[9] K. Chandy and C. Sauer, Computational algorithms for product-form queueing networks, Comm. ACM, 23 (1980).

[10] E. Coffmann and P. Denning, Operating System Theory, Prentice-Hall, 1973.

[11] P. Courtois, Decomposability: Queueing and Computer System Applications, Academic Press, 1977.

[12] D. Cox, Renewal theory, Metheun's monographs on applied probability and statistics, London, 1962.

[13] S. Daniel and R. Geist, VSCAN: an adaptive disk scheduling routine, Duke University, submitted to Comm. ACM.

[14] P. Denning, Virtual memory, Computing Surveys, 2(1970).

[15] W. Feller, Introduction to Probability Theory and Its Applications, vol. 1, 3rd ed., Wiley, 1968.

[16] E. Fuchs and P. Jackson, Estimates of distributions of random variables for certain computer communications traffic models, Proc. ACM Symp. on Opt. of Data Comm. Sys., 1969.

[17] R. Geist and K. Trivedi, Optimal design of multilevel storage hierarchies, IEEE Trans. Comput., to appear.

[18] W. Gordon and G. Newell, Closed queueing systems with exponential servers, Oper. Res., 15 (1967).

[19] R. Hogg and E. Tanis, Probability and Statistical Inference, Macmillan, 1977.

[20] J. Jackson, Networks of waiting lines, Oper. Res., 5 (1957).

[21] L. Kleinrock, Queueing Systems, Vol. I: Theory, Wiley, 1975.

[22] ______, Queueing Systems, Vol. II: Computer Applications, Wiley, 1976.

[23] H. Kobayashi, Modeling and Analysis, Addison-Wesley, 1978.

[24] A. Konheim and M. Reiser, A queueing model with finite waiting room and blocking, J. Assoc. Comput. Mach., 23 (1976).

[25] J. Labetoulle, G. Pujolle, and C. Soulla, Rapport de recherche no. 348, IRIA Laboria, France, 1979.

[26] J. Little, A proof of the queueing formula $L=\lambda W$, Oper. Res., 9 (1961).

[27] D. Maki and M. Thompson, Mathematical Models and Applications, Prentice-Hall, 1973.

[28] R. Muntz, Networks of queues models: applications to computer system modeling, unpublished notes.

[29] E. Parzen, Stochastic Processes, Holden Day, 1962.

[30] T. Price, Probability models of multiprogrammed computer systems, Ph.D. dissertation, Stanford, 1974.

[31] M. Reiser and S. Lavenberg, Mean-value analysis of closed multichain queueing networks, J. Assoc. Comput. Mach., 27 (1980).

[32] K. Sevcik, A. Levy, S. Tripathi, and J. Zahorjan, Improving approximations of aggregated queueing network subsystems, Computer Performance, K. M. Chandy and M. Reiser (eds.), North Holland, Amsterdam, 1977.

[33] W. Smith, Renewal theory and its ramifications, J. Roy. Statis. Soc., 20 (1958).

[34] S. Tripathi, On approximate solution techniques for queueing network models of computer systems, Ph.D. dissertation, Univ. of Toronto, 1979.

[35] K. Trivedi and R. Wagner, A decision model for closed queueing networks, IEEE Trans. Software Engrg., SE-5 (1979).

[36] K. Trivedi, R. Wagner, and T. Sigmon, Optimal selection of CPU speed, device capacities, and file assignments, J. Assoc. Comput. Mach., 27 (1980).

[37] K. Trivedi, Probability and Statistics with Reliability, Queueing, and Computer Science Applications, Prentice-Hall, 1982.

[38] E. Walter and V. Wallace, Further analysis of a computing center environment, Comm. ACM, 15 (1967).

[39] A. Williams and R. Bhandiwad, A generating function approach to queueing network analysis of multiprogrammed computers, Networks, 6 (1976).

Mathematics and Literature

Mathematics provides writers with a rich source of themes, images and metaphors.

D. O. KOEHLER
Miami University
Oxford, OH 45056

... with pencil words on your page only
Δt from the things they stand for ...

Gravity's Rainbow

When thinking of disciplines which use mathematics in a meaningful way it is easy to overlook literature. And yet the interface between mathematics and literature is a natural and interesting one, nurtured by the fact that mathematics is a language which evolves from our need to describe the world in which we live. This point is made by John Stark in his discussion of the works of Thomas Pynchon:

> By drawing nonmathematical conclusions from mathematics and by incorporating them into his literary works, he (Pynchon) points out a feature that its technological applications might easily obscure: mathematics is metaphoric because it describes universals. [**48**, p. 69]

That is, mathematics speaks to our human experience and as such provides writers with a means for creating interesting themes and images.

There are diverse types of mathematics that are used in literature as well as diverse ways in which they are used. However, writers most often draw upon areas like geometry, probability and statistics since they are most closely related to the world in which we live. A story may ridicule a mathematically-based solution to a societal problem that fails to take important issues into account. Or it may use mathematics to help characters impose meaning on events which they seem helpless to control. In the eerie world of fantasy and science fiction, mathematical metaphors heighten the intensity of the surreal surroundings. And in the strange worlds of Moebius bands and Klein bottles there exist a multitude of exotic plots and eccentric characters. The possibilities are endless. In this paper we shall explore some of the things writers have done with mathematics as well as illustrate what they might tell us about mathematics.

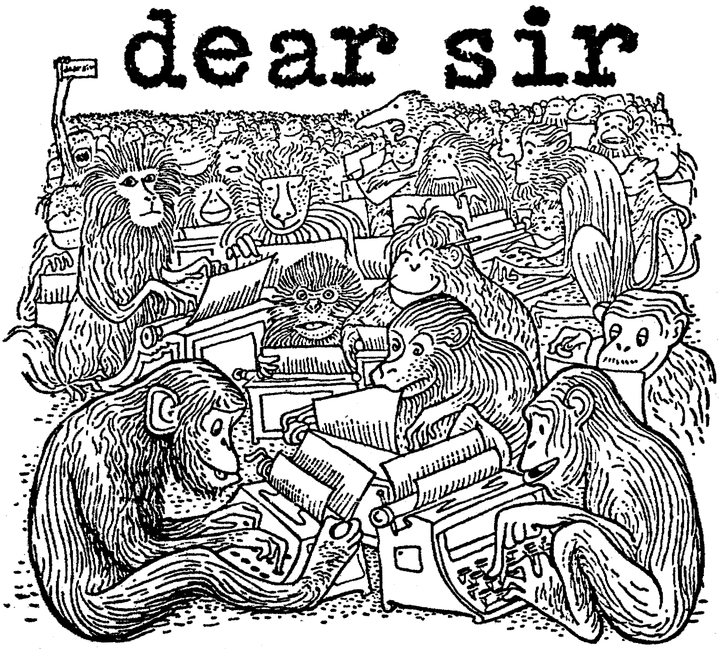
Satire—some variations on "A Modest Proposal"

Satire is often used as a form of political and social criticism in which the writer ridicules apparent stupidity, folly, or vice. There are numerous examples in which the object of this ridicule is the use of some overly-simplified mathematical law or model to solve a complex problem. Most people in literature would cite Jonathan Swift's *A Modest Proposal* [**49**] as an example of this type. In it, Swift proposed that the problem of poverty could easily be solved if poor people sold their children for food. His matter-of-fact suggestion, supported by carefully evaluated "environmental impact statements," provided a biting and eloquent indictment of the British authorities who administered the Irish people as though they were numbered objects to be manipulated according to appropriate mathematical rules.

There are numerous other examples in which the mathematics involved is more substantial. A good illustration is Ralph Schoenstein's essay "60 Million Projections Can't Be Wrong" [**41**] in which he ridicules the vast wasteland of television programming that results from the use of sophisticated viewer profiles obtained from statistical samples. He suggests the real problem is

By chance, one monkey should average one "d" a minute.

Three monkeys should achieve a "dear" every ten weeks.



10,000 of 'em will average one "dear sir" per 150 years.

The mathematician's obsession in extending a simple thought (having chimpanzees reproduce great literature by chance using typewriters) to absurd limits (calculating how long it would take for "Dear Sir" to appear) is the source of ridicule in Russell Maloney's *Inflexible Logic* [28]. This illustration is reproduced from *How to Take a Chance* by Darrell Huff, illustrated by Irving Geis, with the permission of the publisher, W. W. Norton & Company, Inc., New York, N.Y. Copyright © 1959 by W. W. Norton & Company, Inc.

that the samples are much too large and should be replaced by one consisting of a single typical family he calls the Retinas. Schoenstein describes them as follows:

> The Retinas live in a trailer camp in Dubuque, where their greatest dream is to understand subtle jokes before their neighbors do. They have two television sets, which are side by side and often play together to accommodate diverse tastes. Claude Retina is a 46.3-year-old sometime Presbyterian farmer who sells insurance, votes both Democratic and Republican as a way to fight Communism, loves LBJ but hates what he's doing, and puts his faith in God and a man's deodorant. His wife, Georgia, is a 38.4-year-old junior high school graduate who dreams of someday getting new brakes for the trailer, eats 2.7 pieces of dietetic bread a day, and sneezes 6.5 times whenever she uses the low-sudser that doesn't clog her automatic. She and Claude have 2.6 children: a boy of 9.8 who uses greasy kid stuff, a girl of 8.9 who likes round-the-clock protection, and a basset hound of three. [**41**, p. 280]

Schoenstein goes on to suggest that each day each Retina would be required to complete a questionnaire containing probing questions like "Do you remember any program you saw today?", "Give the plots of three commercials you particularly enjoyed," and "If your children had to live under educational television or Communism, which would you choose and why?" He proposes that the results of these questionnaires "be sent to the heads of networks and advertising agencies, who could interpret them any way they wished, depending on which programs they wanted to justify killing."

A different variation of this theme appears in *The Law* [**14**] by Robert Coates, which describes what happens when the law of averages becomes inoperative. The "meltdown of society" begins as a sequence of isolated variances from the expected: the George Washington Bridge is jammed with drivers out on a moonlight drive; restaurants experience unusual runs on a particular dish; a New York-to-Chicago train leaves with only three passengers aboard; and so forth. In time these perturbations become increasingly widespread, creating catastrophic problems and an inevitable investigation by Congress which is described by Coates as follows:

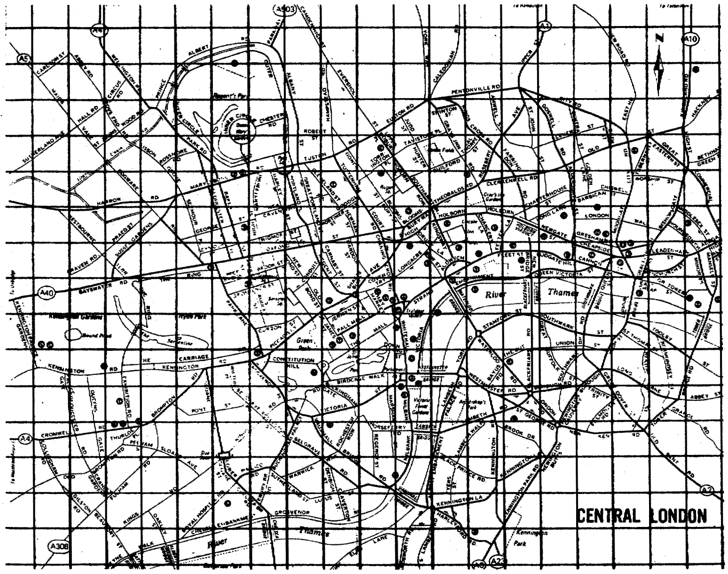
> In the course of the committee's investigations it had been discovered, to everyone's dismay, that the Law of Averages had never been incorporated into the body of federal jurisprudence, and though the upholders of States' Rights rebelled violently, the oversight was at once corrected, both by Constitutional amendment and by a law—the Hills-Slooper Act—implementing it. According to the act, people were *required* to be average, and, as the simplest way of assuring it, they were divided alphabetically and their permissible activities catalogued accordingly. Thus, by the plan, a person whose name began with "G," "N," or "U," for example, could attend the theater only on Tuesdays and he could go to baseball games only on Thursdays, whereas his visits to a haberdashery were confined to the hours between ten o'clock and noon on Mondays. [**14**, p. 19]

Those of us who recall the "odd-even" scheme of rationing gasoline several years ago realize that the "Hills-Slooper" act described by Coates is lurking closer to us than we would like.

Probabilistic imagery in "Gravity's Rainbow"

A different use of mathematics in literature occurs in Thomas Pynchon's *Gravity's Rainbow* [**38**], a complex novel set during the Second World War in which the characters search for meaning in a world dominated by the V-2 rocket. One of the major themes seems to be a debate of whether the universe is deterministic or probabilistic, and Pynchon makes considerable use of probabilistic imagery to personify this debate using a Pavlovian psychologist, Ned Pointsman, and a statistician, Roger Mexico. Mexico has discovered that the distribution of rocket hits about London has a classical Poisson distribution. However this probabilistic model troubles Pointsman.

> "Can't you ... tell from your map here, which places would be safest to go into, safest from attack?"
>
> "No."
>
> "But surely—"
>
> "Every square is just as likely to get hit again. The hits aren't clustering. Mean density is constant."



Nothing on the map to the contrary. Only a classical Poisson distribution, quietly neatly sifting among the squares exactly as it should ... growing to its predicted shape ... .

"But squares that have already had several hits, I mean—"

"I'm sorry. That's the Monte Carlo Fallacy. No matter how many have fallen inside a particular square, the odds remain the same as they always were. Each hit is independent of all the others. Bombs are not dogs. No link. No memory. No conditioning."

Nice thing to tell a Pavlovian. [**38**, pp. 55-6]

For Pointsman the world must be described in terms of cause and effect.

If ever the Antipointsman existed, Roger Mexico is the man. Not so much, the doctor admits, for the psychical research. The young statistician is devoted to number and to method, not table-rapping or wishful thinking. But in the domain of zero to one, not something to something, Pointsman can only possess the zero and the one. He cannot, like Mexico, survive anyplace in between. Like his master I. P. Pavlov before him, he imagines the cortex of the brain as a mosaic of tiny on/off elements. Some are always in bright excitation, others darkly inhibited. The contours, bright and dark, keep changing. But each point is allowed only the two states: waking or sleep. One or zero. "Summation," "transition," "irradiation," "concentration," "reciprocal induction"—all Pavlovian brain-mechanics—assume the presence of these bi-stable points. But to Mexico belongs the domain between zero and one—the middle Pointsman has excluded from his persuasion—the probabilities. A chance of 0.37 that, by the time he stops his count, a given square on his map will have suffered only one hit, 0.17 that it will suffer two ... . [**38**, p. 55]

Other characters in the book try to find themselves somewhere along the deterministic/statistical scale. One of these is Roger's girlfriend, Jessica.

Roger has tried to explain to her the V-bomb statistics: the difference between distribution, in angel's-eye view, over the map of England, and their own chances, as seen from down here. She's almost got it: nearly understands his Poisson equation, yet can't quite put the two together—put her own enforced calm day-to-day alongside the pure numbers, and keep them both in sight. Pieces keep slipping in and out.

"Why is your equation only for angels, Roger? Why can't we do something, down here? Couldn't there be an equation for us too, something to help us find a safer place?"

"Why am I surrounded," his usual understanding self today, "by statistical illiterates? There's no way, love, not as long as the mean density of strikes is constant. Pointsman doesn't even understand that." [**38**, p. 54]

FLYING-BOMB HITS ON LONDON

k	0	1	2	3	4	5 and over
N_k	229	211	93	35	7	1
$p(k; 0.9323)$	226.74	211.39	98.54	30.62	7.14	1.57

As an example of a spatial distribution of random points consider the statistics of flying-bomb hits in the south of London during World War II. The entire area is divided into $N = 576$ small areas of $t = 1/4$ square kilometers each, and the table above records the number N_k of areas with exactly k hits. (The figures are taken from R. D. Clarke, An application of the Poisson distribution, *Journal of the Institute of Actuaries*, vol. 72 (1946) p. 48.)

Pynchon seems to have created Roger Mexico as a mathematician wanting to be understood and appreciated for the insight his model provides, and yet failing to grasp that a statistical interpretation of death is no less comforting than a deterministic one.

> "The Romans," Roger and the Reverand Dr. Paul de la Nuit were drunk together one night, or the vicar was, "the ancient Roman priests laid a sieve in the road, and then waited to see which stalks of grass would come up through the holes."
>
> Roger saw the connection immediately. "I wonder," reaching for pocket after pocket, why are there never any damned—ah here, "if it would follow a Poisson ... let's see ... ."
>
> "Mexico." Leaning forward, definitely hostile. "They used the stalks that grew through the holes to cure the sick. The sieve was a very sacred item to them. What will you do with the sieve you've laid over London? How will you use the things that grow in your network of death?"
>
> "I don't follow you." It's just an equation ... .
>
> Roger really wants other people to know what he's talking about. Jessica understands that. When they don't his face often grows chalky and clouded, as behind the smudged glass of a railway carriage window as vaguely silvered barriers come down, spaces slide in to separate him that much more, thinning further his loneliness. [**38**, p. 56]

Pynchon's use of metaphors based on calculus

In the quote above, which describes Mexico's growing despair, Pynchon seems to be making use of one of several mathematical metaphors based on calculus. These metaphors mostly revolve about differentiation and integration, and, as Lance Ozier suggests [**33**], seem to represent transformations taking place in the novel. At one point Pynchon is describing what makes it possible to become a political activist.

> Franz was never much in the street. Always some excuse. Worried about security, being caught on a stray frame by one of the leather-coated photographers, who will be always at the fringes of the action. Or it was, "What'll we do with Ilse? What if there's violence?" If there's violence, what'll we do with Franz?
>
> She tried to explain to him about the level you reach, with both feet in, when you lose your fear, you lose it all, you've penetrated the moment, slipping perfectly into its grooves, metal-gray but soft as latex,. ...
>
> She even tried, from what little calculus she'd picked up, to explain it to Franz as Δt approaching zero, eternally approaching, the slices of time growing thinner and thinner, a succession of rooms each with walls more silver, transparent, as the pure light of the zero comes nearer ... . [**38**, pp. 158-9]

Ozier has noted this last sentence is suggestive of the refining of partitions in the formal definition of integration. And, just as letting "Δt approach zero" in this definition transforms continuous functions into differentiable ones, for Franz's wife Leni it is a metaphor for the transformation in her life that takes place when she finally commits herself to the revolution. For Mexico the "silvered barriers sliding in to separate him that much more, thinning further his loneliness" is a metaphor for his transformation into a state of paranoia and despair.

Similar sequences of metaphors occur when an apprentice architect by the name of Hupla suddenly recognizes the true shape of the tunnels at the rocket assembly plant.

> "It—it's about the shape of the tunnels here, Master."
>
> "Don't flinch like that. I based that design on the double lightning-stroke, Hupla—the SS emblem."
>
> "But it's also a double integral sign! Did you know that?"
>
> "Ah. Yes: Summe, Summe, as Leibnitz said."
>
> All right. But Etzel Olsch's genius was to be fatally receptive to imagery associated with the Rocket. In the static space of the architect, he might've used a double integral now and then, early in his career, to find volumes under surfaces whose equations were known—masses, moments, centers of gravity. But it's been years since he's had to do with anything that basic. Most of his calculating these days is with marks and pfennigs, not functions of idealistic r and θ, naive x and y.... But in the dynamic space of the living Rocket, the double integral has a different meaning. To integrate here is to operate on a rate of change so that time falls away: change is stilled.... "Meters per second" will integrate to "meters." The moving vehicle is frozen, in space, to become architecture, and timeless. It was never launched. It will never fall. [**38**, pp. 300-01]

Consequently the process of integration transforms the moving vehicle into a permanent part of history, a transformation from which the world will never fully recover. To strengthen the image, Pynchon goes on to describe how the internal guidance system carries out the double integration electrically in order to locate its position in space and determine the Brennschluss point.

> Brennschluss exactly here would make the Rocket go on to hit 1000 yards east of Waterloo Station. At the instant the charge (B_{iL}) accumulating in flight equaled the preset charge (A_{iL}) on the other side, the capacitor discharged. A switch closed, fuel cut off, burning ended. The Rocket was on its own. [**38**, p. 301]

Thus for Pynchon the Brennschluss point seems to represent the point where the living rocket is transformed from a creation of science into a deterministic instrument of death.

Ozier points out how Pynchon reinforces this idea later when describing the approaching death of Klaus Närrisch as he attempts to cover an escape. Närrisch, like the rocket, is about to reach the point where the charge in the flight accumulator closes the switch and sets him on a path of action from which there is no return.

> ... it meant that this was the Last Day—and, too, with only the grim sixth sense, as much faith as clear reception, that the B of Many Subscripts just over the electric horizon was really growing closer,..., another integrating, not of gyro rate but of the raw current flow itself....
>
> B, B-sub-N-for-Narrisch, is nearly here—nearly about to burn through the last whispering veil to equal "A"—to equal the only fragment of himself left by them to go through the moment. [**38**, pp. 517–18]

Ozier concludes that "the process of double integration that calculates the fatal B_N is like a transformation of life into death as well as a transformation of time into eternity" [**33**, p. 200].

In a recent paper [**47**], Lawerence Stahlberg has suggested that Pynchon's Δt metaphor provides the characters with an opportunity for change. Thus they, like words, "can remain Δt away from what they mean" or, like Franz's wife Leni, they can "penetrate the moment." In fact at the end of the book where Pynchon's readers sit in a darkened movie theatre, he offers a final Δt opportunity:

> And it is just here, just at this dark and silent frame, that the pointed tip of the Rocket, falling nearly a mile per second, absolutely and forever without sound, reaches its last unmeasurable gap above the roof of this old theatre, the last delta-t.
>
> There is time, if you need the comfort, to touch the person next to you, or to reach between your own cold legs.... [**38**, p. 760]

Symmetry as a central image in "Death and the Compass"

Because geometry has been an important part of our human experience for so long, it is only natural that it plays a variety of roles in literature. An interesting illustration of one of these occurs in *Death and the Compass* [**9**] by Jorge Luis Borges. In this unusual detective story, Borges creates a character who weaves a labyrinth from which he cannot escape. The principal imagery hinges on mirror symmetry and the geometric fact that an object and its reflected image are in opposite orientation.

The protagonist is a detective Eric Lönrot, who likes to picture himself as a real-life Auguste Dupin, the famous detective created by Edgar Allen Poe. As the story begins Lönrot is discussing the case with Inspector Treviranus.

> "No need to look for a three-legged cat here," Treviranus was saying as he brandished an imperious cigar. "We all know that the Tetrarch of Galilee owns the finest sapphires in the world. Someone intending to steal them, must have broken in here by mistake. Yarmolinsky got up; the robber had to kill him. How does it sound to you?"
>
> "Possible, but not interesting," Lönrot answered. "You'll reply that reality hasn't the least obligation to be interesting. And I'll answer you that reality may avoid that obligation but that hypotheses may not. In the hypothesis that you propose, chance intervenes copiously. Here we have a dead rabbi; I would prefer a purely rabbinical explanation, not the imaginary mischances of an imaginary robber."
>
> Treviranus replied ill-humoredly:
>
> "I'm not interested in rabbinical explanations. I am interested in capturing the man who stabbed this unknown person." [**9**, p. 77]

Curiously enough, in Poe's *The Purloined Letter* [**36**], there is a similar discussion between the detective Dupin and the chief of police. However, in that story, the roles are reversed with Dupin suggesting a simple solution must exist while the chief of police contends the problem is too complex to have a simple solution.

This curious reversal is amplified throughout the story as Lönrot proceeds to discover a crime worthy of his powers of discernment. Obvious symmetry among clues is rejected in favor of more complex symmetry. For example, three murders occur on the evenings of the third day of three successive months and at locations forming the vertices of an equilateral triangle. An anonymous letter suggests the symmetry is complete and the episode finished. However, Lönrot quickly rejects this suggestion. He points out that at the site of each of the three murders there has been a reference to the Tetragrammaton, the unutterable name of God which contains four letters. He has also learned the Jewish day starts at sundown and hence the murders have all occurred on the fourth day of the month. Consequently Lönrot concludes there will be a fourth murder at a villa named Triste-le-Roy whose location forms a rhombus with the other three.

It is at this point, as Lönrot appears at the villa, that Borges makes excellent use of geometric imagery in portraying eerie and surreal feelings of hopelessness and bewilderment.

> Lönrot advanced among the eucalypti treading on confused generations of rigid, broken leaves. Viewed from anear, the house of the villa of Triste-le-Roy abounded in pointless symmetries and in maniacal repetitions: to one Diana in a murky niche corresponded a second Diana in another niche; one balcony was reflected in another balcony; double stairways led to double balustrades. A two-faced Hermes projected a monstrous shadow. Lönrot circled the house as he had the villa. He examined everything; beneath the level of the terrace he saw a narrow Venetian blind.
>
> He pushed it; a few marble steps descended to a vault. Lönrot, who had now perceived the architect's preferences, guessed that at the opposite wall there would be another stairway. He found it, ascended, raised his hands and opened the trap door.
>
> A brilliant light led him to a window. He opened it: a yellow, rounded moon defined two silent fountains in the melancholy garden. Lönrot explored the house. Through anterooms and galleries he passed to duplicate patios, and time after time to the same patio. He ascended the dusty stairs to circular antechambers; he was multiplied infinitely in opposing mirrors; he grew tired of opening or half-opening windows which revealed outside the same desolate garden

> from various heights and various angles; inside, only pieces of furniture wrapped in yellow dust sheets and chandeliers bound up in tarlatan. A bedroom detained him; in that bedroom, one single flower in a porcelain vase; at the first touch the ancient petals fell apart. On the second floor, on the top floor, the house seemed infinite and expanding. The house is not this large, he thought. Other things are making it seem larger: the dim light, the symmetry, the mirrors, so many years, my unfamiliarity, the loneliness. [**9**, pp. 83-4]

Just when Lönrot the discerner has followed the events to their logical conclusion, he begins to become hopelessly lost in a labyrinth he has unwittingly helped to create. In seeing himself "multiplied infinitely in opposing mirrors" we see him, falling as though into a whirlpool, unable to comprehend that the fourth murder will be his own.

Now we learn Lönrot has been lured to the villa by a criminal, Red Scharlach, seeking revenge on Lönrot for his brother's death three years earlier. Scharlach planned the second and third murders after learning Lönrot was dissatisfied with the hypothesis that the first murder was due to chance. Now Lönrot stands before Scharlach.

> Lönrot avoided Scharlach's eyes. He looked at the trees and the sky subdivided into diamonds of turbid yellow, green and red. He felt faintly cold, and he felt, too, an impersonal —almost anonymous—sadness. It was already night; from the dusty garden came the futile cry of a bird. For the last time, Lönrot considered the problem of symmetrical and periodic deaths.
>
> "In your labyrinth there are three lines too many," he said at last. "I know of one Greek labyrinth which is a single straight line. Along that line so many philosophers have lost themselves that a mere detective might well do so, too. Scharlach, when in some other incarnation you hunt me, pretend to commit (or do commit) a crime at A, then a second crime at B, eight kilometers from A, then a third crime at C, four kilometers from A and B, half-way between the two. Wait for me afterwards at D, two kilometers from A and C, again halfway between both. Kill me at D, as you are now going to kill me at Triste-le-Roy."
>
> "The next time I kill you," replied Scharlach, "I promise you that labyrinth, consisting of a single line which is invisible and unceasing."
>
> He moved back a few steps. Then, very carefully, he fired. [**9**, pp. 86–7]

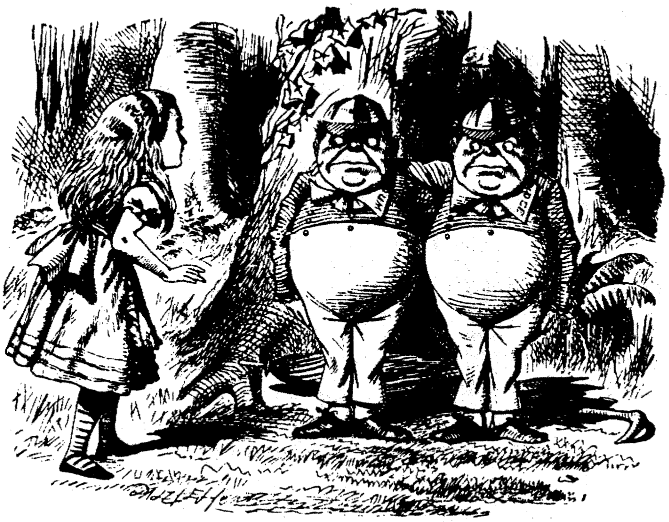
In a postcript Borges suggests that Lönrot and Scharlach may be the same person [**10**, p. 269]. If so, the curious mirror image reversal at the beginning of the story serves to identify the two, and for Borges the story becomes one about the suicide of Lönrot. The suggestion of the straight-line labyrinth at the end of the story (which is Zeno's paradox) becomes Lönrot's acknowledgement that he finally understands what has happened and realizes his reasons for suicide are not as complex as he wants to believe.

Other geometric imagery in literature

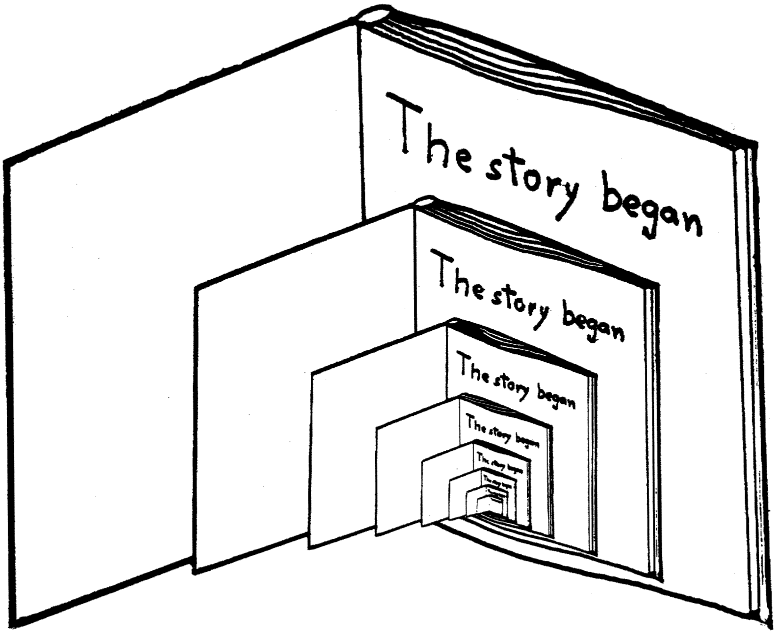
There are many other examples of geometry in literature. For instance, Martin Gardner [**19**] has noted that the geometry of reversal plays an important role in the classic nonsense literature of Lewis Carroll. In *Alice Through the Looking Glass*, Carroll (who in real life was the English mathematician Charles Dodgson) records the adventures of Alice when she enters a world behind the mirror on her mantle. There it is as though she is watching herself in a mirror, for she must continually do the opposite of what seems natural. When she tries to reach the garden by walking toward it she only succeeds in going away from it, and she is only able to approach the Red Queen by walking away from her. The poem "Jabberwocky" must be read in a mirror, and the King's messenger is in jail for a crime he won't commit until next week. Mirror geometry is also present in the characters Tweedledum and Tweedledee who are mirror images of each other. When one extends his right hand to shake hands, the other extends his left, and Tweedledee's favorite word seems to be "contraiwise."

Edwards [**17**] has noted that reversal, which is usually at the heart of nonsense, is deeply imbedded in Carroll's work. In fact, she suggests that in terms of themes, metathemes, structure, and metastructure, *Alice in Wonderland* and *Alice Through the Looking Glass* are reversals of each other. She further suggests that Carroll's stories, where life is forever turning upside down, reflect the effects on Victorian life brought about by industrialization, Darwin's theory of evolution, and other scientific, economic, and political changes. Gardner [**19**, p. 15] has added that the final metaphor of reversal in Carroll's work may be that to accept the Victorian lifestyle rationally and without illusion can only lead to the conclusion that life is irrational. In fact, it may only be when life is viewed irrationally that it can make sense.

The geometric image of a parabola plays a role in *Gravity's Rainbow* [**38**]. In fact, as Ozier notes [**33**, p. 208], underlying the whole novel is the geometric image of the parabolic flight of the rocket, gravity's own rainbow, focusing the attention of the reader and the characters on the deterministic rise and fall of society. This metaphor is suggested by the parabolic cross sections of the rocket assembly tunnels [**38**, p. 298], and repeatedly emphasized by details like the parabolic arches leading to the slums in Berlin [**38**, p. 436], or the abortive German attempt to make a parabolic sonic sound-mirror [**38**, p. 724]. In *Young Archimedes* [**21**], Aldous Huxley uses a young boy's fascination with the Pythagorean theorem to help create a touching story of the exploitation of genius. Finally, in stories like *Flatland* by Edwin Abbot or Norton Juster's delightful *The Dot and the Line*, people are personified by geometric shapes.



Tweedledum and Tweedledee, as drawn by John Tenniel for *Through the Looking Glass* by Lewis Carroll.



Paradoxes of the infinite

Geometry also lends itself to producing images of the infinite, a source of mystery and paradox in literature as well as in mathematics. For example, the parallel mirrors that multiply Lönrot infinitely are an illusion of his life accelerating toward his death. And as Alice shrinks after drinking a magic potion, she describes feeling like a telescope shutting up (much like the terms of a geometric series converging to zero), and fears that it might continue until she goes out like the flame of a candle. This leads her to wonder what the flame of a candle looks like after it goes out. A geometric image of the infinite is also present in Zeno's straight-line labyrinth that is mentioned by Lönrot. Lewis Carroll has invented a clever variation on Zeno in *What the Tortoise said to Achilles* [**64**, pp. 1225–30]. In contrast with Zeno's racecourse where there are an infinite number of distances, each smaller than the previous, the Tortoise describes a three-step racecourse consisting of an infinite number of steps, each larger than the previous. The actual racecourse is a proof:

A. Things that are equal to the same thing are equal to each other.
B. The two sides of this triangle are things equal to the same thing.
Z. The two sides of this triangle are equal to each other.

The Tortoise accepts A and B while suggesting if they don't clearly justify Z one could interpolate

C. If A and B are true, Z must be true.

Then if the argument is still not clear one can continue interpolating

D. If A, B, and C are true, Z must be true.
E. If A, B, C, and D are true, Z must be true, and so forth.

Each of these metaphoric images of the infinite is an example of "infinite regress," something closely related to the "self-referential loops" discussed by Hofstadter in his excellent book *Gödel, Escher, Bach* [**20**]. Examples of these loops occur frequently in literature. For example, at one point in Cervantes' *Don Quixote*, the barber, a creation of Cervantes, questions the merits of Cervantes' writing style. In the middle of *One Thousand and One Nights*, the character Scheherazade begins to tell the story *One Thousand and One Nights*. Lewis Carroll has Alice wonder if her dream is a part of the Red King's dream. And in *Gravity's Rainbow*, Pynchon has one of his characters begin to wonder about the true nature of the laboratory in which the Pavlovian experiments are conducted.

> From overhead, from a German camera-angle, it occurs to Webley Silvernail, this lab here is also a maze, i'n't it now ... behaviorists run these aisles of tables and consoles just like rats 'n' mice. Reinforcement for them is not a pellet of food, but a successful experiment. But who watches from above, who notes their responses? [**38**, p. 229]

Borges is particularly fascinated with the infinite and continuously probes its interminable mysteries seeking some insight into our existence. In her book *Borges: The Labyrinth Maker* [**8**], Ana Maria Barrenechea discusses the infinite as one of five central themes Borges uses in his writing. She suggests he seems to appreciate the profound and subtle ideas inherent in the infinite as well as to enjoy reflecting on their meaning for us. As evidence she refers to his thoughts in *The Language of Argentinians* where he writes

> I suspect that the word infinite was at one time the insipid equivalent of unfinished; today it is one of God's perfections in Theology, a cause for argument in Metaphysics, as well as a popular point of emphasis in Literature, a revived abstract concept in Mathematics—Russell explains the addition, multiplication, and raising of cardinal numbers to infinite powers and the reason for their almost terrible dynasties—and a true intuition on looking at the sky. [**8**, p. 23]

She then goes on to suggest that his models for the infinite include "the vast spatial and temporal reaches, the interminable multiplications, the endless path (be it lineal or cyclical), and the immobilization through a gesture" [**8**, p. 24].

A representation of the infinite as an endless path occurs when Borges describes the infinite hatred of Scharlach for Lönrot. He accomplishes this through simple circular patterns, capable of unending repetitions, thus forming a labyrinth from which escape is impossible. Scharlach is telling Lönrot of an earlier time when he was critically wounded and his brother was in jail because of Lönrot. Lönrot notes in Scharlach's voice

> ... a fatigued triumph, a hatred the size of the universe, a sadness not less than the hatred. "Nine days and nine nights I lay in agony in this desolate, symmetrical villa; fever was demolishing me, and the odious two-faced Janus who watches the twilights and the dawns lent horror to my dreams and to my waking. ... An Irishman tried to convert me to the faith of Jesus; he repeated to me the phrase of the goyim: All roads lead to Rome. At night my delirium nurtured itself on that metaphor; I felt that the world was a labyrinth, from which it was impossible to flee, for all roads, though they pretend to lead to the north or south, actually lead to Rome, which was also the quadrilateral jail where my brother was dying and the villa of Triste-le-Roy. On those nights I swore by the God who sees with two faces and by all the gods of fever and of the mirrors to weave a labyrinth around the man who had imprisoned my brother." [**9**, pp. 84–5]

In *The Garden of the Forking Paths* [**9**] Borges represents the universe metaphorically as $\mathbf{C}^{\mathbf{C}}$ (**C** is the cardinality of the continuum) to illustrate the frail and tenuous nature of the path our lives follow.

> The explanation is obvious: The Garden of Forking Paths is an incomplete, but not false, image of the universe as Ts'ui Pen conceived it. In contrast to Newton and Schopenhauer, your ancestor did not believe in a uniform, absolute time. He believed in an infinite series of times, in a growing, dizzying net of divergent, convergent and parallel times. This network of times which approached one another, forked, broke off, or were unaware of one another for centuries, embraces all possibilities of time.
>
> Once again I felt the swarming sensation. It seemed to me that the humid garden that surrounded the house was infinitely saturated with invisible persons. Those persons were Albert and I, secret, busy and multiform in other dimensions of time. I raised my eyes and the tenuous nightmare dissolved. [**9**, p. 28]

Perhaps Borges' most fascinating metaphoric use of the infinite occurs in *The Library of Babel* [**9**]. Borges, a librarian himself, creates a library containing every possible book that can be obtained by permuting the various letters and punctuation symbols. Such a library, like the infinite, is filled with mystery and paradox for it must contain

> ... the minutely detailed history of the future, the archangels' autobiographies, the faithful catalogue of the library, thousands and thousands of false catalogues, the demonstration of the fallacy of these catalogues, the demonstration of the fallacy of the true catalogue, ... the true story of your death, ... [**9**, p. 54]

A reader familiar with mathematics can easily see this library is literally as well as figuratively filled with mathematics. For example, a Cantor diagonalization argument could be used to show that no discrete catalogue could list all the books. Also the Epimenides paradox must exist in the form of a reference book that only lists reference books that are not self-referential. Such a reference book must be false regardless of whether or not it is self-referential. (Of course the library must contain versions that are self-referential as well as ones that are not self-referential).

In addition there are all the mathematical problems associated with the magnitude of the library—for you could easily spend a lifetime searching for a book containing a meaningful sentence. And consider the problem of a censor wishing to eliminate a particular book, only to discover it is duplicated by an inexhaustible supply of almost perfect copies differing only by a single letter (not to mention those which differ from it by less than a page). Or consider the frustration of those who seek a book A explaining the meanings and secrets of the universe. They know that within the library there must be a book B describing how to find A, as well as a book C describing how to find B, and so forth. Some people, realizing that searching for such a book is fruitless, even attempt to reproduce it probabilistically by the tossing of dice.

It is likely that Borges wrote this story to describe the horrors of working in a library. On the other hand one can read into his account strong mathematical metaphors. Whether or not he intended this is irrelevant, for as he has remarked

> Ana Maria Barrenechea's book has unearthed many secret links and affinities in my own literary output of which I had been quite unaware. I thank her for those revelations of an unconscious process. This means that my best writings are things that were striving to come to life through me, or in spite of me, and not simply allegories where the thought comes before the sign. [**8**, p. viii]

In some sense, this provides ample justification to Stark's claim that mathematics is metaphoric, since the works of authors like Borges, Carroll, and Pynchon allow mathematics to speak to us about nonmathematical things.

Stylistic analysis

Any discussion of mathematics and literature would not be complete without at least mentioning the technical uses of mathematics in analyzing literature. While much of the area of "stylistic analysis" properly belongs to linguistics and automated language processing, some of it is relevant (although not central) to literary criticism.

According to Sally and Walter Sedelow [**42**], stylistic analysis includes indexing, concordancing, author-attribution studies, content analysis, and syntactic analysis. Interestingly enough, Richard Bailey has suggested that the part of stylistic analysis called statistical stylistics originated with a remark by Augustus de Morgan that word-length might be a distinctive trait of an author's style [**7**, p. 217]. Although this hypothesis proved to be insufficient, it paved the way for a variety of quantitative methods for studying the way a writer encodes his or her thoughts. Many of the techniques require a computer in order to detect the subtle patterns of style. For example, Tasman made considerable use of computers to create the indices and concordance necessary to assemble the 40,000 fragments of the Dead Sea Scrolls (representing approximately 400 manuscripts written in three languages) into their original form.

The most frequently cited examples of stylistic analysis seem to be in the area of author-attribution studies. One of the classical examples in this area is the work by Mosteller and Wallace in determining the authorship of the twelve unsigned Federalist Papers [**32**]. This work, which makes use of Bayes' Theorem and a collection of discriminator words, is important because it serves as a model for future work in this area. Another interesting author attribution study by Morton suggests that probably only five of the fourteen epistles attributed to St. Paul were actually written by him [**31**].

Another example of stylistic analysis closer to the content of this paper is the study of thematic structure discussed by Smith [**45**]. He suggests that literary interpretation of a work can be aided by studying its changing patterns of images. To do this, one first decides on the images of interest, and then makes up a list of words and phrases which suggest these images. Then successive 500 word passages of the text are searched (by a computer) to obtain the graphs of the distributions of these images. The changing patterns suggested by these graphs as well as image transition diagrams provide insight into the thematic structure of the work. Smith's own study of Joyce's *Portrait of an Artist as a Young Man* [**44**] found that major developments in Stephen Dedalus's personality correspond to the occurrences of important images, and that the changing personality is mirrored by changing associations between these images. While some may question the importance of such studies, they provide the seeds for future development in this expanding area.

Conclusions

As we have seen, there are many interesting examples of the use of mathematics in literature. It should also be noted that there are a variety of ways these examples and others can be used to enhance the study of mathematics. One way is to incorporate readings in literature as a means of introducing students to ideas in mathematics. Jerry Lenz [**27**] has described the use of science fiction as a means to introduce students to ideas in geometry and topology. This same idea could be applied to other topics such as probability and statistics, logic and foundations, history, codes and ciphers, and so forth.

For advanced students, there are some excellent books that treat a topic in mathematics in an interesting or artistic way, and these could serve as the basis for a seminar or independent study. For example, in the book *Kandelman's Krim* by the British physicist J. L. Synge, the characters develop the real numbers starting from the positive integers. A more demanding book is *Surreal Numbers* by Donald Knuth, in which two students rediscover Conway's development of numbers using a generalization of Dedekind cuts. And a beautiful interdisciplinary seminar could be based on Douglas Hofstadter's Pulitzer prize-winning *Gödel, Escher, Bach: An Eternal Golden Braid.* (A group of students at St. Olaf College took time to listen to recordings of Bach's *Musical Offering* as part of a seminar on this book, and ended their course by producing a puppet play based on the material.)

A more ambitious project would be to create a course on mathematics and literature which could explore different aspects of the two subjects. In seeing how nonmathematical conclusions can be drawn from mathematical ideas, students gain new and interesting insights into mathematics as well as catch a glimpse of how writers view mathematics. By including in the course one or more biographies or autobiographies of mathematicians (such as *Hilbert* by Constance Reid or Norbert Wiener's *I Am a Mathematician*) the students become aware of the lives of famous mathematicians and the milieu in which they worked.

There are undoubtedly many other ways to introduce students to mathematics and literature. Fortunately, no matter which approach is used, most students seem to enjoy it, and are soon discovering examples of their own.

References, with annotations

[1] E. A. Abbott, Flatland, Barnes and Noble, New York, 1963. (A story about life in the plane.)

[2] Isaac Asimov, "Franchise," Earth Is Room Enough, Doubleday and Company, Inc., Garden City, New York, 1957, pp. 58-73. (A sample of size one is used to predict the outcome of the national elections.)

[3] ______, "Gimmicks Three," Earth Is Room Enough, Doubleday and Company, Inc., Garden City, New York, 1957, pp. 74-79. (The fourth dimension is used to break a contract with the Devil.)

[4] ______, "Living Space," Earth is Room Enough, Doubleday and Company, Inc., Garden City, New York, 1957, pp. 98-111. (The problem of overpopulation is solved by having people commute to Earth from different probabilistic versions of the universe.)

[5] ______, "The Feeling of Power," The Mathematical Magpie, Clifton Fadiman, ed., Simon and Schuster, New York, 1962, pp. 3-15. (A computer programmer rediscovers how to do arithmetic in a time when all computations are done by computers.)

[6] Desmond Bagley, The Spoilers, Doubleday and Company, Inc., Garden City, New York, 1970. (The birthday problem, game theory, and the St. Petersburg paradox all play minor roles in this spy thriller.)

[7] Richard W. Bailey and Lubomir Dolezel, editors, Statistics and Style, American Elsevier Publishing Company, Inc., New York, 1969, pp. 217-237.

[8] Ana Maria Barrenechea, Borges: The Labyrinth Maker, New York University Press, New York, 1965.

[9] Jorge Luis Borges, Labyrinths: Selected Short Stories and Other Writings, New Directions, New York, 1964.

[10] ______, The Aleph and Other Stories, 1933-1969, E. P. Dutton and Co., Inc., New York, 1970.

[11] Joan Fisher Box, R. A. Fisher: The Life of a Scientist, John Wiley and Sons, New York, 1978.

[12] Dionys Burger, Sphereland, Thomas Y. Crowell Co., Apollo Editions, 1965.

[13] Arthur C. Clarke, "Wall of Darkness," Other Dimensions, Robert Silverberg, ed., Hawthorn Books, Inc., 1973, pp. 46-67. (The story of an attempt to cross a mysterious wall encircling the entire known world.)

[14] Robert M. Coates, "The Law," The Mathematical Magpie, Clifton Fadiman, ed., Simon and Schuster, New York, 1962, pp. 15-20.

[15] A. J. Deutsch, "A Subway Named Moebius," Fantasia Mathematica, Clifton Fadiman, ed., Simon and Schuster, New York, 1958, pp. 222-236. (Topology helps explain strange events on the Boston MTA.)

[16] Auguste Dick, Emmy Noether, 1882-1935, Birkhäuser Verlag, Boston, 1980.

[17] Shirley M. Edwards, The Scientific Milieu of Lewis Carroll: An Analysis of Theme and Structure in Alice in Wonderland and Alice Through the Looking Glass, M. A. Thesis, Miami University, Oxford, Ohio, 1969.

[18] Bob Elliott and Ray Goulding, "The Day the Computers got Waldon Ashenfelter," 119 Years of the Atlantic, Louise Desaulniers, ed., Atlantic Monthly Company, 1977, pp. 553-57. (The computer discovers a number of unusual correlations and nabs Waldon Ashenfelter.)

[19] Martin Gardner, The Annotated Alice, Bramhall House, New York, 1960.

[20] Douglas R. Hofstadter, Gödel, Escher, Bach: An Eternal Golden Braid, Basic Books, New York, 1979.

[21] Aldous Huxley, "Young Archimedes," Fantasia Mathematica, Clifton Fadiman, ed., Simon and Schuster, New York, 1958, pp. 3-34.

[22] Leopold Infeld, Whom the Gods Love: The Story of Evariste Galois, McGraw-Hill, New York, 1948.

[23] Norton Juster, The Phantom Tollbooth, Random House, New York, 1972. (An interesting children's story that reflects the Newtonian influence on literature during the seventeenth century.)

[24] ______, The Dot and the Line, Random House, New York, 1977. (A delightful story in which a line tries to win the love of a dot who is hopelessly in love with a squiggle.)

[25] Morris Kline, Mathematics in Western Culture, Oxford University Press, New York, 1953, pp. 272-286.

[26] Donald Knuth, Surreal Numbers, Addison-Wesley, Reading, MA, 1974.

[27] Jerry Lenz, Geometry and other science fiction, Math. Teacher 66 (1973) 529.

[28] Russell Maloney, "Inflexible Logic," Fantasia Mathematica, Clifton Fadiman, ed., Simon and Schuster, New York, 1958, pp. 91-98. (A mathematician has a hard time accepting six chimpanzees who start reproducing all the literature in the world by typing on typewriters.)

[29] Harland Manchester, "The Permanent Traffic Solution," The Saturday Review Sampler of Wit and Wisdom, Martin Levin, ed., Simon and Schuster, New York, 1966, pp. 161-163.

[30] George R. McMurray, Jorge Luis Borges, Frederick Ungar Publishing Co., New York, 1980.

[31] A. Q. Morton, A computer challenges the church, The Observer, 3 (1963) 21.

[32] F. Mosteller and D. L. Wallace, Inference in an authorship problem, J. Amer. Statis. Assoc., 58 (1963) 275-309.

[33] Lance W. Ozier, The calculus of transformation: more mathematical imagery in Gravity's Rainbow, Twentieth Century Literature, 21 (1975) 193-210.

[34] ______, Antipointsman/Antimexico: some mathematical imagery in Gravity's Rainbow, Critique, 16 (1974) 73-89.

[35] Alexi Panshin, "The Destiny of Milton Gomrath," Other Dimensions, Robert Silverberg, ed., Hawthorn Books, Inc., 1973, pp. 67-69. ("Probability Central" rectifies an accidental error.)

[36] Edgar Allan Poe, Tales of Mystery and Imagination, The Heritage Press, New York, 1941. (In *The Purloined Letter* Dupin explains why a mathematician could not be a poet. Also *The Goldbug* is a classic story that involves breaking a cipher.)

[37] Arthur Porges, "The Devil and Simon Flagg," Fantasia Mathematica, Clifton Fadiman, ed., Simon and Schuster, New York, 1958, pp. 63-69. (Simon Flagg challenges the Devil to solve Fermat's last theorem.)

[38] Thomas Pynchon, Gravity's Rainbow, The Viking Press, New York, 1973.

[39] Constance Reid, Courant in Göttingen and New York, Springer-Verlag, New York, 1976.

[40] ______, Hilbert, Springer-Verlag, New York, 1970.

[41] Ralph Schoenstein, "60,000,000 Projections Can't Be Wrong," The Saturday Review Sampler of Wit and Wisdom, Martin Levin, ed., Simon and Schuster, New York, 1966, pp. 279-82.

[42] Sally Y. Sedelow and Walter A. Sedelow, Jr., Automated Language Processing, Harold Borko, ed., John Wiley and Sons, Inc., New York, 1968, pp. 181-214.

[43] Walter A. Sedelow and Sally Y. Sedelow, eds., Computers in Language Research: Formal Methods, Mouton Publishers, New York, 1979.

[44] John B. Smith, "Image and Imagery in Joyce's Portrait," Directions in Literary Criticism, Stanley Weintraub and Philip Young, eds., The Penn State Press, 1972, pp. 220-227.

[45] ______, Thematic structure and complexity, Style 9 (1975) 32-54.

[46] ______, Computer criticism, Style, 12 (1978) 326-356.

[47] Lawrence Stahlberg, The calculus of semantics and the possibilities of metaphor: technology and morality in the fiction of Thomas Pynchon, Humanities and Technology, 1 (1980) 18-25.

[48] John O. Stark, Pynchon's Fictions: Thomas Pynchon and the Literature of Information, Ohio University Press, Athens, Ohio, 1980.

[49] Jonathan Swift, "A Modest Proposal," The Prose Works of Jonathan Swift, D. D., vol. 7, Temple Scott, ed., George Bell and Sons, London, 1905, pp. 201-216.

[50] ______, Gulliver's Travels, Louis A. Landa, ed., Houghton Mifflin, Boston, 1960.

[51] John Lighton Synge, Kandleman's Krim, Jonathan Cape, London, 1957.

[52] P. Tasman, Indexing the Dead Sea Scrolls by electronic literary data processing methods, IBM brochure, November, 1958.

[53] S. M. Ulam, Adventures of a Mathematician, Charles Scribner's Sons, New York, 1976.

[54] Bob Vinnicombe, Quiz-mathematics in literature, J. Recreational Math., 10 (1977-1978) 267-269.

[55] Sylvia Townsend Warner, Mr. Fortune's Maggot, The Viking Press, New York, 1927.

[56] Stanley G. Weinbaum, "The Circle of Zero," A Martian Odyssey and Other Science Fiction Tales, Hyperion Press, Inc., Westport, Connecticut, 1974, pp. 271-290.

[57] ______, "The Brink of Infinity," A Martian Odyssey and Other Science Fiction Tales, Hyperion Press, Inc., Westport, Connecticut, 1974, pp. 463-474. (A mathematician solves a mathematical riddle to escape from a madman.)

[58] Norbert Wiener, I Am a Mathematician, Doubleday and Company, Inc., Garden City, New York, 1956.

Collections

[59] Convergent Series, Larry Niven, ed., Ballantine, New York, 1979.

[60] Famous Stories of Code and Cipher, Raymond T. Bond, ed., Rinehart and Company, Inc., New York, 1947.

[61] Fantasia Mathematica, Clifton Fadiman, ed., Simon and Schuster, New York, 1958.

[62] Other Dimensions, Robert Silverberg, ed., Hawthorn Books, Inc., 1973.

[63] The Mathematical Magpie, Clifton Fadiman, ed., Simon and Schuster, New York, 1962.

[64] The Complete Works of Lewis Carroll, Charles Dodgson, Modern Library, Random House, New York, 1976.

NOTES

Tristram's Mathematics Problem

Richard Moore
New England Conservatory
Boston, MA 02115

Editor's note: We publish this short story as a footnote to the article "Mathematics and Literature" by Don O. Koehler. Moore, a poet, writes, "Mathematics has meant a great deal to me as a poet (though it may not have been very good for my poetry career). We seem to live in a civilization where only mathematicians can enjoy mathematics and only poets appreciate poetry. My combination of interests gives me a very strong urge to fill in these gaps somehow. And in both cases the ancient Greeks play a part: the same impulse that made me study the ancient Greek theory of ratio led me to read Homer in the original for the sake of my poetry."

The story is excerpted from [2]; *other recent work by the author appears in the References.*

Without mathematics, there would be no....

"No, Tristram, you are not going to fob off your obsession with mathematics by calling it poetry. Be honest with your readers for a change. Make the right comparison for once. You are like a man with a drinking problem.

"First he is only a social drinker. That is the insidious beginning. You are aware of nothing particularly abnormal, only that he seems to have a drink in his hand every time you see him. You are only a little surprised and faintly uneasy at the frequent occasions he can find to let the bottle gurgle into the cups. That's how it was in your earlier chapters when at every conceivable opportunity you wheeled in mathematics, all aglitter and clinking, to illustrate some point. Even back then, we had vaguely begun to suspect that you can't leave the stuff alone. But now in these last two chapters you have dropped all excuses, all pretenses, and gone on a real bender. The truth is out. You're hooked—your once-happy wife, innocent children, blooming career, all sacrificed to that thin-lipped, remorseless god of oblivion that you keep capped on your bookshelves. But they—your mathematics books—are not displayed openly and confidently. You are furtive with them. You hide them under chairs and behind other books where your wife, where you yourself, won't see them. When you hear Jessie coming down the hall, you clap the guilty volume closed and hide it under a pile of dirty underwear.

"Your marriage is decaying, not because this old house has oppressed you with burdens and memories (what a trivial idea!), not because your beautiful sensitive self, which you say doesn't exist, is in mortal danger from a pack of Harpies, as you still fatuously imagine, but because of that oblivion sucking at your vitals. The drunk uses drink to numb himself to the pain and problems of his life, and in the hideous process all his warm human relationships are destroyed. Only last week Jessie decided again that she was going to leave you; and for two whole days, while she clung to her determination, what did you do? A problem in number theory! You fiddle like Nero while your world goes up in flames."

I am stunned. What can I say? I've been found out... so let's just have one more for the road, one more little hair of the dog that bit us. The ideal world of mathematics is an area of experience with respect to which, for civilized man, the state of selflessness is possible. I glide noiselessly through my thicket of theorems as a Pigmy through his forest. They are me. I am them. When I contemplate those timeless forms, time stands still. A point on a line can never move to another point.

"What?"

Never!

"Nonsense. Mathematicians talk about moving points all the time."

They lie through their teeth when they do. No point can ever move. Motion is impossible. Suppose a point moved from point A to point B. Then it would have to traverse *a path* from A to B; this path would be a line, and our moving point would have to pass through all the points on this line between A and B in succession. But it cannot possibly do this. If it is going to pass through all the points between A and B in succession, there has to be a point C, different from A, that it passes through first. Suppose, then, that C is the first point it reaches after leaving A. Then, no matter how close C is to A, there are an infinite number of points on the line between C and A, and our traveling point has already passed through them before it reached C. Therefore, C is *not* the first point our traveling point passes through. We have deduced a contradiction from our assumption of motion. Therefore, motion is impossible.

"You mean *continuous* motion is *inconceivable* in geometric space. That is Zeno's paradox: old stuff. Why repeat it now?"

Because it is never stated adequately in all the talk about Achilles and the tortoise. We want to pretend that it is what it isn't, so that we can trot out our limits and progressions and pretend that we know better. In spite of all the scholarship, there is still a tacit cultural prejudice, which amounts to an insidious plot, to misunderstand the Greek Mind.

"No, Tristram, there is an insidious plot in your own mind to misunderstand the conditions and problems of your own life, to divert yourself from those real and important matters into these elaborate games...".

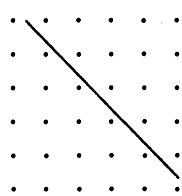
References

[1] Richard Moore, Tristram on mathematics: a poet's view, the Midwest Quart., to appear.

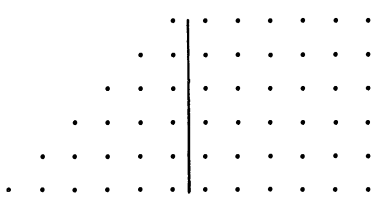
[2] ______, Tristram's rhapsody, two reviews, The Unborn Book, A Magazine of Discovery, vol. 2, no. 1, 5.

[3] ______, Empires (Poems), Ontario Review Press, 1982.

Proofs Without Words: Count the Dots



$$\sum_{k=1}^{n} k + \sum_{k=1}^{n-1} k = n^2$$



$$\sum_{k=1}^{n} k + n^2 = \sum_{k=n+1}^{2n} k$$

—WARREN PAGE
New York City Technical College, CUNY
Brooklyn, NY 11201

"Nonsense. Mathematicians talk about moving points all the time."

They lie through their teeth when they do. No point can ever move. Motion is impossible. Suppose a point moved from point A to point B. Then it would have to traverse *a path* from A to B; this path would be a line, and our moving point would have to pass through all the points on this line between A and B in succession. But it cannot possibly do this. If it is going to pass through all the points between A and B in succession, there has to be a point C, different from A, that it passes through first. Suppose, then, that C is the first point it reaches after leaving A. Then, no matter how close C is to A, there are an infinite number of points on the line between C and A, and our traveling point has already passed through them before it reached C. Therefore, C is *not* the first point our traveling point passes through. We have deduced a contradiction from our assumption of motion. Therefore, motion is impossible.

"You mean *continuous* motion is *inconceivable* in geometric space. That is Zeno's paradox: old stuff. Why repeat it now?"

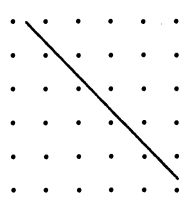
Because it is never stated adequately in all the talk about Achilles and the tortoise. We want to pretend that it is what it isn't, so that we can trot out our limits and progressions and pretend that we know better. In spite of all the scholarship, there is still a tacit cultural prejudice, which amounts to an insidious plot, to misunderstand the Greek Mind.

"No, Tristram, there is an insidious plot in your own mind to misunderstand the conditions and problems of your own life, to divert yourself from those real and important matters into these elaborate games...".

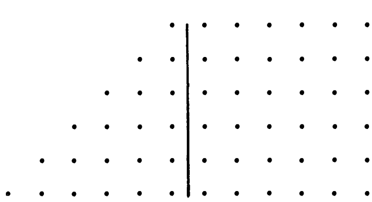
References

[1] Richard Moore, Tristram on mathematics: a poet's view, the Midwest Quart., to appear.
[2] ______, Tristram's rhapsody, two reviews, The Unborn Book, A Magazine of Discovery, vol. 2, no. 1, 5.
[3] ______, Empires (Poems), Ontario Review Press, 1982.

Proofs Without Words: Count the Dots



$$\sum_{k=1}^{n} k + \sum_{k=1}^{n-1} k = n^2$$



$$\sum_{k=1}^{n} k + n^2 = \sum_{k=n+1}^{2n} k$$

—Warren Page
New York City Technical College, CUNY
Brooklyn, NY 11201

On the Value of Mathematics (Books)

G. L. ALEXANDERSON
L. F. KLOSINSKI
University of Santa Clara
Santa Clara, CA 95053

Leonardo's *Codex Leicester* sold at auction at Christie's recently for $5.8 million. Experts viewed this as a bargain price and it may well be, for one has little on which to base a comparison. Few Leonardo manuscripts come up for sale. But of greater interest to mathematicians might be the question of what has happened over the years to the prices of scientific, or specifically mathematical, books and manuscripts less exalted than something in Leonardo's own hand. Have prices for this type of material risen as rapidly as those for art and antiques, as investors search for something of lasting value in times of rapid inflation? Along with increasing demand from investors one sees the parallel phenomenon of decreasing supply as more and more desirable material finds its way into museums and libraries, often donated by private collectors seeking tax relief.

There are several reasons why one might expect prices for paintings or other types of books to rise more rapidly than those for books on mathematics. For a start, though mathematicians see great beauty in a proof, most people get more aesthetic enjoyment from a Monet, or from a book with a lavish binding or with literary associations. Further, it has been noted that while almost everyone will know about Rembrandt, Bach and Shakespeare and the educated will know about Fabritius, Buxtehude and Marlowe, few educated people could identify Euler, Lagrange and Fermat. Yet these are mathematicians of the stature of the former group instead of the latter. So it might seem that great books and manuscripts in mathematics would be of interest only to collectors who are also mathematicians, and mathematicians are notorious for not caring about their history.

In spite of all of this, prices of rare mathematical material have been climbing steadily over the past 20 years. A case in point is the valuable collection of Robert Honeyman. Mr. Honeyman, a graduate of Lehigh University (1920), laid the foundations for his science and literature collections in the 1930's when he was based in New York. His collections continued to grow after his move to California until the collector was nearly evicted from his house by his collection. The collection of English and American literature was given to Lehigh University in 1959 but the scientific collection continued to grow until purchased by Sotheby Parke Bernet & Co. and sold at auction in a series of seven sales between October 1978 and May 1981. These auctions have set new records for many seldom seen items in the history of mathematics. At this sale, Euler's important *Introductio ad analysin infinitorum* of 1748 (in which he introduced the partition function and in which he evaluated the zeta function at some positive integral values, among other things) was sold for $660, not a surprisingly high price until one realizes that less than 20 years ago a West Coast dealer had two copies of this two volume set on his shelves, each set priced at $75. (It should also be noted that if the copy sold at Sotheby's was purchased by a dealer for resale it will almost certainly be priced at about twice the amount paid, so the collector may end up paying roughly $1200-$1300 for this set, a price in line with recent catalogue prices.) Another great Euler, the *Methodus inveniendi lineas curvas* (1744), viewed by some as his greatest work and often called the first work in the calculus of variations, was sold for $2000. Fifteen years ago this book could be found in catalogues for roughly $500. (All comparisons here are between copies in similar condition.) In a somewhat earlier auction of books from the same collection, the *Opera omnia* (1742) of Johann Bernoulli sold for $2000; less than 20 years ago a London dealer had several copies on his shelves, each marked £ 28, roughly $75 at that time.

The books just mentioned are important books but they are not books that would have as great appeal to the collector untrained in mathematics as, for instance, the first editions of Copernicus'

Johann Bernoulli, *Opera Omnia*, 1742.

De revolutionibus (1543), Newton's *Principia* (1687), or the first printed Euclid of 1482. These books are great monuments in the history of science, indeed, in the history of Western thought, and the prices reflect this. They are well beyond the means of all but a few collectors. The Copernicus brought $58,000 at the Honeyman sale; one could buy a copy in 1963 for about $900. The Honeyman copy of the 1482 Euclid brought $39,600, but in the early 1940's one could buy this book for around $200. In 1957 a first edition of Newton's *Principia* sold for $250, and in recent years it passed $10,000, then $15,000, and a nice copy was recently offered by a London dealer for $35,000.

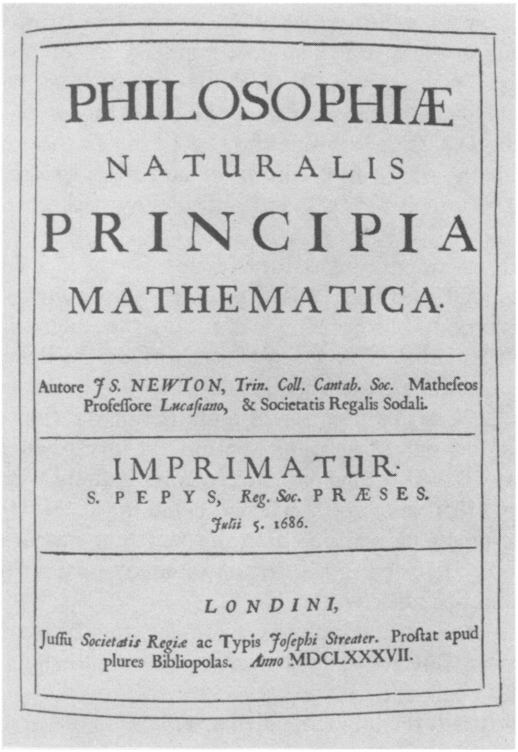
Other books, not nearly as important as the Newton or the Copernicus, nevertheless, have recently commanded astonishing prices. Pascal's *Traité du triangle arithmétique* (1665), for example, is an unusual little book but certainly not of great mathematical importance. After all, what we know as Pascal's triangle was known to Nasir-Ad-Din, the Persian, in 1265, according to [4], and was known to the Chinese even earlier, around 1100 [3]. Even in Europe it appeared in a work of Apianus in 1527, so Pascal's treatise certainly is not the first appearance of this array (though, to give Pascal his due, one must admit that he did investigate properties of elements of the array, both as elements defined recursively and as combinations). Experts estimated that this book might bring between $1850 and $2400 when it came up at the Honeyman sale. Instead, it brought $9500. This price might be attributed to the fact that Pascal has a reputation outside mathematics, but even those interested in his literary or theological writing would surely hesitate to pay prices like that for his scientific work.

These examples attest to the fact that many investors feel that books, like art and antiques, are a good hedge against inflation. But not all those who buy old mathematics books are investors—a collector buys books because collecting is fun.

A collector enjoys owning and reading a book that represents the first appearance in print of a great idea. And then there are the dealers, who are often remarkably interesting, well-informed people, so visiting a good bookshop is almost a social occasion. One never knows what will turn up in a shop or in a dealer's catalogue. For example, in the mid 1960's, one of the authors found in an English dealer's catalogue Hardy's copy (signed by Hardy) of Waring's *Meditationes Algebraicae* (1770). This is an interesting book on two counts: it belonged to Hardy who, with Littlewood, worked for many years on Waring's problem, and it contains the first appearance in print of what is now known as Wilson's Theorem (a theorem neither first observed by nor proved by Wilson!). The book cost less than $35.

On another occasion the experience was more expensive: the discovery in San Francisco of the published version of Gauss' doctoral dissertation in which he gave the first generally accepted proof of the fundamental theorem of algebra. For this, the price paid may have been too much or it may have been a bargain, but one may never know because the previous recorded sale of this book (scarcely more than a pamphlet) was in 1928.

One summer in an Oslo bookstore we asked the usual question: Do you have anything interesting in mathematics? The staff assured us that they did not. Then the owner remembered having something downstairs, but he was sure we would not be interested in such a thing. We pressed him and finally he produced a run of nine issues of the Norwegian journal *Magazin for Naturvidenskaberne* for 1823–25. These contained four papers by Abel written when he was a student, only three of which we could find later in the *Oeuvres Complètes*, edited by Holmboe in 1839. That Holmboe's edition was less than complete is not surprising. Abel's great work on elliptic functions, submitted to the Paris Academy in 1826 and lost by Cauchy and Legendre, was not published until 1841, two years after Holmboe's edition. But surely Holmboe would have known all of the early works in the *Magazin for Naturvidenskaberne*, since he contributed to the same journal. This mystery was cleared up only by searching through the 1881 edition of the *Oeuvres Complètes* where the editors, L. Sylow and S. Lie, explain the omission of this paper "dans lequel il s'était glissé, par inadvertance, une faute grave." So the Oslo visit had turned up a paper of Abel's that contained an error and which, Sylow and Lie go on to explain, Abel had, in


PHILOSOPHIÆ

NATURALIS

PRINCIPIA

MATHEMATICA.

Autore *J S. NEWTON*, *Trin. Coll. Cantab. Soc.* Matheseos Professore *Lucasiano*, & Societatis Regalis Sodali.

IMPRIMATUR.

S. PEPYS, *Reg. Soc.* PRÆSES.

Julii 5. 1686.

LONDINI,

Jussu *Societatis Regiæ* ac Typis *Josephi Streater*. Prostat apud plures Bibliopolas. *Anno* MDCLXXXVII.

Isaac Newton, *Principia Mathematica*, 1687.

fact, withdrawn!

The journal turned out to be interesting in other ways. It was founded in Oslo (Christiania) in 1823 to provide an opportunity to Norwegian scientists to publish their work. One of the editors was Hansteen, who contributed numerous articles, some on magnetism and the location of the earth's magnetic poles. Hansteen must have anticipated some objections to the inclusion of Abel's mathematical papers because to two of the papers he adds short apologetic essays, explaining the value of mathematics, presumably to readers who wanted general scientific articles on rocks, birds and physical phenomena. Hansteen says: "It may seem that in a periodical intended for the natural sciences, a memoir in pure mathematics is not in its right place. But mathematics is nature's doctrine of pure form. For the scientist it is similar to the dissecting knife of the anatomist, an absolutely necessary tool without whose aid one cannot penetrate the surface. Over the curtain which hides the entrance to the inner sanctum, the master builder of nature has placed the same motto as the Greek philosopher above the entrance to his lecture hall: 'Let no one ignorant of geometry enter.'" The complete essays (translated, fortunately, from the Norwegian) are included in Ore's biography of Abel [**5**].

The Oslo bookseller made a sale that day, whether he wanted to or not. The price of the journals was approximately $70.

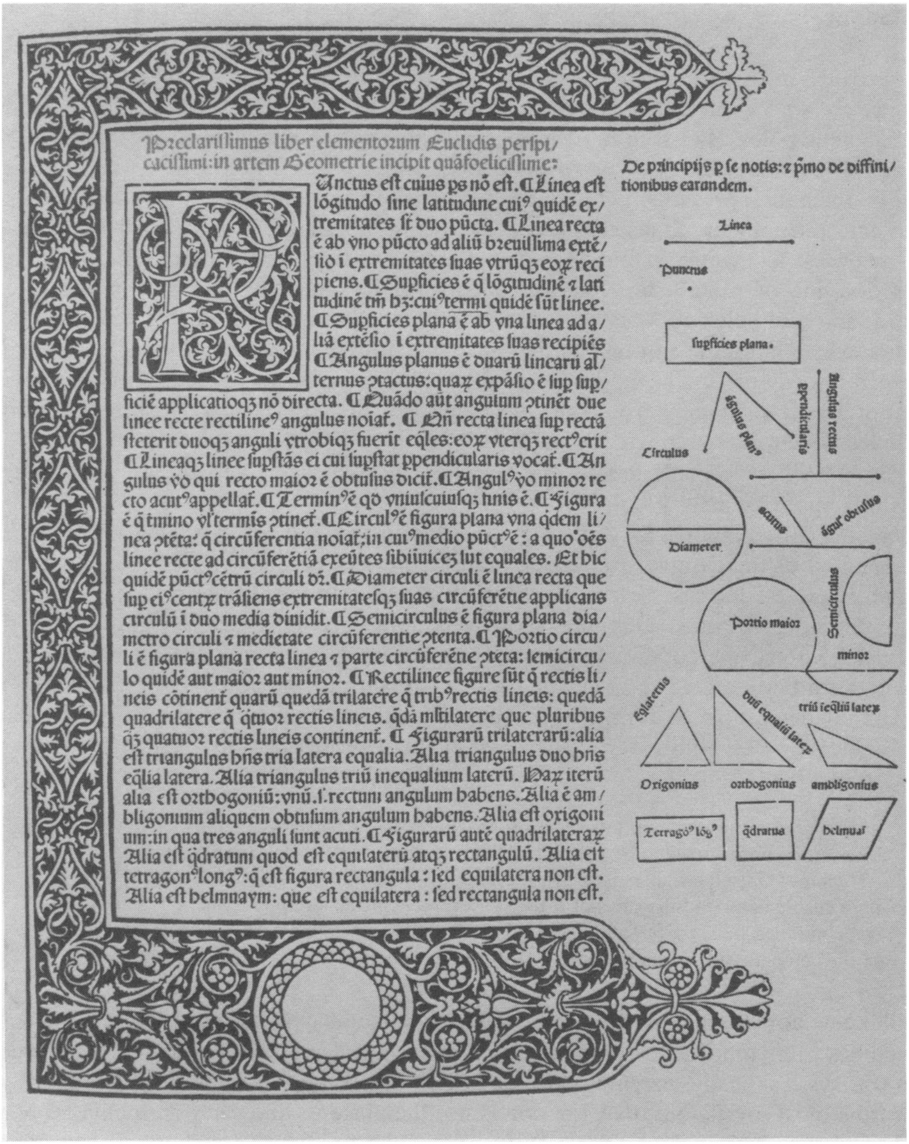
In perusing early editions, the collector sometimes marvels at the mathematics, and sometimes at the ingenuity of the author, often made more impressive because of inadequate notation. And then again, there are the cases where one marvels at the notation. In the first appearance in print of the differential calculus, Leibniz' "Nova methodus pro maximis et minimis" in the *Acta Eruditorum* for 1684, the notation is the same as that used in the most recent books on calculus.

Sometimes a book affords a glimpse of the personality of the author. For example, following his successful derivation of formulas for the Bernoulli polynomials, the following paragraph occurs in Jacques Bernoulli's *Ars Conjectandi* (1713) (the paragraph is taken from the English translation of 1795 [**1**]):

> I cannot but observe on this occasion, that the learned Ismael Bullialdus, or Bouillaud, has been rather unfortunate in his manner of treating this subject, in his *Treatise on the Arithmetick of Infinites*; since the whole of the folio volume which he has written upon it does nothing more than enable us to find the sums of the first six powers of the natural numbers, 1, 2, 3, 4, 5, 6, 7, etc. continued to any given number n; which is only a part of what we have here accomplished in the compass of a dozen pages.

Bernoulli knew how to put someone in his place.

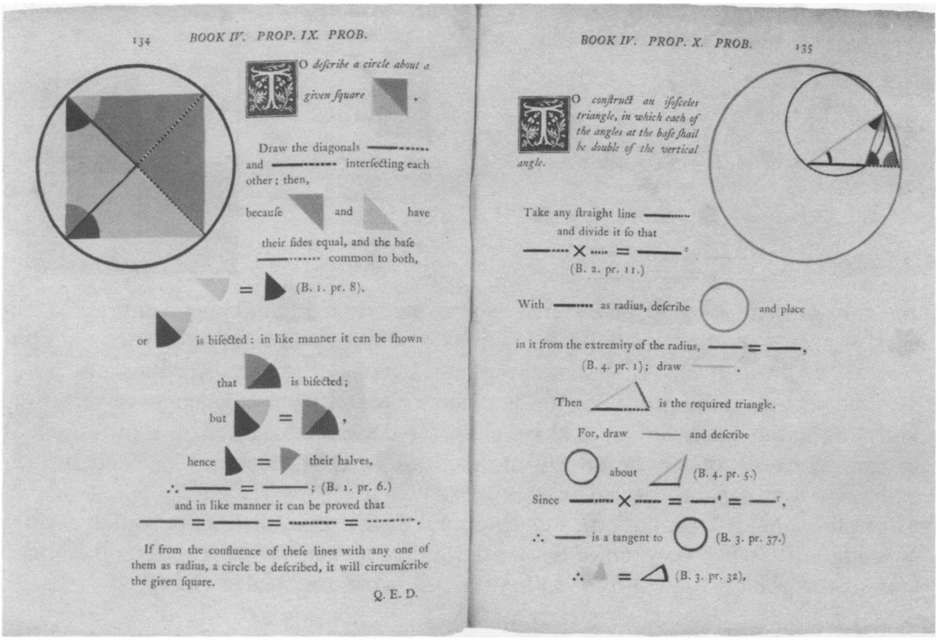
Sometimes one comes across something of pedagogical interest. Some years ago one of the authors was systematically searching through London bookshops (where, incidentally, he was told in one of the most prestigious, that London is not the place to find early mathematics books—he would have better luck in Los Angeles!). In one of the tiny cluttered shops in Cecil Court, the owner recalled that there was, indeed, a Euclid somewhere in the shop. He went down into the basement and after a considerable time emerged clutching a dusty copy of the 1847 edition of Euclid, redone by Oliver Byrne, "Surveyor of Her Majesty's Settlements in the Falkland Islands." Byrne must have had many long evenings to kill in that remote spot, so he devised a way of teaching geometry using colors. When one line is to equal another, or areas are to be equal, they are given the same color. The book, although published by Pickering in London, was printed by the Chiswick Press and is certainly one of the most spectacular Euclids ever. The colors (black, red, blue and yellow) are flat and extremely vivid—they were laid on by wood block printing—and the paper is as crisp and as clean as if it were printed last week. In the mid-19th century, it was clearly too expensive a process to be used widely, but Byrne anticipated the color printing used in modern texts. Collecting editions of Euclid alone could be a lifetime hobby, from the 1482 incunable mentioned earlier to the first in a modern language (Italian, 1543), the first in English (1570), with the remarkable preface of John Dee [**6**], all the way up to the elegant edition designed for Random House in 1944 by Bruce Rogers and containing an essay on geometry by Paul Valéry.


Preclarissimus liber elementorum Euclidis perspi/
cacissimi: in artem Geometrie incipit quãfoelicissime:

Punctus est cuius ps nõ est. Linea est
lõgitudo sine latitudine cui⁹ quidẽ ex/
tremitates sũt duo pũcta. Linea recta
ẽ ab vno pũcto ad aliũ breuissima extẽ/
siõ i extremitates suas vtrũq3 eoꝝ reci
piens. Supficies ẽ q̃ lõgitudinẽ ⁊ lati
tudinẽ tm̃ hz: cui⁹ termi quidẽ sũt linee.
Supficies plana ẽ ab vna linea ad a/
liã extẽsio i extremitates suas recipiẽs
Angulus planus ẽ duarũ linearũ al/
ternus ɔtactus: quaꝝ expãsio ẽ sup sup/
ficiẽ applicatioq3 nõ directa. Quãdo aũt angulum ɔtinẽt due
linee recte rectiline⁹ angulus noiat. Qñ recta linea sup rectã
steterit duoq3 anguli vtrobiq3 fuerit eq̃les: eoꝝ vterq3 rect⁹ erit
Lineaq3 linee supstãs ei cui supstat ppendicularis vocat. An
gulus vo qui recto maior ẽ obtusus dicit. Angul⁹ vo minor re
cto acut⁹ appellat. Termin⁹ ẽ qd vniuscuiusq3 finis ẽ. Figura
ẽ q̃ tmino vl' termis ɔtinet. Circul⁹ ẽ figura plana vna q̃dem li/
nea ɔtẽta: q̃ circũferentia noiat; in cui⁹ medio pũct⁹ ẽ: a quo oẽs
linee recte ad circũferẽtiã exeũtes sibiinuicez sũt equales. Et hic
quidẽ pũct⁹ cẽtrũ circuli dr̃. Diameter circuli ẽ linea recta que
sup ei⁹ centꝝ trãsiens extremitatesq3 suas circũferẽtie applicans
circulũ i duo media diuidit. Semicirculus ẽ figura plana dia/
metro circuli ⁊ medietate circũferentie ɔtenta. Portio circu/
li ẽ figura plana recta linea ⁊ parte circũferẽtie ɔteta: semicircu/
lo quidẽ aut maior aut minor. Rectilinee figure sũt q̃ rectis li/
neis cõtinent quarũ quedã trilatere q̃ trib⁹ rectis lineis: quedã
quadrilatere q̃ q̃tuor rectis lineis. q̃dã mltilatere que pluribus
q̃3 quatuor rectis lineis continent. Figurarũ trilaterarũ: alia
est triangulus hñs tria latera equalia. Alia triangulus duo hñs
eq̃lia latera. Alia triangulus triũ inequalium laterũ. Haꝝ iterũ
alia est orthogoniũ: vnũ .s. rectum angulum habens. Alia ẽ am/
bligonium aliquem obtusum angulum habens. Alia est oxigoni
um: in qua tres anguli sunt acuti. Figurarũ autẽ quadrilateraꝝ
Alia est q̃dratum quod est equilaterũ atq3 rectangulũ. Alia est
tetragon⁹ long⁹: q̃ est figura rectangula: sed equilatera non est.
Alia est helmuaym: que est equilatera: sed rectangula non est.

Euclid, *Elements*, 1482.

Sometimes in searching for good mathematics books one finds something quite different but which one cannot pass up. For example, in a San Francisco bookshop one of the authors turned up a copy of J. J. Sylvester's poem, *Spring's Debut: A Town Idyll*, privately published and, no doubt, at the author's expense. It is corrected in the author's own hand and there is a presentation inscription to a fellow professor at Johns Hopkins. The poem consists of 113 lines—every single line ending in the sound "in", surely one of the more stringent constraints on poetry since the decline of Alexandrine verse! It is an incredible tour de force, but it was the sort of thing that may have led to the terse comment on Sylvester's poetry in the *Dictionary of American Biography*: "Most of Sylvester's original verse showed more ingenuity than poetic feeling." Cajori [**2**] remarks that "at the reading, at the Peabody Institute in Baltimore, of his [Sylvester's] Rosalind poem, consisting of about 400 lines all rhyming with 'Rosalind', he first read all his explanatory footnotes, so as not to interrupt the poem; these took one hour and a half. Then he read the poem itself to the remnant of his audience."



Euclid, by Oliver Byrne, 1847.

Those who contemplate collecting mathematics books should be aware of the diminishing supply. The great books of Napier, Kepler, Copernicus, Newton, and Fermat have practically disappeared from the market. (Newton's theological discourses are still to be had at modest cost!) The later classics—Gauss, Euler, Lagrange, Laplace, and so on—are now priced out of the reach of many. First editions of modern classics are now appearing in dealers' catalogues: Hardy and Wright's *Introduction to the Theory of Numbers*, Pólya and Szegö's *Aufgaben und Lehrsätze*, books by Hilbert and Minkowski. These books sell in the $40-$100 range. These and classic texts, as well as minor books from earlier centuries, may be the standard fare for collectors in the future.

For those interested in seeing the great books in the history of mathematics, visits to any large university rare book room will be worthwhile. There are some collections that are particularly notable, and displays of materials are often scheduled: The Burndy Library in Norwalk, Connecticut; the De Golyer Collection at the University of Oklahoma; the Brasch Isaac Newton Collection at Stanford University; the Stanitz Collection at Kent State University; the Milestones of Science Collection at the Buffalo Society of Natural Sciences; the Burndy Collection at the Smithsonian Institution; the William A. Cole Collection at the University of Wisconsin; the Herbert McLean Evans Collection and the Prandtl Collection at the University of Texas at Austin.

References

[1] James Bernoulli, The Doctrine of Permutations and Combinations, Being an Essential and Fundamental Part of the Doctrine of Chances, Maseres, London, 1795, p. 197.

[2] Florian Cajori, A History of Mathematics, 2nd ed., Macmillan, New York, 1919, p. 343.

[3] Lam Lay-Yong, The Chinese connection between the Pascal Triangle and the solution of numerical equations of any degree, Historia Math., vol. 7, no. 4, pp. 407–24.

[4] Nasir-Ad-Din At-Tusi, Handbook of arithmetic using board and dust, (Russian translation by S. A. Ahmedov and B. A. Rosenfeld) Istor.-Mat. Issled., vol. 15, pp. 431–44.

[5] Oystein Ore, Niels Henrik Abel, Mathematician Extraordinary, University of Minnesota Press, Minneapolis, 1957, pp. 62–63.

[6] John L. Thornton and R. I. J. Tully, Scientific Books, Libraries and Collectors, 2nd ed., The Library Association, London, 1962, pp. 11–12.

On a Family of Polygons

PAUL R. SCOTT
University of Adelaide
Adelaide, South Australia 5001
Australia

Because polygons are easy to define, it is sometimes thought that they only possess trivial properties. However, sometimes even quite "obvious" properties turn out to be tricky to establish.

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be n given vectors in the plane, having different directions and satisfying $\Sigma_{i=1}^{n}\mathbf{v}_i = \mathbf{0}$. Each vector can be translated in the plane without affecting its length or direction, and so by performing suitable translations, a polygon having the n vectors as edges is produced. Such a polygon may be convex, nonconvex, or self-intersecting (see FIGURES 1 and 2). There are $(n-1)!$ distinct polygons that can be formed in this way (two polygons are equivalent if one can be obtained from the other by translation). We denote by $\mathbf{S}_n$ the set of distinct polygons having the n vectors $\mathbf{v}_i$ as edges. If P_n is a polygon of $\mathbf{S}_n$, we define the **area of** P_n, $A(P_n)$, to be the area of the finite region of the plane bounded by P_n. We establish two properties of $\mathbf{S}_n$:

(a) *There exist just two convex polygons* P_n^* *and* $P_n^{*\prime}$ *in* $\mathbf{S}_n$, *where each is obtained from the other by a rotation of* 180° *about a point.*

(b) *If* $P_n \in \mathbf{S}_n$, *then* $A(P_n) \leqslant A(P_n^*)$, *with strict inequality unless* $P_n = P_n^*$ *or* $P_n = P_n^{*\prime}$.

The proof of (a) is easy. We order the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ so that the positive (anticlockwise) angles made with the positive x-axis are increasing. Now the polygons with successive edges $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ and $\mathbf{v}_n, \mathbf{v}_{n-1}, \ldots, \mathbf{v}_1$ are clearly convex, and each is obtained from the other by a 180° rotation about a point. Further, no other polygon of $\mathbf{S}_n$ can be convex, since any rearrangement of the vectors $\mathbf{v}_i$ destroys the monotonic change of angle, producing a reentrant angle in the polygon.

Property (b) is more interesting, and we give a proof by mathematical induction. We let P_n^* be the convex polygon with successive edges $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, and let P_n be some other polygon of $\mathbf{S}_n$.

The result is trivially true for $n = 3$, since then $\mathbf{S}_n$ contains only the two congruent triangles P_3^* and $P_3^{*\prime}$ (see FIGURE 1). Let us assume the result to be true for any family $\mathbf{S}_{n-1}$ of polygons determined by $n-1$ given vectors having vector sum zero.

If P_n is a polygon of $\mathbf{S}_n$, we can subdivide P_n into an $(n-1)$-gon P_{n-1} and triangle T by drawing in a segment $\mathbf{u}$, the vector sum of two adjacent edges of P_n. Since P_{n-1} is a closed polygon, its edges are given by vectors having zero vector sum. Thus P_{n-1} belongs to a certain set of polygons $\mathbf{S}_{n-1}$, containing just two convex polygons, by Property (a). Let P_{n-1}^* denote one of these convex polygons. Then our inductive hypothesis gives

$$A(P_{n-1}) \leqslant A(P_{n-1}^*), \tag{1}$$

with strict inequality unless P_{n-1} is convex.

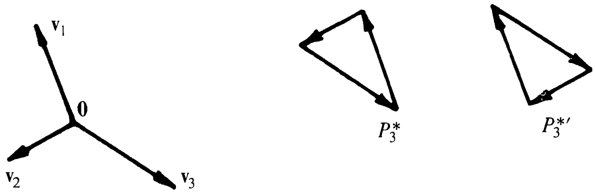


FIGURE 1

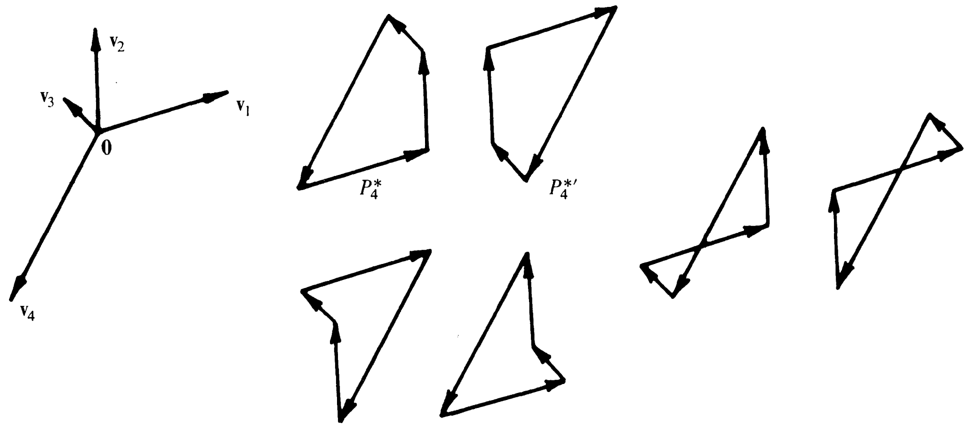


FIGURE 2

Also, since the region bounded by P_n is a subset of the union of the regions bounded by P_{n-1} and T,

$$A(P_n) \leq A(P_{n-1}) + A(T), \tag{2}$$

with equality when and only when P_{n-1} and T have no interior points in common.

If we can now show that

$$A(P_{n-1}^*) + A(T) \leq A(P_n^*), \tag{3}$$

our required result will follow from the inequalities (1), (2) and (3).

We place T and P_{n-1}^* so that they have edge $\mathbf{u}$ in common and T is external to P_{n-1}^*. Using the "free" edges of T and the edges of P_{n-1}^* adjacent to $\mathbf{u}$, we begin to construct two chains of parallelograms bordering P_{n-1}^* as in FIGURE 3. The parallelograms have no interior points in common with P_{n-1}^*, each chain ending when the next parallelogram to be constructed has such an interior point.

By the construction, the boundary of the resulting figure (containing P_{n-1}^*, T, and the parallelograms) is a polygon which is convex and which has edges determined by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Thus we have constructed P_n^* or $P_n^{*\prime}$. Since this polygon contains P_{n-1}^* and T, inequality (3) follows. Strict inequality holds here unless the constructed chains contain no parallelograms. For this to happen, the union of the regions bounded by P_{n-1}^* and T must be convex.

Thus, for all n, if $P_n \in \mathbf{S}_n$, then $A(P_n) \leq A(P_n^*)$ with strict inequality unless $P_n = P_n^*$ or $P_n^{*\prime}$.

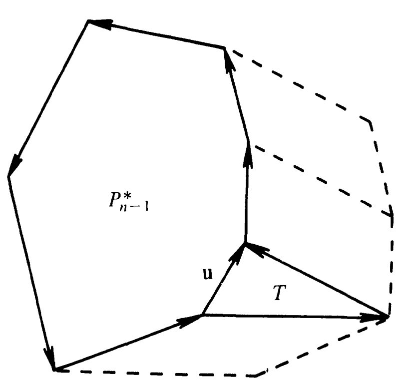


FIGURE 3

Symmetric Matrices with Prescribed Eigenvalues and Eigenvectors

KONRAD J. HEUVERS
Michigan Technological University
Houghton, MI 49931

It is a well-known result that *if* $\mathbf{S}$ *is a real* $n \times n$ *symmetric matrix then it is diagonalizable, it has n mutually orthogonal eigenvectors, and there exists an orthogonal matrix* $\mathbf{P}$ *such that its columns are eigenvectors of* $\mathbf{S}$ *and such that* $\mathbf{P}^T\mathbf{S}\mathbf{P} = \mathbf{D}$ *is the diagonal matrix of corresponding eigenvalues* [**1**, **2**, **3**, **4**, and **5**].

In matrix theory courses it is customary to look at a particular real $n \times n$ symmetric matrix for some small $n \geqslant 2$ and to try to find its eigenvalues and corresponding eigenvectors. Naturally it is up to an instructor to provide the students with "nice" symmetric matrices, i.e., matrices possessing integer (or rational) eigenvalues and eigenvector with integer (or rational) components before normalization. One of the goals of this note is to show how to start with an arbitrary orthogonal basis $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ of $\mathbb{R}^n$ and find a real $n \times n$ symmetric matrix having the $\mathbf{p}_i$ for eigenvectors and having either "nice" eigenvalues or having n arbitrarily prescribed real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ for its eigenvalues.

Recall that a real $n \times n$ symmetric matrix $\mathbf{S}$ is **positive** (**nonnegative**) **definite** if and only if all of its eigenvectors are positive (nonnegative). It is a well-known result that *a matrix* $\mathbf{S}$ *is symmetric positive definite if and only if there exists a nonsingular matrix* $\mathbf{B}$ *such that* $\mathbf{S} = \mathbf{B}\mathbf{B}^T$ [1, 3, and 5]. Another goal here is to sharpen this result and generalize it to characterize real symmetric matrices (Theorems 2 and 3). Our last goal is to provide a simple method for constructing "nice" symmetric matrices. We first present our theorems and then give several examples which illustrate the ideas and methods from the theorems and their proofs.

Our first theorem will be useful in two ways; it gives an explicit method for constructing a symmetric matrix with prescribed eigenvectors and eigenvalues and it also is useful in establishing further results. Throughout our discussion, vectors will be considered as column vectors, and we will denote a matrix $\mathbf{A}$ with column vectors $\mathbf{a}_i$ as $\mathbf{A} = (\mathbf{a}_1 | \mathbf{a}_2 | \ldots | \mathbf{a}_n)$.

THEOREM 1. *Let* $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ *be an arbitrary orthonormal basis for* $\mathbb{R}^n$. *Given* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *n arbitrary real numbers, and* τ *any real number satisfying* $\tau \leqslant \lambda_j$ *for* $j = 1, 2, \ldots, n$, *define* $\mu_j = \sqrt{\lambda_j - \tau}$ *and* $\mathbf{b}_j = \mu_j \mathbf{p}_j$, $j = 1, 2, \ldots, n$. *If* $\mathbf{B}$ *is the matrix* $\mathbf{B} = (\mathbf{b}_1 | \mathbf{b}_2 | \ldots | \mathbf{b}_n)$, *and* $\mathbf{S}$ *is the matrix* $\mathbf{S} = \mathbf{B}\mathbf{B}^T + \tau \mathbf{I}$, *then* $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ *are eigenvectors of* $\mathbf{S}$ *and* $\lambda_j = \mu_j^2 + \tau = \|\mathbf{b}_j\|^2 + \tau$, $j = 1, 2, \ldots, n$ *are the corresponding eigenvalues of* $\mathbf{S}$.

Proof. The rows of $\mathbf{B}^T$ are transposed columns of $\mathbf{B} = (\mu_1 \mathbf{p}_1 | \ldots | \mu_n \mathbf{p}_n)$ so let $\mathbf{B}^T = (\mu_i \mathbf{p}_i^T)$ express $\mathbf{B}^T$ in terms of its rows. Now, since $\mathbf{B}^T \mathbf{p}_j = (\mu_i \mathbf{p}_i^T)\mathbf{p}_j = (\mu_j \delta_{ij})$, it follows that $\mathbf{S}\mathbf{p}_j = (\mathbf{B}\mathbf{B}^T + \tau \mathbf{I})\mathbf{p}_j = \mathbf{B}\mathbf{B}^T \mathbf{p}_j + \tau \mathbf{p}_j = (\mu_1 \mathbf{p}_1 | \ldots | \mu_n \mathbf{p}_n)(\mu_j \delta_{ij}) + \tau \mathbf{p}_j = \mu_j^2 \mathbf{p}_j + \tau \mathbf{p}_j = (\mu_j^2 + \tau)\mathbf{p}_j = \lambda_j \mathbf{p}_j$.

Now we can characterize real symmetric and symmetric positive (nonnegative) definite matrices.

THEOREM 2. *Let* $\mathbf{S}$ *be a real* $n \times n$ *matrix. Then* $\mathbf{S}$ *is symmetric nonnegative definite if and only if there exists an* $n \times n$ *matrix* $\mathbf{B}$ *such that* $\mathbf{S} = \mathbf{B}\mathbf{B}^T$ *and the columns of* $\mathbf{B}$ *are mutually orthogonal.*

Proof. If $\mathbf{S}$ is symmetric nonnegative definite, let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be its eigenvalues and let $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ be an orthonormal basis of corresponding eigenvectors. If $\mathbf{P} = (\mathbf{p}_1 | \ldots | \mathbf{p}_n)$, then $\mathbf{P}$ is an orthogonal matrix such that $\mathbf{P}^T\mathbf{S}\mathbf{P} = \mathbf{D}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \mathbf{D}$ is a diagonal matrix and $\mathbf{S} = \mathbf{P}\mathbf{D}\mathbf{P}^T$. Let $\mu_j = \sqrt{\lambda_j}$ for $j = 1, 2, \ldots, n$; then the diagonal matrix $\mathbf{M} = \mathbf{M}(\mu_1, \mu_2, \ldots, \mu_n)$ is a real matrix, since each $\lambda_j \geqslant 0$, and $\mathbf{M}^2 = \mathbf{D}$. Now, $\mathbf{S} = \mathbf{P}\mathbf{D}\mathbf{P}^T = \mathbf{P}\mathbf{M}^2\mathbf{P}^T = \mathbf{P}\mathbf{M}(\mathbf{P}\mathbf{M})^T = \mathbf{B}\mathbf{B}^T$ where $\mathbf{B} = \mathbf{P}\mathbf{M} = (\mathbf{p}_1 | \ldots | \mathbf{p}_n)\mathbf{M}(\mu_1, \ldots, \mu_n) = (\mu_1 \mathbf{p}_1 | \ldots | \mu_n \mathbf{p}_n)$.

Now suppose that $\mathbf{S}=\mathbf{B}\mathbf{B}^T$ and the columns $\mathbf{b}_i$ of $\mathbf{B}$ are mutually orthogonal. We may assume that $\mathbf{b}_1,\ldots,\mathbf{b}_r$ are the nonzero columns of $\mathbf{B}$, for if the columns of $\mathbf{B}$ are not so arranged then an orthogonal matrix $\mathbf{Q}$ can be found such that the columns of $\mathbf{BQ}$ have that arrangement. It would then follow that $\mathbf{BQ}(\mathbf{BQ})^T=\mathbf{BQQ}^T\mathbf{B}^T=\mathbf{BB}^T=\mathbf{S}$. Now let $\mathbf{p}_j=\|\mathbf{b}_j\|^{-1}\mathbf{b}_j$ for $j=1,2,\ldots,r$, and select $\mathbf{p}_j$ for $r<j$ so that $\{\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n\}$ is an orthonormal basis of $\mathbb{R}^n$. From our definition of $\mathbf{p}_j$ and since $\mathbf{b}_j=\mathbf{0}$ for $r<j$, it follows that $\mathbf{b}_j=\|\mathbf{b}_j\|\mathbf{p}_j$ for $j=1,2,\ldots,n$. But now by Theorem 1 with $\tau=0$ and $\lambda_j=\|\mathbf{b}_j\|^2$, it follows that $\lambda_1,\lambda_2,\ldots,\lambda_n$ are the eigenvalues of $\mathbf{S}$ and that they are all nonnegative.

COROLLARY 1. *Let* $\mathbf{S}$ *be a real nonzero* $n\times n$ *matrix. Then* $\mathbf{S}$ *is symmetric nonnegative definite with at least one zero eigenvalue if and only if there exists an* $n\times r$ *matrix* $\mathbf{B}$ *with nonzero orthogonal columns such that* $\mathbf{S}=\mathbf{BB}^T$ *and* $r<n$.

Proof. Suppose $\mathbf{B}=(\mathbf{b}_1|\ldots|\mathbf{b}_r)$ where $\mathbf{b}_j$, $j=1,2,\ldots,r$, are mutually orthogonal and nonzero, and define the $n\times n$ matrix $\mathbf{A}=(\mathbf{b}_1|\ldots|\mathbf{b}_r|\mathbf{0}|\ldots|\mathbf{0})$. Then $\mathbf{BB}^T=\mathbf{AA}^T=\mathbf{S}$ satisfies the theorem.

Conversely, if $\mathbf{S}$ is symmetric nonnegative definite, then by Theorem 2 there exists a matrix $\mathbf{A}$ such that $\mathbf{S}=\mathbf{AA}^T$ and the columns of $\mathbf{A}$ are mutually orthogonal. By Theorem 1, if $\mathbf{A}=(\mathbf{a}_1|\ldots|\mathbf{a}_n)$ then $\|\mathbf{a}_j\|^2=\lambda_j$, $j=1,2,\ldots,n$, are the eigenvalues of $\mathbf{S}$. As in the proof of Theorem 2, we may assume that the λ_i are ordered so that $\lambda_j>0$ for $1\leqslant j\leqslant r$ and $\lambda_j=0$ for $r<j$. Then $\mathbf{a}_j=\mathbf{0}$ for $r<j$ and $\mathbf{a}_j$, $j=1,2,\ldots,n$, are mutually orthogonal. Define $\mathbf{B}=(\mathbf{a}_1|\ldots|\mathbf{a}_r)$; then we have $\mathbf{A}=(\mathbf{a}_1|\ldots|\mathbf{a}_r|\mathbf{0}|\ldots|\mathbf{0})$ and $\mathbf{BB}^T=\mathbf{AA}^T=\mathbf{S}$.

COROLLARY 2. *Let* $\mathbf{S}$ *be a real* $n\times n$ *matrix. Then* $\mathbf{S}$ *is symmetric positive definite if and only if there exists a nonsingular matrix* $\mathbf{B}$ *such that* $\mathbf{S}=\mathbf{BB}^T$ *and the columns of* $\mathbf{B}$ *form an orthogonal basis of* $\mathbb{R}^n$.

Proof. This is just a special case of Theorem 2. Note that the eigenvalues of $\mathbf{S}$ are $\lambda_j=\|\mathbf{b}_j\|^2$ where the $\mathbf{b}_j$ are the columns of $\mathbf{B}$. These $\mathbf{b}_j$ form an orthogonal basis of $\mathbb{R}^n$ if and only if all the λ_j are positive.

COROLLARY 3. *Let* $\mathbf{S}$ *be a real* $n\times n$ *matrix. Then* $\mathbf{S}$ *is symmetric if and only if there exists a nonsingular matrix* $\mathbf{B}$ *and a number* τ *such that* $\mathbf{S}=\mathbf{BB}^T+\tau\mathbf{I}$ *and the columns of* $\mathbf{B}$ *form an orthogonal basis of* $\mathbb{R}^n$.

Proof. If $\mathbf{S}$ is a symmetric matrix with eigenvalues $\lambda_1,\lambda_2,\ldots,\lambda_n$, then $\mathbf{S}=(\mathbf{S}-\tau\mathbf{I})+\tau\mathbf{I}$ where $\mathbf{S}-\tau\mathbf{I}$ is a positive definite symmetric matrix with eigenvalues $\lambda_j-\tau$, $j=1,2,\ldots,n$, provided that $\tau<\lambda_j$ for all j. By Theorem 2, there exists a nonsingular matrix $\mathbf{B}$ such that $\mathbf{S}-\tau\mathbf{I}=\mathbf{BB}^T$ and the columns of $\mathbf{B}$ form an orthogonal basis of $\mathbb{R}^n$.

THEOREM 3. *Let* $\mathbf{S}$ *be a real* $n\times n$ *matrix. Then* $\mathbf{S}$ *is symmetric if and only if for some* $r<n$, *there exists an* $n\times r$ *matrix* $\mathbf{B}$ *and a number* σ *such that* $\mathbf{S}=\mathbf{BB}^T+\sigma\mathbf{I}$ *and the columns of* $\mathbf{B}$ *are mutually orthogonal.*

Proof. If $\mathbf{S}$ is symmetric, let $\lambda_1,\lambda_2,\ldots,\lambda_n$ be its eigenvalues and let $\sigma=\min(\lambda_1,\ldots,\lambda_n)$. Then $\mathbf{S}=(\mathbf{S}-\sigma\mathbf{I})+\sigma\mathbf{I}$ where $\mathbf{S}-\sigma\mathbf{I}$ is a symmetric nonnegative definite matrix with eigenvalues $\lambda_j-\sigma$, $j=1,2,\ldots,n$. From Corollary 1 we have $\mathbf{S}-\sigma\mathbf{I}=\mathbf{BB}^T$ where $\mathbf{B}$ is an $n\times r$ matrix for some $r<n$ and $\mathbf{B}$ has mutually orthogonal columns. It follows that $\mathbf{S}=\mathbf{BB}^T+\sigma\mathbf{I}$.

We can illustrate the usefulness of the above results through several examples. The first example will illustrate the method used in Theorem 1 and Corollary 3 to construct a real symmetric 2×2 matrix having prescribed numbers for eigenvalues and a prescribed orthogonal basis for its eigenvectors.

EXAMPLE 1. Choose real numbers α,β such that the vectors $\mathbf{a}_1=[\alpha\quad\beta]^T$ and $\mathbf{a}_2=[-\beta\quad\alpha]^T$ form an orthogonal basis of $\mathbb{R}^2$, and let λ_1, λ_2 and τ be real numbers satisfying $\tau<\lambda_1,\tau<\lambda_2$. Thus $\mu_1=\sqrt{\lambda_1-\tau}$ and $\mu_2=\sqrt{\lambda_2-\tau}$ are positive. If $\mathbf{a}_1$ and $\mathbf{a}_2$ are normalized then

$$\mathbf{p}_1 = \|\mathbf{a}_1\|^{-1}\mathbf{a}_1 = (\alpha^2+\beta^2)^{-1/2}[\alpha \quad \beta]^T$$

and

$$\mathbf{p}_2 = \|\mathbf{a}_2\|^{-1}\mathbf{a}_2 = (\alpha^2+\beta^2)^{-1/2}[-\beta \quad \alpha]^T$$

will be an orthonormal basis of $\mathbb{R}^2$. Take $\mathbf{b}_1 = \mu_1\mathbf{p}_1$ and $\mathbf{b}_2 = \mu_2\mathbf{p}_2$, and define $\mathbf{B} = (\mathbf{b}_1|\mathbf{b}_2)$ and $\mathbf{S} = \mathbf{B}\mathbf{B}^T + \tau\mathbf{I}$. Then $\mathbf{B}$ is nonsingular and we have

$$\mathbf{S} = \frac{1}{\alpha^2+\beta^2}\begin{pmatrix} \sqrt{\lambda_1-\tau}\,\alpha & -\sqrt{\lambda_2-\tau}\,\beta \\ \sqrt{\lambda_1-\tau}\,\beta & \sqrt{\lambda_2-\tau}\,\alpha \end{pmatrix}\begin{pmatrix} \sqrt{\lambda_1-\tau}\,\alpha & \sqrt{\lambda_1-\tau}\,\beta \\ -\sqrt{\lambda_2-\tau}\,\beta & \sqrt{\lambda_2-\tau}\,\alpha \end{pmatrix} + \tau\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \frac{1}{\alpha^2+\beta^2}\begin{pmatrix} \lambda_1\alpha^2+\lambda_2\beta^2 & (\lambda_1-\lambda_2)\alpha\beta \\ (\lambda_1-\lambda_2)\alpha\beta & \lambda_1\beta^2+\lambda_2\alpha^2 \end{pmatrix}$$

which has eigenvalues λ_1 and λ_2, and has $\mathbf{a}_1 = \sqrt{\alpha^2+\beta^2}\,\mathbf{p}_1$ and $\mathbf{a}_2 = \sqrt{\alpha^2+\beta^2}\,\mathbf{p}_2$ for the corresponding eigenvectors.

The second example shows how to obtain the same symmetric matrix using the method developed in the proof of Theorem 3.

EXAMPLE 2. Let $\mathbf{a}_1 = [\alpha \quad \beta]^T$ and $\mathbf{a}_2 = [-\beta \quad \alpha]^T$ be an arbitrary orthogonal basis of $\mathbb{R}^2$, and let λ_1 and λ_2 be any real numbers with $\lambda_1 \geqslant \lambda_2$. Let $\mathbf{p}_1 = \|\mathbf{a}_1\|^{-1}\mathbf{a}_1$, and define $\mathbf{b}_1 = \sqrt{\lambda_1-\lambda_2}\,\mathbf{p}_1$; finally, let $\mathbf{S} = \mathbf{b}_1\mathbf{b}_1^T + \lambda_2\mathbf{I}$. Then we have

$$\mathbf{S} = \frac{\lambda_1-\lambda_2}{\alpha^2+\beta^2}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}(\alpha \quad \beta) + \lambda_2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{\alpha^2+\beta^2}\begin{pmatrix} \lambda_1\alpha^2+\lambda_2\beta^2 & (\lambda_1-\lambda_2)\alpha\beta \\ (\lambda_1-\lambda_2)\alpha\beta & \lambda_1\beta^2+\lambda_2\alpha^2 \end{pmatrix},$$

the same result obtained in Example 1.

Example 3 provides a specific numerical illustration of the general constructions we have demonstrated in Examples 1 and 2. Using the method from Example 2 first and then the method from Example 1, the reader can decide which method seems easier.

EXAMPLE 3. Let $\mathbf{a}_1 = [3 \quad 1]^T$, $\mathbf{a}_2 = [-1 \quad 3]^T$, $\lambda_1 = 13$, and $\lambda_2 = 3$ be the given eigenvectors and eigenvalues. Normalize $\mathbf{a}_1$ and $\mathbf{a}_2$ to obtain $\mathbf{p}_1 = [3/\sqrt{10} \quad 1/\sqrt{10}]^T$ and $\mathbf{p}_2 = [-1/\sqrt{10} \quad 3/\sqrt{10}]^T$, an orthonormal basis of $\mathbb{R}^2$.

If the method from Example 2 is used, then (since $\lambda_1 = 13 > 3 = \lambda_2$) we let $\mathbf{b}_1 = \sqrt{10}\,\mathbf{p}_1 = [3 \quad 1]^T$ and define

$$\mathbf{S} = \mathbf{b}_1\mathbf{b}_1^T + 3\mathbf{I} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}(3 \quad 1) + 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 12 & 3 \\ 3 & 4 \end{pmatrix}.$$

If the method from Example 1 is used, some number $\tau < \lambda_2$ must be chosen (obviously we wish to keep calculations as simple as possible). Let $\tau = -3$; then $\mu_1 = \sqrt{\lambda_1-\tau} = 4$ and $\mu_2 = \sqrt{\lambda_2-\tau} = \sqrt{6}$. Define $\mathbf{b}_1 = \mu_1\mathbf{p}_1 = [12/\sqrt{10} \quad 4/\sqrt{10}]^T$ and $\mathbf{b}_2 = \mu_2\mathbf{p}_2 = [-\sqrt{6}/\sqrt{10} \quad 3\sqrt{6}/\sqrt{10}]^T$; let $\mathbf{B} = (\mathbf{b}_1|\mathbf{b}_2)$ and let $\mathbf{S} = \mathbf{B}\mathbf{B}^T - 3\mathbf{I}$. Then we have

$$\mathbf{S} = \frac{1}{10}\begin{pmatrix} 12 & -\sqrt{6} \\ 4 & 3\sqrt{6} \end{pmatrix}\begin{pmatrix} 12 & 4 \\ -\sqrt{6} & 3\sqrt{6} \end{pmatrix} - 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 12 & 3 \\ 3 & 4 \end{pmatrix}.$$

It can be checked directly that this matrix $\mathbf{S}$ has eigenvalues $\lambda_1 = 13$ and $\lambda_2 = 3$ corresponding to the eigenvectors $\mathbf{a}_1 = [3 \quad 1]^T$ and $\mathbf{a}_2 = [-1 \quad 3]^T$.

The purpose of our next example is to construct a real symmetric 3×3 matrix $\mathbf{S} = (\mathbf{s}_{ij})$ having arbitrary real numbers $\lambda_1, \lambda_2, \lambda_3$ for eigenvalues and an arbitrary orthonormal basis of $\mathbb{R}^3$ for its eigenvectors, and to express the entries $\mathbf{s}_{ij}$ in terms of the λ_j and the coordinates of the eigenvectors. Theorem 1 and Corollary 3 are utilized in the construction.

EXAMPLE 4. Let $\mathbf{a}=[\alpha_1 \quad \alpha_2 \quad \alpha_3]^T$, $\mathbf{b}=[\beta_1 \quad \beta_2 \quad \beta_3]^T$, and $\mathbf{c}=[\gamma_1 \quad \gamma_2 \quad \gamma_3]^T$ be an arbitrary orthonormal basis for $\mathbb{R}^3$, let $\lambda_1,\lambda_2,\lambda_3$ be three real numbers and τ a real number satisfying $\tau<\lambda_j$, $j=1,2,3$. Since $\{\mathbf{a},\mathbf{b},\mathbf{c}\}$ is an orthonormal basis, the matrix $\mathbf{P}=(\mathbf{a}|\mathbf{b}|\mathbf{c})$ is orthogonal, and thus its rows $[\alpha_i \quad \beta_i \quad \gamma_i]$, $i=1,2,3$, are also orthonormal. Let

$$\mu_1=\sqrt{\lambda_1-\tau}, \quad \mu_2=\sqrt{\lambda_2-\tau} \quad \text{and} \quad \mu_3=\sqrt{\lambda_3-\tau},$$

$$\mathbf{b}_1=\mu_1\mathbf{a}, \quad \mathbf{b}_2=\mu_2\mathbf{b} \quad \text{and} \quad \mathbf{b}_3=\mu_3\mathbf{c},$$

$$\mathbf{B}=(\mathbf{b}_1|\mathbf{b}_2|\mathbf{b}_3) \quad \text{and} \quad \mathbf{S}=\mathbf{B}\mathbf{B}^T+\tau\mathbf{I}.$$

Thus

$$\mathbf{S}=\begin{pmatrix} \mu_1\alpha_1 & \mu_2\beta_1 & \mu_3\gamma_1 \\ \mu_1\alpha_2 & \mu_2\beta_2 & \mu_3\gamma_2 \\ \mu_1\alpha_3 & \mu_2\beta_3 & \mu_3\gamma_3 \end{pmatrix}\begin{pmatrix} \mu_1\alpha_1 & \mu_1\alpha_2 & \mu_1\alpha_3 \\ \mu_2\beta_1 & \mu_2\beta_2 & \mu_2\beta_3 \\ \mu_3\gamma_1 & \mu_3\gamma_2 & \mu_3\gamma_3 \end{pmatrix}+\tau\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiplying out, we can obtain equations for the entries of $\mathbf{S}=(\mathbf{s}_{ij})$. We have

$$\mathbf{s}_{ii}=\mu_1^2\alpha_i^2+\mu_2^2\beta_i^2+\mu_3^2\gamma_i^2+\tau=\lambda_1\alpha_i^2+\lambda_2\beta_i^2+\lambda_3\gamma_i^2,$$

since

$$\mu_i^2=\lambda_i-\tau \quad \text{and} \quad \alpha_i^2+\beta_i^2+\gamma_i^2=1,$$

for $i=1,2,3$. Also,

$$\mathbf{s}_{ij}=\mathbf{s}_{ji}=\lambda_1\alpha_i\alpha_j+\lambda_2\beta_i\beta_j+\lambda_3\gamma_i\gamma_j \quad \text{for } i\neq j,$$

since $\alpha_i\alpha_j+\beta_i\beta_j+\gamma_i\gamma_j=0$ for $i\neq j$. The matrix $\mathbf{S}$, with eigenvalues $\lambda_1,\lambda_2,\lambda_3$ and eigenvectors $\mathbf{a},\mathbf{b},\mathbf{c}$ is then

$$\mathbf{S}=\begin{pmatrix} \lambda_1\alpha_1^2+\lambda_2\beta_1^2+\lambda_3\gamma_1^2 & \lambda_1\alpha_1\alpha_2+\lambda_2\beta_1\beta_2+\lambda_3\gamma_1\gamma_2 & \lambda_1\alpha_1\alpha_3+\lambda_2\beta_1\beta_3+\lambda_3\gamma_1\gamma_3 \\ \lambda_1\alpha_1\alpha_2+\lambda_2\beta_1\beta_2+\lambda_2\gamma_1\gamma_2 & \lambda_1\alpha_2^2+\lambda_2\beta_2^2+\lambda_3\gamma_2^2 & \lambda_1\alpha_2\alpha_3+\lambda_2\beta_2\beta_3+\lambda_3\gamma_2\gamma_3 \\ \lambda_1\alpha_1\alpha_3+\lambda_2\beta_1\beta_3+\lambda_3\gamma_1\gamma_3 & \lambda_1\alpha_2\alpha_3+\lambda_2\beta_2\beta_3+\lambda_3\gamma_2\gamma_3 & \lambda_1\alpha_3^2+\lambda_2\beta_3^2+\lambda_3\gamma_3^2 \end{pmatrix}.$$

We next provide a numerical illustration of the construction of a symmetric 3×3 matrix with prescribed eigenvalues and eigenvectors a given orthonormal basis; we parallel the method used in Example 2.

EXAMPLE 5. Let

$$\mathbf{p}_1=[1/3 \quad 2/3 \quad 2/3]^T, \quad \mathbf{p}_2=[2/3 \quad 1/3 \quad -2/3]^T \quad \text{and} \quad \mathbf{p}_3=[2/3 \quad -2/3 \quad 1/3]^T,$$

and let $\lambda_1=6$, $\lambda_2=3$ and $\lambda_3=-3$ be the given eigenvectors and eigenvalues. Since the smallest eigenvalue is $\lambda_3=-3$ and $\lambda_1-\lambda_3=9$ and $\lambda_2-\lambda_3=6$, we take $\mathbf{b}_1=3\mathbf{p}_1$ and $\mathbf{b}_2=\sqrt{6}\,\mathbf{p}_2$. Define $\mathbf{B}=(\mathbf{b}_1|\mathbf{b}_2)$ and let $\mathbf{S}=\mathbf{B}\mathbf{B}^T-3\mathbf{I}$. Thus

$$\mathbf{B}=\frac{1}{3}\begin{pmatrix} 3 & 2\sqrt{6} \\ 6 & \sqrt{6} \\ 6 & -2\sqrt{6} \end{pmatrix}$$

and

$$\mathbf{S}=\frac{1}{9}\begin{pmatrix} 3 & 2\sqrt{6} \\ 6 & \sqrt{6} \\ 6 & -2\sqrt{6} \end{pmatrix}\begin{pmatrix} 3 & 6 & 6 \\ 2\sqrt{6} & \sqrt{6} & -2\sqrt{6} \end{pmatrix}-3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}=\frac{1}{3}\begin{pmatrix} 2 & 10 & -2 \\ 10 & 5 & 8 \\ -2 & 8 & 11 \end{pmatrix}.$$

It can be readily verified that this matrix $\mathbf{S}$ has the given eigenvalues 6, 3, -3 corresponding to the given eigenvectors $\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3$.

Our remaining examples are all specific numerical illustrations of the applicability of the theorems and corollaries, and generally show how easy it is to use orthogonal vectors with integer components to generate symmetric matrices with integer eigenvalues. Example 6 illustrates Corollary 2; an orthogonal basis of $\mathbb{R}^3$ is given and a symmetric positive definite matrix $\mathbf{S}$ is constructed.

EXAMPLE 6. Let $\mathbf{b}_1=[1 \quad 1 \quad -2]^T$, $\mathbf{b}_2=[1 \quad 1 \quad 1]^T$, and $\mathbf{b}_3=[1 \quad -1 \quad 0]^T$. Set $\mathbf{B}=(\mathbf{b}_1|\mathbf{b}_2|\mathbf{b}_3)$ and $\mathbf{S}=\mathbf{B}\mathbf{B}^T$; then

$$\mathbf{S}=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ -2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}=\begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{pmatrix}$$

which has eigenvalues $\lambda_1=\|\mathbf{b}_1\|^2=6$, $\lambda_2=\|\mathbf{b}_2\|^2=3$, and $\lambda_3=\|\mathbf{b}_3\|^2=2$.

Example 7 illustrates Corollary 3; an orthogonal basis of $\mathbb{R}^3$ and a number τ are given and a symmetric matrix $\mathbf{S}$ is constructed.

EXAMPLE 7. Let $\mathbf{b}_1=[1 \quad 2 \quad 3]^T$, $\mathbf{b}_2=[1 \quad 1 \quad -1]^T$, $\mathbf{b}_3=[-5 \quad 4 \quad -1]^T$, and $\tau=-25$. If $\mathbf{B}=(\mathbf{b}_1|\mathbf{b}_2|\mathbf{b}_3)$ and $\mathbf{S}=\mathbf{B}\mathbf{B}^T-25\mathbf{I}$, then

$$\mathbf{S}=\begin{pmatrix} 1 & 1 & -5 \\ 2 & 1 & 4 \\ 3 & -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ -5 & 4 & -1 \end{pmatrix}-25\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} 2 & -17 & 7 \\ -17 & -4 & 1 \\ 7 & 1 & -14 \end{pmatrix}$$

which has eigenvalues $\lambda_1=\|\mathbf{b}_1\|^2-25=-11$, $\lambda_2=\|\mathbf{b}_2\|^2-25=-22$, and $\lambda_3=\|\mathbf{b}_3\|^2-25=17$.

Example 8 illustrates Theorem 2 and Corollary 1. An orthogonal basis of $\mathbb{R}^3$ is given and then two nonnegative definite symmetric matrices are constructed, the first having two repeated zero eigenvalues and the second having just one zero eigenvalue.

EXAMPLE 8. Let $\mathbf{a}_1=[1 \quad 2 \quad 3]^T$, $\mathbf{a}_2=[3 \quad 0 \quad -1]^T$, and $\mathbf{a}_3=[-1 \quad 5 \quad -3]^T$. In order to use Theorem 2 to find a nonnegative definite matrix with two repeated zero eigenvalues, let $\mathbf{b}_1=\mathbf{a}_1$, $\mathbf{b}_2=\mathbf{0}$ and $\mathbf{b}_3=\mathbf{0}$ and define $\mathbf{B}=(\mathbf{b}_1|\mathbf{b}_2|\mathbf{b}_3)$. Then

$$\mathbf{S}=\mathbf{B}\mathbf{B}^T=\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$$

has eigenvalues $\lambda_1=\|\mathbf{b}_1\|^2=14$, $\lambda_2=\|\mathbf{b}_2\|^2=0$, and $\lambda_3=\|\mathbf{b}_3\|^2=0$. Note that if Corollary 1 is used to find $\mathbf{S}$, then $\mathbf{S}=\mathbf{a}_1\mathbf{a}_1^T$ which gives the same result.

In order to use Theorem 2 to find a nonnegative definite matrix with just one zero eigenvalue, let $\mathbf{b}_1=\mathbf{a}_1$, $\mathbf{b}_2=\mathbf{a}_2$ and $\mathbf{b}_3=\mathbf{0}$. Set $\mathbf{B}=(\mathbf{b}_1|\mathbf{b}_2|\mathbf{b}_3)$; then

$$\mathbf{S}=\mathbf{B}\mathbf{B}^T=\begin{pmatrix} 10 & 2 & 0 \\ 2 & 4 & 6 \\ 0 & 6 & 10 \end{pmatrix}$$

has eigenvalues $\lambda_1=\|\mathbf{b}_1\|^2=14$, $\lambda_2=\|\mathbf{b}_2\|^2=10$, and $\lambda_3=\|\mathbf{b}_3\|^2=0$. Using Corollary 1 instead, $\mathbf{S}$ is defined as $\mathbf{S}=(\mathbf{a}_1|\mathbf{a}_2)(\mathbf{a}_1|\mathbf{a}_2)^T$, which gives the same result.

Our last example illustrates Theorem 3; an orthogonal basis of $\mathbb{R}^3$ and a twice repeated least eigenvalue σ are given and a symmetric matrix $\mathbf{S}$ is constructed.

EXAMPLE 9. Let $\mathbf{b}_1=[1 \quad -1 \quad 1]^T$, $\mathbf{b}_2=[1 \quad 2 \quad 1]^T$, $\mathbf{b}_3=[1 \quad 0 \quad -1]^T$, and $\sigma=2$. Define $\mathbf{S}=\mathbf{b}_1\mathbf{b}_1^T+2\mathbf{I}$; then

$$\mathbf{S}=\begin{pmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

which has eigenvalues $\lambda_1=\|\mathbf{b}_1\|^2+2=5$, $\lambda_2=2$, and $\lambda_3=2$.

References

[1] G. Birkhoff and S. Mac Lane, A Survey of Modern Algebra, 4th ed., Macmillan, 1977.
[2] R. Bronson, Matrix Algebra: An Introduction, Academic Press, 1970.
[3] Daniel T. Finkbeiner, Introduction to Matrices and Linear Transformations, 3rd ed., Freeman, 1978.
[4] Felix R. Gantmacher, Theory of Matrices, Chelsea, 1959.
[5] D. C. Murdoch, Linear Algebra, John Wiley and Sons, New York, 1970.

A Number-Theoretic Sum

JOHN D. BAUM
Oberlin College
Oberlin, OH 44074

A topic which comes up in many elementary number theory texts, and one which I like to cover in a first course in number theory, is the Möbius inversion formula. Recall that the Möbius function $\mu(n)$ assigns -1, 0, or 1 to each positive integer n by the rule

$$\mu(n)=\begin{cases}1 & \text{if } n=1\\ 0 & \text{if } a^2|n,\text{ for some integer } a>1\\ (-1)^r & \text{if } n=p_1p_2\cdots p_r,\ p_i \text{ distinct primes.}\end{cases}$$

The Möbius inversion formula asserts that if $F(n)$ and $f(n)$ are functions with domain the set of positive integers, and range in the set of complex numbers, (i.e., number-theoretic functions), with $F(n)=\Sigma_{d|n}f(d)$ for every positive integer n, then $f(n)=\Sigma_{d|n}\mu(d)F(n/d)$, where $\Sigma_{d|n}$ indicates that the sum is to be taken over all divisors d of n [**1**, p. 245], [**2**, p. 88].

The number of exercises which use this formula and which are interesting is small, and it was in an attempt to add to this small number that I discovered the theorem of this note. First a bit of notation. As usual, (m,n) designates the greatest common divisor of integers m and n. If T is a set of positive integers, ΣT designates the sum of the elements of T (for example, $\Sigma\{2,4,7\}=13$). For positive integers n and k, define the following sets of positive integers:

$$R_k(n)=\{x^k\,|\,1\leqslant x\leqslant n,\ (x,n)=1\},$$
$$R'_k(n)=\left\{x^k\,|\,1\leqslant x\leqslant \frac{n}{2},\ (x,n)=1\right\},$$

where we write $R(n)$ for $R_1(n)$ and $R'(n)$ for $R'_1(n)$. We are interested in calculating the sum of the elements in these sets; accordingly we let $S_k(n)=\Sigma R_k(n)$ and $S'_k(n)=\Sigma R'_k(n)$. (We write $S(n)$ for $S_1(n)$ and $S'(n)$ for $S'_1(n)$.)

It is a nice exercise in the use of the Möbius inversion formula to calculate $S_2(n)$, and this calculation is the content of Exercises 9–14 on page 90 of [**2**]. It is not a difficult matter to extend these ideas to calculate $S_3(n)$, $S_4(n)$ and so on. Note that the number of elements in $R(n)$ is $\phi(n)$, the Euler ϕ-function, and that it is a simple matter to compute $S(n)$. For if $(x,n)=1$, then also $(n-x,n)=1$, and since the sum of x and $n-x$ is n, it follows that

$$S(n)=\frac{n\phi(n)}{2} \tag{1}$$

(see also [**1**, p. 150] or [**2**, p. 36]). It is natural to wonder if the calculation of $S'(n)$ might also be interesting. My first conjecture was that

References

[1] G. Birkhoff and S. Mac Lane, A Survey of Modern Algebra, 4th ed., Macmillan, 1977.
[2] R. Bronson, Matrix Algebra: An Introduction, Academic Press, 1970.
[3] Daniel T. Finkbeiner, Introduction to Matrices and Linear Transformations, 3rd ed., Freeman, 1978.
[4] Felix R. Gantmacher, Theory of Matrices, Chelsea, 1959.
[5] D. C. Murdoch, Linear Algebra, John Wiley and Sons, New York, 1970.

A Number-Theoretic Sum

JOHN D. BAUM
Oberlin College
Oberlin, OH 44074

A topic which comes up in many elementary number theory texts, and one which I like to cover in a first course in number theory, is the Möbius inversion formula. Recall that the Möbius function $\mu(n)$ assigns -1, 0, or 1 to each positive integer n by the rule

$$\mu(n)=\begin{cases}1 & \text{if } n=1\\ 0 & \text{if } a^2|n, \text{ for some integer } a>1\\ (-1)^r & \text{if } n=p_1p_2\cdots p_r, p_i \text{ distinct primes.}\end{cases}$$

The Möbius inversion formula asserts that if $F(n)$ and $f(n)$ are functions with domain the set of positive integers, and range in the set of complex numbers, (i.e., number-theoretic functions), with $F(n)=\Sigma_{d|n}f(d)$ for every positive integer n, then $f(n)=\Sigma_{d|n}\mu(d)F(n/d)$, where $\Sigma_{d|n}$ indicates that the sum is to be taken over all divisors d of n [**1**, p. 245], [**2**, p. 88].

The number of exercises which use this formula and which are interesting is small, and it was in an attempt to add to this small number that I discovered the theorem of this note. First a bit of notation. As usual, (m,n) designates the greatest common divisor of integers m and n. If T is a set of positive integers, ΣT designates the sum of the elements of T (for example, $\Sigma\{2,4,7\}=13$). For positive integers n and k, define the following sets of positive integers:

$$R_k(n)=\{x^k|1\leq x\leq n, (x,n)=1\},$$
$$R_k'(n)=\left\{x^k|1\leq x\leq \frac{n}{2}, (x,n)=1\right\},$$

where we write $R(n)$ for $R_1(n)$ and $R'(n)$ for $R_1'(n)$. We are interested in calculating the sum of the elements in these sets; accordingly we let $S_k(n)=\Sigma R_k(n)$ and $S_k'(n)=\Sigma R_k'(n)$. (We write $S(n)$ for $S_1(n)$ and $S'(n)$ for $S_1'(n)$.)

It is a nice exercise in the use of the Möbius inversion formula to calculate $S_2(n)$, and this calculation is the content of Exercises 9–14 on page 90 of [**2**]. It is not a difficult matter to extend these ideas to calculate $S_3(n)$, $S_4(n)$ and so on. Note that the number of elements in $R(n)$ is $\phi(n)$, the Euler ϕ-function, and that it is a simple matter to compute $S(n)$. For if $(x,n)=1$, then also $(n-x,n)=1$, and since the sum of x and $n-x$ is n, it follows that

$$S(n)=\frac{n\phi(n)}{2} \tag{1}$$

(see also [**1**, p. 150] or [**2**, p. 36]). It is natural to wonder if the calculation of $S'(n)$ might also be interesting. My first conjecture was that

$$S'(n) = \left(\frac{n}{2}\right)\phi\left(\frac{n}{2}\right)/2 = \frac{1}{8}n\phi(n).$$

Indeed, $S'(4) = 1 = \frac{1}{8}4\phi(4)$ gave confirmation for one example. But I found to my surprise, in trying other values of n, that although the sums $S(n)$, $S_2(n)$, and in general, $S_k(n)$, do not depend upon the form of n, the sum $S'(n)$ does. The residue modulo 4 of n determines the value of $S'(n)$. The method of calculation of $S'(n)$ is the content of the theorem below.

Several well-known facts that we shall use are collected in the following lemma. In this, and in the theorem, the function $\psi(n)$ is defined

$$\psi(n) = \Pi_{p|n}(1-p).$$

LEMMA. 1. $\phi(n) = n \cdot \Pi_{p|n}\left(1 - \frac{1}{p}\right)$, *the product being taken over all primes which divide n* [**1**, p. 139].

If $\mu(n)$ *is the Möbius function, then*

2. $\Sigma_{d|n}\mu(d)d = \psi(n)$ [**1**, p. 125]

3. $\Sigma_{d|n}\frac{\mu(d)}{d} = \frac{\phi(n)}{n}$ [**1**, p. 151].

Furthermore we take it as well known that $\Sigma_{i=1}^{n} i = n(n+1)/2$ and that $\Sigma_{i=1}^{n}(2i-1) = n^2$.

THEOREM. *Let n be an integer greater than* 2. *If* $n \equiv r$ mod 4, *where r is one of the residues*: -1, 0, 1, *or* 2, *then*

$$S'(n) = \frac{1}{8}\left[n\phi(n) - |r|\psi(n)\right].$$

Proof. CASE I. $n \equiv 0$ mod 4. Observe that if $(x, n) = 1$, then also $\left(x, \frac{n}{2}\right) = 1$ and conversely. $S'(n)$ therefore equals $S'\left(\frac{n}{2}\right)$, which by (1), equals $\frac{1}{2}\left(\frac{n}{2}\phi\left(\frac{n}{2}\right)\right)$. Since $n \equiv 0$ mod 4 it is easily seen that $\phi\left(\frac{n}{2}\right) = \frac{1}{2}\phi(n)$, so that $S'(n) = \frac{1}{8}n\phi(n)$.

CASE II. $n \equiv \pm 1$ mod 4. Let $N_d = \{x \mid 1 \leq x \leq (n-1)/2, (x, n) = d\}$. Now

$$\sum_{i=1}^{\frac{n-1}{2}} i = \sum_{d|n} \sum N_d. \tag{2}$$

Let x be a summand in ΣN_d; then $(x, n) = d$, whence $\left(\frac{x}{d}, \frac{n}{d}\right) = 1$. Since $1 \leq x \leq (n-1)/2$, it follows that $1 \leq x/d \leq (n-1)/2d < n/2d$. Thus x/d occurs as a summand in $S'\left(\frac{n}{d}\right)$, or equivalently, x occurs as a summand in $dS'\left(\frac{n}{d}\right)$. The argument works just as well conversely, so that

$$\sum N_d = dS'\left(\frac{n}{d}\right) = \frac{n}{d^*}S'(d^*), \text{ where } d^* = \frac{n}{d}.$$

Thus from (2),

$$n\sum_{d|n}\frac{S'(d)}{d} = \sum_{i=1}^{\frac{n-1}{2}} i = \frac{1}{2}\left(\frac{n-1}{2}\right)\left(\frac{n+1}{2}\right) = \frac{n^2-1}{8}$$

or

$$\frac{n - \frac{1}{n}}{8} = \sum_{d|n}\frac{S'(d)}{d}.$$

Note that this last result holds only if n is odd, but that is guaranteed by the hypothesis of this case. Applying the Möbius inversion formula to this last sum, we see that

$$\begin{aligned}\frac{S'(n)}{n} &= \frac{1}{8}\sum_{d|n}\mu(d)\left(\frac{n}{d}-\frac{d}{n}\right)\\ &= \frac{1}{8}\left(n\sum_{d|n}\frac{\mu(d)}{d}-\frac{1}{n}\sum_{d|n}d\mu(d)\right)\\ &= \frac{1}{8}\left(\phi(n)-\frac{\psi(n)}{n}\right),\end{aligned}$$

by 2. and 3. of the Lemma. Thus, $S'(n)=\frac{1}{8}(n\phi(n)-\psi(n))$.

CASE III. $n\equiv 2 \bmod 4$. In this case $n=2m$, where m is odd. Thus $(x,n)=1$ is equivalent to $(x,m)=1$ and x is odd. From this observation, we have

$$\begin{aligned}S'(n) &= \sum\{x\,|\,1\leqslant x\leqslant m,\,(x,m)=1,\,x \text{ odd}\}\\ &= \sum\{x\,|\,1\leqslant x\leqslant m,\,(x,m)=1\}-\{2x\,|\,1\leqslant 2x\leqslant m,\,(x,m)=1\}.\end{aligned}$$

In this case, $\phi(n)=\phi(m)$ and $\psi(n)=-\psi(m)$, and since m is odd we may apply Case II and (1) to our last equation to obtain

$$\begin{aligned}S'(n) &= \frac{m\phi(m)}{2}-2\left(\frac{1}{8}[m\phi(m)-\psi(m)]\right)\\ &= \frac{n\phi(n)}{4}-\frac{1}{4}\left(\frac{n}{2}\phi(n)+\psi(n)\right)\\ &= \frac{1}{8}(n\phi(n)-2\psi(n)).\end{aligned}$$

This completes the proof.

As an exercise, the reader may verify that computations similar to those in the above proof lead to the following formulae for $S_2'(n)$:

$$S_2'(n)=\begin{cases}\dfrac{n^2\phi(n)+2n\psi(n)}{24} & \text{if } n\equiv 0 \bmod 4\\[2ex] \dfrac{n^2\phi(n)-4n\psi(n)}{24} & \text{if } n\equiv 2 \bmod 4\\[2ex] \dfrac{n^2\phi(n)-n\psi(n)}{24} & \text{if } n\equiv \pm 1 \bmod 4.\end{cases}$$

We note by way of contrast the formula for $S_2(n)$ [**2**, p. 91]:

$$S_2(n)=\frac{2n^2\phi(n)+n\psi(n)}{6}.$$

Finally, it is interesting to observe that $S'(n)/S(n)$ tends to the limit $(1/2)^2$ as n tends to infinity and that $S_2'(n)/S_2(n)$ tends to $(1/2)^3$. This is what one might expect if the elements of $R(n)$ were uniformly distributed in the interval $0\leqslant x\leqslant n$.

The author wishes to thank both the editor and the referee for many helpful suggestions.

References

[1] David M. Burton, Elementary Number Theory, Allyn-Bacon, Boston, 1976.
[2] I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, 3rd ed., Wiley, New York, 1972.

PROBLEMS

Leroy F. Meyers, Editor
G. A. Edgar, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before August 1, 1982.

1140. Show that, if $0<x<y$, then

$$x+\ln(e^{y}-y-1)\geqslant x^{3/4}(x+y)^{1/4}+\ln y\left(\sqrt{x(x+y)}-x\right).$$

[*Mihály Bencze, Braşov, Romania.*]

1141. Let $X_1, X_2, \ldots, X_k$ be independent binomial random variables with parameters (n_1, p_1), $(n_2, p_2), \ldots, (n_k, p_k)$, respectively. Let $S = X_1 + X_2 + \cdots + X_k$, $n = n_1 + n_2 + \cdots + n_k$, and $p = \max(p_1, p_2, \ldots, p_k)$. If j is a fixed integer chosen from $0, 1, 2, \ldots, n$, show that $P[S \geqslant j] \leqslant P[X \geqslant j]$, where X is a binomial random variable with parameters (n, p). (Recall that a binomial random variable with parameters (n, p) is a random variable X taking only the values $0, 1, \ldots, n$ and satisfying $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$.) [*Joe Dan Austin, Rice University.*]

1142. If a two-terminal series-parallel circuit built with n unit resistors has total resistance p/q, where $(p, q) = 1$, prove that p does not exceed the $(n+1)$th Fibonacci number. [*Jeremy D. Primer, student, Princeton University.*]

1143. (a) Show that a ring R is a Boolean ring if and only if $x^{20} = x$ for all elements x of R.

(b*) For which positive integers k is it true that a ring R satisfying $x^k = x$ for all x in R must be Boolean? [*Thomas Hand, Indiana State University, Terre Haute.*]

ASSISTANT EDITORS: Daniel B. Shapiro *and* William A. McWorter, Jr., *The Ohio State University. We invite readers to submit problems believed to be new. Proposals should be accompanied by solutions, when available, and by any information that will assist the editors. Solutions to published problems should be submitted on separate, signed sheets. An asterisk (**) *will be placed by a problem to indicate that the proposer did not supply a solution. A problem submitted as a Quickie should be one that has an unexpected succinct solution. Readers desiring acknowledgement of their communications should include a self-addressed stamped card. Send all communications to this department to Leroy F. Meyers, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.*

Quickies

Solutions to Quickies appear at the conclusion of the Problems section.

Q671. Show that 111 cannot represent a square in any number base. [*L. Kuipers*, *Mollens*, *Switzerland*.]

Solutions

The Function Exists **January 1981**

1114. Prove or disprove: There exists a function f defined on $[-1,1]$ with f' continuous such that $\Sigma_{n=1}^{\infty} f(1/n)$ converges but $\Sigma_{n=1}^{\infty}|f(1/n)|$ diverges. (This is a relaxing of the condition "f'' continuous" to "f' continuous" in Problem 1060, this MAGAZINE, January 1979 and March 1980.) [*Robert Clark*, *student*, *Temple University*.]

Solution: The function defined by $g(x)=(x\sin(\log x))/\log x$ for $x>0$ and $g(0)=0$ is continuously differentiable on $[0,\frac{1}{2}]$. Let f be any continuously differentiable extension of g to $[-1,1]$, and let $F(x)=f(1/x)$ for $0<x\leqslant 1$. We show that $\Sigma_{n=1}^{\infty} f(1/n)=\Sigma_{n=1}^{\infty} F(n)$ converges, but not absolutely.

For $N\geqslant 2$, the Euler-Maclaurin summation formula

$$\sum_{n=2}^{N} F(n)=\int_2^N F(x)\,dx+\frac{1}{2}(F(2)+F(N))+\int_2^N\left(x-[x]-\frac{1}{2}\right)F'(x)\,dx$$

is applied to $F(x)=(\sin(\log x))/(x\log x)$. Now $F(N)\to 0$ as $N\to\infty$, and the substitution $u=\log x$ transforms the first integral on the right into $\int_{\log 2}^{\log N}\frac{\sin u}{u}\,du$, which is known to converge as $N\to\infty$. Furthermore, the absolute value of the integrand in the second integral on the right is easily seen not to exceed $3/x^2$ for $x\in[2,\infty)$, and so this second integral also converges as $N\to\infty$. It follows that $\Sigma_{n=1}^{\infty}F(n)$ converges.

For each positive integer k let n_k and m_k be integers (they exist!) such that

$$\begin{gathered}\left(k+\frac{1}{6}\right)\pi<\log(n_k-1)<\log n_k<\left(k+\frac{1}{3}\right)\pi,\\ \left(k+\frac{2}{3}\right)\pi<\log m_k<\log(m_k+1)<\left(k+\frac{5}{6}\right)\pi\end{gathered} \tag{1}$$

so that $|\sin(\log x)|>1/2$ for $x\in[n_k,m_k]$. Let $A_k=\Sigma_{n=n_k}^{m_k}|F(n)|$. Then

$$\sum_{n=1}^{\infty}|F(n)|\geqslant\sum_{k=1}^{\infty}A_k \tag{2}$$

and

$$\begin{aligned}A_k&>\frac{1}{2}\sum_{n=n_k}^{m_k}\frac{1}{n\log n}>\frac{1}{2}\int_{n_k}^{m_k}\frac{1}{x\log x}\,dx=\frac{1}{2}\log\frac{\log m_k}{\log n_k}\\ &>\frac{1}{2}\log\frac{\left(k+\frac{2}{3}\right)\pi}{\left(k+\frac{1}{3}\right)\pi}=\frac{1}{2}\log\left(1+\frac{1}{3k+1}\right)\geqslant\frac{1}{2}\log\left(1+\frac{1}{4k}\right)>\frac{1}{2}\cdot\frac{1}{8k}.\end{aligned}$$

Since $\Sigma_{k=1}^{\infty} 1/(16k)$ is known to diverge, it follows from (2) that $\Sigma_{k=1}^{\infty} |F(n)|$ diverges.

CHICO PROBLEM GROUP
California State University

Also solved by George Bridgman and the proposer, and partially by Victor Hernandez (Spain).

Elliptical Probability **January 1981**

1115. Let D be the disc $x^2+y^2<1$. Let points A and B be selected at random in D. Find the probability that the open disc whose center is the midpoint of $\overline{AB}$ and whose radius is $AB/2$ is a subset of D. [*Roger L. Creech, East Carolina University.*]

Solution: The answer is 2/3. After selecting the two points, orient the axes so that the positive x-axis passes through A, and call the coordinates of A and B $(r,0)$ and (x,y), respectively. If we label the event in the problem E, then $\Pr(E)=\int_0^1 f(r)\Pr(E|r)\,dr$ where $f(r)$ is the density function for the abscissa of A. But

$$f(r)=\frac{d}{dr}\Pr(OA\leqslant r)=\frac{d}{dr}\left(\frac{\pi r^2}{\pi}\right)=2r.$$

To find $\Pr(E|r)$, we first need to describe the set of points B in D such that the open disc whose center is the midpoint of $\overline{AB}$ and whose radius is $AB/2$ is a subset of D. But this is precisely Problem 1092 (this MAGAZINE, January 1980 and March 1981), and that set consists of the points on or within the ellipse $x^2+y^2/(1-r^2)=1$, with the exception of the points $(\pm 1,0)$. So $\Pr(E|r)$ is the ratio of the area of this ellipse to the area of D, or $\Pr(E|r)=\sqrt{1-r^2}$. Hence we have $\Pr(E)=\int_0^1 2r\sqrt{1-r^2}\,dr=2/3$.

ROGER B. NELSEN
Lewis and Clark College

Also solved by J. E. Chance, Chico Problem Group of California State University, Delmar Crabill, Delmar Crabill & Lawrence Ringenberg, Jordi Dou (Spain), P. Ehlers (Canada), Milton P. Eisner, Edilio Escalona & Adolfo Quiroz (Venezuela), Nick Franceschine III, John A. Gillespie, Klaus Grünbaum (Denmark), Victor Hernandez (Spain), G. A. Heuer, J. R. Hilditch (England), Dana L. Mabbott, Jerry Metzger, W. W. Meyer, Bruce L. Montgomery, Robert Patenaude, Michael I. Ratliff, C. Ray Rosentrater, Ken Yocom, and the proposer.

In several incorrect solutions it was assumed that the abscissa of A rather than A itself is distributed uniformly. Ehlers and Meyer stated generalizations to higher dimensions.

Answers

Solutions to the Quickies which appear near the beginning of the Problems section.

Q671. If $g^2=a^2+a+1$, then $(a+\frac{1}{2})^2<g^2<(a+1)^2$, or $a+\frac{1}{2}<g<a+1$, which is impossible in positive integers.

REVIEWS

Paul J. Campbell, Editor
Beloit College

Pierre J. Malraison, Jr., Editor
MDSI, Ann Arbor

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

Golomb, Solomon W., *Rubik's cube and a model of quark confinement*, American Journal of Physics 49 (November 1981) 1030-1031.

Designating a +120° rotation ("twist") of a corner cell as a (mathematical) quark, and a -120° rotation as an antiquark, the same restrictions that confine physical quarks are observed: isolated quarks cannot arise, and combinations can occur only if the excess of one type over the other is a multiple of three.

Stigler, Stephen M., *Stigler's Law of Eponymy*, Transactions of the New York Academy of Sciences (II) 39 (24 April 1980) 147-158.

The Law, in its simplest form: "No scientific discovery is named after its original discoverer." Explanations, with examples mostly from mathematics and statistics, are accompanied by a statistical study of the names used for the "normal" distribution.

Stigler, Stephen, *Gauss and the invention of least squares*, Annals of Statistics 9 (1981) 465-474.

The most famous priority dispute in statistics was between Gauss and Legendre over discovery of the method of least squares. New evidence, both documentary and statistical, is presented that suggests Gauss had the method earlier but did not succeed in impressing contemporaries with its importance as Legendre did.

Pavelle, Richard, *et al.*, *Computer algebra*, Scientific American 245:6 (December 1981) 136-152, 186.

Exhibits strengths and successes of automatic systems of algebraic manipulation, with some hint at how computers are programmed to carry out the manipulations.

Kahneman, Daniel, and Tversky, Amos, *The psychology of preferences*, Scientific American 246:1 (January 1982) 160-173, 180.

Offers mathematical description and psychological explanation of apparently paradoxical behavior in making risky choices: subjective over-valuation of small probabilities, preference for risk-taking reverses from gains to losses, and regret over action taken is more intense than over inaction with the same undesirable result.

Hofstadter, Douglas R., *Metamagical Themas: Strange attractors: mathematical patterns delicately poised between order and chaos*, Scientific American 245:5 (November 1981) 22-43, 202.

A simple parabola provides the setting for remarkable recent discoveries about nonlinear systems, such as turbulent flow, erratic population fluctuations in predator-prey systems, and instability of laser modes. A surprising feature is that different curves all lead to the same "universal" numerical constants, including one called Feigenbaum's number and approximately equal to 4.669. See also "Feigenbaum's number," *Scientific American* 245:3 (October 1981) 116-118.

Gardner, Martin, *Mathematical Games: The Laffer curve and other laughs in current economics*, Scientific American 245:6 (December 1981) 18-31C, 186.

In his last column for *Scientific American*, Gardner returns to debunking pseudoscience, this time in economics. "What do professional economists make of supply-side theory? Most of them, including the most conservative, regard it in much the same way as astronomers regard the theories of Immanuel Velikovsky....As the Yale economist William Nordhaus put it....'We can only hope that supply-side economics turns out to be laetrile rather than thalidomide.'"

Morris, Scott, *Interview: Martin Gardner*, Omni 4:4 (January 1982) 66-69, 80-86.

A curtain call as Gardner retires from the stage of his *Mathematical Games* column. What does he think about teaching creative thinking? What does he want to be remembered for? Fans will also appreciate his remarks on pseudoscience, whose debunking is one of his pastimes.

Kim, Scott, Inversions: A Catalog of Calligraphic Cartwheels, BYTE Books, 1981; 122 pp, $8.95(P).

"Like most books, this book is made of words. The words in this book are meant to be seen and not read. Each word has a special visual trick, something to fool your eye. Some of the words read the same upside down, some read the same in a mirror, some repeat off to infinity....Always I have tried to make the style of the lettering reflect the meaning of the word. Thus the form echoes the content."

Finerty, James Patrick, The Population Ecology of Cycles in Small Mammals, Yale U Pr, 1980; xiv + 234 pp, $18.50.

Predator-prey models are a staple of modeling and differential equations courses, and the interaction between lynx and the snowshoe rabbit is often cited as an application. Finerty thoroughly investigates the data and theories for periodicity in those and related animal populations. Results: strong evidence for a ten-year cycle of snowshoe rabbits, muskrats, and their predators, and for a four-year cycle of lemmings and their predators. He uses time-series, Fourier, and path analyses to investigate why; but concludes that the theoretical tools "are far in excess of what we know."

Cavalli-Sforza, L.L., and Feldman, M.W., Cultural Transmission and Evolution: A Quantitative Approach, Princeton U Pr, 1981; xiv + 388 pp, $25, $10.50(P).

The authors classify and systematize modes of transmitting "culture" and develop a mathematical theory of nongenetic transmission of cultural traits, employing calculus, probability, linear algebra, and statistics. A subsequent volume will take account of individual, inherited differences in learning ability.

Konheim, Alan G., Cryptography: A Primer, Wiley, 1981; xiv + 432 pp, $36.95.

Splendid up-to-date introduction to the mathematical methods of cryptography; there is no comparable book. Applications to the Data Encryption Standard (DES), public key systems, and digital signature authentication are all covered, in addition to the standard topics of substitution ciphers, rotor systems, and block ciphers. Not covered are transposition ciphers.

Herman, Gabor T., Image Reconstruction from Projections: The Fundamentals of Computerized Tomography, Academic Pr, 1980; xii + 316 pp, $29.50.

Basis for the author's two-semester graduate computer science course on image reconstruction. Emphasis is on the algorithms for reconstruction; computer science undergraduates should note the essential use of vector calculus, linear algebra, probability, and especially real analysis. Lots of illustrations and real data keep the treatment tied to reality.

Committee on the Undergraduate Program in Mathematics, Recommendations for a General Mathematical Sciences Program, MAA, 1981; 102 pp, (P).

Not a prescription for "mathematics for the 80s," but guidelines to a flexible refocusing of attention toward mastery of the in-depth reasoning and useful tools needed for a career of different jobs and continuing education. The main panel's report offers a program philosophy, advice on how to teach mathematical reasoning and how much theory to teach, sample majors, and new course descriptions for discrete methods, applied algebra, and numerical analysis. Separate subpanels report on calculus, core mathematics, computer science, modeling and operations research, and statistics.

Frauenthal, James C., Smallpox: When Should Routine Vaccination Be Discontinued?, UMAP Expository Monograph Series, Birkhäuser, 1981; xii + 50 pp, $7.95(P).

The author builds mathematical models for the stages of a smallpox epidemic, then analyzes them to examine the tradeoff between competing risks of death from vaccination and infection. The result is satisfying in its simplicity--only discrete probability models (Poisson, exponential) are used--and in its significance.

UMAP Modules 1977-1979: Tools for Teaching, Birkhäuser, 1981; xii + 727 pp, $35.

First of an annual series of collections of self-contained, lesson-length instructional units in undergraduate mathematics and its applications. The 27 modules here pertain mostly to calculus and mathematical modeling, and application fields include medicine, economics, politics, psychology, geoscience, and computing.

Hinton, Charles H., Speculations on the Fourth Dimension: Selected Writings of Charles H. Hinton, with an introduction by Rudolf v.B. Rucker, Dover, 1980; xix + 204 pp, $4(P).

Popular essays by a 19th-century mathematician whose aim in life was to teach everyone to see 4-D space and who anticipated some of Einstein's discoveries.

Hunter, J.A.H., Challenging Mathematical Teasers, Dover, 1980; 101 pp, $2.75(P).

A new collection of 100 elementary problems from Hunter's newspaper column "Fun with Figures," plus 40 new alphametics.

NEWS & LETTERS

AWARD FOR DISTINGUISHED SERVICE

The MAA's 1982 Award for Distinguished Service honors Dr. Thornton C. Fry for his many contributions as an industrial mathematician. The citation, prepared by G. Baley Price, was read at the Business meeting of the MAA in Cincinnati, on January 17, 1982, ten days after Dr. Fry's 90th birthday. Dr. Fry's very active mathematical career included industrial research in war problems during both world wars, and the establishment and leadership of the mathematics research group at Western Electric which eventually grew into the present Mathematics and Research Center of Bell Laboratories. The full text of the citation appears in the February 1982 *American Mathematical Monthly*.

ELECTION REVERSAL

James Ward's "Probability of Election Reversal" (this *Magazine*, November 1981) omits an important consideration. The probability of reversal when an election is repeated is very much larger than his estimates, because there can be a change in circumstances between the original election and the repeat election. Voters may change their minds or people who didn't think their votes would count for much in the first election may decide to vote.

If we knew a probability distribution of events that might shift n votes, we could compute the probability of a shift of given size when an election is held. I would conjecture that the expected shift from an event influencing public opinion is ordinarily much larger than the expected shift from an accidental bias on the part of voters casting invalid ballots. This makes holding repeat elections unfair to the winner of the original election if the court or other body that may require the repeat may be biased in favor of the original loser. An extreme unfairness would occur if the side favored by election authorities could always demand a revote if the results were at all close. Therefore, there needs to be a fixed rule stated before the election to determine how many invalid ballots are needed to require a new election.

John McCarthy
Stanford University
Stanford, CA 94305

The author replies:

Election reversal as defined in my article (this *Magazine*, November 1981) means reversal of the results of an election by the removal of invalid ballots, not reversal of the outcome of one election by the outcome of a second. There is ample historical evidence to show that the outcome of a second election may very well be different. Professor McCarthy cites some possible reasons for this phenomenon. Their essence is that voters react to the outcome of the first election and to the very fact that a new election is held. Because a second election is a different election, held under different circumstances at a different time, legal bodies wishing to use the formula given in the article should probably decide that a rather large probability of reversal of an election is necessary to warrant a new one. And, of course, to be fair to all sides, this decision should be made before the first election. However, in the last analysis, such decisions are political rather than mathematical ones.

James E. Ward
Bowdoin College
Brunswick, ME 04011

GRAPHICAL CONSTRUCTION

The major part of the Note "Making Connections: A Graphical Construction" by Dorothy Wolfe (this *Magazine*, November 1981) is devoted to a construction that will form the adjacency matrix of a graph having a given sequence satisfying her condition as its degree sequence. Such a construction is of interest because it provides another proof of the Erdös-Gallai theorem. Wolfe's construction, however, is not described for an arbitrary degree sequence but rather for a particular example. Although her algorithm is sufficient for the example she considers, it does not work in all cases. Specifically, the second strategy described on pp. 254-55 (the one dealing with the case when all the inequalities in her condition (1) are strict) is not always adequate. For example, if her procedure is applied to a Ferrers diagram F with degree sequence (6,6,5,4,4,4,4,1), the resulting derived Ferrers diagram F^* has degree sequence $D^* = (6,6,4,3,3,3,1)$. For this diagram $d_1^* = d_2^* = 6$, $k_1^* = 7$, and $k_2^* = 6$, so that $d_1^* + d_2^* > k_1^* + k_2^* - 2$ in violation of condition (1).

I offer an algorithm that remedies the defect in Wolfe's construction; the notation and terminology that I use are essentially the same as hers. My algorithm differs from hers principally in routine B (her second strategy), but I have also made a minor change in her first strategy (my routine A) so that after my routines are used the size of the unfilled submatrix of the adjacency matrix is $k_1^* \times k_1^*$, where k_1^* is the number of rows in the derived Ferrers diagram F^*.

Algorithm for Constructing the Adjacency Matrix Corresponding to a Ferrers Diagram

The construction process uses two routines to form rows (and the corresponding symmetric columns) of the adjacency matrix. After using each routine the current Ferrers diagram is changed into a new diagram with which the construction process can be continued. In the course of modifying a Ferrers diagram, dots will be moved to the right into new positions in the same row. If, in the current Ferrers diagram F,

$$\sum_1^r d_i < \sum_1^r (k_i - 1) \text{ for } 1 \leq r \leq h,$$

use routine B; otherwise use routine A.

ROUTINE A. Let $r \leq h$ denote the largest integer for which

$$\sum_1^r d_i = \sum_1^r (k_i - 1).$$

Modify F as described in steps (a), (b), and (c) below.

(a) Move the dot at $a_{r,r}$ into the first empty position in row r, that is, into position $a_{r,s}$, where $s = d_r + 1$.

(b) Note that condition (1) and the definition of r imply that the number d_r of dots in row r is not less than the number $k_r - 1$ of dots now in column r. If $d_r = k_r - 1$, we can skip directly to the construction following step (c). Otherwise, if $d_r > k_r - 1$, we will move $d_r - k_r + 1$ dots into column r from columns to the left. These dots will be taken, one dot per row, from row m or below, where m is the smallest integer for which $d_m < d_{k_r}$. (The existence of such dots follows from the fact that $k_1 - 1 \geq d_1 \geq d_r$.) The dots will be moved first from column d_m and then, if necessary, from columns to the left of d_m; in each column the lowermost dot should be moved before a dot in a higher row.

(c) Make row r symmetric with column r by moving dots as necessary in row r.

Now use row r of the modified Ferrers diagram to construct row r and column r of the unfilled submatrix of the adjacency matrix by substituting 1's for dots and 0's for blanks. In addition, if in step (b) any dots from rows containing only one dot were moved into column r, fill in the corresponding rows and columns of the adjacency matrix with 0's. Finally, form a new Ferrers diagram F^* by deleting row r and column r from the modified version of F and closing up the spaces left by the deleted dots.

ROUTINE B. Modify F as described in steps (a), (b), and (c) below.

(a) Move the dot at $a_{h,h}$ into the first empty position in row h.

(b) If $k_h \leq d_h + 1$, proceed as in routine A with $r = h$. (Routine A works in the case $k_h = d_h + 1$ because both row h and column h contain d_h dots after step (a).) Otherwise, if $k_h \geq d_h + 2$, let s denote the largest integer such that $d_s > d_h$, or if $d_1 = d_h$, set $s = 0$. Retain the first s dots at the top of column h, and the last $d_h - s$ dots at the bottom of column h, and move the remaining dots in column h to the first empty position in their respective rows.

(c) Make row h symmetric with column h by moving dots in row h as necessary.

Now construct row h and column h of the unfilled submatrix of the adjacency matrix by substituting 1's for dots and 0's for blanks using row h of the modified Ferrers diagram. Then form a new Ferrers diagram F^* by deleting row h and column h from the configuration produced by step (c) and closing up the spaces left by the deleted dots.

This algorithm, when applied to the degree sequence (6,6,5,4,4,4,4,1) mentioned above, produces the adjacency matrix

$$\begin{pmatrix} 0&1&1&1&1&1&1&0\\ 1&0&1&1&1&1&0&1\\ 1&1&0&1&1&0&1&0\\ 1&1&1&0&0&0&1&0\\ 1&1&1&0&0&1&0&0\\ 1&1&0&0&1&0&1&0\\ 1&0&1&1&0&1&0&0\\ 0&1&0&0&0&0&0&0 \end{pmatrix}$$

Lawrence E. Spence
Illinois State Univ.
Normal, IL 61761

Editor's note: Limitation of space did not permit publication of the justification of Spence's algorithm or his step-by-step computer printout of its application to the example mentioned.

MUSEUM SHOW

An exhibition, "Creativity--The Human Resource," developed by Chevron will tour several museums in the next six months. The show opened at the Miami Museum of Science, and is currently at the New Orleans Museum of Art (until May 16). It is scheduled for the Atlanta Memorial Arts Center (June 19-August 15) and the Fort Worth Museum of Science and History (September 11-November 28). The creative process is discussed on videotape by several contemporary American scientists and artists, and visitors to the exhibit are encouraged to discover their own creativity through interactive computer games. One question posed is as follows.

Draw a square. Inside it place 25 dots spread out in five rows. Draw lines through the dots in such a way as to divide the square into four equal parts. In how many ways can this be done? (The answer supplied is 104.)

42ND ANNUAL WILLIAM LOWELL PUTNAM COMPETITION (DECEMBER 5, 1981)

A-1. Let $E(n)$ denote the largest integer k such that 5^k is an integral divisor of the product $1^1 2^2 3^3 \ldots n^n$. Calculate

$$\lim_{n\to\infty} \frac{E(n)}{n^2}.$$

A-2. Two distinct squares of the 8 by 8 chessboard C are said to be adjacent if they have a vertex or side in common. Also, g is called a C-gap if for every numbering of the squares of C with all the integers 1,2,...,64 there exist two adjacent squares whose numbers differ by at least g. Determine the largest C-gap g.

A-3. Find

$$\lim_{t\to\infty} [e^{-t} \int_0^t \int_0^t \frac{e^x - e^y}{x - y} \, dx \, dy]$$

or show that the limit does not exist.

A-4. A point P moves inside a unit square in a straight line at unit speed. When it meets a corner it escapes. When it meets an edge its line of motion is reflected so that the angle of incidence equals the angle of reflection.

Let $N(T)$ be the number of starting directions from a fixed interior point P_0 for which P escapes within T units of time. Find the least constant a for which constants b and c exist such that

$$N(T) \leq aT^2 + bT + c$$

for all $T > 0$ and all initial points P_0.

A-5. Let $P(x)$ be a polynomial with real coefficients and form the polynomial

$$Q(x) = (x^2 + 1)P(x)P'(x) + x([P(x)]^2 + [P'(x)]^2).$$

Given that the equation $P(x) = 0$ has n distinct real roots exceeding 1, prove or disprove that the equation $Q(x) = 0$ has at least $2n - 1$ distinct real roots.

A-6. Suppose that each of the vertices of ΔABC is a lattice point in the (x,y)-plane and that there is exactly one lattice point P in the interior of the triangle. The line AP is extended to meet BC at E. Determine the largest possible value for the ratio of lengths of segments

$$\frac{|AP|}{|PE|}.$$

(A lattice point is a point whose coordinates x and y are integers.)

B-1. Find

$$\lim_{n\to\infty} \left[\frac{1}{n^5} \sum_{h=1}^{n} \sum_{k=1}^{n} (5h^4 - 18h^2k^2 + 5k^4)\right].$$

B-2. Determine the minimum value of

$$(r-1)^2 + \left(\frac{s}{r} - 1\right)^2 + \left(\frac{t}{s} - 1\right)^2 + \left(\frac{4}{t} - 1\right)^2$$

for all real numbers r, s, t with $1 \leq r \leq s \leq t \leq 4$.

B-3. Prove that there are infinitely many positive integers n with the property that if p is a prime divisor of $n^2 + 3$ then p is also a divisor of $k^2 + 3$ for some integer k with $k^2 < n$.

B-4. Let V be a set of 5 by 7 matrices, with real entries and with the property that $rA + sB \in V$ whenever $A, B \in V$ and r and s are scalars (i.e., real numbers). Prove or disprove the following assertion: If V contains matrices of ranks 0, 1, 2, 4, and 5 then it also contains a matrix of rank 3. (The rank of a nonzero matrix M is the largest k such that the entries of some k rows and some k columns form a k by k matrix with a nonzero determinant.)

B-5. Let $B(n)$ be the number of ones in the base two expression for the positive integer n. For example, $B(6) = B(110_2) = 2$ and $B(15) = B(1111_2) = 4$. Determine whether or not

$$\exp\left(\sum_{n=1}^{\infty} \frac{B(n)}{n(n+1)}\right)$$

is a rational number. Here $\exp(x)$ denotes e^x.

B-6. Let C be a fixed unit circle in the Cartesian plane. For any convex polygon P each of whose sides is tangent to C, let $N(P,h,k)$ be the number of points common to P and the unit circle with center at (h,k). Let $H(P)$ be the region of all points (x,y) for which $N(P,x,y) \geq 1$ and $F(P)$ be the area of $H(P)$. Find the smallest number u with

$$\frac{1}{F(P)} \iint N(P,x,y) \, dx \, dy < u$$

for all polygons P, where the double integral is taken over $H(P)$.

THE CHAUVENET PAPERS

A Collection of Prize-Winning Expository Papers in Mathematics

James C. Abbott, editor

This two-volume collection of the twenty-four prize winning Chauvenet Papers contains the finest collection of expository articles in mathematics ever assembled!

THE CHAUVENET PRIZE for special merit in mathematical exposition was first awarded by the Mathematical Association of America in 1925 to G. A. Bliss. Since that time, twenty-four Chauvenet Prizes have been awarded and the Prize has become the Association's most prestigious award for mathematical exposition. The list of authors is a veritable WHO's WHO of mathematical expositors, and contains some of the more prominent mathematicians of the past fifty years.

Clearly written, well organized, expository in nature, these papers are the jewels of mathematical writing during our times. They were selected by peer juries of experts judged for their presentation as well as their content. They are a sheer joy to read for those who delight in the beauty of mathematics alone.

Volume I—xviii + 312 pages. Hardbound. List: $21.00—MAA Member: $16.00

Volume II—vii + 282 pages. Hardbound. List: $21.00—MAA Member: $16.00

Special Package Price for Both Volumes: List $36.00—MAA Member: $27.00

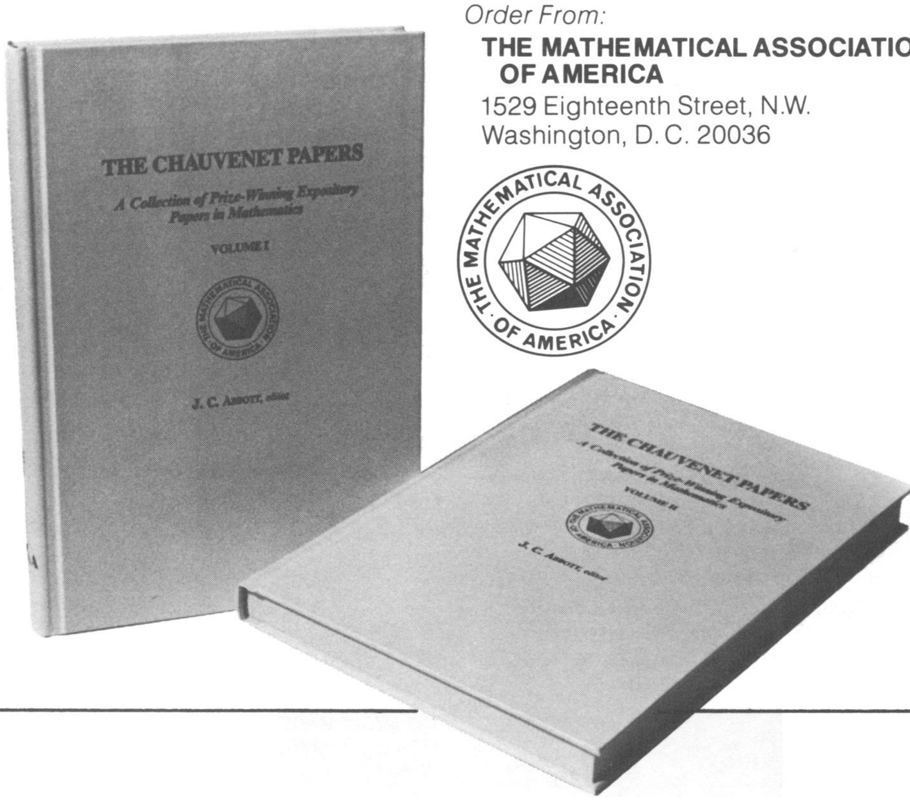
Order From:

THE MATHEMATICAL ASSOCIATION OF AMERICA

1529 Eighteenth Street, N.W.

Washington, D.C. 20036

THE MATHEMATICAL ASSOCIATION OF AMERICA



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

MATHEMATICS MAGAZINE VOL. 55, NO. 2, MARCH 1982