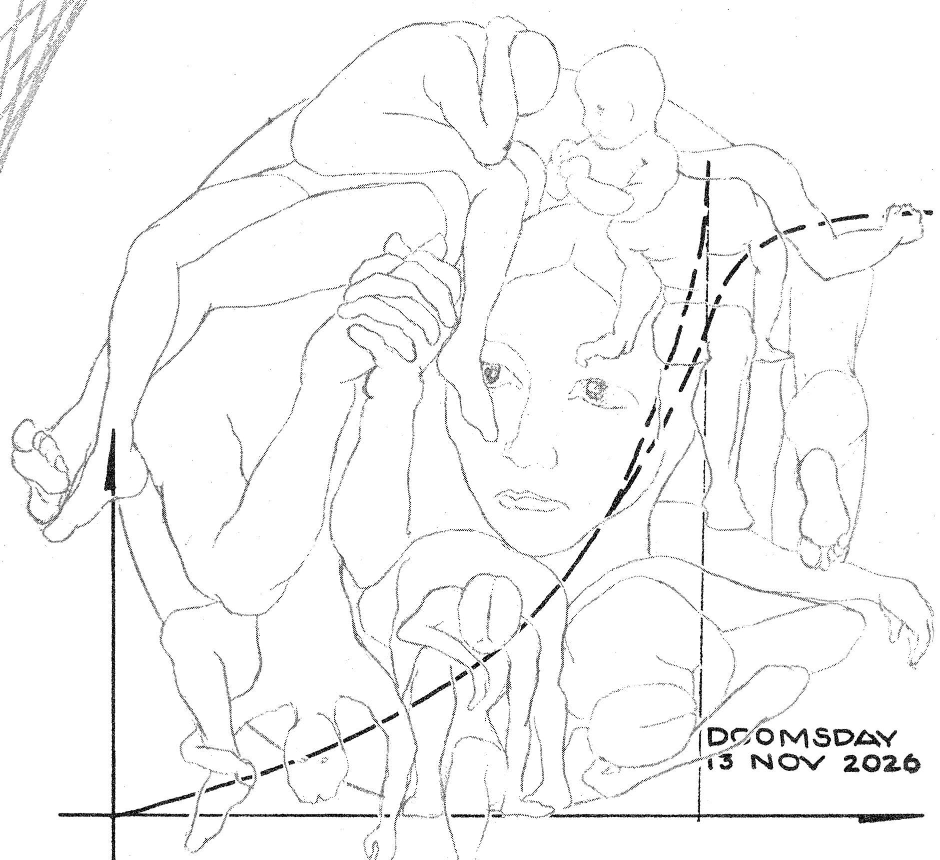
MATHEMATICS MAGAZINE



Vol. 50, No. 4
September, 1977
CODEN: MAMGA8

POPULATION MODELS • DISTRIBUTIONS OF DIGITS
PREVENTION VS PENALTY • LINEAR INEQUALITIES

MATHEMATICS MAGAZINE

Volume 50, Number 4, September 1977

COVER: Which way should the graph go? Some models of world population growth predict a stable maximum while others indicate an explosion. Vigorous discussion of alternative models (see page 186) illustrates the complexity of mathematical modelling in an area of global concern.

ARTICLES

NOTES

PROBLEMS

REVIEWS

NEWS AND LETTERS

EDITORIAL POLICY

Mathematics Magazine is a journal of collegiate mathematics designed to enrich undergraduate study of the mathematical sciences. The *Magazine* should be an inviting, informal journal emphasizing good mathematical exposition of interest to undergraduate students. Manuscripts accepted for publication in the *Magazine* should be written in a clear and lively expository style. The *Magazine* is not a research journal, so papers written in the terse "theorem-proof-corollary-remark" style will ordinarily be unsuitable for publication. Articles printed in the *Magazine* should be of a quality and level that makes it realistic for teachers to use them to supplement their regular courses. The editors especially invite manuscripts that provide insight into applications and history of mathematics. We welcome other informal contributions, for example, brief notes, mathematical games, graphics and humor.

Editorial correspondence should be sent to: Mathematics Magazine, Department of Mathematics, St. Olaf College, Northfield, Minnesota 55057. 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; the printers will insert printed letters on the illustration in the appropriate locations.

Authors planning to submit manuscripts may find it helpful to obtain the more detailed statement of guidelines available from the editorial office.

ABOUT OUR AUTHORS

Martin H. Pearl and Alan J. Goldman ("Policing the Market Place") are mathematicians with the National Bureau of Standards. Pearl holds a Ph.D. in mathematics from the University of Wisconsin and is now a professor at the University of Maryland. After writing a Ph.D. thesis in topology at Princeton, Goldman joined the Applied Mathematics Division of NBS where he is currently chief of the Operations Research Section. Pearl and Goldman's work in inspection-system performance was sponsored by the Office of Weights and Measures of NBS, which acts as a technical secretariat for the almost 800 regional weights and measures agencies throughout the country.

David A. Smith ("Human Population Growth: Stability or Explosion?") was born in New York City in 1938 and received a Ph.D. in algebra from Yale University in 1963. He has been at Duke University since 1963, except for a sabbatical year (1975–76) at Case Western Reserve University, during which the present article was written. Professor Smith's interest in population dynamics was stimulated by research for his book, *Interface: Calculus and the Computer* (Houghton Mifflin, 1976). The present article grew from notes prepared for a calculus/differential equations class of Western Reserve sophomores.

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. Ordinary subscriptions are $12 per year. Members of the Mathematical Association of America or of Mu Alpha Theta may subscribe at special reduced rates. Colleges and university mathematics departments may purchase bulk subscriptions (5 or more copies to a single address) for distribution to undergraduate students.

Subscription correspondence and notice of change of address should be sent to A. B. Willcox, Executive Director, Mathematical Association of America, Suite 310, 1225 Connecticut Avenue, N.W., Washington, D.C. 20036. Back issues may be purchased, when in print, from P. and H. Bliss Co., Middletown, CT 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

Policing the Market Place

A simple game-theoretic model yields insight into the optimal use of available resources for inspecting commercial measuring devices.

MARTIN H. PEARL
ALAN J. GOLDMAN
National Bureau of Standards
Washington, D.C. 20234

A retail establishment containing a measuring device which meters its transactions with customers has an economic incentive to "cheat" in connection with that device. The ugly word "cheat" is used here as an abbreviation for any of the following: deliberately causing the device to malfunction in the direction advantageous to the establishment ("short weighing"), or knowingly permitting such a malfunction (originating through natural causes) to go uncorrected, or employing such a malfunctioning device, which though not explicitly recognized as such, was not properly checked for malfunction. It is the responsibility of the Weights and Measures agencies of the various states to inspect these measuring devices regularly in order to minimize the loss to consumers due to such cheating. This paper presents a simple game-theoretic mathematical model of the weights-and-measures inspection process.

The model we will develop is based on several general observations. Different establishments with different dollar-flows of transactions will have economic incentives of different magnitudes. Moreover, the inspection agency typically has only a limited quantity of inspection resources (manpower, time, money, etc.) and this quantity is likely to be too small to cover all devices in its jurisdiction as often as it would like. Finally, we assume that the detection of cheating always leads to the imposition of a penalty on the cheater.

Our purpose is to further the design of more efficient responses by society to the existence of cheating. These responses include both the penalty P and the policing resources which we will represent by m. By analyzing how the level of illicit activity might depend on P and m, we hope to contribute to better understanding of the effectiveness of this two-element response. By exhibiting the tradeoffs between P and m, we hope to contribute to better understanding of the efficiency with which that response is allocated between the two elements. These particular objectives remain relevant in a broader context of regulatory and general criminal-justice activity.

The model described in this paper is based on the mathematical theory of games. For our purposes it is sufficient to restrict our attention to games with just two players: the inspection agency (personified as "the inspector") and the firms subject to inspection (consolidated and personified as the "inspectee"). The objective of the inspectee is to maximize his gain. This gain is the amount he obtains by cheating less the penalty he must pay when he is caught. On the other hand, the aim of the inspector is not as clear-cut. He may wish to minimize the inspectee's gain, or, he may wish to minimize the consumer's loss. Since the consumer does not profit directly from the penalty paid by the inspectee when he is caught cheating, these two situations may not be identical. Still another possible aim of the inspector might be to minimize undetected cheating, since someone who is caught cheating once is unlikely to cheat again. These various possibilities will be considered as options in the general model described below.

Inspector vs. Inspectee

Our mathematical model takes the form of a 2-player, zero-sum game. The players are the **inspector** (an aggregate representing the inspection agency) and the **inspectee** (an aggregate representing the establishments in which the measuring devices are used). (While there might appear to be some loss in reality through regarding these establishments as forming a single player with a single interest, it can be shown that for the present model this does not affect the solution.) The inspectee maintains a set of devices, $D_1, D_2, \dots, D_n$, on each of which he can cheat or not. The inspector selects devices for inspection up to the limit of his resources. We assume that if a device on which cheating takes place is inspected then detection is certain. The data for the model are:

n = number of devices;

V_i = the payoff to the inspectee from cheating on device D_i;

P = the penalty levied against the inspectee for each detection of cheating;

m = the number of devices the inspector can examine ($m < n$).

A strategy c for the inspectee can be described by the n-tuple $c = (c_1, c_2, \dots, c_n)$ in which $c_i = 1$ if there is cheating on device D_i, and $c_i = 0$ if there is no cheating on D_i. Analysis in [**1**] justifies the natural step of defining a strategy for the inspector as a rule which tells him how often (i.e., with what probability) each device is to be inspected, subject only to the requirement that he not exceed his resources. Thus, strategy for the inspector is defined by specifying an n-tuple of probability values, $p = (p_1, p_2, \dots, p_n)$, $0 \leqq p_i \leqq 1$, where p_i is the probability that device D_i will be inspected. Since m is the total amount of inspection resources (and the inspector cannot improve his performance by using less than all of his resources), we must have $\Sigma_{i=1}^{n} p_i = m$.

The net expected payoff to the inspectee from device D_i is 0 if there is no cheating on that device (i.e., $c_i = 0$). If there is cheating on D_i (i.e., $c_i = 1$) then the inspectee's net expected payoff is given by the gain from cheating (V_i) minus the expected penalty (the amount P of the penalty multiplied by the probability p_i of inspection). These two possibilities are simultaneously represented by the formula $V_ic_i - Pp_ic_i = [V_i - Pp_i]c_i$. Thus the total net expected payoff to the inspectee, if the two players choose respective strategies c and p, is

$$F(c,p) = \sum_{i=1}^{n} (V_i - Pp_i)c_i. \tag{1}$$

From the zero-sum assumption that the interests of the two players are diametrically opposed, it follows that $-F(c,p)$ is the expected payoff to the inspector.

The common assumption of game theory is that both the inspector and the inspectee are (perhaps just intuitively) aware of (1). Neither the inspector nor the inspectee can control the values of

P, $V_1, V_2, \dots, V_n$. These numbers are part of the rules of the game. On the other hand, the inspectee chooses $c_1, c_2, \dots, c_n$ in such a way as to maximize $F(c, p)$ but has no control over the choices of $p_1, p_2, \dots, p_n$. In the same way, the inspector uses his opportunity to choose $p_1, p_2, \dots, p_n$, so as to minimize $F(c, p)$ but has no control over the choices of $c_1, c_2, \dots, c_n$. A solution of the game must specify these choices for both players, together with the resulting value of $F(c, p)$. Before proceeding (in the next section) to present the solution given by this model, we note some of the model's limitations and imperfections:

— The zero-sum assumption of diametrically opposed interests is not quite right unless one thinks of society, whose agent is the inspector, as seeking vengeance rather than deterrence. Perhaps one of the alternative goals for the inspector suggested earlier is more realistic.

— The "cheat-or-no-cheat" dichotomy is a severe idealization, ignoring as it does the possibility of introducing different degrees of bias into the devices. To consider this situation would also require formulating some mathematical representation of how the probability of detection depends on the degree of cheating.

— Even with the above restriction it seems odd to assume that cheating will be detected with certainty whenever the offending device is inspected. This, for example, seems to rule out the use of the model to examine the relative merits of training, or selecting, or otherwise encouraging inspectors to work more rapidly (in effect, increasing m), versus stressing the thoroughness or quality of their work (in effect, increasing the probability of cheat-detection). Fortunately this limitation is only apparent. A detection probability d can be represented in the model simply by replacing the fixed penalty P in (1) by the average penalty $P' = Pd$.

— A quite natural extension of the model would be the replacement of P in (1) by device-specific penalties, P_i. (For example, the presence of site-dependent detection probabilities d_i would lead, as above, to the use of Pd_i in place of P.) This would raise no real problem were we content with solving the model numerically, but it does interfere with achieving the kind of nearly closed-form solution sought here for the sake of insight, and it definitely contradicts the desire to maintain in the model a simple clear-cut scalar quantity representing the intensity of society's sanction against cheating. This complication will be omitted from the present paper.

— If a serious level of cheating is detected, the government may react with measures which are onerous to all merchants, even those with properly functioning devices. Similarly, if detected cheating is publicized, honest as well as dishonest establishments may suffer from loss of public confidence. These considerations suggest an inadequacy in the way (1) associates penalty specifically to those devices at which cheating is discovered. Perhaps there should be a penalty P^* which is "activated" if cheating is discovered at *any* device (or at more than some "threshold" number of devices). This possibility is also left for future investigation.

— The inspector must decide which subset of the n devices he will inspect, and the family of subsets which represent "allowable" outcomes of this decision are limited by the amount of inspection resources available. In the present model, that limitation is expressed by specifying the number of inspections which can be undertaken. This is clearly an idealization of the real-world situation in which some inspections may (predictably) require more time than others, some inspection sites are remote from the majority so that visiting them substantially reduces the number of other devices which can be inspected during a given time period, etc. One would like to improve the model by incorporating a more realistic representation of the family of "allowable" subsets from which the inspector can choose.

The preceding list suggests a number of directions for further analysis. Some of them are pursued in **[1]** and in our current research. But we feel the present model, despite its evident deficiencies, represents a suitable first step in focusing on the issues of interest.

The Optimal Strategies

Before discussing the general results for our model, we will consider a simple example. Recall that among the data for the problem, two items can be considered society's response to cheating. These are P, the penalty imposed when cheating is discovered, and, m, the resources available to the inspector. We shall see how various combinations of P and m affect the inspectee's gain from cheating, $F(c, p)$, arising when optimal strategies are used.

Example 1: *Equal Sized Firms.* For this example we assume that the value of cheating is the same for all devices, i.e., $V_i = V$ for all i. Then (1) gives

(2) $$F(c, p) = \sum_{i=1}^{n} (V - Pp_i)c_i.$$

It will be convenient to distinguish three cases based on the size of the penalty P, namely

$$P < V \qquad \text{(Case I)}$$

$$V \leqq P < nV/m \qquad \text{(Case II)}$$

$$nV/m \leqq P \qquad \text{(Case III)}$$

That $nV/m > V$ follows from our assumption that $m < n$.

Case I represents the least response by society to cheating. Since $0 \leqq p_i \leqq 1$, it follows that $V - Pp_i > 0$. Thus the inspectee maximizes his gain $F(c, p)$ by cheating on every device, that is, $c_i = 1$ for all i. In this case $F(c, p) = nV - Pm$ for every choice of strategy for the inspector. The fact that the inspectee's gain does not depend on the strategy chosen by the inspector means that with such a small penalty, inspection is not a threat to the inspectee.

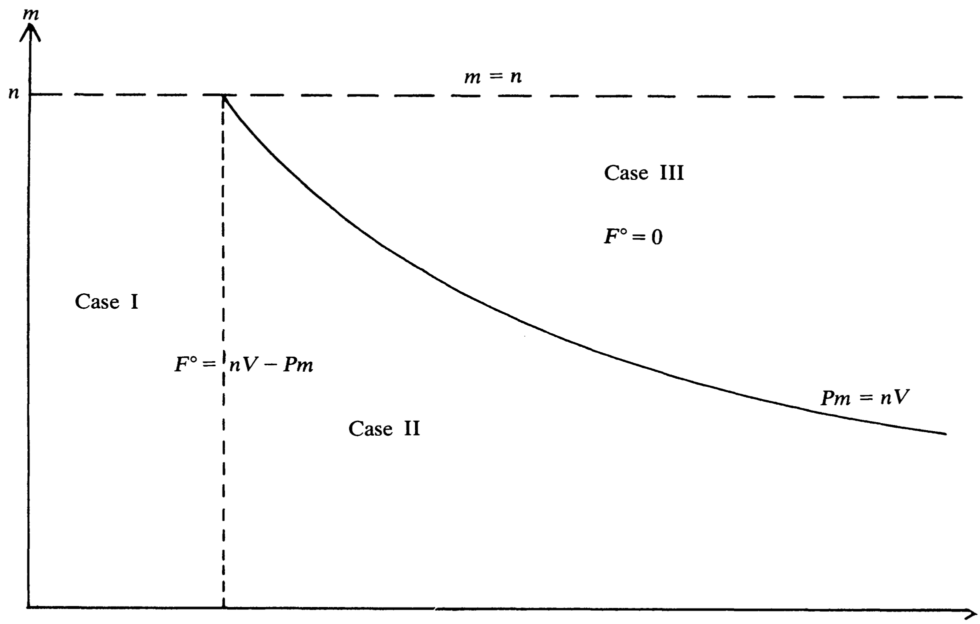
For Cases II and III, as might be guessed from the uniform way the devices figure in the situation, an optimal strategy for the inspector is to inspect all devices with equal likelihood, that is, $p_i = m/n$ for all i. Then for Case II, as in Case I, $V - Pp_i = V - Pm/n > 0$ and again, the inspectee's optimal strategy is to cheat on every device. As in Case I, $F(c, p) = nV - Pm$. Note that in Case II, because of the increased size of the penalty, the inspector is able to deter cheating (or, at least, to make it unprofitable) on any single device D_i by always inspecting it, that is, by setting $p_i = 1$. However, by choosing such a strategy, he creates an imbalance on other devices which would permit the inspectee, by choosing a suitable strategy, to obtain a net gain.

In Case III the penalty is large enough to make cheating unprofitable. Since $V - Pp_i = V - Pm/n \leqq 0$, an optimal strategy for the inspectee is never to cheat, i.e., $c_i = 0$ for all i. In this case, $F(c, p) = 0$, completing Example 1.

Our main interest is not in the optimal strategies themselves but in the dependence of the optimized payoff $F(c, p)$ as a function of P and m. We will refer to this value as the "net illicit gain" F° a measure of (imperfect) performance by the inspection system. Its dependence is summarized in FIGURE 1. The value of P is measured along the horizontal axis and the value of m is measured along the vertical axis. Since we have assumed that $m < n$ (n being a fixed number beyond the control of the inspection agency), the point on the m-axis at which $m = n$ is specially marked and the region above this is not relevant to our study. Since the case in which $P = V$ is of special interest, that point is marked on the P-axis.

The curve in FIGURE 1 is one branch of the hyperbola $Pm = nV$. The quantity Pm is the maximum total penalty, if the inspector knew exactly where cheating was taking place and if there was enough of it to build the penalty up to that total. Note that F° is continuous across the curve.

The situation of Example 1 represents the simplest special case of our model, and FIGURE 1 illustrates the kind of information it yields. Later we will consider a second special case, the "next simplest one", to provide a further concrete example of the application of our model. But first we will describe what this model's solution looks like in the general case. This requires the introduction of some additional concepts and notations, to which we now turn.



Net Illicit Gain — Equal Sized Firms.

FIGURE 1.

Example 1 illustrates the need for distinguishing those devices D_i for which $P < V_i$ from those for which $P \geqq V_i$. We say that a device D_i is **tempting** for cheating if $P < V_i$. It will be convenient to denote the totality of those devices which are tempting by T and to denote the number of devices by $|T|$. In the same way we say that D_i is **untempting** if $P \geqq V_i$ and denote the set of untempting devices by U. The number of untempting devices is denoted by $|U|$. Since each device is either tempting or untempting we have $|T| + |U| = n$.

Further, we denote the sum of the cheating-payoff values of all the tempting devices by V_T and the sum of these values for all the untempting devices by V_U. Finally, the sum of these values for all devices is called V_Σ. Then $V_T + V_U = V_\Sigma$. For instance, in Case I of Example 1 every device is tempting, and so $|T| = n$, $|U| = 0$ and $V_T = V_\Sigma = nV$, $V_U = 0$. In Cases II and III of Example 1 every device is untempting and so $|T| = 0$, $|U| = n$, $V_T = 0$ and $V_U = V_\Sigma = nV$.

The solution of the general problem can be stated in terms of $|T|$ and V_U. The solution takes different forms according as

$$m < |T| + V_U/P \qquad \text{(Case A)}$$

or its opposite

$$m \geqq |T| + V_U/P \qquad \text{(Case B)}$$

holds.

How do the three cases of Example 1 fit into this classification scheme? We pointed out above that in Case I of Example 1, $V_U = 0$ and $|T| = n$. Thus $|T| + V_U/P = n$ and so Case A holds since $n > m$. For Cases II and III, $|T| + V_U/P = nV/P$. Thus Case II also belongs to Case A whereas Case III belongs to Case B.

If the number n of devices and the values V_i of cheating on them are thought of as fixed, then Cases A and B partition the (P, m)-area into two regions. The Case A region corresponds to situations where the inspectee's optimal strategy is to cheat on every device, i.e., $c_i = 1$ for all i. The condition stated in Case A seems to have the form: "m is too small". This might be misleading; the equivalent form $P(m - |T|) < V_U$ exhibits the *joint* smallness of P and m more clearly.

In Case A, any strategy p for the inspector is optimal if it satisfies the "no overkill" condition

$p_i \leqq V_i/P$ for all i in U. In other words, the inspector does not use up more inspection resource on any one untempting device than the minimum necessary to make cheating uprofitable on that device. The illicit gain is given by $F^\circ = V_\Sigma - Pm$, which is the same as the value $nV - Pm$ for Cases I and II of Example 1. Thus, for the Case A region, the major results for the special case in Example 1 carry over without essential change to the general case of our model.

Within the Case A region there is a trade-off between society's two modes of response to cheating, namely, the values of P and m. Since the illicit gain is given by $V_\Sigma - Pm$, any two choices for the pair P and m (within the Case A region) for which the product Pm is the same will yield the same gain. As we shall soon see, the situation in the Case B region is quite different.

Case B represents a more adequate societal response to cheating. In fact, the inspectee can no longer expect to profit by cheating on any untempting device. Of course he will still gain by cheating on all tempting devices. An optimal strategy for the inspectee is $c_i = 1$ if $i \in T$ and $c_i = 0$ if $i \in U$. In Case B the optimal strategies for the inspector require always inspecting the tempting devices and satisfying the "no underkill" condition $p_i \geqq V_i/P$ for all i in U. In other words, the inspector uses at least enough inspection resource on each untempting device to make cheating unprofitable. It now follows from (1) that $F^\circ = V_T - P|T|$. This implies that within the B region there is resource-sufficiency, that is, increasing the value of m without changing the value of P does not affect the gain F°. (On the border between Cases A and B — when $m = |T| + V_u/P$ — the inspectee has some additional optimal strategies. Details are given in [**1**].)

We now have a complete solution for the general case of our model. In order to further illustrate this solution, we will provide another example.

Example 2: *Big Firms, Small Firms.* In this example we assume that there are n_b "big firms" with high incentive for cheating (each with the same value V_b for V_i), and n_s "small firms" facing a lesser temptation (each has $V_i = V_s$, where $V_s < V_b$). Since we assume that there are no other devices, we have $n_b + n_s = n$. Also, $V_\Sigma = n_b V_b + n_s V_s$. As in Example 1, it will be convenient to consider three cases, depending on the size of the penalty P:

$$P < V_s, \qquad \text{(Case I)}$$

$$V_s \leqq P < V_b, \qquad \text{(Case II)}$$

$$V_b \leqq P, \qquad \text{(Case III)}.$$

In Case I we have $|T| = n$, $|U| = 0$ and so $|T| + V_U/P = n$. Because we assume throughout that $m < n$, Case I in this example falls into the category in the general problem of Case A. Accordingly, an optimal strategy for the inspectee is to cheat on all devices, i.e., $c_i = 1$ for all i, and the illicit gain is given by $F^\circ = V_\Sigma - Pm$.

For Case II we have $|T| = n_b$, $|U| = n_s$, $V_U = n_s V_s$, and $V_T = n_b V_b$. It then follows that $|T| + V_U/P = n_b + n_s V_s/P$. Thus Case A or Case B of the general model holds according as

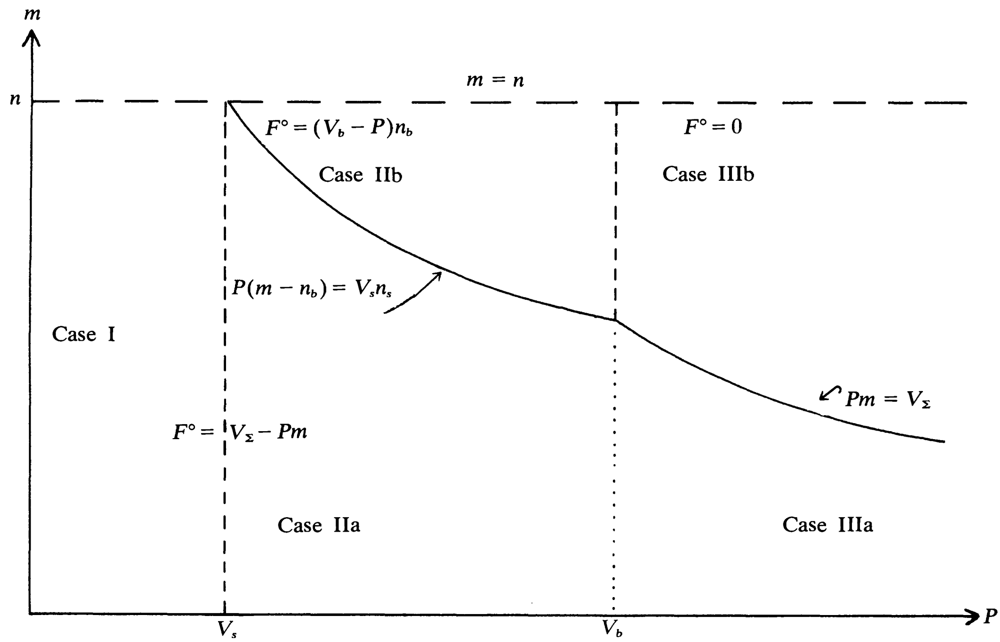
$$m < n_b + n_s V_s/P \qquad \text{(Case IIa)}$$

or

$$m \geqq n_b + n_s V_s/P \qquad \text{(Case IIb)}.$$

Case IIa falls into the general Case A category and so an optimal strategy for the inspectee is to cheat on every device. Also, in Case IIa, every strategy for the inspector is optimal if it satisfies the no overkill condition $p_i \leqq V/P$ for all i in U. The inspectee's gain is, as in the general Case A, $F^\circ = V_\Sigma - Pm$.

Case IIb falls into the general Case B category and so an optimal strategy for the inspectee is $c_i = 1$ for big firms and $c_i = 0$ for small firms. An optimal strategy for the inspector is to always inspect every big firm's devices and to avoid underkill on the devices of the small firms. In other words, $p_i = 1$ for big firms while $p_i \geqq V_s/P$ for small firms. It then follows that $F^\circ = n_b V_b - P n_b = n_b(V_b - P)$.



Net Illicit Gain — Big Firms, Small Firms.

FIGURE 2.

It remains to consider Case III. Here $|T| = 0$, $|U| = n$, $V_T = 0$, and $V_U = V_\Sigma$. Since $|T| + V_u/P = V_\Sigma/P$, Case A or Case B of the general model holds according as

$$m < V_\Sigma/P \qquad \text{(Case IIIa)}$$

or

$$m \geqq V_\Sigma/P \qquad \text{(Case IIIb).}$$

Case IIIa yields the same results as Case IIa, namely, an optimal strategy for the inspectee is to cheat on every device and every strategy for the inspector is an optimal strategy if it satisfies the no overkill condition. The inspectee's illicit gain is $F^\circ = V_\Sigma - Pm$.

Finally, in Case IIIb, it follows from $m \geqq V_\Sigma/P$ that an optimal strategy calls for the inspectee never to cheat. An optimal strategy for the inspector requires only that he satisfy the no underkill condition $p_i \geqq V/P$ for all i in U. Since the inspectee never cheats, his illicit gain is $F^\circ = 0$.

These results are summarized in FIGURE 2. Now two points are marked on the P axis, namely, V_s and V_b, and these are used to distinguish Case I, II and III. The areas corresponding to the subcases IIa, IIb, and IIIb are also indicated. It is easily verified that F° is continuous across each of the three boundary curves. Note that the region corresponding to Case IIb is inspection saturated, that is, increasing m while leaving P fixed does not affect F°. Thus a stiffer penalty P, rather than more inspection activity, is needed to reduce the inspectee's expected payoff F° from his optimal strategy (cheat at the big firms, but not at the small ones).

The model described above is one of three inspector-inspectee models studied in [**1**]. That paper contains the optimal strategies for both players and the illicit gain F° for all three models. In addition, proofs that the described strategies are indeed optimal are also given in full detail. Readers who are interested are invited to write to either of the authors for a copy.

Reference

[1] A. J. Goldman and M. H. Pearl, The dependence of inspection-system performance on levels of penalties and inspection resources, J. Res. Nat. Bur. Standards vol. 80B, 2 (1976).

Human Population Growth: Stability or Explosion?

An historical survey of various models of population growth which gives attention to their character, derivations, and flaws.

DAVID A. SMITH
Duke University
Durham, NC 27706

1. Deterministic population models

In this paper we will examine several simple, but historically interesting, models for human population growth, all of which are special cases of the following general population model. We suppose that the size N of the population under consideration is a function of time t, and that its rate of change dN/dt is a "known" function of N and t. (In practice, the rate of change is usually *conjectured* rather than known, and the resulting differential equation is studied with regard to the reasonableness of its consequences.) To be a little more precise, we suppose that the population has a birth rate (per individual) B and a mortality rate M, both of which may be functions of N and t, expressed as fractions of the total population, so that the *net* growth rate may be written as

$$\frac{dN}{dt} = BN - MN = (B - M)N = PN, \tag{1}$$

where $P = B - M$ is the **production** rate, that is, the net rate at which individuals in the population are reproducing.

One notices immediately that there is something wrong here: N is an *integer*-valued function, so it doesn't make much sense for it to have a nonzero derivative at every point, as suggested by equation (1).

For biological populations with discrete generations (e.g., cicadas), it is more appropriate to use a difference equation for the growth rate, for example,

$$\frac{\Delta N}{\Delta t} = \frac{N(t + \Delta t) - N(t)}{\Delta t} = PN,$$

where Δt is the length of a generation. However, large human populations vary "almost continuously" (births and deaths are taking place all the time), so it is reasonable to *model* the true N by letting the discrete generation time Δt tend to zero. This yields equation (1).

A model such as equation (1), expressed solely in terms of differential equations that ignore random fluctuations in the population, is called **deterministic.** Under modest assumptions about the niceness of the function P, equation (1) together with a single known population size, say $N = N_0$ at time $t = t_0$, uniquely determines N at any future time [**22**, Section 56]. Since we know that biological populations are subject to many random effects (ranging from chance encounters of lovers or enemies to natural disasters such as earthquakes), any single deterministic model will be at best a crude approximation, and then possibly only for a limited time or place. Nevertheless, even the simplest such models may contribute to our understanding of changing phenomena, as we shall see.

2. The Malthusian or exponential model

About the turn of the 19th century, the British economist Thomas Malthus (1766–1834) observed that biological populations (including human ones) tended to increase at a rate proportional to the population size **[15]**. In terms of our general model (1) this is equivalent to the assumption that the production rate P is constant, which of course would be the case if both birth and mortality rates were constant. One learns in freshman calculus (see any of **[8]**, **[9]**, **[24]**, **[25]**, or **[26]**) how to solve equation (1) with P constant by "separating the variables" to obtain

$$N = N_0 e^{Pt}, \tag{2}$$

where N_0 is the population at time $t = 0$. Thus the assumption of constant production rate leads to the conclusion that the population must grow exponentially. Malthus reasoned that such an exponential growth of the world's population could not go on indefinitely, and therefore some sort of catastrophe (war, plague, or famine, for example) must intervene from time to time to interrupt the inexorable working of equation (1) by artificially (and rapidly) reducing N.

The Malthusian model leaves a lot to be desired as a description of human population growth, since it takes into account hardly any of the important characteristics of human reproductive behavior — nothing that would distinguish us from laboratory colonies of bacteria, say. The other models to be considered here are similarly unsophisticated, and differ only from the Malthusian model by changing the assumption that the production rate is constant.

3. The Verhulst or logistic model

In the 1840's, a Belgian mathematician, P. F. Verhulst (1804–1849), proposed the following alternative to the Malthusian model (rediscovered and popularized in the 1920's by Pearl and Reed **[17]**). Assume the birth rate B is constant, but the mortality rate M is proportional to the population, say $M = mN$ for some constant m. Then the growth model (1) becomes

$$\frac{dN}{dt} = (B - mN)N. \tag{3}$$

Such a mortality rate might result, for example, from competition for available resources, such as living space, food supply, air, water, etc. Using available census data from the period 1790–1840 to determine the constants B and m, Verhulst predicted the population of the United States in 1940 and was off by less than one percent. Of course, the success of the prediction was the result of the averaging out of several important short-term fluctuations in population (caused by World War I and the depression, for example), and as it turned out, 1940 was a lucky choice for a stopping point. (The original Verhulst model also had a small, positive, constant term on the right-hand side of (3) to account for net immigration to the U.S. Such a term complicates the mathematical analysis without adding any enlightenment about population models. Furthermore, our objective is to compare simple models for *world* population growth, which rules out immigration, as far as we know.)

Equation (3) raises a possibility that does not occur with the Malthusian model, namely that there might be a non-zero **equilibrium** population, one for which $N' = 0$. Specifically, this would be the case if $N = B/m$. The possibility of a non-zero equilibrium population is very important, as we shall see, so we introduce the abbreviation $\lambda = B/m$ for this special number. Observe from equation (3) that, if $0 < N < \lambda$, then $N' > 0$, so the population is increasing, and if $N > \lambda$, then $N' < 0$, so the population is decreasing. This is an example of a **stable** equilibrium: If the population is not *at* equilibrium, it is moving *toward* it.

By factoring out the birth rate B, the Verhulst model may be rewritten as

$$\frac{dN}{dt} = B\left(1 - \frac{N}{\lambda}\right)N. \tag{4}$$

The factor $1 - N/\lambda$ may be interpreted as an "environmental resistance" factor, with λ interpreted as the "maximum supportable population". When N is small compared to λ, the resistance factor is close to 1, and the model resembles the Malthusian case. However, when N becomes sufficiently large, the resistance factor causes the growth rate to tend to zero and the population to stabilize.

Equation (4) may be differentiated implicitly to obtain

$$N'' = B\left(1 - \frac{2N}{\lambda}\right)N', \tag{5}$$

from which we see that there is an inflection point at $N = \lambda/2$, and the growth rate is maximal when the population reaches half its maximum supportable level. For $N < \lambda/2$, the graph of N is concave upward, resembling the exponential growth curve, but for $N > \lambda/2$, the graph is concave downward and leveling off. This S-shape is called the **logistic** growth curve. It has been quite successful as a model for laboratory populations (for example, bacteria or fruit flies) with limited resources such as space or food supply, and in at least one case, as noted above, for describing the gross changes in a large human population over a century. However, we know from many examples that the Verhulst model does not adequately describe either short-range changes or very long-range trends in human population growth, whether for a given geographical region or for the earth as a whole. Pearl and Reed [**17**] predicted a *maximum* world population of about two billion, which was exceeded by 1930.

We have by now discovered the important qualitative features of the solution of (4) without actually solving the equation. Its solution is not difficult to obtain, however, by separation of variables:

$$N = \frac{\lambda N_0}{N_0 + (\lambda - N_0)e^{-Bt}}. \tag{6}$$

The details may be found in [**8**], [**9**], or [**24**], or worked out as an exercise. The integration step provides a good reason for studying partial fraction decompositions.

4. The von Foerster or Doomsday model

A little less than a generation ago, H. von Foerster, P. M. Mora, and L.W. Amiot published an article [**27**] entitled "Doomsday: Friday, 13 November, A.D. 2026", which suggested quite a different approach to the problem of formulating a gross model for world population growth. They argued from a historical perspective that as human population has increased, improvements in technology and in mass communication have had the effect of welding the population into a more and more effective "coalition" in a vast "game against nature", rendering natural environmental hazards less effective, improving living conditions, and extending the average life span. It follows (they said) that the net production rate P might actually be an *increasing* function of N rather than a decreasing one. It should, of course, be a very slowly increasing function, and they tentatively proposed

$$P = P_0 N^{1/k}, \tag{7}$$

for some constants P_0 and k to be determined from historical data. This leads to the differential equation

$$\frac{dN}{dt} = P_0 N^{1+1/k}. \tag{8}$$

As with the previous models, equation (8) may be solved by separation of variables. In this case the integration step requires nothing deeper than the power rule, and the resulting equation is easily solved for N:

$$N = \frac{k^k}{(C - P_0 t)^k} \tag{9}$$

where C is a constant of integration.

Now equation (9) leads to a very disturbing conclusion: There is a finite time t, namely $t = C/P_0$, at which N becomes infinite, hence the phrase "population explosion". That is a far more striking prediction of future catastrophe than even Malthus could come up with. Hence it is important to know whether there is any possibility that such a model could accurately reflect world population trends, and if so, how far off is "Doomsday". If we abbreviate C/P_0 as t_0, then (9) may be simplified to

$$N = \frac{K}{(t_0 - t)^k}, \tag{10}$$

for appropriate constants K, k, t_0. The question of whether (10), and hence (8), is plausible as a model may be resolved by "fitting" the equation to available historical data to determine the "best" values for K, k, and t_0, and then observing whether the "best fit" is indeed a good fit. This was done by von Foerster and his colleagues, using the least squares method to fit equation (10) to all of the independent estimates of world population they could find, ranging over the span from 0 to 1958, when their article was written. The best values of the parameters were found to be:

$$\begin{aligned} t_0 &= \text{A.D. } 2026.87 \pm 5.50 \text{ years}, \\ K &= (1.79 \pm 0.14) \times 10^{11}, \\ k &= 0.990 \pm 0.009. \end{aligned} \tag{11}$$

In particular, t_0 is the Doomsday of their title. The root mean square (rms) error of the fit was 7%; this is a very good fit, considering the uncertainties and inconsistencies of estimates of world population prior to 1900.

If we take logs on both sides of (10) we have

$$\ln N = \ln K - k \ln(t_0 - t); \tag{12}$$

so $\ln N$ is a linear function of $\ln(t_0 - t)$ with negative slope $-k$. Thus a further test of the plausibility of the model is to make a double logarithmic plot of the data and see whether they appear to lie on a straight line. Specifically, $\ln N$ is to be plotted against $\ln \tau$, where $\tau = t_0 - t$ is "doomstime", the time left before the population grows without bound. The results of such a plot are shown in FIGURE 1, adapted from [**27**].

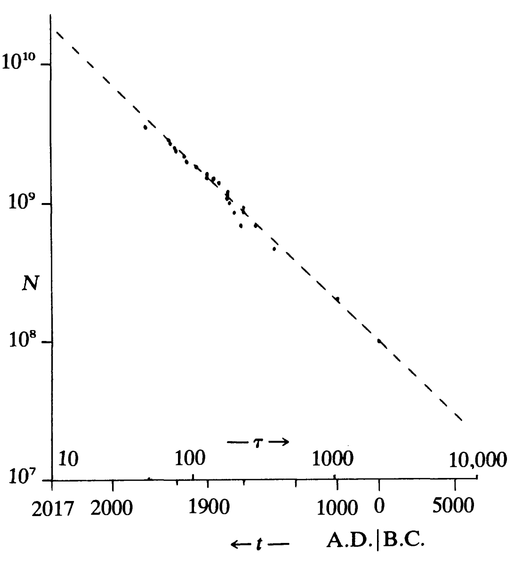


FIGURE 1.

Date, A.D.	Population in millions
0	100
1000	200
1650	545
1750	728
1800	906
1850	1171
1900	1608
1920	1834
1930	2070
1940	2295
1950	2517
1960	3005

Selected estimates of world population

TABLE 1.

In their analysis, von Foerster and his colleagues point out that one cannot discredit the population "optimists" who argue that technology has always succeeded in producing enough food to keep up with increasing population. Distribution problems aside, that is correct. "Our great-great-grandchildren will not starve to death, they will be squeezed to death." (If you are of college age now, delete one "great".) Their conclusion was that what is needed is a "population servomechanism", a feedback control device which will control P on the basis of the current value of N. That means worldwide control of population production, as yet an unsolved problem. But what is the alternative?

The Doomsday model is, of course, controversial. It has been attacked on a variety of grounds, represented by subsequent letters and articles in *Science* (**[5]**, **[10]**, **[19]**). We will examine these objections and the answers by von Foerster, Mora, and Amiot (**[5]**, **[19]**, **[28]**) in Section 6.

5. Fitting historical data by least squares

It should be clear from the previous sections that one crucial test of a conjectured model is the closeness of its fit to observable data, from which one may also infer values for the parameters in the model. If the fit is reasonable, one may then justify statements of the form "If past and present trends continue, then ...".

The least squares principle, which dates back to Gauss, may be formulated as follows. Suppose n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ (or "observations") are given to which we want to fit a functional relationship $y = f(x)$, where the form of f is known but there are one or more unknown parameters $a, b, c, \ldots$ in the description of f. It would be unrealistic to suppose that all, or perhaps any, of the data points would lie exactly on the graph of f, due to errors of observation and/or approximations in the modelling process. Thus we define "deviations"

$$d_i = y_i - f(x_i), \qquad 1 \leqq i \leqq n, \tag{13}$$

and in order to keep all of these reasonably small, we attempt to choose the parameters $a, b, c, \ldots$ so as to minimize

$$S = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2. \tag{14}$$

The variables in equation (14) are just the parameters, since the x_i's and y_i's are known, as is the form of f. Thus we have a multivariable minimization problem, and we know from calculus that a necessary condition for minimizing S is to find "critical" values for the parameters. In short, we must solve

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = \cdots = 0. \tag{15}$$

For many cases of practical interest, condition (15) turns out to be sufficient as well.

Minimizing the sum of squares of deviations (whence the name "least squares") is not the only criterion for "best fit", but there is a simple geometric reason why it is a good one. The sequence of observations $(y_1, y_2, \ldots, y_n)$ of the dependent variable may be thought of as a point in n-dimensional Euclidean space. For each choice of the parameters $a, b, c, \ldots$, the sequence $(f(x_1), f(x_2), \ldots, f(x_n))$ also represents a point in n-space. The totality of all such points constitutes a "surface" in n-space whose dimension is the number of parameters. The expression S given by (14) is the squared distance from $(y_1, y_2, \ldots, y_n)$ to a point on the surface, and the least squares principle selects a point on the model surface which is closest to the observed point.

We will illustrate the least squares process by fitting the Doomsday model (10) to the data shown in Table 1, taken from Austin and Brewer **[1]**. We should emphasize that the data shown in TABLE 1 are not those used by von Foerster, *et al.*, and therefore the fitted parameters need not agree with (11). The Doomsday paper **[27]** gives references to 24 data points from widely scattered sources, but the actual numbers are not given. The principle of selection in **[27]** was to reject only those estimates thought to be copied from an earlier source. Thus their data includes estimates of widely varying reliability, as well as a number of inconsistencies. Austin and Brewer, on the other hand, used only what they considered to be the most authoritative and consistent estimates, and the data in TABLE 1 are from just three different sources. This distinction will play an important role in Section 7 below, where we consider the work of Austin and Brewer in more detail. Needless to say, one would expect to obtain a better fitting curve from a given family of curves if the data is consistent than if it is not, but whether the result is more "correct" is purely conjectural.

The least squares process is easiest to carry out if the parameters appear linearly in the formal description of $y = f(x)$. We will illustrate this first with equation (12), which we write as

$$y = a + bx, \tag{16}$$

where $y = \ln N$, $a = \ln K$, $b = -k$, and $x = \ln(t_0 - t)$. If we suppose t_0 is known (an unjustified assumption, but one that is convenient for illustrating the process), then we can use $x_i = \ln(t_0 - t_i)$ and $y_i = \ln N_i$ as our data, where (t_i, N_i) are the data in TABLE 1. Substitution of (16) and (13) into (14) leads to

$$S = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2 \tag{17}$$

as the function to be minimized. S is a quadratic function of a and b with a unique minimum and no maximum, so it suffices to find the critical point as in (15). Setting the partial derivatives equal to zero, we arrive at the **normal equations** (see **[26]**, Section 14–10, Problem 1):

$$\begin{aligned} na + (\Sigma x_i)b &= \Sigma y_i, \\ (\Sigma x_i)a + (\Sigma x_i^2)b &= \Sigma x_i y_i. \end{aligned} \tag{18}$$

This illustrates the fact that if the parameters appear linearly, the optimal selection can be found by solving simultaneous linear equations.

To apply (18) to TABLE 1, we first scale the data by taking the unit of time to be 1000 years and the unit of population to be 100 million. This permits us to work with "reasonable size" numbers and avoids possible loss-of-significance problems. Using $t_0 = 2030$ (a closer guess than we would be able to make if we didn't already know the "answer"), $x_i = \ln(2.03 - t_i/1000)$, and $y_i = \ln(N_i/100)$, equations (18) become

$$\begin{aligned} 12.000\,a - 18.830\,b &= 27.517, \\ -18.830\,a + 41.745\,b &= -55.412, \end{aligned} \tag{19}$$

for which the solution is

$$a = 0.71925, \qquad b = -1.0029. \tag{20}$$

Substitution of these answers in equation (10) gives

$$N = \frac{e^{0.71925} \times 10^8}{[(2030 - t) \times 10^{-3}]^{1.0029}} = \frac{2.0529 \times 10^{.0087}}{(2030 - t)^{1.0029}} \times 10^{11} = \frac{2.0944}{(2030 - t)^{1.0029}} \times 10^{11}. \tag{21}$$

Equation (21) represents the best fit of (10) to the data of TABLE 1 *if* Doomsday is 2030 A.D.

That brings us to the sticky point of how to determine parameters that appear nonlinearly, such as t_0 in either (10) or (12). Whichever form is used, the partial derivative of S with respect to t_0 is an extremely complicated nonlinear expression. One approach to nonlinear least squares fits is to guess at starting values, perhaps by plotting the data, then expand the desired function in a Taylor series around the guessed parameters (**[26]**, page 827), discard the terms of degree higher than the first, and fit to the resulting approximation, which is linear in all the parameters. The result of this fit gives new, hopefully improved, values of the parameters. The process is then repeated, each time using the new values as guesses, until convergence is obtained.

In the case of the Doomsday model, where we have a way of treating two of the three parameters linearly, it is convenient to apply the approximation process only to t_0. Thus, the computation above provides values of a and b (hence of k and K) to go with the guess $t_0^* = 2030$. Treating k and K as constants, we can expand (10) in a Taylor series in powers of $(t_0 - t_0^*)$:

$$N = K(t_0 - t)^{-k} = K[(t_0^* - t)^{-k} - k(t_0^* - t)^{-k-1}(t_0 - t_0^*) + \cdots],$$

or, as a first order approximation,

$$N = \frac{K}{(t_0^* - t)^k}\left[1 - k\,\frac{t_0 - t_0^*}{t_0^* - t}\right]. \tag{22}$$

(Note that N is being considered as a function of t_0, with t considered constant. Thus the differentiation is with respect to t_0, and the derivatives are evaluated at $t_0 = t_0^*$.) Since t_0 appears linearly in (22), the corresponding sum of squares S given by (17) is quadratic in t_0; setting $dS/dt_0 = 0$ leads to the following linear equation in t_0, with $\tau_i = t_0^* - t_i$:

$$\Sigma N_i \tau_i^{-k-1} - K\Sigma\tau_i^{-2k-1} + (Kk\Sigma\tau_i^{-2k-2})(t_0 - t_0^*) = 0. \tag{23}$$

We may solve equation (23) for t_0 explicitly to get our next estimate of Doomsday. That value of t_0 may then be used to recompute the x_i's in (17) and (18), leading to improved values of a and b, hence of K and k. Then we use (23) again to further refine t_0. The whole process repeats until convergence to optimal values is obtained.

TABLE 2 shows selected output from a computer run of this process, starting with a crude guess of $t_0^* = 2050$ A.D., or 2.05 in scaled value. The first column gives the iteration number. The last two columns give the root mean square error and percentage error for the fitted curve (10) with current values of the parameters. The root mean square error is the square root of the average squared deviation, or $\sqrt{S/n}$, where S is given by equation (14). This is an appropriate expression for the "average" deviation of fitted points from data points because it is minimized when S is minimized. The percentage error expresses root mean square error as a percentage of the average population in TABLE 1 (after scaling).

As in equation (21), K has to be rescaled by a factor of $10^{3(k-1)} = 1.0413$ to give

$$N = \frac{2.139}{(2030.9 - t)^{1.006}} \times 10^{11} \tag{24}$$

as the best function of the form (10) to fit the world population data in TABLE 1. If the coefficients are compared with those given by von Foerster, *et al.*, we see the effects of using different historical data: Doomsday is postponed about four years, k turns out to be slightly more than 1 instead of slightly less, and K, which represents the population one year before Doomsday, is over 21 billion rather than about 18 billion.

J	a	$k(=-b)$	$K(=e^a)$	t_0	rms	%
1	0.7358	1.079	2.087	2.0477	0.513	3.6
2	0.7335	1.070	2.082	2.0457	0.481	3.4
3	0.7315	1.063	2.078	2.0439	0.453	3.2
4	0.7299	1.057	2.075	2.0424	0.430	3.0
5	0.7284	1.051	2.072	2.0410	0.408	2.9
10	0.7240	1.030	2.063	2.0362	0.343	2.4
20	0.7208	1.013	2.056	2.0310	0.304	2.1
40	0.7199	1.007	2.054	2.0309	0.295	2.1
60	0.7198	1.006	2.054	2.0309	0.295	2.1

Selected computer output for fitting equation (10) to data in TABLE 1

TABLE 2.

6. Discussion of the Doomsday model

A number of objections have been raised to the Doomsday model by various commentators, most of which have been answered by von Foerster, Mora, and Amiot [**19**], [**5**], [**28**]. We will summarize the criticisms and responses in this section and discuss an alternate model in the next.

OBJECTION 1. A model that predicts that an obviously finite quantity will become infinite in finite time is clearly at variance with the facts and therefore irrelevant (Howland, Shinbrot [**19**]; Coale [**5**]; Austin and Brewer [**20**]).

Response. First, the model provides remarkably good fit with all available observations, the usual scientific criterion for (tentatively) accepting a hypothesis, until there is evidence for rejection. [However, we shall see that other models can fit just as well.] Second, there are many examples of accepted models of finite systems with singularities. Von Foerster and his colleagues cite the following physical models: Pressure as a function of velocity approaching the speed of sound, current approaching breakdown voltage in gaseous conduction, index of refraction in optical absorption bands, and magnetic susceptibility at Curie temperature in ferromagnetism. The usual interpretation of such a model is that one expects the system to be highly unstable in the vicinity of a singularity. It is reasonable to expect the same of a social structure subjected to extreme overpopulation.

OBJECTION 2. Not only the population but also the production rate is predicted to go to infinity on Doomsday. However, there are obvious biological factors that keep the production rate bounded. Specifically, the mortality rate cannot be reduced below zero, and the birth rate is limited by the gestation period; it could not possibly be higher than 0.75 babies per year per female of child bearing age. It is possible that biologists might learn to extend the child bearing years, shorten the gestation period, and lengthen the life span (to infinity?), but the production rate must still remain bounded. [None of the participants in the debate mentioned the possibilities of "test tube babies" or cloning, but there would still be bounds on the production rate, perhaps available glass for labware. In any case, scientists would have little incentive to drive the production rate up artificially, even if it were technically feasible.] (Robertson, Bond, and Cronkite [**19**]; Coale [**5**]; Austin and Brewer [**1**].)

Response. Different critics received different responses depending on the assumptions made about an upper bound on the production rate, but all the responses lead to very large populations in relatively short time. The boundedness of the production rate is not challenged by von Foerster, Mora, and Amiot. The most conservative bound is proposed by Coale, who notes that the world birth rate has been relatively constant at 3.9% (while the mortality rate has been declining), and in no case could the production rate conceivably exceed 6%. The Doomsday authors note that the assumption of constant birth rate of 3.9% and production rate given by (7) leads to the conclusion that the mean life span tends to infinity (i.e. the immortality problem is solved) in the year 2001. Assuming a constant 6% growth rate thereafter (i.e. reverting to the Malthusian model), one finds a world population of 30 billion in the year 2027. Knowing that it is finite is small comfort. [One might quibble about a sudden jump in birth rate from 3.9% to 6%, especially since the Doomsday model itself doesn't reach a 6% production rate until 2010, but all such computations lead to uncomfortably large population in the projected lifetimes of most who will read this.]

OBJECTION 3. There must be an upper limit on the earth's life support capabilities, and therefore the population cannot grow without bound (Hutton, Howland [**19**]; Austin and Brewer [**1**]).

Response. This is essentially the Malthusian argument pointing to impending disaster as the supportable limit is approached, a position supported by von Foerster, Mora, and Amiot. Malthus was wrong only in that he did not foresee that the maximum supportable population would be an increasing function of time, due to agricultural technology. Howland proposed that the constant λ in the Verhulst-Pearl model (4) be replaced by a linear function of t, and claimed that such a model gives a good fit to U.S. population data over the span 1790–1960. The Doomsday authors respond that no such model has been found to fit the world population over a very long period without *ad hoc* adjustments to λ from time to time. Furthermore, λ is not an observable parameter, and there is no well-developed theory of supportable population to resolve the question of the type of function λ should be. They note that one possibility for an increasing supportable population is:

$$\lambda = \frac{N}{1-\beta N^{1/k}}, \tag{25}$$

where β is a constant. This choice has the advantages of providing a good fit over one hundred generations and of eliminating unobservable parameters from the model, since substitution of (25) into (4) leads to (8). [Of course, this evades the question of boundedness of λ.]

The question of supportable population has been dealt with in a somewhat different way by Austin and Brewer [**1**], whose model will be considered in the next section. (To our knowledge, von Foerster has not commented in print on the Austin-Brewer model.) The matter is often treated superficially. For example, Stein ([**24**] or [**25**]) suggests a limit of 40 billion, based on 10 billion acres of arable land and a requirement of one quarter acre per person for food production. He shows that this number will be reached in 2109 A.D. if the growth rate remains constant at 1.8%. (In an exercise he asks the student to show that the Malthusian model predicts a "standing room only" population in about 700 years.)

OBJECTION 4. Population forecasting by fitting mathematical curves is notably unreliable because it ignores so many important factors of demography (Dorn [**10**]). Actually, Dorn dismisses the Doomsday model in even stronger terms: "... this forecast probably will set a record, for the entire class of forecasts prepared by the use of mathematical functions, for the short length of time required to demonstrate its unreliability."

Response. Dorn's article is primarily about the "component" or "analytical" method of demography, which attempts to consider all the factors ignored in simple deterministic models, such as distributions of the population by age, sex, and geography, standards of living, and so on. The article itself shows that the analytical method has constantly led to underestimates of future population and growth rates. The Doomsday authors respond with a comparison of the projections for A.D. 2000 made by the United Nations, which Dorn considers "most authoritative", with their own projection given by (10) and (11), plus or minus 7 percent. The results are given in Table 3 (adapted from [**28**]) from which

> "... it appears that the 'most unreliable' values are just the asymptotes, at the moment of truth, to the 'most authoritative projections'; we might mention in passing that the 'most authoritative' projectors changed their minds in the last decade by roughly a factor of 2, while the 'most unreliable' values... are almost independent of the time of their derivation...".

On the matter of record short time for demonstrating the unreliability of the Doomsday model, it was observed in 1970 [**1**] and again in 1975 [**21**] that the actual world population is slightly ahead of the Doomsday projection, nearly a generation after it was made.

	From U.N. Estimate in Year				From (10) and (11)
	1950	1957	1958	1959	
Low			4.88		6.44
Medium	3.20	5.00	5.70	6.20	6.91
High			6.90	~ 7.00	7.40

Projections of population in A.D. 2000 (in billions).

TABLE 3.

Objection 5. A differential equation model is inappropriate for N, since it can take only integer values. If assumption (7) about the production rate is granted, the appropriate model is a difference equation:

$$N(n)-N(n-1)=P_0N(n-1)^{1+1/k}, \qquad n=1,2,3,\ldots, \tag{26}$$

where n numbers generations. It is easy to see that equation (26) implies that N is finite for all n (Shinbrot [19]). (Indeed, this is clear for any equation of the form $N(n)=N(n-1)+$ any finite quantity.)

Response. It is agreed by von Foerster and colleagues that "it is unkind to perform a Dedekind cut on a man". However, equation (26) is not appropriate as a model because there are hardly any integer triples $N(n)$, $N(n-1)$, P_0 satisfying it, unless $1/k$ is an integer. "Obviously he must know such triples, and thus his suggested relationship will remain forever 'Shinbrot's last theorem'."

Shinbrot's objection to the model is easily the weakest, and it drew the weakest response, cuteness notwithstanding. In the first place, there is no reason why P_0 should be an integer, nor is there any *a priori* reason why $1/k$ should not be an integer (indeed, it is very close to 1). On the other hand, we have already observed in Section 1 why a differential equation may be an appropriate model for the growth of an integer-valued function. Human population growth is a nearly "continuous" process, not one that jumps by discrete generations. Nor is there any evidence that the production rate remains constant over a generation, as implied by (26). Thus, Shinbrot's objection betrays a fundamental lack of understanding of the modelling process.

7. The Austin-Brewer or modified coalition model

The rationale for the Doomsday model was the increasingly powerful human "coalition" against environmental forces. For this reason, the Doomsday model is also called the **coalition** model. But we noted in the previous section that both the production rate for world population and earth's capability for supporting population are bounded. Technology has been able to increase the bounds, and may continue to do so, but the bounds cannot be removed entirely. Austin and Brewer [1] have attempted to formulate a variant of the Doomsday or coalition model that would take these bounds into account and still fit the historical data. First they interpreted von Foerster's production rate (7) as a "fertility rate", corresponding to the factor B in equation (4). They then replaced the production rate (7) by the fertility form

$$B=A\left[1-\exp\left(-\frac{P_0}{A}N^{1/k}\right)\right]. \tag{27}$$

By replacing the exponential function by its Maclaurin series, one can see that B is approximately $P_0N^{1/k}$ for relatively small values of N, in accordance with (7). On the other hand, $B\to A$ as $N\to\infty$, so A represents the maximal fertility rate.

If the constant λ in the Verhulst–Pearl model (4) is interpreted as a time-independent maximum supportable population (i.e., an upper bound for supportable populations at any given time, no matter how advanced our technology becomes), then one may substitute (27) in (4) to obtain the model

$$\frac{dN}{dt}=A\left[1-\exp\left(-\frac{P_0}{A}N^{1/k}\right)\right]\left(1-\frac{N}{\lambda}\right)N, \tag{28}$$

which Austin and Brewer call the **modified coalition** model. Note that by wedding the "coalition" and "environmental resistance" concepts, stable equilibrium has reentered the picture. Specifically, $N=\lambda$ is an equilibrium solution of (28), and for $N<\lambda$, we have $N'>0$, so the population increases toward equilibrium.

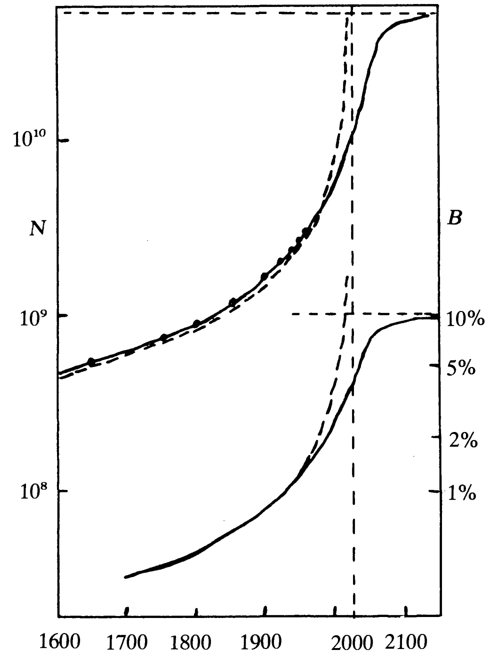
Further discussion of the properties of (28) diverges at this point from the route taken earlier. In particular, it is not possible to carry out the solution of the differential equation in closed form, and hence there is no way to apply the least squares method directly to obtain values for the parameters.

The differential equation can be solved numerically, but only when specific numerical values for the parameters are provided. Austin and Brewer followed an *ad hoc* approach of estimating the parameters, solving (28) numerically, measuring the fit with the data in TABLE 1, and then adjusting the parameters as appropriate. They reported "an excellent data fit" with the following values:

(29) $$A = 0.1, \qquad P_0 = 5.0 \times 10^{-12}, \qquad k = 1.0, \qquad \lambda = 50 \times 10^9.$$

These values imply a maximal production rate of 10%, considerably larger than the 6% Coale considered the largest conceivable [**5**], and a maximum supportable population of 50 billion, considerably larger than most other estimates. Austin and Brewer note, however, that a change of 1% in k can change λ by as much as 20%.

The results of the curve fit with parameters (29) are shown graphically in FIGURE 2, adapted from [**1**]. The figure appears to show a better fit to the data points than that given by the coalition or Doomsday model. However, that is a misleading conclusion, resulting from the fact that Austin and Brewer used a different set of data. In fact, the percentage error of their fit is approximately the same as that given in Section 5 for the Doomsday model, and equation (24) would fit just as well as the solid curve in FIGURE 2.



The graphs of fertility rate *B* (lower two curves) and total population *N* (upper curves) according to the Doomsday model (dotted) and the Austin and Brewer model (solid) with data points for population taken from TABLE 1. The vertical asymptote represents "doomsday" (2027 AD) and the horizontal asymptotes represent the more hopeful stable bounds of the modified coalition model.

FIGURE 2.

Schwartz [**20**] has noted that various models can be fitted to the same data with λ ranging anywhere from 9 billion to infinity with *rms* fractional errors of no more than 2%, and therefore no conclusion can be drawn from such data that would distinguish between the coalition and modified coalition models, or that would indicate an upper bound on the supportable population. Thus we are brought back to the conclusion stated earlier, and not disputed by Austin and Brewer: Past and present trends of world population growth, no matter how they are projected into the future, point to a disastrously large population within the life span of many already born, unless a way is found to alter these trends dramatically on a worldwide scale.

We will give von Foerster, Mora, and Amiot the last words:

"... while we were displaying our wits and know-how in more or less learned discussions about the perennial question of how many angels can dance on a pin point, over ten million real people of flesh and bone, with hopes and desires, with sorrows and pain, have been added to our family of man. Our responsibility demands that we be ready with an answer when these millions ask for their right to live the span of their human condition in dignity." **[28]**.

"The real problem is that today we have to prepare each single member in a family of 3 billion to face soon a decision — namely, either to persist in enjoying his children and to pay for it by having no more than two and remaining mortal, or to reach for immortality and remain childless forever. In 20 years [from 1961], of course, 4 billion will have to make this decision." **[5]**.

To emphasize the conservative nature of that last statement, we add the following footnote: The Population Reference Bureau announced in March, 1976, that world population had passed the 4 billion mark.

References

[1] A. L. Austin and J. W. Brewer, World Population Growth and Related Technical Problems, IEEE Spectrum, 7 (Dec., 1970) 43–54; correction 8 (March, 1971) 10. Reprinted in Technological Forecasting and Social Change, 3 (1971) 23–49.

[2] N. T. J. Bailey, The Mathematical Approach to Biology and Medicine, Wiley, New York, 1967, Chapter 1.

[3] R. Bellman and K. L. Cooke, Modern Elementary Differential Equations, second ed., Addison-Wesley, Reading, 1971, especially Chapter 1.

[4] C. Clark, Some socially relevant applications of elementary calculus, Two-Year College Mathematics Journal, 4 (Spring, 1973) 1–15.

[5] A. J. Coale, Population density and growth (Letter regarding **[25]** and response by von Foerster, Mora, and Amiot), Science, 133 (1961) 1931–1937.

[6] A. J. Coale, The history of the human population, Scientific American, 231 (September, 1974) 40–51.

[7] W. G. Costello and H. M. Taylor, Deterministic population growth models, Amer. Math. Monthly, 78 (1971) 841–855.

[8] R. F. Coughlin, Applied Calculus, Allyn and Bacon, Boston, 1976, especially Section 8–2.

[9] A. B. Cruse and M. Granberg, Lectures on Freshman Calculus, Addison-Wesley, Reading, 1971, especially Section 71.

[10] H. F. Dorn, World population growth: An international dilemma, Science, 135 (1962) 283–290.

[11] R. Freedman and B. Berelson, The human population, Scientific American, 231 (September, 1974) 30–39.

[12] N. Keyfitz, Introduction to the Mathematics of Population, Addison-Wesley, Reading, 1968.

[13] N. Keyfitz and W. Flieger, Population: Facts and Methods of Demography, Freeman, San Francisco, 1971.

[14] D. P. Maki and M. Thompson, Mathematical Models and Applications, Prentice-Hall, Englewood Cliffs, New Jersey, 1973, especially Chapter 8.

[15] T. R. Malthus, Mathematics of population and food, The World of Mathematics, ed. by J. R. Newman, Simon and Schuster, New York, 1956, pp. 1189–1199.

[16] W. Page, Population Forecasting, Chapter 11 in Models of Doom, ed. by H. S. D. Cole and others, Universe Books, 1973, pp. 159–174.

[17] R. Pearl and L. J. Reed, The Growth of Human Population, in Studies in Human Biology, ed. by R. Pearl, Williams and Wilkins, 1924.

[18] E. C. Pielou, An Introduction to Mathematical Ecology, Wiley-Interscience, New York, 1969, Chapter 1.

[19] J. S. Robertson and others, Doomsday (Letters regarding **[25]** and response by von Foerster, Mora, and Amiot), Science, 133 (1961) 936–946.

[20] C. L. Schwartz, World Population Growth (Letter regarding **[1]** and response by Austin and Brewer), IEEE Spectrum, 9 (March, 1972) 10–16.

[21] J. Serrin, Is 'Doomsday' on target? (Letter), Science, 189 (1975) 86–88.

[22] G. F. Simmons, Differential Equations with Applications and Historical Notes, McGraw-Hill, New York, 1972.

[23] D. A. Smith, Interface: Calculus and the Computer, Houghton Mifflin, Boston, 1976, Chapter 48.

[24] S. K. Stein, Calculus in the First Three Dimensions, McGraw-Hill, New York, 1967, especially Chapter 21.

[25] ———, Calculus and Analytic Geometry, McGraw-Hill, New York, 1973, especially Section 6.4.

[26] G. B. Thomas, Jr., Calculus and Analytic Geometry, Alternate ed., Addison-Wesley, Reading, 1972.

[27] H. von Foerster, P. M. Mora and L. W. Amiot, Doomsday, Friday, 13 November, A.D. 2026, Science, 132 (1960) 1291–1295.

[28] ———, 'Projections' versus 'forecasts' in human population studies (Letter regarding [10]), Science, 136 (1962) 173–174.

NOTES

Distribution of Digits in Integers

RALPH P. BOAS, JR.
Northwestern University
Evanston, IL 60201

If you look at a long run of the decimal digits of an irrational number like e, π or $\sqrt{2}$, you will probably get the impression that each digit occurs about as often as any other digit, so that about one-tenth of the digits are (say) 2's. Except for numbers deliberately constructed to have lop-sided distributions of digits, you would find the same phenomenon for pairs of digits (which are single digits in base 100) and generally for digits in any base. To make statements like these precise, we need a little terminology. Given a number N, written in base b, let there be d_n occurrences of the digit d among the first n digits of N. Then if $d_n/n \to 1/b$ as $n \to \infty$, and if this holds for every base b and every digit we might select, we call the number N **normal.** As a matter of fact, nobody has ever proved that e, π or $\sqrt{2}$ is normal, but we do know that there are lots of normal numbers. More precisely, we can find a sequence of intervals, of total length as small as we like, in which all the non-normal numbers lie; hence there must be many more normal numbers than non-normal ones. In the terminology of measure theory, the set of non-normal numbers has measure zero. This situation is usually described briefly by saying that almost all real numbers are normal. For a proof see, for example, [**2**] or [**4**].

The fact that most numbers are normal has some startling consequences. We can translate English text into a numerical code by assigning numbers to the letters of the alphabet and the punctuation marks, for example by using the ASCII code, which can be implemented on many computers. Suppose then that we translate the whole *Encyclopaedia Britannica* into numerical code and read the result as a single (rather large) integer B. This is a single digit in base $B+1$ (or any larger base), and so for almost every real number (all the normal ones) the digit B occurs infinitely often in the representation of that number in any base greater than B. If we take the base to be a power of 10, saying that the "digit" B occurs infinitely often in the expansion of a given number to that base is equivalent to saying that the numerical representation of the Encyclopaedia occurs infinitely often in the ordinary decimal expansion of the same number. We cannot, of course, observe even a single occurrence, since the length of the decimals required to show it would be unreasonably large. Notice, by the way, that we do not know that the numerical representation of the Encyclopaedia occurs in the decimal expansion of e, π, $\sqrt{2}$, or similar numbers, since none of the standard numbers that arise in elementary mathematics is known to be normal.

Explicit examples of normal numbers are not easy to find. One is the number $0.12345678910111213\ldots$ obtained by writing the consecutive integers, in base 10, in a row and putting a decimal point in front of it ([**4**], p. 112).

We cannot expect the decimal representations of integers to be normal, since each one has only finitely many non-zero digits. However, if you look at some large integers (factorials, for example) written out in base 10, you will quite likely again get the impression that the digits are fairly uniformly distributed. One might guess that this should be the case for most integers, with some interpretation of "most". Now it is known (see, for example, [**5**]) that the integers whose decimal representations contain no zeros are rare enough so that $\Sigma 1/n$, summed over these integers, converges; whereas we learn in elementary calculus that $\Sigma 1/n$, summed over *all* integers, diverges. This suggests a way of describing the size of a set S of integers. Let us say that S is "small" if $\Sigma 1/n$, summed over S, converges, and otherwise that S is "large". The set of integers that contain no zeros at all is small; can we say that a set of integers must be small if each integer in the set contains either too few or too many zeros in its decimal representation (or representation in base b)? Since we expect intuitively that about one-tenth of the digits (in base 10) should usually be equal to each of the possible numbers $0, 1, 2, \ldots, 9$, we need a way of saying that the average proportion of a particular digit in a given integer is something other than 1/10. Let us do this for an arbitrary base b.

We would like to give meaning to statements like "one-third of the digits in the n-digit number N are d's". This ought to mean, if N contains k digits d among its n digits, that $n = 3k$; but n will not always be divisible by 3. We shall say, for a given λ $(0 < \lambda < 1)$ that the proportion of d's in the n-digit integer N is λ, or that d occurs in N with frequency λ, if the number of d's in the base b representation of N is $[\lambda N]$, the integral part of λN. (It would agree more closely with intuition if we used $[\lambda N + \frac{1}{2}]$, the closest integer to λN, but the former definition is easier to work with and the difference has no effect on the results.) This definition gives an unambiguous answer to the question of whether or not d occurs in N with frequency λ; but if N and the number of d's are specified these data do not determine λ uniquely. For example, $N = 99^{40}$ has 80 digits of which 11 are 8's. Since $\frac{1}{7} \times 80 = 11.4$, we can say that one-seventh of the digits of N are 8's. We could also say that $\frac{4}{29}$ or $\frac{5}{36}$ of the digits of N are 8's. In the statements that follow, it has to be understood that λ is chosen first.

We are going to show that the set of integers in base b in which a given digit occurs either with frequency less than $1/b$ or with frequency greater than $1/b$ is a small set (in the sense defined above), whereas the set of integers in which a given digit occurs with frequency $1/b$ is large. This seems to reinforce the idea that a specified digit usually occurs the "right" number of times in a randomly chosen integer. Moreover, when $b \geqq 4$, two specified digits usually occur with frequency $1/b$ each. However, if we specify three or more digits, we find that they do *not* usually occur with frequency $1/b$ each: the set of integers in which three specified digits occur with frequency $1/b$ each is a small set. For example, in base 10 most integers have a tenth of their digits 0; most integers have a tenth of their digits 0 and another tenth 1; but rather few have a tenth each of 0, 1 and 2. In fact, it appears that most integers do not have their digits very evenly distributed after all.

Suppose first that $0 < \lambda < 1/b$ and specify a digit d. Let $m = [\lambda n] > 0$ and consider the n-digit integers with precisely m digits equal to d. If we disregard the condition that the left-hand digit cannot be zero, then if precisely m digits are to be equal to d, there are $\binom{n}{m} = n!/m!(n-m)!$ ways to place them and $(b-1)^{n-m}$ ways to fill the remaining places. Consequently the total number of n-digit integers with m or fewer digits equal to d would be

$$(1) \qquad \binom{n}{m}(b-1)^{n-m} + \binom{n}{m-1}(b-1)^{n-m+1} + \cdots + (b-1)^{m}.$$

Since we cannot have 0 as left-hand digit, (1) overestimates the number of n-digit integers with at most m digits equal to d. Since we want to show that $\Sigma 1/N$ converges for the integers N in which d occurs with frequency λ, we shall certainly succeed if we show that the series converges for the integers enumerated by (1). The reciprocals of these integers are all less than b^{-n+1}, and so the sum of their reciprocals does not exceed

$$\binom{n}{m}(b-1)^{n-m}b^{-n+1} + \binom{n}{m-1}(b-1)^{n-m+1}b^{-n+1} + \cdots + (b-1)^{n}b^{-n+1}.$$

If we write $p = 1/b$, $q = (b-1)/b$, this is

$$(2) \qquad b\left\{\binom{n}{m}p^{m}q^{n-m} + \binom{n}{m-1}p^{m-1}q^{n-m+1} + \cdots + q^{n}\right\} = b\sum_{j=0}^{m}\binom{n}{j}p^{j}q^{n-j}.$$

This sum can be recognized as the last $m+1$ terms of the expansion of $(p+q)^n$ by the binomial theorem. Fortunately such sums have been thoroughly studied by probabilists: the sum in (2) is the probability of m or fewer successes in n independent trials (so-called Bernoulli trials) of an experiment with probability p of success and $q=1-p$ of failure. (See books on probability, for example [**1**], pp. 146 ff.) It is known (cf. [**1**], pp. 182 ff.) that the largest term in the expansion of $(p+q)^n$ is the one for which m is as close as possible to np; and that the terms decrease steadily as we leave this value and proceed toward either end. Since $m=[\lambda n]$ and $\lambda<1/b=p$, we have $n\lambda<np$ and the largest term (at least for large n) is outside the sum in (2); hence the terms in (2) decrease from left to right, and there are fewer than n of them. Consequently the sum in (2) does not exceed n times the first term. We have therefore found that the contribution of the n-digit integers to the sum of the reciprocals of the numbers counted in (1) is less than

$$nb\binom{n}{m}p^m q^{n-m}=nb\frac{n!}{m!(n-m)!}p^m q^{n-m}. \tag{3}$$

We want to show that we get a convergent series by summing (3) over $n=1,2,\ldots$. Again, fortunately, the probabilists have provided the answer: when n is large, (3) does not exceed Λ^n for some Λ, where $0<\Lambda<1$, and so we do get a convergent series. Alternatively, we could attack (3) directly by using Stirling's formula (see, for example, [**3**], p. 531).

We have now shown the convergence of the sum of the reciprocals of the integers that have the digit d, in base b, with frequency at most λ, when $\lambda<1/b$. When the frequency is at least $\mu>1/b$, the argument is the same except that we have the other end of the binomial expansion.

Next suppose that S is the set of integers in which the digit d occurs in the proportion $1/b$; we want to show that $\Sigma 1/k$ diverges for $k\in S$. This does not follow directly from what we have proved, since we want to consider, not integers with the frequency of digit d between some $\lambda<1/b$ and some $\mu>1/b$, but the integers with the frequency of digit d exactly $1/b$.

It will be enough to show that $\Sigma 1/k$ diverges over the smaller set S' of those n-digit integers k in S for which n is divisible by b. Then $m=n/b$ of the digits are d's. Since we require a lower bound for the number of integers in S', we have to underestimate rather than overestimate this number. When $b>2$ we can do this by counting the $(n-1)$-digit sequences (possibly beginning with 0) that contain n/b digits b, since we get an integer in S' by prefixing any base b digit except 0 or d to one of these. The number of such sequences is

$$\binom{n-1}{m}(b-1)^{n-m-1},$$

and each element of S' is less than b^n, so the sum of the reciprocals of the integers in S' exceeds

$$\binom{n-1}{m}(b-1)^{n-m-1}b^{-n}=b^{-1}\binom{n}{m}q^{n-m}p^m. \tag{4}$$

For $b=2$ we can get the same estimate by somewhat different reasoning. The last expression is a multiple of the central term of the binomial expansion, which is known to be of order $n^{-\frac{1}{2}}$ (or we could use Stirling's formula). The sum of the terms (4) over n thus diverges.

Finally, when $b\geqq 4$, consider the integers in which digits $d_1,d_2,\ldots,d_j$ have frequency $1/b$. Assume that no d_i is 0 (otherwise the frmulas have to be modified) and write $m=[n/b]$. The number of n-digit integers under consideration is between two multiples of

$$\binom{n}{m}\binom{n-m}{m}\cdots\binom{n-m(j-1)}{m}(b-j)^{n-mj},$$

and each reciprocal is between b^{-n} and b^{-n+1}. If we sum over all n we find that the sum of all the reciprocals is between two multiples of

$$\sum_{n=1}^{\infty}\frac{n!}{(m!)^j(n-mj)!}\frac{(b-j)^{n-mj}}{b^n}. \tag{5}$$

I have not found the asymptotic form of (5) in works on probability, but by applying Stirling's formula and doing some algebra we can show that the terms of the series (5) turn out to be between two constant multiples of $n^{-j/2}$, so that (5) converges when $j>2$ and diverges when $j=1$ or 2.

References

[1] W. Feller, An Introduction to Probability Theory and Its Applications, vol. 1, 3d ed., Wiley, New York, 1968.

[2] M. Kac, Statistical Independence in Probability, Analysis and Number Theory, Mathematical Association of America, 1959.

[3] K. Knopp, Theory and Application of Infinite Series, Blackie, London, 1928.

[4] I. Niven, Irrational Numbers, Mathematical Association of America, 1956.

[5] A. D. Wadhwa, An interesting subseries of the harmonic series, Amer. Math. Monthly, 82 (1975) 931–933.

Computation of Constrained Plane Sets

ROBERT H. LOHMAN
Kent State University
Kent, Ohio 44242

TERRY J. MORRISON
Hiram College
Hiram, Ohio 44234

Consider a simplified energy consumption model in which u units of imported crude oil and v units of domestic coal are consumed each day. Suppose that each of the energy sources of imported oil and domestic coal are assigned economic effect weights x and y, respectively. For instance, the determination of the (possibly negative) number x would depend upon such factors as the economic effects of relying on foreign sources of oil, inflationary effects of the price of the oil, etc. On the other hand, the determination of the (possibly negative) number y would depend upon such factors as the effect that the use of coal has on the environment, increased employment due to domestic production, etc. As these factors change, so does the pair (x, y). Under short term, reasonably controlled circumstances, the pair (x, y) will ordinarily be constrained to belong to some set U of pairs of possible economic effect weights.

The total economic effect of energy consumed under this model is most conveniently measured by the linear expression $ux+vy$. For reasons of economic stability, it is often desirable to subject the number $|ux+vy|$ to some constraint, say $|ux+vy|\leqq c$ for all $(x,y)\in U$. By choosing units appropriately, we may assume without loss of generality that $-1\leqq x\leqq 1$, $-1\leqq y\leqq 1$ and $c=1$. It then becomes interesting to examine those consumption pairs (u, v) which are amenable to the set U of pairs of possible economic effect weights. This leads to a general geometric problem that we investigate in this paper.

For any subset U of the vector space R^2 of pairs of real numbers let

$$U^0=\{(u,v): |ux+vy|\leqq 1 \text{ for all } (x,y)\in U\}.$$

In this note we are concerned with the geometric problem of exhibiting a method whereby U^0 can be precisely computed, provided a reasonable description of U is given. If U is a convex symmetric polygon, so is U^0, and the authors have shown in an earlier note **[2]** how to find the vertices of U^0. In the more general case, however, no simple computational method for determining these sets seems to appear in the literature. Instead, these sets are usually determined in an *ad hoc* manner, using special inequalities when they are available. The simple method we present below does not require the use of these special inequalities and is applicable in a very general setting. In fact, it requires only a rudimentary knowledge of the calculus and a dash of geometric intuition. Moreover, the method is algorithmic in nature and hence allows the user to "plug in" to a standard formula.

The assumed range of the variables x and y implies that

(i) U is contained in the "unit square"

$$S = \{(x, y) \mid |x| \leqq 1, |y| \leqq 1\}.$$

To simplify computations we will assume further that

(ii) U is symmetric with respect to both coordinate axes;

(iii) U is convex;

(iv) The boundary of U is determined by a smooth curve in each quadrant which cuts off corners of the boundary of the unit square S. The curves pass through $(0, \pm 1)$.

(A relaxation of some of these conditions is discussed after considering the above case.)

To solve our problem we need only study the boundary of U in the first quadrant. We see that this boundary is determined by a continuous function $y = f(x)$, $0 \leqq x \leqq 1$, satisfying $f(0) = 1$. Typical sets U are illustrated in FIGURE 1. To insure the downward concavity of f we require that $f' < 0$ and $f'' < 0$ on $(0,1)$. It follows that if $0 < x_1 < x_2 \leqq 1$,

$$(*) \qquad \frac{f(x_2) - f(x_1)}{x_2 - x_1} < f'(x_1).$$

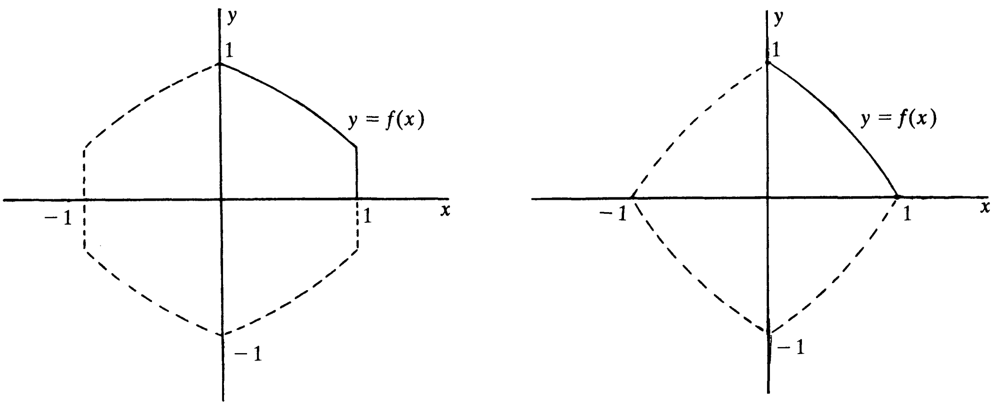


FIGURE 1.

Simple reasoning with the fundamental inequality $|xu + yv| \leqq 1$ shows that U^0 shares much of the general geometric structure of U: U^0 is contained in the unit square, it is symmetric with respect to both coordinate axes, it is convex, and it contains $(0, \pm 1)$. Also, its boundary in the upper half-plane must be a convex curve $v = g(u)$ satisfying $g(0) = 1$. By symmetry we need only determine the shape of the boundary of U^0 in the first quadrant.

A point (u, v) in the first quadrant of the u, v-plane belongs to U^0 if and only if $0 \leqq v \leqq g(u)$ and $ux + g(u)f(x) \leqq 1$ for all $x \in [0,1]$. It follows from the continuity of f that $g(u)$ *is the minimum value of the function* $h_u(x) = (1 - ux)/f(x)$ on its domain, i.e., on $[0,1]$ if $f(1) > 0$ and on $[0,1)$ if $f(1) = 0$. So our task reduces to determination of this minimum. We do this by a consideration of cases:

Case 1. Let $M = -\text{Lim}_{x \to 0^+} f'(x)$. Then $M = 0$ if the graph of f has slope 0 at the point $(0,1)$; otherwise $M > 0$. (See FIGURE 2 for an illustration of these possibilities.) Now if $0 \leqq u \leqq M$, the line $y = 1 - ux$ lies above or on the line $y = 1 - Mx$ on $[0,1]$. Since the latter line lies above the curve $y = f(x)$ on $[0,1]$, we have $f(x) \leqq 1 - ux$ for $x \in [0,1]$. Thus $1 \leqq h_u(x)$ for all x while $h_u(0) = 1$. Hence the minimum of $h_u(x)$ is 1, so $g(u) = 1$. (In case $M = 0$, this case is degenerate and applies only to the single point $u = 0$.)

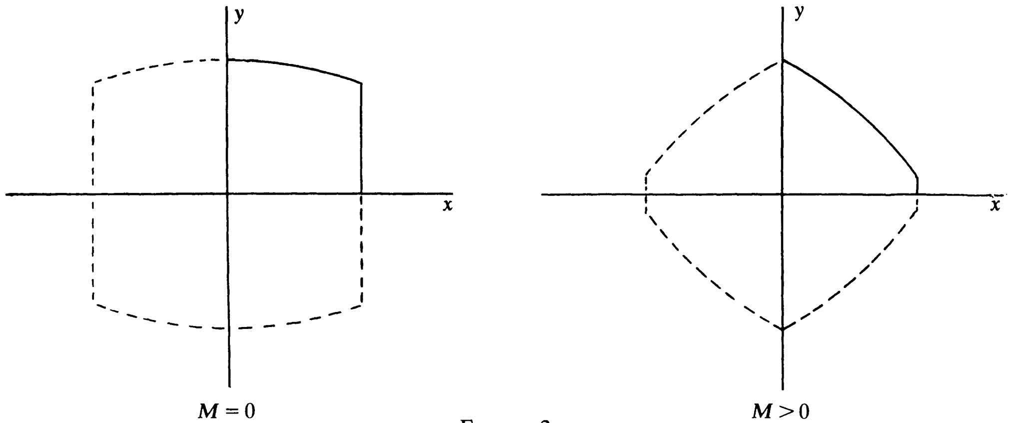


FIGURE 2.

Case 2. Let $L = \text{Lim}_{x\to 1^-} f'(x)/[f'(x) - f(1)]$. Then $L = 1$ if $f(1) = 0$ or if the graph of f approaches the line $x = 1$ with slope tending to $-\infty$; otherwise $L < 1$. Assume the latter for this case, and let $L \leqq u < 1$. Therefore $f(1) > 0$ and $\lim_{x\to 1^-} f'(x)$ is finite; we will denote this limit by N. Then $L = N/(N - f(1))$. Note that if $A > 0$, the function $x/(x - A)$ is decreasing on $(-\infty, 0)$. Since there exists $x' \in (0,1)$ with $f'(x') = f(1) - 1$ and since $f'(x') > N$, we have $L > 1 - f(1)$.

We claim that h_u attains its minimum value at $x_0 = 1$. Assume, to the contrary, that h_u attains its minimum at $x_0 \in [0,1)$. Since $h(0) = 1$ and since $h_u(1) \leqq (1 - L)/f(1) < 1$, it follows that $x_0 \neq 0$. Therefore $h'_u(x_0) = 0$, implying that $-f'(x_0) + u(x_0 f'(x_0) - f(x_0)) = 0$. Also, (*) shows that $x_0 f'(x_0) - f(x_0) < f'(x_0) - f(1)$. By combining these last two facts, we obtain $f'(x_0)/(f'(x_0) - f(1)) > u$. On the other hand, $N < f'(x_0)$ implies that $L > f'(x_0)/(f'(x_0) - f(1)) > u$. The contradiction shows that h_u attains its minimum at $x_0 = 1$ so that $g(u) = h_u(1) = (1 - u)/f(1)$.

Case 3. Assume $M < u < L$. Then the line $y = 1 - ux$ lies below the line $y = 1 - Mx$ on $(0,1]$. Hence h_u must attain its minimum value at a point $x_0 \in (0,1]$. We claim that $x_0 < 1$. If $f(1) = 0$, this is clear. Thus we may assume $f(1) > 0$. If $x_0 = 1$, then we have $(1 - u)/f(1) \leqq (1 - ux)/f(x)$ for $x \in [0,1]$. Observing that $f(x) - xf(1) > 0$ for $x \in [0,1)$, the latter inequality can be solved for u to show that for $x \in [0,1)$

$$[f(x) - f(1)]/[f(x) - f(1) - (x - 1)f(1)] \leqq u.$$

For each such x there exists $\xi_x \in (0,1)$ such that $f(x) - f(1) = f'(\xi_x)(x - 1)$. Consequently, $f'(\xi_x)/[f'(\xi_x) - f(1)] \leqq u$. Letting $x \to 1^-$ forces $\xi_x \to 1^-$. This yields $L \leqq u$, a contradiction. It follows that h_u must attain its minimum at a point $x_u \in (0,1)$. Therefore $h'_u(x_u) = 0$ which can be written as $(ux_u - 1)f'(x_u) = uf(x_u)$.

We summarize this analysis in the following way. Let

$$M = -\lim_{x\to 0^+} f'(x),$$

$$L = \lim_{x\to 1^-} f'(x)/[f'(x) - f(1)],$$

and let x_u be a solution of the equation $(ux - 1)f'(x) = uf(x)$. Then

$$g(u) = \begin{cases} 1, & \text{if } 0 \leqq u \leqq M, \\ \dfrac{1 - ux_u}{f(x_u)}, & \text{if } M < u < L, \\ \dfrac{1 - u}{f(1)}, & \text{if } L \leqq u \leqq 1. \end{cases}$$

Moreover, $g(u)$ is a continuous function on $[0,1]$.

Let us illustrate the application of this formula by returning to the example of the introduction. Suppose we are working with a very simple model in which it has been determined empirically that the set U of pairs of possible economic effect weights is symmetric with respect to the coordinate axes and that in the first quadrant it consists of those points (x, y) that lie below the graph of $f(x) = (3x-4)/(2x-4)$, see FIGURE 3. First note that $M = -f'(0) = 1/4$ and $L = \text{Lim}_{x\to 1^-} f'(x)/[f'(x) - f(1)] = 2/3$. The values of $g(u)$ for $0 \leqq u \leqq 1/4$ and $2/3 \leqq u \leqq 1$ are then easy to obtain from the first and third parts of the formula for g. To get the equation of $g(u)$ for $1/4 < u < 2/3$, one first calculates the value of x_u by solving the equation $(ux-1)f'(u) = uf(x)$ for x. This yields $x_u = (4u - \sqrt{6u - 8u^2})/3u$. It now follows that

$$g(u) = \frac{1 - ux_u}{f(x_u)} = \frac{4}{9}\sqrt{6u - 8u^2} - \frac{4}{9}u + \frac{2}{3}$$

where u ranges through the interval $(1/4, 2/3)$. U^0 is now completely described; it is graphed in FIGURE 4.

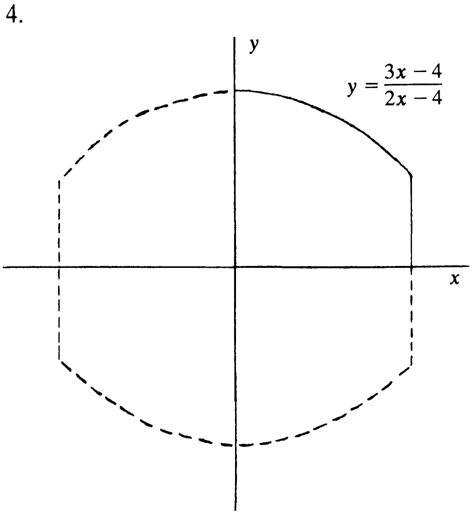


FIGURE 3.

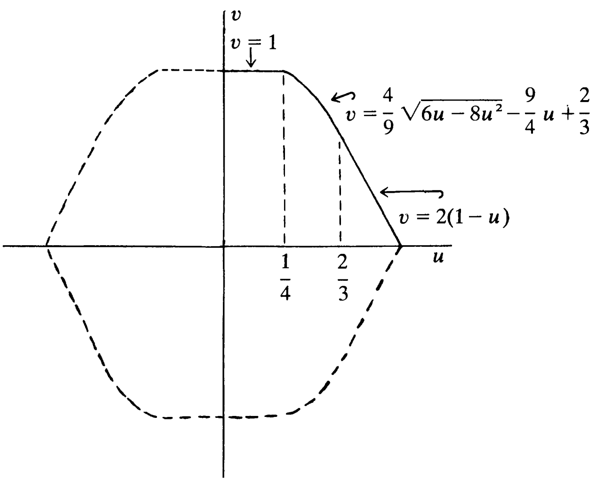
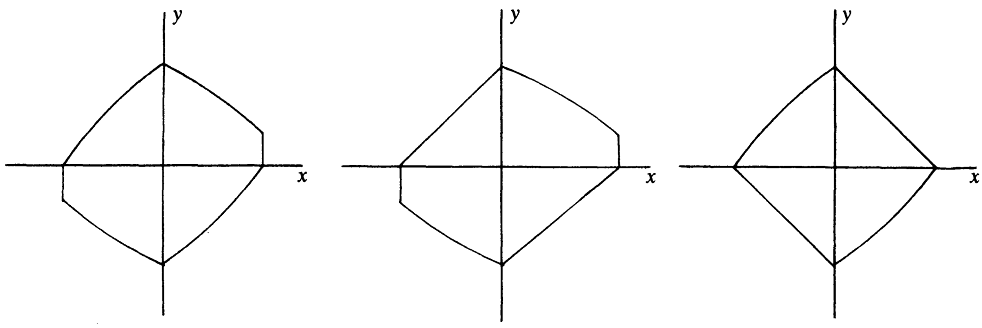


FIGURE 4.

Let us suppose that the units for oil are millions of barrels and that the units for coal are tens of thousands of tons. Thus if we determine that 400,000 barrels ($u = .4$) of imported crude oil will be consumed each day, then we must have $v \leqq g(.4) = .96$ in order to satisfy the constraint $|ux + vy| \leqq 1$ for all $(x, y) \in U$. Therefore, at most 9,600 tons of coal should be consumed.

The reader who has some familiarity with the notion of a norm has by now probably recognized that the sets U we have been considering are sets for which there exists a norm $\|\cdot\|$ on R^2 for which U is the unit ball; that is, $U = \{(x, y) \colon \|(x, y)\| \leqq 1\}$. Recall that a norm on R^2 is a non-negative, real-valued function $\|\cdot\|$ satisfying (a) $\|(x, y)\| = 0$ if and only if $x = y = 0$; (b) $\|(ax, ay)\| = |a|\,\|(x, y)\|$; (c) $\|(x_1 + x_2, y_1 + y_2)\| \leqq \|(x_1, y_1)\| + \|(x_2, y_2)\|$. Standard arguments from linear algebra show that the dual space X of R^2 is isomorphic to R^2 under an isomorphism that associates (u, v) to a linear functional f such that $f(x, y) = ux + vy$ for all $(x, y) \in R^2$. Moreover, the equality $\|f\|^* = \sup\{|ux + vy| : \|(x, y)\| \leqq 1\}$ defines a norm $\|\cdot\|^*$ on X. Thus the problem we are actually considering is one of computing the unit ball of the dual space, commonly called the dual unit ball. This is of importance since the shape of the unit ball determines the "geometry" of the space X when it is equipped with the distance function determined by $\|\cdot\|^*$. (The interested reader can find an elementary exposition of these concepts in [1].)

In closing, we would like to point out that conditions (ii)–(iv) can be somewhat relaxed. For instance, suppose that U satisfies (iii) and (iv) but is symmetric only with respect to the origin, not both



U U_1 U_2

FIGURE 5.

axes. We then write $U = U_1 \cup U_2$, where U_1 is the convex hull of that part of U that lies in the first and third quadrants and U_2 is the convex hull of that part of U that lies in the second and fourth quadrants; the relation between U, U_1 and U_2 is illustrated in FIGURE 5. If we let V_1 denote the points of U_1 which lie in the first quadrant and V_2 the reflection through the y-axis of the points of U_2 which lie in the second quadrant, then the first quadrant boundary functions g_1 and g_2 for V_1^0 and V_2^0, respectively, can be computed using the algorithm we have presented. Straightforward calculations now show that the boundary function g for U^0 in the upper half-plane is given by

$$g(u) = \begin{cases} g_1(u), & 0 \leqq u \leqq 1, \\ g_2(-u), & -1 \leqq u \leqq 0. \end{cases}$$

Also, given any set U in R^2, U^0 is the same as V^0, where V is the smallest convex set containing U. Replacing U by V, we see that condition (iii) is automatically satisfied. One could also consider sets U whose boundaries consist of piecewise smooth curves rather than the smooth curves to which we have restricted our attention. The function g would then be constructed in pieces. We have not emphasized these generalizations here because the added notation and technical considerations would tend to obscure the underlying simplicity of the general case presented in this note. We do hope, however, that the preceding remarks will encourage the interested reader to fill in these details and explore other ways in which the algorithm might be extended. In particular, we invite the reader to find a workable algorithm for computing dual unit balls in R^3.

References

[1] C. Goffman and G. Pedrick, First Course in Functional Analysis, Prentice-Hall, Englewood Cliffs, 1965.
[2] R. H. Lohman and T. J. Morrison, On polars of convex polygons, Amer. Math. Monthly, 81 (1974) 1016–1018.

Subgroups and Equivalence Relations

PIERRE J. MALRAISON, JR.
Control Data Corp.
St. Paul, MN 55112

Equivalence relations have many faces. One face is the set of equivalence classes. This is what one usually studies in abstract algebra: form the set of equivalence classes and try to impose an algebraic structure on that set. Another face appears in the definition and is seldom seen again: an equivalence relation R on a set G is a subset of the product $G \times G$.

If the set G has a group structure, $G \times G$ is also a group with the structure given by the direct product. So an equivalence relation R in this setting is now a subset of a group, and a candidate for being a subgroup. In this note we will look at R as a subobject of $G \times G$ and examine the relation between algebraic structures on R and algebraic objects in G which give rise to R. The questions we shall answer about groups make sense in any algebraic setting where a direct product is defined. We will mention rings in passing at the conclusion of this note; the interested reader may wish to examine the problems in the broader context of universal algebra or category theory.

First we fix notation. G will be a group, written multiplicatively, with identity e. The set R will be an equivalence relation on G. Thus R is a subset of $G \times G$ satisfying the following three conditions:

Reflexivity. For all g in G, (g, g) is in R.

Symmetry. If (g, g') is in R, (g', g) is also in R.

Transitivity. If (g, g') and (g', g'') are both in R, then (g, g'') is also in R.

The set of all elements in G equivalent to an element g is called the equivalence class of g and will be denoted $[g]$. Multiplication in the group $G \times G$ is defined by $(a, b)(c, d) = (ac, bd)$; in particular, in $G \times G$, $(a, b)^{-1} = (a^{-1}, b^{-1})$. Finally, if $a \in G$ and if H is a subgroup of G, a (left) coset aH of H is the set of all elements of the form ah, where h is an element of H.

The most immediate question (and the easiest to answer) is this: when is R a subgroup of $G \times G$? Two conditions must be met: R must be closed under products and inverses. In fact, since R is an equivalence relation, it suffices to have R closed under products. (Suppose that R is closed under products: If (a, b) is in R then $(a^{-1}, a^{-1})(a, b) = (e, a^{-1}b)$ is also in R, by reflexivity and closure under products. Similarly, $(e, a^{-1}b)(b^{-1}, b^{-1}) = (b^{-1}, a^{-1})$ is in R and, by symmetry, $(a^{-1}, b^{-1}) = (a, b)^{-1}$ is also in R.) The following proposition describing R shows that closure under products is both a necessary and a sufficient condition:

PROPOSITION 1. *For G a group and R an equivalence relation on G, the following are equivalent*:

(i) *R is closed under products in $G \times G$;*

(ii) *R is a subgroup of $G \times G$;*

(iii) *R is the relation of belonging to the same coset of some normal subgroup of G.*

Proof. That (i) implies (ii) was shown above. To show (ii) implies (iii), we must first exhibit an appropriate subgroup. So assume R is a subgroup of $G \times G$. We claim that $[e]$ is then a subgroup of G. For if a is in $[e]$, then (a, e) is in R. Product closure for R says that (a, e) and (b, e) in R implies that $(ab, e^2) = (ab, e)$ is in R. Inverse closure says that $(a^{-1}, e^{-1}) = (a^{-1}, e)$ is in R. So $[e]$ is a subgroup of G. Moreover, if (a, e) is in R so is $(c, c)(a, e)(c^{-1}, c^{-1}) = (cac^{-1}, e)$, and thus $[e]$ is a normal subgroup. Finally, if (a, b) is in R, $(e, a^{-1}b)$ is also in R and conversely. In other words, (a, b) is in R if and only if ab^{-1} is in $[e]$; this means precisely that a and b are in the same coset of $[e]$. The implication (iii) implies (i) is just the statement that multiplication of cosets is well defined for normal subgroups.

There are two natural questions raised by Proposition 1. First, the set of cosets of any subgroup of G is a set of equivalence classes. What are the conditions on R that force it to be the relation of belonging to the same left coset of a (not necessarily normal) subgroup of G? Second, if R is a subgroup, when is it normal? The remainder of this note will be devoted to these two questions.

To obtain necessary and sufficient conditions for R to arise from an arbitrary subgroup in the sense described above, we need something a little weaker than the notion of a subgroup. The requisite notion is that of a sub-G-set. Let X be a set and G a group. A **G-action on X** is a map from $G \times X$ to X, $(g, x) \to gx$, such that $g(g'x) = (gg')x$ and $ex = x$. A set X together with a particular G action is called a **G-set.** A subset Y of X such that gy is in Y for every y in Y and g in G is called a **sub-G-set.** The group G itself is a G-set under the action given by multiplication. The direct product $G \times G$ is also a G-set via the "scalar multiplication" $g(a, b) = (ga, gb)$. We use these actions in the following proposition.

Proposition 2. *For G a group and R an equivalence relation on G, the following are equivalent:*
(i) *R is closed under scalar multiplication;*
(ii) *R is the relation of belonging to the same left coset of some subgroup of G.*

Proof. First, let H be a subgroup of G, and let R be the relation of belonging to the same coset of H. Then (a, b) is in R if and only if $a^{-1}b$ is an element of H. We must show that R is closed under the G-action on $G \times G$. But if (a, b) is in R and g is in G, $(ga)^{-1}(gb) = a^{-1}g^{-1}gb = a^{-1}b$ is in H and so $(ga, gb) = g(a, b)$ is in R. Thus (ii) implies (i).

Conversely, suppose R is closed under the G-action of scalar multiplication on $G \times G$. If x and y are in $[e]$, then (x, e) and (y, e) are in R. Act on (y, e) by x to get (xy, x) in R. Transitivity then yields (xy, e) in R so that $[e]$ is closed under products. Act on (x, e) by x^{-1} and apply symmetry to show $[e]$ is closed under inverses as well and is thus a subgroup. We leave to the reader the easy proof that (a, b) is in R if and only if a and b are in the same coset of $[e]$.

We return now to the setting of Proposition 1, where R is a subgroup of $G \times G$ and $[e]$ is a normal subgroup of G. The following example shows that R need not be normal. Let G be the group $Z_2 \times S_3$, the direct product of the cyclic group of order 2 and the symmetric group on three elements. The subgroup $Z_2 \times e$ is normal, but the corresponding R is not normal in $G \times G$. We will see shortly that if $[e]$ is a "nice" normal subgroup, then R must be normal. Even for arbitrary normal subgroups of G, however, the cosets of R are well behaved, as the following proposition indicates:

Proposition 3. *If the relation R is a subgroup of $G \times G$, then there is a 1–1 correspondence between cosets of R in $G \times G$ and cosets of $[e]$ in G.*

Proof. Define a map F from left cosets of $[e]$ to left cosets of R in $G \times G$ by $F(g[e]) = (g, e)R$. The map F is well defined. For suppose $g[e] = k[e]$. Then $g^{-1}k$ is in $[e]$ and so $(g^{-1}k, e)$ is in R. But $(g^{-1}k, e) = (g, e)^{-1}(k, e)$ and thus $(g, e)R = (k, e)R$. The preceding argument can be reversed to show that F is 1–1. Finally, let $(a, b)R$ be a coset of R. Since R is reflexive, (b^{-1}, b^{-1}) is an element of R and so $(a, b)^{-1}(ab^{-1}, e) = (b^{-1}, b^{-1})$ belongs to R and $(a, b)R = (ab^{-1}, e)R = F(ab^{-1}[e])$ and F is onto and thus a 1–1 correspondence.

It follows immediately that if R is normal, $(G \times G)/R = G/[e]$, since the map F defined in the proof is then a homomorphism:

$$F((g[e])(k[e])) = F(gk[e]) = (gk, e)R = (g, e)(k, e)R = ((g, e)R)((k, e)R) = F(g[e])F(k[e]).$$

If G is abelian, R is always normal. If $[e]$ has index 2, R is normal since it also has index 2. To state the next theorem, which gives a necessary and sufficient condition for R to be normal, we recall that an element of G which can be written in the form $gkg^{-1}k^{-1}$ is called a **commutator.** The set of all commutators is not generally a subgroup, so the (normal) subgroup G' generated by that set is defined to be the **commutator subgroup** of G.

Theorem. *If the equivalence relation R is a subgroup of $G \times G$, then the following are equivalent:*
(i) *R is normal in $G \times G$;*
(ii) *$G/[e]$ is abelian;*
(iii) *G' is a subgroup of $[e]$.*

Proof. We begin by showing that (ii) implies (i). Define a mapping P from $G \times G$ to $G/[e]$ by $P(a, b) = a^{-1}b[e]$. Since $G/[e]$ is abelian, P is a homomorphism:

$$P((a, b)(c, d)) = P(ac, bd) = (ac)^{-1}(bd)[e] = (c^{-1}[e])(a^{-1}[e])(b[e])(d[e])$$
$$= (a^{-1}[e])(b[e])(c^{-1}[e])(d[e]) = (a^{-1}b[e])(c^{-1}d[e]) = P(a, b)P(c, d).$$

The kernel of P is precisely R, so R is normal in $G \times G$.

Normality of R means that if (a, b) is in R and x and y are arbitrary elements of G, then (xax^{-1}, yay^{-1}) is also in R. Let g be an element of G. Then (g, g) and $(kgk^{-1}, ege^{-1}) = (kgk^{-1}, g)$ are both in R, and thus $kgk^{-1}g^{-1}$ is in $[e]$. So $[e]$ contains the commutator of any two elements of G: this means that G' is a subgroup of $[e]$, so (i) implies (iii). The equivalence of (ii) and (iii) is standard [**1**, Theorem 12.6, p. 108].

In closing we note that everything up to Proposition 3 has analogs for rings. Let S be a ring, and let an S-set be an S-module without the abelian group structure. Then an equivalence relation R considered as a subobject of $S \times S$ arises from a subring if and only if R is a sub-S-set of $S \times S$, and R arises from an ideal iff R is a subring of $S \times S$. The analog of the theorem is less interesting: R is an ideal in $S \times S$ iff $R = S \times S$.

Reference

[1] J. B. Fraleigh, A First Course in Abstract Algebra, sec. ed. Addison-Wesley, Reading, Mass., 1976.

Powers of $x^2 + 1$

Harold N. Shapiro
Courant Institute of Mathematical Sciences
New York, NY 10012

It can be easily verified that the polynomial $f(x) = x^2 + 1$ satisfies the functional equation

$$f(x)f(x+1) = f(x^2 + x + 1). \tag{1}$$

A natural question which arises is that of determining all non-constant polynomials $f(x)$, with complex coefficients, which satisfy (1). The amusing answer is that *they are precisely the positive integral powers of* $x^2 + 1$.

Note first that the above assertion follows once we show that (1) implies that the roots of $f(x)$ are precisely i and $-i$. For then $f(x) = c(x-i)^{r_1}(x+i)^{r_2}$ where $c \neq 0$ is complex. Comparing the coefficient of the highest power of x on both sides of (1) yields $c^2 = c$ and hence $c = 1$. Inserting this $f(x)$ in (1) gives

$$\begin{aligned}(x-i)^{r_1}&(x+i)^{r_2}(x+1-i)^{r_1}(x+1+i)^{r_2}\\ &= (x^2+x+1-i)^{r_1}(x^2+x+1+i)^{r_2}\\ &= (x-i)^{r_1}(x+1+i)^{r_1}(x+i)^{r_2}(x+1-i)^{r_2}\end{aligned}$$

which implies $r_1 = r_2$ and hence that $f(x) = (x^2+1)^{r_1}$.

We turn now to the task of determining the roots of $f(x)$. If θ is such a root, (1) implies that $\theta^2 + \theta + 1$ is also a root of $f(x)$. Also, since $\theta - 1$ is a root of $f(x+1)$, $(\theta-1)^2 + (\theta-1) + 1 = \theta^2 - \theta + 1$ must also be a root of $f(x)$. In other words, introducing $T(x) = x^2 + x + 1$, both $T(x)$ and $T(-x)$ must map the set of roots of $f(x)$ into itself. The desired result is then a consequence of the following general fact:

Lemma. *If $\mathcal{S}$ is a non-empty finite set of complex numbers such that both $T(x)$ and $T(-x)$ map $\mathcal{S}$ into itself, then $\mathcal{S}$ consists of i and $-i$.*

Proof. We observe that $\mathrm{Re}(\theta^2+1)>0$ implies that

$$\mathrm{Re}\, T(\pm\theta)=\mathrm{Re}(\theta^2+1)\pm\mathrm{Re}\,\theta > \pm\mathrm{Re}\,\theta,$$

and $\mathrm{Re}(\theta^2+1)<0$ implies that

$$-\mathrm{Re}\, T(\mp\theta)= -\mathrm{Re}(\theta^2+1)\pm\mathrm{Re}\,\theta > \pm\mathrm{Re}\,\theta.$$

Thus we have that $\mathrm{Re}(\theta^2+1)\neq 0$ implies that

$$\max\{|\mathrm{Re}\, T(\theta)|,|\mathrm{Re}\, T(-\theta)|\}>|\mathrm{Re}\,\theta|. \tag{2}$$

Similarly, $\mathrm{Im}(\theta^2+1)\neq 0$ implies that

$$\max\{|\mathrm{Im}\, T(\theta)|,|\mathrm{Im}\, T(-\theta)|\}>|\mathrm{Im}\,\theta|. \tag{3}$$

Note also that $\mathrm{Re}(\theta^2+1)=0$ implies that $|\mathrm{Re}\, T(\pm\theta)|=|\mathrm{Re}\,\theta|$; and $\mathrm{Im}(\theta^2+1)=0$ implies that $|\mathrm{Im}\, T(\pm\theta)|=|\mathrm{Im}\,\theta|$.

Starting with any $\theta_0\in\mathscr{S}$, suppose $\mathrm{Re}(\theta_0^2+1)\neq 0$, so that using (2) we have for $\theta_1=T(\varepsilon\theta_0)$, $\varepsilon=\pm 1$, that $|\mathrm{Re}\,\theta_1|>|\mathrm{Re}\,\theta_0|$. Continuing in this way, as long as we have $\mathrm{Re}(\theta_k^2+1)\neq 0$, $k=0,\dots,m-1$, we obtain $\theta_0,\theta_1,\dots,\theta_m$ in $\mathscr{S}$ such that

$$|\mathrm{Re}\,\theta_0|<|\mathrm{Re}\,\theta_1|<\cdots<|\mathrm{Re}\,\theta_m|. \tag{4}$$

Thus these θ_k are distinct, and since $\mathscr{S}$ in finite this process must stop, i.e., we reach a θ_m such that $\mathrm{Re}(\theta_m^2+1)=0$. If $\mathrm{Im}(\theta_m^2+1)\neq 0$, use (3) to obtain $\theta_{m+1}=T(\eta\theta_m)$, $\eta=\pm 1$, such that $|\mathrm{Im}\,\theta_{m+1}|>|\mathrm{Im}\,\theta_m|$. Then θ_{m+1} is distinct from θ_m, and since $|\mathrm{Re}\,\theta_{m+1}|=|\mathrm{Re}\,\theta_m|$ (4) insures that θ_{m+1} is distinct from the θ_k, $k=0,\dots,m-1$. We now continue this new procedure so long as $\mathrm{Re}(\theta_k^2+1)=0$ and $\mathrm{Im}(\theta_k^2+1)\neq 0$, obtaining thereby $\theta_{m+1},\dots,\theta_{m'}$, such that

$$|\mathrm{Re}\,\theta_m|=|\mathrm{Re}\,\theta_{m+1}|=\cdots=|\mathrm{Re}\,\theta_{m'}| \tag{5}$$

and

$$|\mathrm{Im}\,\theta_m|<|\mathrm{Im}\,\theta_{m+1}|<\cdots<|\mathrm{Im}\,\theta_{m'}|. \tag{6}$$

Again, since $\mathscr{S}$ is finite this stops either with $\mathrm{Re}(\theta_{m'}^2+1)\neq 0$, or with $\mathrm{Re}(\theta_{m'}^2+1)=\mathrm{Im}(\theta_{m'}^2+1)=0$. In the former case we resume again with the previous process with the real parts which continues to produce elements of $\mathscr{S}$ with larger and larger real parts. When the absolute value of these real parts stays constant the processing with imaginary parts is resumed. The only thing which brings this process to a halt (as it must) is when one first reaches a θ_n such that $\mathrm{Re}(\theta_n^2+1)=\mathrm{Im}(\theta_n^2+1)=0$. Then, of course, $\theta_n^2+1=0$ and $\theta_n=\pm i$. If this is not our very initial step (as we've presumed) we will produce a contradiction. For then either $T(\theta_{n-1})=\pm i$ or $T(-\theta_{n-1})=\pm i$; and this implies that θ_{n-1} is either $\pm i$ or $\pm 1\pm i$. The possibility that θ_{n-1} is one of $\pm 1\pm i$ is untenable. Since $|\mathrm{Im}(\pm 1\pm i)|=|\mathrm{Im}(\pm i)|$, we see from (6) that this termination could not have occurred during a processing of the imaginary parts; and since $|\mathrm{Re}(\pm 1\pm i)|>|\mathrm{Re}(\pm i)|$ we see from (4) that it could not have occurred during a processing of the real parts. Thus the only possibility is that θ_0 was i or $-i$. Finally, noting that $T(-(\pm i))=\mp i$, the lemma follows.

While the lemma completes our characterization of polynomial solutions to equation (1), there remain further questions worth investigating. For example, what are the continuous solutions to this functional equation? What is the most general solution? Interested readers are referred to **[1, 2]** for further information.

References

[1] J. Aczél, Lectures on Functional Equations and Their Applications, Academic Press, New York, 1966.

[2] M. Luczma, Functional Equations in a Single Variable, Polish Scientific Publishers, Warszawa, 1968.

A Talmudic Approach to the Area of a Circle

SHELDON EPSTEIN
Seton Hall University
South Orange, NJ 07079

MURRAY HOCHBERG
Brooklyn College (CUNY)
Brooklyn, NY 11210

Most elementary approaches to proving that the area of a circle is πr^2 employ various sorts of approximation techniques. For example, the circle is repeatedly bisected by diameters until the resulting sections resemble triangles which when placed together yield the desired result [**1**, pp. 466–467], or by showing that a circle becomes indistinguishable from a regular polygon as the number of sides increases [**2**, p. 22]. The primary objection to these methods lies in the fact that a certain degree of mathematical sophistication is required to appreciate the concept of the limiting triangle or limiting regular polygon, and this is especially difficult to convey to students who are not familiar with the limit concept in calculus. An imaginative technique which avoids this pitfall and is readily reproduceable with simple props in a classroom can be found in the commentaries of the Talmud (Tosfos Pesachim 109a, Tosfos Succah 8a, Marsha Babba Bathra 27a), where the need for explaining geometrical formulas to people lacking in mathematical sophistication was paramount.

Consider a circle of radius r as being composed of a large number of smaller concentric circles (strings), evenly spaced from each other, starting from the center of the circle (a point) and emanating outwards. If one were then to slice the circle from the top to the origin along a radius and pull each of the strings taut parallel to the bottom tangent of the circle (FIGURE 1) while still maintaining the

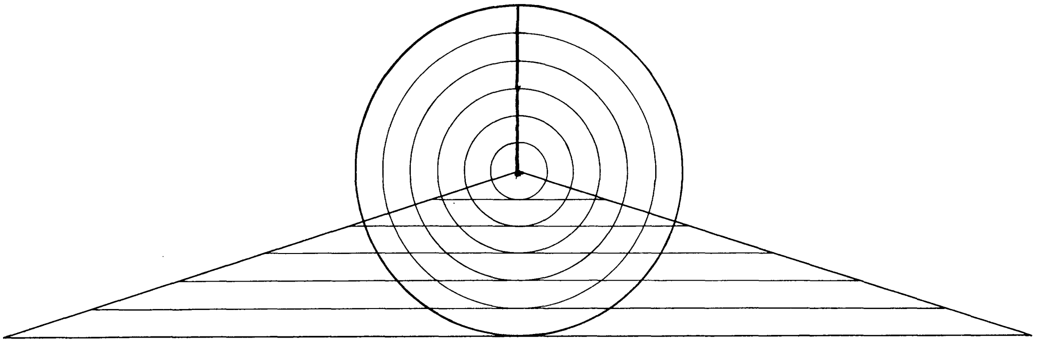


FIGURE 1.

distances between them, the result would be a succession of straight lines with a constant decrease in length as we approach the center. These lines form an isosceles triangle whose base is the circumference of the circle, since it was constructed from the outermost string, and whose altitude can plainly be seen to equal the radius of the circle. Employing the formula for the area of a triangle, we obtain the area of circle to be $\frac{1}{2}(2\pi r)r = \pi r^2$.

References

[1] J. W. Heddens, Today's Mathematics, third ed., Science Research Associates, Chicago, 1974.
[2] J. Stepelman, Milestones in Geometry, Macmillan, New York, 1970.

PROBLEMS

DAN EUSTICE, Editor
LEROY F. MEYERS, Associate Editor
The Ohio State University

Proposals

To be considered for publication, solutions should be mailed before April 1, 1978

1021. Prove or disprove that a countably infinite set of positive real numbers with a finite non-zero cluster point can be arranged in a sequence, $\{a_n\}$, so that $\{(a_n)^{1/n}\}$ is convergent. (See SOLUTIONS, Problem 972, this MAGAZINE.) [*Peter Ørno, The Ohio State University.*]

1022. We have n cards numbered 1 though n. Find the expected number of drawings needed to put the cards in order by each of the following strategies:

(a) The shuffled cards are drawn without replacement until card 1 is drawn. The remaining $n-1$ cards are shuffled and drawn without replacement until card 2 is drawn. This process is continued until all the cards are drawn and put in linear order.

(b) A card, say card k, is drawn from the shuffled deck. The remaining cards are shuffled and drawn without replacement until either card $k-1$ or card $k+1$ is drawn. We identify card $1-1$ as card n and card $n+1$ as card 1. This process is continued until all the cards are drawn and put in circular order. [*Joe Dan Austin, Emory University.*]

1023.* Call a triangle *super-Heronian* if it has integral sides and integral area, and the sides are consecutive integers. Are there infinitely many distinct super-Heronian triangles? [*Steven R. Conrad, Benjamin N. Cardozo H. S., Bayside, N. Y.*]

1024. In many athletic leagues the progress of teams is reported both in terms of winning percentage and in terms of "games behind" the league leader, defined as the difference in games won minus the difference in games lost, divided by 2. Sports fans often observe, especially early in the season, that the league leader in percentage (the official standard) is behind some other team in games.

Suppose team A is the percentage leader, but team B is ahead of Team A in games. Assume no ties.

a. Which team has played more games?

b. What is the minimum difference in number of games played?

ASSISTANT EDITORS: DON BONAR, *Denison University*: 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 Dan Eustice, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210.*

c. Characterize possible won/lost records for the two teams if the difference in number of games played is minimal.

d. Is it possible for this anomaly to occur late in the season?
[*David A. Smith, Duke University.*]

Solutions

A Geometric Inequality: Completed **November 1975**

959. Let P be a point in the interior of the triangle ABC and let r_1, r_2, r_3 denote the distances from P to the sides of the triangle. Let R denote the circumradius of ABC. Show that

$$\sqrt{r_1}+\sqrt{r_2}+\sqrt{r_3}\leqq 3\sqrt{R/2},$$

with equality if and only if ABC is equilateral and P is the center of ABC. [*L. Carlitz, Duke University.*]

Comment. The n-dimensional extension of this problem (Jan. 1977) is not entirely complete. The verification of the extreme point was said to be easy and consequently was not done. However, since this is a maximum problem subject to the constraints $\Sigma\, e_i/h_i = 1$, $e_i \geqq 0$, one has to check for extrema on all the boundaries of the constraint domain, which consists of very many faces of dimensions 0 to $n-1$. Here we give a still further extension with a simple (non-calculus) proof using Hölder's inequality.

We will show that

$$\left\{\sum x_i^{2p/(2p-3)}\right\}^{(2p-3)/2p}\cdot\left\{\frac{R^2(n+1)^3}{n^2}\right\}^{1/2p}\geqq \sum x_i r_i^{1/p} \tag{1}$$

where x_i ($i = 1,2,\dots,n+1$) are arbitrary non-negative numbers, r_i are the distances from an interior point P of an n-simplex to the $n-1$ dimensional faces, R is the circumradius of the simplex, and p is any number greater than 3/2. Letting $x_i = 1$ and $p \to 3/2$, we recapture the extension given previously by Gerber. For $n = 2$ and $p = 2$, we obtain

$$\frac{27R^2}{4}(x_1^4+x_2^4+x_3^4)\geqq\{x_1\sqrt{r_1}+x_2\sqrt{r_2}+x_3\sqrt{r_3}\}^4. \tag{2}$$

There is equality in (1) if and only if the simplex is regular, P is the centroid and the x_i are equal.

Proof. By Hölder's inequality,

$$\left\{\sum r_i/h_i\right\}^{1/p}\left\{\sum x_i^q h_i^{q/p}\right\}^{1/q}\geqq\sum x_i r_i^{1/p} \tag{3}$$

and

$$\left\{\sum x_i^{2p/(2p-3)}\right\}^{q(2p-3)/2p}\cdot\left\{\sum h_i^2\right\}^{q/2p}\geqq\sum x_i^q h_i^{q/p}, \tag{4}$$

where $1/p+1/q=1$ and $p>3/2$. Combining (3) and (4), using $\Sigma\, r_i/h_i = 1$ ($r_i = e_i$ in Gerber's notation and h_i = altitude of simplex from vertex i), we get

$$\left\{\sum x_i^{2p/(2p-3)}\right\}^{(2p-3)/2p}\cdot\left\{\sum h_i^2\right\}^{1/2p}\geqq\sum x_i r_i^{1/p}. \tag{5}$$

Finally, using

(6) $$\sum h_i^2 \leqq \sum m_i^2, \ (m_i = \text{median of the simplex from vertex } i)$$

and

(7) $$n^2 \sum m_i^2 \leqq R^2(n+1)^3,$$

we obtain (1).

Although Gerber notes that (7) is an immediate consequence of Lagrange's identity (which may have been known to Leibniz), we include a proof for completeness.

Let $\mathbf{V}_i, \mathbf{G}$ denote vectors from the circumcenter O to the vertices V_i and to the centroid, respectively, of the simplex. Then

$$\sum \mathbf{V}_i^2 = \sum \{(\mathbf{V}_i - \mathbf{G}) + \mathbf{G}\}^2$$

$$= \sum |V_iG|^2 + (n+1)|OG|^2.$$

Since $|\mathbf{V}_i|^2 = R^2$ and $|V_iG| = nm_i/(n+1)$, we obtain (7). Further applications of this polar moment of inertia identity are given in this MAGAZINE, 48 (1975), 44–46.

MURRAY S. KLAMKIN
University of Alberta

Probability of Sums **March 1976**

970. A plus or minus sign is assigned randomly to each of the numbers $1, 2, 3, \ldots, n$. What are the probabilities that the sum of the signed numbers is positive, negative, and zero? [*Martin Berman, Bronx Community College.*]

Solution: Let P, N, and Z denote the events that the sum of the signed numbers is positive, negative, and zero, respectively. Clearly, $\Pr(P) = \Pr(N)$, since there is a one-to-one correspondence between the elements of P and the elements of N given by changing the signs of all the numbers.

Now, if x and y denote the absolute values of the sums of the positive and negative numbers, respectively, taken from an element of Z, then $x - y = 0$ and $x + y = n(n+1)/2$. Thus, since x and y must be integers, Z is the null event when $n(n+1)/2$ is odd. That is $\Pr(Z) = 0$ and $\Pr(P) = \Pr(N) = 1/2$ when $n \equiv 1$ or $n \equiv 2 \pmod 4$.

When $n \equiv 0$ or $n \equiv 3 \pmod 4$, $x = y = n(n+1)/4$. Hence, if $d_m(n)$ denotes the number of partitions of n into distinct parts, none larger than m, then

$$\Pr(Z) = 2^{-n} d_n\left(\frac{n(n+1)}{4}\right) \quad \text{and} \quad \Pr(P) = \Pr(N) = \frac{1 - \Pr(Z)}{2}.$$

The function $d_m(n)$ may be determined from the generating function $\sum_{m=0}^{\infty} d_m(n) t^n = \prod_{j=1}^{m} (1 + t^j)$ or from the recurrence relation $d_m(n) = d_{m-1}(n) + d_{m-1}(n-m)$ with boundary conditions $d_m(0) = 1$ and $d_m(n) = 0$ if $n < 0$ or $n > m(m+1)/2$.

F. G. SCHMITT, JR.
Berkeley, California

Also solved by Michael W. Ecker, Richard A. Gibbs, J. M. Gil (Portugal), Pambuccian Victor (Romania), and the proposer.

971. In designing pipes and other conduits it is usually desirable to enclose the maximum cross-sectional area for a given weight of pipe. Mathematically, this may be simplified by enclosing the maximum area for a given perimeter.

Dual ducts are often used to convey fluids in two directions. They have a portion of their perimeter in common. For example, two equal squares, each of side s are placed to share a common side. The total perimeter is $7s$ and the total cross-sectional area is $2s^2$. Thus, the ratio of the area to the square of the perimeter is 2/49. Assume equal cross-sectional area of the two ducts.

(i) Which regular polygon is the most efficient for use as a dual duct?

(ii)* Which contour is the most efficient for use as a dual duct?

[*Sidney Kravitz, Dover, New Jersey.*]

Solution. (i) Let $A(n)$ be the area of an n-sided regular polygon, each side having length $s = P/(2n-1)$ where P is the fixed perimeter of the dual duct. Thus

$$A(n) = \frac{P^2}{4}\frac{n}{(2n-1)^2}\cot\frac{\pi}{n}.$$

We first show that the function $A(x)$ is decreasing for $x \geqq a = \pi/\cot^{-1}\pi$ (approximately 10.2). Replacing x by π/y, this is equivalent to showing that

$$f(y) = \frac{P^2\pi y}{4(2\pi-y)^2}\cot y = \frac{P^2\pi}{4}\frac{y}{\sin y}\frac{\cos y}{(2\pi-y)^2}$$

is increasing for $0 < y \leqq \pi/a = \cot^{-1}\pi$. But

$$\frac{d}{dy}\frac{y}{\sin y} = \frac{\sin y - y\cos y}{\sin^2 y} = \frac{\int_0^y t\sin t\,dt}{\sin^2 y} > 0$$

for $0 < y \leqq \pi/a$, and

$$\frac{d}{dy}\frac{\cos y}{(2\pi-y)^2} = \frac{2\cot y + (y-2\pi)}{(2\pi-y)^3}\sin y > \frac{2\cot(\pi/a) - 2\pi}{(2\pi-y)^3}\sin y = 0$$

for $0 < y \leqq \pi/a$, so that the assertion follows.

Finally, direct computation of $A(n)$ for $3 \leqq n \leqq 11$ shows that $A(7)$ is the largest (approximately $.0215\,P^2$). Thus the most efficient polygon for use as a dual duct is the one with seven sides.

(ii) Let P be the fixed perimeter, l the length of the common side, and C the length of each of the two outside contours. Then $P = l + 2C$. Since a circle gives the maximum area for a given perimeter, it follows that the most efficient contour for use as a dual duct is the circular arc of length $C = (P-l)/2$.

It is necessary then to find the circular arc which maximizes the enclosed area. If r is the radius of the circle and 2θ is the central angle determined by l $(0 \leqq \theta < \pi)$, then $l = 2r\sin\theta$, $(P-l)/2 = C = 2(\pi-\theta)r$, and $A = r^2(\pi - \theta + \cos\theta\sin\theta)$. Eliminating r and l gives

$$A = \frac{P^2(\pi-\theta+\cos\theta\sin\theta)}{4(\sin\theta + 2(\pi-\theta))^2},$$

and letting $\theta = \pi - \phi$ yields

$$A = \frac{P^2(\phi - \cos\phi\sin\phi)}{4(\sin\phi + 2\phi)^2}.$$

Differentiating A with respect to ϕ, we find

$$\frac{dA}{d\phi} = \frac{P^2(\sin\phi - \phi\cos\phi)(1+2\cos\phi)}{2(\sin\phi+2\phi)^3} = \frac{P^2\left(\int_0^\phi t\sin t\,dt\right)(1+2\cos\phi)}{2(\sin\phi+2\phi)^3}.$$

Hence A is increasing for $0<\phi<2\pi/3$ and decreasing for $2\pi/3<\phi\leqq\pi$, so that A is maximal for $\phi=2\pi/3$, or equivalently, for $\theta=\pi/3$. It follows that the most efficient contour for one side of a dual duct is two thirds of a circle. Its length is $C=(P-l)/2$ where $l=(1+(8\sqrt{3}/9)\pi)^{-1}P$.

H. T. SEDINGER
University of Portland

Also solved by Jordi Dou (Spain), and Michael Goldberg. Partial solutions by Thomas E. Elsner, Lew Kowarski, and the proposer. Sedinger also provided a generalization to a conduit of n identical subducts, $n\geqq 2$.

Converges to One **March 1976**

972. Prove or disprove that the set of all positive rational numbers can be arranged in an infinite sequence, $\{a_n\}$, such that $\{(a_n)^{1/n}\}$ is convergent. [*Marius Solomon, Student, University of Pennsylvania.*]

Solution: From the familiar diagram,

1	1/2	1/3	1/4	1/5	$\cdots$
2	2/2	2/3	2/4	2/5	$\cdots$
3	3/2	3/3	3/4	3/5	$\cdots$
4	4/2	4/3	4/4	4/5	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

we use the usual serpentine method to form the sequence $\{a_n\}=\{1,1/2,2,3,1/3,\ldots\}$, where we omit all fractions not reduced to lowest terms. Since every element in the nth row is less than or equal to n, and every element in the nth column is greater than or equal to $1/n$, we see that $1/n\leqq a_n\leqq n$.

Then $(1/n)^{1/n}\leqq a_n^{1/n}\leqq n^{1/n}$. But $\lim_{n\to\infty}n^{1/n}=\lim_{n\to\infty}(1/n)^{1/n}=1$, thus $\lim_{n\to\infty}a_n^{1/n}=1$, and $\{a_n^{1/n}\}$ converges.

JORDAN I. LEVY, student
University of Delaware

Also solved by Daniel R. Buskirk, Michael W. Ecker, Donald C. Fuller, Eli Leon Isaacson, Peter Lindstrom, James McKim, Gerhard Metzen (Canada), Hugh Noland, Denny Noto, Adam Riese, A. Solomon, J. M. Stark, Pambuccian Victor (Romania), Western Maryland College Problems Seminar, and the proposer.

The Period of N^{-1} **March 1976**

973*. Let N be an odd number; if the period of N^{-1} is P in base b, and if $N^2\nmid b^P-1$, then the period of N^{-n} in base b is PN^{n-1}. [*Robert Cranga, University of Paris.*]

Counterexample. Let $N=21$. The period of 21^{-1} is 6 in base 10 and $21^2\nmid 10^6-1$. But the period of 21^{-2} is 42 and not $126=6\cdot 21$.

KAY P. LITCHFIELD
Clearfield, Utah

Solution. The proposed result becomes a true theorem if N is prime.

Proof. Let N^{-1} have period P base b. Then b has order P (mod N); i.e., $b^P-1\equiv 0\pmod N$. Now suppose b has order e (mod N^2). Then $b^e-1\equiv 0\pmod{N^2}$ and $b^e-1\equiv b^P-1\equiv 0\pmod N$ which implies $e = Pr$ for some integer r. Furthermore r is the smallest positive integer such that $b^{Pr} - 1 \equiv 0 \pmod{N^2}$. Then since $b^P = 1 + kN$ for some integer k such that $N \nmid k$ we have $b^{Pr} - 1 = (1 + kN)^r - 1 \equiv krN \equiv 0 \pmod{N^2}$. Hence $N \mid r$. Finally since r is the smallest positive integer such that $b^{Pr} - 1 \equiv 0 \pmod{N^2}$ we have $r = N$ so that b has order $e = PN \pmod{N^2}$ and N^{-2} has period PN.

By induction this can be extended so that N^{-n} has period PN^{n-1}. In the event that $N^2 \mid b^P - 1$ let α be the exponent of the highest power of the prime N which divides $b^P - 1$. Then a similar argument establishes that N^{-m} has period P if $n \leqq \alpha$ and N^{-n} has period $PN^{\alpha - n}$ if $n > \alpha$.

Finally suppose R is an odd number base b such that $(R, b) = 1$. Let $R = N_1^{a_1} N_2^{a_2} \cdots N_s^{a_s}$ be the prime decomposition of R and let p_i denote the period base b of $N_i^{-a_i}$ for $i = 1, 2, \ldots, s$. Then R has period $P^* = [P_1, P_2, \ldots, P_s]$ where $[P_1, P_2, \ldots, P_s]$ denotes the least common multiple of $P_1, P_2, \ldots, P_s$. This result and the "corrected theorem" are generalizations of discoveries of Thibault (p. 164, Vol. I of Dickson's *History of the Theory of Numbers*).

Kenneth M. Wilke
Topeka, Kansas

The Limit is One **March 1976**

974. Let $n^{[i]} = n(n-1)\cdots(n-i+1)$. For k a positive integer, evaluate

$$\underset{n\to\infty}{\text{Limit}} \sum_{i=1}^{n} \frac{n^{[i]}}{(2n+k)^{[i]}}.$$

[*John P. Hoyt, Indiana, Pennsylvania.*]

Solution I: Fix k and define $a_{in} = n^{[i]}/(2n+k)^{[i]}$. Then

$$a_{in} = \frac{n}{2n+k} \cdot \frac{n-1}{2n+k-1} \cdot \cdots \cdot \frac{n-i+1}{2n+k-i+1} \leqq \left(\frac{1}{2}\right)^i$$

since $(n-j)/(2n+k-j) \leqq \frac{1}{2}$. Thus, $\sum_{i=1}^n a_{in} \leqq \sum_{i=1}^n (\frac{1}{2})^i \leqq 1$. Also, $\lim_{n\to\infty} a_{in} = (\frac{1}{2})^i$ since $\lim_{n\to\infty} (n-j)/(2n+k-j) = \frac{1}{2}$. Therefore, if $N \geqq 1$ is fixed

$$\lim_{n\to\infty} \sum_{i=1}^{n} a_{in} \geqq \lim_{n\to\infty} \sum_{i=1}^{N} a_{in} = \sum_{i=1}^{N} \lim_{n\to\infty} a_{in} = \sum_{i=1}^{N} \left(\frac{1}{2}\right)^i.$$

It follows immediately that the sum is 1.

Eli Leon Isaacson
New York University

Solution II: By induction or otherwise one can prove that

$$S_n = \sum_{i=1}^{n} \frac{n^{[i]}}{(2n-k)^{[i]}} = \frac{n}{n+k+1},$$

whence it follows that the limit is one.

Henry W. Gould
Morgantown, West Virginia

Also solved by Barry C. Arnold, J. M. Brown & D. A. Voss, P. G. Chauveheid (Belgium), Romae J. Cormier, Irving Allen Dodes, Donald C. Fuller, M. G. Greening (Australia), Richard A. Groeneveld, Eldon Hansen, H. Kappus (Germany), Jordan I. Levy, Peter W. Lindstrom, J. M. Stark, Pambuccian Victor (Romania), Edward T. Wang (Canada), and the proposer.

REVIEWS

PAUL J. CAMPBELL, Editor
Beloit College

PIERRE MALRAISON, Editor
Control Data Corp.

Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of the mathematics literature. Reviews of books are adapted from the Telegraphic Reviews in the American Mathematical Monthly.

Stewart, Ian, *Gauss*, Scientific American 237 (July 1977) 122-131, 154.

Sketch of Gauss's life and work, emphasizing his use of inductive reasoning. A construction method for a regular polygon of 17 sides is illustrated in detail on p. 122.

Mandelbrot, Benoit, Fractals: Form, Chance, and Dimension, Freeman, 1977, xvi + 365 pp, $14.95.

A unique enterprise in natural philosophy accompanied by striking computer graphics. Technically, a fractal is a set whose Hausdorff-Besicovitch (or fractal) dimension exceeds its topological dimension; the difference is a measure of irregularity and fragmentation. Familiar fractals include Brownian motions, Cantor sets, and snowflake and Peano curves. But fractals model such diverse phenomena as Swiss cheese, the earth's surface, clusterings of stars, shapes of clouds, windings of rivers and coastlines, turbulence, curdling, liquid crystals, and word frequencies. The book is written in informal essay style; but readers would do well first to read P. Morrison's review of the 1975 French edition (*Scientific American* 233 (November 1975) 143-144) and M. Gardner's column on it (*Scientific American* 235 (December 1976) 124-128, 133, 152).

Kolata, Gina Bari, *The Calabi conjecture: a proof after 25 years*, Science 196 (17 June 1977) 1308.

S.T. Yau (Stanford) has closed the remaining gap to prove a 1954 conjecture of Eugenio Calabi: if the volume of a higher-dimensional surface is known, then a particular kind of metric (the Kähler metric) can be found which reflects the geometry of the surface. Plugging the gap involved the solution of especially difficult non-linear partial differential equations.

Graham, Ronald L. and Garey, Michael R., *The limits to computation*, 1978 Yearbook of Science and the Future, Encyclopedia Britannica, 1977, pp. 170-185.

An informal discussion of computational complexity of algorithms and the problems they solve. "Nice" problems admit a polynomial algorithm: its time complexity grows no faster than a constant power of the size of the problem instance. The bin-packing problem, the minimal network problem, and the famous travelling salesman problem are given as illustrations of a large class of equivalent problems (called NP-complete, for "nondeterministic polynomial time") for which it is unknown if polynomial algorithms exist.

Davis, Philip J., *Proof, completeness, transcendentals, and sampling*, Journal of the Association for Computing Machinery 24 (April 1977) 298-310.

Can a few examples prove a theorem? Surprisingly, the answer is yes! Davis examines several possible methodologies of proof in the context of Pappus's theorem on collinearity of crisscross points: 1) computer proof, by brute-force analytic geometry in a formal algebraic language (leading to an identity in several thousand monomials); 2) "finite sampling"--proof by exact arithmetical verification for a "complete" set of cases; 3) reduction to *one* numerical verification, of a "transcendental" configuration; and 4) verification of random configurations.

Unprovable problem in arithmetic, Science News 111 (11 June 1977) 373-374.

In 1931 Gödel established that there are elementary statements true about the natural numbers but not deducible from the Peano axioms. This April, Jeffrey Paris (Manchester) exhibited the first such statement of substantial mathematical interest in its own right. It asserts the existence of certain Ramsey numbers; they concern the ways objects can be arranged in patterns. Paris found a non-standard model of the Peano axioms for which the statement is false, yet its truth for the natural numbers can be proved using "extra-Peano" means (namely, Ramsey's Theorem for infinite sets). "The Ramsey numbers grow very large very quickly, so rapidly in fact that they outstrip the power of the axioms of arithmetic to keep pace with them."

Brams, Steven J. and Davis, Morton D., *A game-theory approach to jury selection*, Trial 12 (December 1976) 47-49.

The argument, based in part on a mathematical model: all but one of the various procedures prescribed for exercise of peremptory challenges in selection of a trial jury may be contrary to the U.S. Constitution and Supreme Court rulings. The questioned procedures involve sequential exercise of challenges over time, while the sole unobjectionable procedure--the "struck jury system"--provides that challenges be exercised all at once.

Brams, Steven J. and Muzzio, Douglas, *Game theory and the White House tapes case*, Trial 13 (May 1977) 48-53.

A use of simple game theory to explain why U.S. Supreme Court Justices Burger and Blackmun departed from their apparent personal preferences to side with the majority in an important July 1974 unanimous decision upholding subpoena of tapes and documents from Richard Nixon.

Clark, Colin W., Mathematical Bioeconomics: The Optimal Management of Renewable Resources, Wiley, 1976; xi + 352 pp, $21.95.

A systematic survey of the developing theory of conservation of productive resources (e.g., fish, forests, orchards), using only basic techniques of calculus. Replete with examples of current interest and sometimes surprising conclusions (e.g., immediate extermination of a species may be the most "profitable" long-range policy). Reviewed at length in *Science* 196 (3 June 1977) 1082.

May, Robert M., *Simple mathematical models with very complicated dynamics*, Nature 261 (10 June 1976) 459-467.

"First-order difference equations arise in many contexts in the biological, economic and social sciences. Such equations, even though simple and deterministic, can exhibit a surprising array of dynamical behavior..." For a slightly more sophisticated treatment, see also Robert M. May and George F. Oster, "Bifurcations and dynamic complexity in simple ecological models", *American Naturalist* 110 (1976) 573-599.

Cook, Thomas M. and Alprin, Bradley S., *Snow and ice removal in an urban environment*, Management Science 23 (November 1976) 227-234.

Case history of a successful operations research project. The authors devised a simple dynamic routing heuristic to minimize the time required to salt snow routes and then demonstrated its value to city officials with a simulation. Real data is given; the improvement was 36%.

Harville, David, *The use of linear-model methodology to rate high school or college football teams*, J. American Statistical Association 72 (June 1977) 278-289.

Raup, David M., *Probabilistic models in evolutionary paleobiology*, American Scientist 65 (January-February 1977) 50-57.

An intriguing investigation of the consequences of modelling natural selection as a random process: "[A] selection-free model will inevitably produce certain patterns that we have always assumed were possible only as a result of natural selection."

Weinhold, Frank, *Thermodynamics and geometry*, Physics Today 29 (March 1976) 23-26, 28-30.

Shows that the Gibbsian equilibrium surfaces of thermodynamics possess an intrinsic Euclidean metric structure, thereby fulfilling Gibbs's vision of a geometrization of thermodynamics.

Thom, René, *Structural stability, catastrophe theory, and applied mathematics*, SIAM Review 19 (April 1977) 189-201.

The 1976 John von Neumann lecture, in which Thom compares catastrophe theory with control theory and investigates the importance of quantitative and qualitative aspects of catastrophe theory in applied mathematics. "[Catastrophe theory] may give quite a lot of insight without involving the need of a specific information on the substrate itself. In that respect, catastrophe theory may seem very arrogant to specialists, who have taken great pains to acquire a rich and detailed information on this substrate itself."

Senechal, Marjorie, Lewis, Marc, Rosen, Robert and Deakin, Michael A.B., *Letters: catastrophe theory*, Science 196 (17 June 1977) 1270, 1272.

Four letters commenting on Gina Bari Kolata's earlier article "Catastrophe theory: the emperor has no clothes", *Science* 196 (15 April 1977) 287, 350-351. Deakin also reveals that he has found "a relatively simple proof of the theorem of the seven [elementary catastrophes]", to appear in *Bull. Math. Biol.*.

Kolata, Gina Bari, *Mathematical games: are they bona fide research?*, Science 197 (5 August 1977) 546.

Hugh Montgomery (Michigan) suggests that the thought processes involved and the gratification derived from mathematical games are the same as from serious mathematical research, although J.H. Conway (Cambridge) says he would not feel intellectually satisfied with only the games, as serious mathematics offers a greater depth of argument.

Mowshowitz, Abbe (ed), Inside Information, Computers in Fiction, A-W, 1977; xxiii + 345 pp, $7.95 (P).

This entertaining and provocative "study-anthology" of computers in fiction focusses on social issues. Pieces are grouped by theme, with each group preceded by a reflective essay by the editor. Extensive bibliography of other fictional works and criticisms, and good indexes.

Péter, Rózsa, Playing with Infinity, Mathematical Explorations and Excursions, Trans: Z.P. Dienes, Dover, 1976; xiii + 268 pp, $3 (P).

A popularization of mathematics touching on number theory, Galois theory, geometry, calculus, symbolic logic, Gödel's undecidability theory and much more. Paperback reprint.

Reed, Ronald C., Tangrams--330 Puzzles, Dover, 1965; 152 pp, $1.50 (P).

Reprint of a 1965 volume. Hundreds of Tangram patterns, a bit of history, and some variations. Solutions in the back.

Elffers, Joost, Tangram, The Ancient Chinese Shapes Game, Trans: R.J. Hollingdale, Penguin, 1976; 169 pp, $5.95 (P).

1,600 shapes, with solutions, supplemented by a brief history, bibliography, and mathematical approach to counting and classifying Tangrams. Comes with a plastic Tangram set.

Barrington, John, *15 new ways to catch a lion*, Manifold 18 (Spring 1976) 10-14.

Mathematical developments of the last decade have contributed to improvements(?) in the methodology of big game hunting.

Raudsepp, Eugene, *Games that stimulate creativity*, Machine Design 49 (1977) 88-94.

Some old and new brain teasers, presented as a way to generate new ways of looking at problems.

Davis, Martin, Applied Nonstandard Analysis, Wiley, 1977; xii + 181 pp, $16.95.

A beautiful development stressing the Transfer Principle. Leads the beginner through the necessary logic and, assuming only minimal background in algebra and analysis, proceeds to apply the "nonstandard" method to real analysis, topology, and Hilbert space.

Stroyan, K.D. and Luxemburg, W.A.J., Introduction to the Theory of Infinitesimals, Pure and Appl. Math., V. 72, Acad Pr, 1976; xiii + 326 pp, $24.50.

A thorough and charming exposition of Robinson's infinitesimals. The first part covers much more than just a foundation for analysis (e.g., logic, rings, and even categories); the second part does more advanced topics in functional analysis.

Spanier, Jerome, *The Claremont Mathematics Clinic*, SIAM Review 19 (July 1977) 536-549.

An expanded version of a recent report by the same author ("The Mathematics Clinic: an innovative approach to realism within an academic environment", *Amer. Math. Monthly* 83 (1976) 771-774). Added are the details of two consulting projects undertaken by the Clinic.

Steen, Lynn Arthur, *Mathematics*, 1978 Yearbook of Science and the Future, Encyclopedia Britannica, 1977, pp. 353-356.

A report on "the major news from the world of mathematics" for 1976, devoted entirely to popular exposition of the resolution of the four-color problem; adapted from *Math. Magazine* 49 (1976) 219-222.

Zadeh, Norman and Kobliska, Gary, *On optimal doubling in backgammon*, Management Science 23 (April 1977) 853-858.

NEWS & LETTERS

DELAYS ...

Recent issues of *Mathematics Magazine* have been reaching subscribers about a month late due to delays at the compositors and a lengthy strike at the printers. We apologize for the inconvenience caused by these delays, and hope to be back on schedule by the beginning of 1978.

...ERRATA

Problem proposal No. 1014 in the May 1977 issue should have included among its hypotheses the condition $AD = BE = CF$. Due to the publication delays, problem solutions are being accepted past the announced deadlines, so readers who find problem 1014 easier to solve with the complete hypotheses may still submit a solution.

LINE PRINTER GRAPHICS

Readers interested in "Inexpensive Computer Graphing of Surfaces" (this *Magazine*, May 1977, pp. 143-147) may also be interested in the following article from the proceedings of the 1977 SIGGRAPH meeting: "A practical approach to implementing line printer graphics" by J.R. Rumsey and R.S. Walker, *Computer Graphics* 11:2 (1977) 102-106.

Pierre J. Malraison, Jr.
Control Data Corp.
St. Paul
Minnesota 55112

CORRESPONDENCE CORRESPONDENCE

The major result obtained in "Counting by Correspondence" (this *Magazine*, September 1976, pp. 181-186) was exactly the result obtained by G.W. Walker in his solution to problem E927 in the April, 1951 issue of the *Amer. Math. Monthly*.

Steven R. Conrad
Benjamin Cardozo H.S.
Bayside
New York 11364

ALLENDOERFER, FORD, POLYA AWARDS

Authors of nine expository papers published in 1976 issues of journals of the Mathematical Association of America received awards at the August meeting at the University of Washington. This year three sets of awards were established for each of the Association's three journals. The 1976 awards, each in the amount of $100, are:

Carl B. Allendoerfer Awards:

Joseph A. Gallian, The Search for Finite Simple Groups, *Mathematics Magazine* 49 (1976) 163-179.

B.L. van der Waerden, Hamilton's Discovery of Quaternions, *Mathematics Magazine* 49 (1976) 227-234.

Lester R. Ford Awards:

S.S. Abhyankar, Historical Ramblings in Algebraic Geometry and Related Algebra, *Amer. Math. Monthly* 83 (1976) 409-448.

J.H. Ewing, W.H. Gustafson, P.R. Halmos, S.H. Moolgavkar, W.H. Wheeler, and W.P. Ziemer, American Mathematics from 1940 to the Day before Yesterday, *Amer. Math. Monthly* 83 (1976) 503-516.

James P. Jones, Daihachiro Sato, Hideo Wada, and Douglas Wiens, Diophantine Representation of the Set of Prime Numbers, *Amer. Math. Monthly* 83 (1976) 449-464.

Joseph B. Keller, Inverse Problems, *Amer. Math. Monthly* 83 (1976) 107-118.

D.S. Passman, What is a Group Ring?, *Amer. Math. Monthly* 83 (1976) 173-185.

George Polya Awards:

Anneli Lax, Linear Algebra, a Potent Tool, *Two Year College Math J.* 7:2 (1976) 3-15.

Julian Weissglass, Small Groups: An Alternative to the Lecture Method, *Two Year College Math J.* 7:1 (1976) 15-20.

1978-79 CONGRESSIONAL FELLOWSHIP IN MATHEMATICAL SCIENCE

Applications are invited from candidates in the mathematical sciences for a Congressional Science Fellowship to be supported jointly by AMS, MAA and SIAM for the twelve-month period beginning 1 September 1978. The AMS-MAA-SIAM Fellow will serve, along with three or four Fellows selected by the American Association for the Advancement of Science and half a dozen Fellows sponsored by other scientific societies, under an annual program coordinated by the AAAS. The stipend for the Fellowship is $17,000, which may be supplemented by a small amount toward relocation and travel expenses.

[This is the second year of competition for the AMS-MAA-SIAM Congressional Science Fellowship. Last year the program was authorized very late, so their were few applicants. Consequently, the selection committee decided to make no award for 1977-78.]

Congressional Science Fellows spend their fellowship year working on the staff of an individual congressman or a congressional committee or in the congressional Office of Technology Assessment, the objective of the program being to enhance science-government interaction, the effective use of science in government, and the training of persons with scientific background for careers involving such use.

In addition to demonstrating exceptional competence in some areas of the mathematical sciences, an applicant for the AMS-MAA-SIAM Fellowship should have

1977 INTERNATIONAL MATHEMATICAL OLYMPIAD

The nineteenth International Mathematical Olympiad was held in Belgrade, Yugoslavia, on July 3 and 4, 1977. The United States team, coached by Samuel Greitzer and Murray Klamkin, finished first among the seventeen teams that entered the competition; two members of the U.S. team, Randall Dougherty of Fairfax, Va. and Michael Larson of Lexington, Mass., achieved perfect scores. The exam consists of the following six problems:

1. Equilateral triangles ABK, BCL, CDM, DAN are constructed inside the square $ABCD$. Prove that the midpoints of the four segments KL, LM, MN, NK and the midpoints of the eight segments AK, BK, BL, CL, CM, DM, DN, AN are the twelve vertices of a regular dodecagon. *(Holland)*

2. In a finite sequence of real numbers the sum of any seven successive terms is negative and the sum of any eleven successive terms is positive. Determine the maximum number of terms in the sequence. *(Viet Nam)*

3. Let n be a given integer > 2, and let V_n be the set of integers $1 + kn$, where $k = 1, 2, \ldots$. A number $m \in V_n$ is called *indecomposable in* V_n if there do not exist numbers $p, q \in V_n$ such that $pq = m$. Prove that there exists a number $r \in V_n$ that can be expressed as the product of elements indecomposable in V_n in more than one way. (Expressions which differ only in the order of the elements of V_n will be considered the same.) *(Holland)*

4. a, b, A, B are given constant real numbers and $f(\theta) = 1 - a \cos \theta - b \sin \theta - A \cos 2\theta - B \sin 2\theta$. Prove that if $f(\theta) \geq 0$ for all real θ, then $a^2 + b^2 \leq 2$ and $A^2 + B^2 \leq 1$. *(Gt. Britain)*

5. Let a and b be positive integers. When $a^2 + b^2$ is divided by $a + b$, the quotient is q and the remainder is r. Find all pairs (a,b), given that $q^2 + r = 1977$. *(W. Germany)*

6. Let $f(n)$ be a function defined on the set of all positive integers and taking on all its values in the same set. Prove that if $f(n+1) > f(f(n))$ for each positive integer n, then $f(n) = n$ for each n. *(Bulgaria)*

a broad scientific and technical background, a strong interest in the uses of the mathematical and other sciences in the solution of societal problems, and should be articulate, literate, flexible and able to work effectively with a wide variety of people.

The AMS-MAA-SIAM Congressional Science Fellowship is to be awarded competitively to a mathematically trained person at the postdoctoral to mid-career level without regard to sex, race, or ethnic group. Selection will be made by a panel of the AMS-MAA-SIAM Joint Projects Committee for Mathematics, a nine-member committee consisting of three representatives from each of these organizations.

Applications should include a resume, a summary of qualifications appropriate to the position, and a statement explaining why the applicant wants to be a Congressional Science Fellow. Applicants should solicit three letters from knowledgeable persons about the applicant's competence and suitability for the award.

Applications and letters should be sent to the Conference Board of the Mathematical Sciences, 2100 Pennsylvania Ave., N.W., #832, Washington, D.C. 20037. The deadline for receipt of applications has been set at 15 February 1978.

MORE ON MERLIN

Merlin, the travelling mathematician in "The gnome and the pearl of wisdom: a fable" (this *Magazine*, May 1977, pp. 141-143), could have taught the greedy gnome a lesson without incurring the expense of so many boxes. He could have challenged the gnome with the task of jumping into and out of a single box infinitely often within one minute of eleven o'clock, with the jumps following Merlin's customary time intervals. Each jump, of course, would be required to begin either from the floor of the box or the floor outside the box. As Merlin would agree, by eleven the gnome would not be found inside the box, nor outside the box, nor anywhere else.

To see why Merlin would agree, we need to look closely at the argument on page 143 concerning the "second task." Since marble n could not be in box k at any time greater than eleven minus $1/2^{n+k}$ minutes and less than eleven, Merlin concludes that the marble could not be in box k at exactly eleven. Merlin is relying on the fact that the gnome could not do something in "no time at all" (page 141); thus if marble n were in box k at eleven, the act of putting it there would have consumed an entire time interval containing eleven, and this was not possible. A similar continuity argument can be applied to the task we are suggesting. For the gnome to appear inside the box at eleven, the act of getting there would require an entire time interval containing eleven, which is not permitted. Likewise, the gnome could not appear outside the box at eleven or at any intermediate position.

Thus Merlin would be free to share all the gnome's marbles. In fact, each gnome in the castle, assuming just countably many, could then become as rich in marbles as the vanished greedy gnome had been.

Arthur Charlesworth
Univ. of Richmond
Richmond
Virginia 23173

COUNTERFEIT NOTE

I was surprised to see the paper "Counterfeit Coin Problems" by Bennet Manvel in the March 1977 issue of *Mathematics Magazine* since all the results (and proofs) in this paper have been known for a long time and can be found, e.g., in §2, Problèmes de détermination de monnaies fausses à l'aide de pesées of chapter 3 of the book Probabilité et information by A. M. Yaglom and I.M. Yaglom (translated by G. Roos, published by Dunod, Paris, 1969).

Edward T.H. Wang
Wilfrid Laurier Univ.
Waterloo, Ontario
Canada N2L 3C5

Coverage you can count on– everytime!

SELLERS: Elementary Statistics

This innovative, one-semester text covers all basic topics including probability and non-parametric statistics. Numerous charts, diagrams and newspaper clippings illustrate key points. Each chapter contains many real-world exercises and problems, behavioral objectives and progress tests. A valuable **Student Study Guide** is also available and a 100 page **Test Manual** is free upon adoption.

By **Gene Sellers,** Sacramento City College. 433 pp. 364 ill. $12.50. April 1977.

NOSAL: Basic Probability and Applications

This comprehensive text provides a sound treatment of the foundation of probability using a large number of applied problems. Numerous definitions, examples and exercises sets reinforce key concepts.

By **Miloslav Nosal,** Univ. of Calgary, Alberta, Canada. 370 pp. $13.50. June 1977.

SETEK: Algebra:
A Fundamental Approach

Using an intuitive approach to intermediate algebra, this lucid text features extensive explanations of key concepts and their modern applications. An **Instructor's Manual** is also available.

By **William M. Setek,** Monroe Community College. 707 pp. Illustd. $12.95. March 1977.

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, DC 20036

MATHEMATICS MAGAZINE VOL. 50, NO. 4, SEPTEMBER 1977