MATHEMATICS MAGAZINE

CONTENTS

VOLUME 43 · MAY 1970 · NUMBER 3

MATHEMATICS MAGAZINE

STEPHEN A. JENNINGS, *EDITOR*

EDITORIAL CORRESPONDENCE should be sent to the EDITOR, S. A. JENNINGS, Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada. Articles should be typewritten and double-spaced on $8\frac{1}{2}$ by 11 paper. The greatest possible care should be taken in preparing the manuscript, and authors should keep a complete copy. Figures should be drawn on separate sheets in India ink and of a suitable size for photographing.

NOTICE OF CHANGE OF ADDRESS and other subscription correspondence should be sent to the Executive Director, A. B. WILLCOX, Mathematical Association of America, Suite 310, 1225 Connecticut Avenue, N. W., Washington, D. C. 20036.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214.

The MATHEMATICS MAGAZINE is published by the Mathematical Association of America at Washington D. C., bi-monthly except July–August. Ordinary subscriptions are: 1 year \$4.00. Members of the Mathematical Association of America and of Mu Alpha Theta may subscribe at the special rate of 2 years for \$6.00. Single copies are 80¢.

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

ON IMPROPER MULTIPLE INTEGRALS

PAUL R. BEESACK

1. Introduction. Although a number of old and new textbooks in analysis give an adequate treatment of multiple Riemann integrals of bounded functions f over bounded sets $A \subset R^n$ which have content, (see [1]–[8]), none of these sources seem to deal adequately with improper multiple Riemann integrals. For example, [1] does not introduce the concept at all, while [3] and [5] deal only with some special cases. In the treatment given in [7], proper and improper multiple integrals (in our sense) are dealt with simultaneously. The texts [2], [4], [6] and [8] deal with the general case where either f or A may be unbounded, but only [2, p. 149 footnote] gives any indication that a convergent improper multiple integral is necessarily absolutely convergent. (In [5, p. 278], the author *states* that if $\int\int |f(x, y)| dxdy$ becomes infinite, then $\int\int f(x, y)dxdy$ is indeterminate.) There is, of course, some justification for the suggestion made in [3, p. 423] that a proper treatment be postponed to the time that Lebesgue integrals are introduced, in which case the difficulties essentially disappear!

Despite the author's agreement with the last sentiments expressed, he also believes that as long as the topic of improper multiple integrals is dealt with, it would be useful to have an adequate treatment of this topic available (to both teachers and students) in a readily accessible source such as this MAGAZINE. In Section 3 of this article, we give a general definition of such improper integrals and prove a few elementary results concerning them. Section 4 is devoted to the technically more difficult task of proving that convergent improper multiple integrals are absolutely convergent. In Section 2 we give a fairly full review of background material on Jordan content and (proper) Riemann integrals. Rather more is given here than is required for the following sections, but it is hoped that its inclusion makes the entire paper more readable than would otherwise be the case.

2. Jordan content and Riemann integrals. In this section we recall certain facts concerning Jordan content and (proper) Riemann integrals in Euclidean n-dimensional space R^n, some of which we will require in the sequel. The definitions and theorems stated below are more or less standard, and the reader may refer to [1, Ch. 10], [2, Ch. 3], [4, Ch. 3] and especially [7, Ch. 2] for details.

If A is a subset of R^n, we denote the *interior*, the *closure*, and the *boundary* of A by A^0, $\overline{A}$, ∂A respectively, and recall that A is *open* iff $A = A^0$, A is *closed* iff $A = \overline{A}$, and that $\overline{A} = A^0 \cup \partial A$, $\partial A = \overline{A} \cap \overline{C(A)}$ where $C(A)$ is the complement of A. A is *compact* iff A is both closed and *bounded*. A *cell* in R^n is any set of the form

$$(2.1) \qquad I = \{x = (x_1, \cdots, x_n) \mid a_i \prec x_i \prec b_i, 1 \leqq i \leqq n\},$$

where $-\infty < a_i \leqq b_i < +\infty$ for $1 \leqq i \leqq n$, and $\prec$ denotes either $\leqq$ or $<$ in any one of the $2n$ places where it occurs. If ξ is the class of all subsets of R^n which are the union of a finite number of cells, then ξ is a *ring*. That is, ξ is a nonempty class of sets with the property that $(A \cup B) \in \xi$ and $(A - B) \in \xi$ whenever $A, B \in \xi$; in addition, every ring ξ has the property that $(A \cap B) \in \xi$ whenever $A, B \in \xi$. The sets of ξ are called the *elementary* subsets of R^n, and we note that each such set is bounded.

The elementary *content* of the cell I defined by (2.1) is defined to be the product of the lengths of the sides of I, and is denoted by $m(I)$. In order to define the content of more complicated subsets of R^n, we note that one can prove that every set $A\in\xi$ can be written as the union of a finite number of mutually disjoint cells, and that if $A=\bigcup_{i=1}^{p} I_i$ is any such representation, then the definition

$$m(A) = \sum_{i=1}^{p} m(I_i) \tag{2.2}$$

is unambiguous. That is, the right side of (2.2) can be shown to have the same value for any representation $A=\bigcup_{i=1}^{p} I_i$ having $I_i\cap I_j=\emptyset$ when $i\neq j$. Now, if B is any *bounded* subset of R^n, then the *Jordan inner content*, $m_*(B)$, and the *Jordan outer content* $m^*(B)$, of B is defined by

$$\begin{aligned} m_*(B) &= \sup\{m(A) \mid A \subset B,\ A \in \xi\}, \\ m^*(B) &= \inf\{m(a) \mid B \subset A,\ A \in \xi\}. \end{aligned} \tag{2.3}$$

The class $\zeta=\{B \mid m_*(B)=m^*(B)\}$ is the class of all (Jordan) *measurable* subsets of R^n, and it can be shown that ζ is a ring. The common value of m_* and m^* on ζ is called the *Jordan content* in R^n. It is true that $m_*(A)=m(A)=m^*(A)$ for all $A\in\xi$, so no ambiguity can result by also using m to denote Jordan content on ζ. In common with every *content* on a ring, m has the following two properties: (i) $m(B)\geqq 0$ for all $B\in\zeta$; (ii) if B_1, $B_2\in\zeta$ with $B_1\cap B_2=\emptyset$, then $m(B_1\cup B_2)=m(B_1)+m(B_2)$. In addition, however, for the *Jordan* content m on ζ, it can be shown that (iii) if $A\in\zeta$, then $A^0\in\zeta$, $\overline{A}\in\zeta$, and $m(A^0)=m(A)=m(\overline{A})$, (iv) a bounded set A is in ζ iff $\partial A=\overline{A}-A^0$ has zero content. Using properties (i) and (ii), it follows that

$$A_1, A_2 \in \zeta, A_1 \subset A_2 \Rightarrow m(A_2 - A_1) = m(A_2) - m(A_1), m(A_1) \leqq m(A_2). \tag{2.4}$$

$$A_1, \cdots, A_k \in \zeta \Rightarrow m\left(\bigcup_{i=1}^{k} A_i\right) \leqq \sum_{i=1}^{k} m(A_i). \tag{2.5}$$

Finally, using properties (iii) and (iv) one can prove that if $A_1, \cdots, A_k$ are *nonoverlapping* measurable sets (that is, $m(\overline{A}_i\cap\overline{A}_j)=0$ when $i\neq j$), then

$$m\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} m(A_i). \tag{2.6}$$

It is worth noting that most authors say that two sets are nonoverlapping if they have no interior points in common. However, for sets $A, B\in\zeta$ it is easy to prove that $A^0\cap B^0=\emptyset$ iff $m(\overline{A}\cap\overline{B})=0$. The latter condition is somewhat easier to work with in practice so we have adopted this equivalent definition.

We now turn to an outline of the theory of the Riemann integral of bounded functions over measurable sets. If A is a measurable subset of $R^n (A\in\zeta)$, then by a *partition* of A we mean any finite collection, $p=\{A_1, \cdots, A_k\}$ of nonempty, nonoverlapping measurable sets whose union is A. The class of all partitions of A

is denoted by $\mathcal{P}(A)$, and the class of all bounded functions f on A is denoted by $\mathcal{B}(A)$. If $f\in\mathcal{B}(A)$, then the *lower* and the *upper Darboux integrals* of f over A are defined by

$$\underline{\int}_A f = \sup_{p\in\mathcal{P}(A)} \sum_{A_j\in p} \{\inf_{x\in A_j} f(x)\} m(A_j) = \sup_{p\in\mathcal{P}(A)} s(f,p),$$

(2.7)
$$\overline{\int}_A f = \inf_{p\in\mathcal{P}(A)} \sum_{A_j\in p} \{\sup_{x\in A_j} f(x)\} m(A_j) = \inf_{p\in\mathcal{P}(A)} S(f,p),$$

respectively. We always have

(2.8)
$$\underline{\int}_A f \leqq \overline{\int}_A f, \qquad f\in\mathcal{B}(A),$$

and if $\alpha\leqq f(x)\leqq\beta$ for all $x\in A$, then

(2.9)
$$\alpha m(A) \leqq \underline{\int}_A f \leqq \overline{\int}_A f \leqq \beta m(A).$$

The *norm* of a partition $p=\{A_1,\cdots,A_k\}$ is defined to be the maximum of the diameters of the sets $A_1,\cdots,A_k$ and is denoted by $\|p\|$. One of the principal results concerning Darboux integrals is

(2.10)
$$\underline{\int}_A f = \lim_{\|p\|\to 0} s(f,p), \qquad \overline{\int}_A f = \lim_{\|p\|\to 0} S(f,p).$$

A function $f\in\mathcal{B}(A)$ is said to be *Riemann integrable* over A iff $\underline{\int}_A f=\overline{\int}_A f$; the common value of the lower and upper integrals is called the *Riemann integral* of f over A, and is denoted by $\int_A f$. The class of all functions which are Riemann integrable over A will be denoted by $\mathcal{R}(A)$. We emphasize that the notation $f\in\mathcal{R}(A)$ implies that $A\in\zeta$ (so in particular, A is bounded), and that $f\in\mathcal{B}(A)$.

There are several important results concerning Riemann integrals that we shall require and which we list below. In (2.15), the functions f^+, f^- are the positive and negative parts of f defined by $f^+(x)=\max\{f(x), 0\}$, $f^-(x)=\max\{-f(x), 0\}$.

(2.11)
$$f\in\mathcal{R}(A), B\subset A, B\in\zeta\Rightarrow f\in\mathcal{R}(B).$$

(2.12)
$$f\in\mathcal{R}(A), f(x)\geqq 0 \text{ on } A, B\subset A, B\in\zeta\Rightarrow 0\leqq\int_B f\leqq\int_A f.$$

(2.13)
$$f\in\mathcal{R}(\overline{A})\Rightarrow\int_{A^0} f=\int_A f=\int_{\overline{A}} f.$$

(2.14)
$$f, g\in\mathcal{R}(A), f(x)\leqq g(x) \text{ on } A\Rightarrow\int_A f\leqq\int_A g.$$

(2.15)
$$f\in\mathcal{R}(A)\Rightarrow |f|, f^+, f^-\in\mathcal{R}(A).$$

(2.16)
$$f, g\in\mathcal{R}(A)\Rightarrow(af+bg)\in\mathcal{R}(A) \text{ and } \int_A(af+bg)=a\int_A f+b\int_A g.$$

(2.17) $f \in \mathcal{R}(A),\ A = \bigcup_{i=1}^{k} A_i$ (nonoverlapping $A_i \in \zeta$) $\Rightarrow \int_A f = \sum_{i=1}^{k} \int_{A_i} f.$

We also note that if $f \in \mathcal{B}(A)$, then (2.17) remains valid if either lower or upper integrals are used throughout.

3. Improper multiple integrals. According to the definitions given in the preceding section, in order that $\int_A f$ be defined, it is necessary (but not sufficient) that (a) A be a *bounded* measurable subset of R^n, and (b) f be defined and *bounded* on A. As to (a) we need merely note that m is only *defined* on (certain) bounded subsets of R^n. As to (b), if f is unbounded on A, and $p = \{A_1, \cdots, A_k\}$ is any partition of A, then for at least one j we will have either $\inf\{f(x) \mid x \in A_j\} = -\infty$ (if f is not bounded below), or $\sup\{f(x) \mid x \in A_j\} = +\infty$ (if f is not bounded above). According to the definitions (2.7) this would imply that $\underline{\int}_A f = -\infty$ or $\overline{\int}_A f = +\infty$, except perhaps in certain degenerate cases.

There is, of course, an extension of the definition of the Riemann integral (due to Cauchy—who preceded Riemann!) which allows us to assign a sensible meaning to the symbol $\int_A f$ in many cases when one, or both, of the conditions (a), (b) is not satisfied. Such integrals are called *improper integrals*, or sometimes *Cauchy-Riemann integrals.* We formulate a definition of improper integrals over subsets $A \subset R^n$ $(n \geqq 2)$ which covers both cases simultaneously.

DEFINITION 1. *Let A be a subset of R^n ($n \geqq 2$, A bounded or unbounded), and suppose that $f \in \mathcal{R}(A')$ for each compact measurable subset $A' \subset A$. A sequence $\{A_k\}$ of compact measurable sets satisfying the conditions*

(3.1) $$A_k \subset A_{k+1} \subset A \quad \textit{for each } k \geqq 1,$$

(3.2) *for each compact measurable $A' \subset A$, $\exists k$ such that $A' \subset A_k$, is called an admissible sequence for A. If the limit*

(3.3) $$\lim \int_{A_k} f = I$$

exists (finite), and has the same value I for every admissible sequence $\{A_k\}$ for A, we say that $\int_A f$ exists (or converges) and has the value I. If $\int_A f$ does not exist, we also say that the improper integral diverges, or is divergent.

In the following we will always assume that A is a set which has at least one admissible sequence. It is easy to prove that if $\{A_k\}$ is a sequence of compact measurable sets such that $A_k \subset A_{k+1}$ for all $k \geqq 1$, and $A_k \subset A \subset \bigcup_{n=1}^{\infty} A_n^0$ for all $k \geqq 1$, then A is necessarily an open set, and $\{A_k\}$ is an admissible sequence for A. On the other hand, A need not be an open set; for example in R^2 the set $A = \{(x, y) \mid 0 < x^2 + y^2 \leqq 1\}$ is neither open nor closed but obviously has admissible sequences.

It is also easy to see that if $\lim \int_{A_k} f$ exists for *every* admissible sequence $\{A_k\}$, then all of these limits must have the same value, so that we need not have required this in Definition 1. We are, of course, interested in the case that either A is unbounded, or f is unbounded in every neighborhood of one or more points of ∂A, or both. In case A is measurable (hence bounded) and f is bounded on A, it is true that the (proper) integral of f over A exists and has the value I iff (3.3) holds for every admissible sequence $\{A_k\}$ for A. To prove this, we note first that *given any such* $\{A_k\}$ *we have* $\lim m(A_k)=m(A)$. For, given $\epsilon>0$, it follows from the first of definitions (2.4) that $\exists A'_\epsilon \in \xi$, which we may assume to be *closed* by property (iii) of Jordan content, such that $A'_\epsilon \subset A$ and $m(A)-\epsilon<m(A'_\epsilon)\leqq m(A)$. Since A'_ϵ is a compact subset of A it now follows from (2.4), (3.2) and (3.1) that $\exists k_1=k_1(\epsilon)$ such that $m(A)-\epsilon<m(A_k)\leqq m(A)$ whenever $k\geqq k_1$, proving the assertion. Next, by the remark following (2.17),

$$(3.4)\qquad \underline{\int}_A f=\underline{\int}_{A_k} f+\underline{\int}_{A-A_k} f,\quad \overline{\int}_A f=\overline{\int}_{A_k} f+\overline{\int}_{A-A_k} f.$$

Since $f\in \mathcal{R}(A_k)$, it follows from this and (2.8) that

$$0\leqq \overline{\int}_A f-\underline{\int}_A f=\overline{\int}_{A-A_k} f-\underline{\int}_{A-A_k} f.$$

Now since $f\in\mathcal{B}(A)$, there are constants α, β such that $\alpha\leqq f(x)\leqq\beta$ for all $x\in A$, whence by (2.9) we have

$$(3.5)\qquad \alpha m(A-A_k)\leqq \underline{\int}_{A-A_k} f\leqq \overline{\int}_{A-A_k} f\leqq \beta m(A-A_k),$$

so that

$$0\leqq \overline{\int}_A f-\underline{\int}_A f\leqq(\beta-\alpha)m(A-A_k)=(\beta-\alpha)\{m(A)-m(A_k)\},$$

using the first of (2.4). Since $m(A_k)\to m(A)$, it follows that $f\in\mathcal{R}(A)$. But then, $f\in\mathcal{R}(A-A_k)$ by (2.11) so all the integrals appearing in (3.4), (3.5) are proper Riemann integrals and we have

$$\int_{A_k} f=\int_A f-\int_{A-A_k} f\to\int_A f=I\qquad \text{as } k\to\infty.$$

This proof shows, in fact, that whenever $A\in\zeta$, $f\in\mathcal{B}(A)$, and $f\in\mathcal{R}(A')$ for every compact measurable $A'\subset A$, then $f\in\mathcal{R}(A)$, and the (proper) integral of f over A is given by the limit (3.3).

A slight modification of the above proof shows that if $A\in\zeta$ and $f\in\mathcal{B}(A)$ but $f\notin\mathcal{R}(A)$, then for each admissible sequence $\{A_k\}$ for A, $f\notin\mathcal{R}(A_k)$ for all sufficiently large k.

THEOREM 1. *If $f(x)\geqq 0$ for all $x\in A$ and f is integrable over each compact measurable subset of A, then*

$$(3.6)\qquad \lim\int_{A_k} f=\lim\int_{A'_k} f$$

for any two admissible sequences $\{A_k\}$, $\{A'_k\}$ *for* A. (*The common limit in* (3.6) *may be* $+\infty$, *in which case* $\int_A f$ *diverges*.)

Proof. Since f is nonnegative on A and $A_k \subset A_{k+1} \subset A$, it follows from (2.12) that the sequence $\{\int_{A_k} f\}$ is monotone increasing, so the limit (3.3) exists, finite or infinite. Similarly, $\lim \int_{A_k'} f$ exists. For each $k \geqq 1$ there is an integer $n_k \geqq 1$ such that $A_k' \subset A_{n_k}$. Hence, by (2.12)

$$\int_{A_k'} f \leqq \int_{A_{n_k}} f \leqq \lim \int_{A_n} f \qquad \text{for each } k \geqq 1,$$

and so $\lim \int_{A_k'} f \leqq \lim \int_{A_n} f$. Reversing the roles of the two sequences, we obtain the opposite inequality, proving the theorem.

COROLLARY 1. *Under the hypothesis of Theorem* 1

$$\int_A f = \lim \int_{A_n} f$$

where $\{A_n\}$ *is any admissible sequence for* A.

The situation for nonnegative functions is thus particularly simple: one selects the most convenient admissible sequence $\{A_n\}$ and computes the above limit. If this limit is finite the improper integral converges; if it is $+\infty$, the improper integral is divergent.

COROLLARY 2. (Comparison Test for Improper Integrals) *Suppose that* $0 \leqq f(x) \leqq g(x)$ *for all* $x \in A$, *and both* f *and* g *are integrable over every compact measurable subset of* A. *Then*

(i) *if* $\int_A g$ *converges, so does* $\int_A f$, *and* $\int_A f \leqq \int_A g$;

(ii) *if* $\int_A f = +\infty$ *(diverges), then* $\int_A g = +\infty$.

This is an immediate consequence of the preceding corollary and (2.14).

We now turn to a consideration of functions which assume both positive and negative values in A.

DEFINITION 2. *Let* f *be integrable over every compact measurable subset of* A. *We say that the improper integral* $\int_A f$ *is absolutely convergent iff* $\int_A |f|$ *converges.*

Since $f \in \mathcal{R}(A')$ implies $|f| \in \mathcal{R}(A')$ by (2.15), for every compact measurable $A' \subset A$, $\int_A |f|$ has meaning.

THEOREM 2. *If* $\int_A f$ *is absolutely convergent, then it is convergent.*

Proof. By hypothesis, $\int_A |f| = I < \infty$. Let $\{A_n\}$ be any admissible sequence for A. Since

$$0 \leqq f(x) + |f(x)| \leqq 2|f(x)| \qquad \text{for all } x \in A,$$

it follows from (2.14) and (2.16) that

$$0 \leqq \int_{A_n} f + \int_{A_n} |f| \leqq 2 \int_{A_n} |f| \qquad \text{for } n \geqq 1.$$

Hence the monotone increasing sequence $\{\int_{A_n}(f+|f|)\} = \{\int_{A_n} f + \int_{A_n} |f|\}$ is bounded above by $2I$, so that

$$\lim\left(\int_{A_n} f + \int_{A_n} |f|\right) = K \text{ exists, and } K \leqq 2I.$$

Moreover, the limit K is independent of the admissible sequence $\{A_n\}$ by Theorem 1. Since $\lim \int_{A_n}|f| = I$, it follows that $\lim\int_{A_n} f$ exists and has the value $K-I$ for any admissible sequence $\{A_n\}$, proving the theorem.

COROLLARY 3. *If $\int_A f$ is absolutely convergent, then*

$$\left|\int_A f\right| \leqq \int_A |f|.$$

For, if $\int_A f = J$, then we have $0 \leqq J+I = K \leqq 2I$ above, so that $-I \leqq J \leqq I$, or $|J| \leqq I$.

4. Convergent improper integrals are absolutely convergent. In this section we shall prove the possibly surprising fact that any convergent improper integral is necessarily absolutely convergent. The basic reason for this is the requirement that the limit (3.3) exist for *every* admissible sequence. The proof is essentially that suggested in the footnote in [2, p. 149].

THEOREM 4. *If $\int_A f$ is convergent, then it is absolutely convergent.*

Proof. If f^+ and f^- denote the positive and negative parts of f, then $f = f^+ - f^-$, $|f| = f^+ + f^-$, and $f^+(x) \geqq 0$, $f^-(x) \geqq 0$ for all $x \in A$. Moreover, f^+, f^- and $|f|$ are, together with f, integrable over every compact measurable subset A' of A by (2.15). To prove the theorem we shall prove that if $\int_A |f|$ diverges, then $\int_A f$ also diverges. First, if $\int_A |f| = +\infty$ (that is, $\int_A |f|$ diverges), then since

$$\int_{A_n} |f| = \int_{A_n} f^+ + \int_{A_n} f^-, \qquad n \geqq 1$$

holds for every admissible sequence $\{A_n\}$ for A, it follows that at least one of the improper integrals $\int_A f^+$, $\int_A f^-$ must diverge (to $+\infty$). If only one of these integrals diverged, say the second, then since

$$\int_{A_n} f = \int_{A_n} f^+ - \int_{A_n} f^-,$$

it would follow that $\lim\int_{A_n} f = -\infty$, so that $\int_A f$ would be divergent in this case.

Suppose that $\int_A f^+ = \int_A f^- = +\infty$. Let $\{A_n\}$ be any admissible sequence for A. We shall prove by induction that there exist two sequences $\{A_n^+\}$, $\{A_n^-\}$ of compact measurable subsets of A such that $A_n = A_n^+ \cup A_n^-$, $A_n^+ \subset A_{n+1}^+$ and $A_n^- \subset A_{n+1}^-$ for all $n \geqq 1$, the sets A_n^+, A_n^- are nonoverlapping, and for each $n > 1$,

$$\int_{A_n^-} f^- > \int_{A_n} f^- - \sum_{k=1}^{n} 2^{-k}, \tag{4.1}$$

$$\int_{A_n^+} f^+ = \int_{A_n} f^+. \tag{4.2}$$

Assuming these results have been proved for now, set $I_n = \int_{A_n^-} f^-$ so $\lim I_n = +\infty$ by (4.1). Since $\lim \int_{A_n^+} f^+ = +\infty$ by (4.2), to each $n \geqq 1$ there corresponds an integer k_n (where we may assume $k_{n+1} > k_n$ since $\{\int_{A_n^+} f^+\}$ is monotone increasing) such that

$$(4.3) \qquad \int_{A_{k_n}^+} f^+ > n + I_n.$$

Let $A_n' = A_{k_n}^+ \cup A_n^-$ for $n \geqq 1$. Then $\{A_n'\}$ is an admissible sequence for A. For, each A_n' is a compact measurable set, and $A_n' \subset A_{n+1}' \subset A$ clearly holds for each $n \geqq 1$. Moreover, if A' is any compact measurable subset of A, then there exists n such that $A' \subset A_n$ and hence $A' \subset A_n^+ \cup A_n \subset A_{k_n}^+ \cup A_n^- = A_n'$. Now, using (2.12), (2.17) and (4.3),

$$\int_{A_n'} f = \int_{A_n'} f^+ - \int_{A_n'} f^- \geqq \int_{A_{k_n}^+} f^+ - \left(\int_{A_{k_n}^+} f^- + \int_{A_n^-} f^-\right)$$

$$= \int_{A_{k_n}^+} f^+ - \left(\int_{A_{k_n}} f^- - \int_{A_{k_n}^-} f^- + \int_{A_n^-} f^-\right)$$

$$> n + I_n - \left(\sum_{j=1}^{k_n} 2^{-j} + I_n\right) > n - 1.$$

Hence, $\lim \int_{A_n'} f = +\infty$, so that $\int_A f$ is divergent.

It remains to prove that there are sequences $\{A_n^+\}$, $\{A_n^-\}$ having the stated properties. First, corresponding to $n = 1$, it follows from the first of (2.10) that there is a partition $p = \{A_{11}, \cdots, A_{r1}\}$ of A_1 of nonoverlapping measurable sets with sufficiently small norm so that

$$\sum_{i=1}^{r} \{\inf_{x \in A_{i1}} f^-(x)\} m(A_{i1}) > \int_{A_1} f^- - \frac{1}{2} \cdot$$

Since $p_1 = \{\overline{A}_{11}, \cdots, \overline{A}_{r1}\}$ is also a partition of the compact set A_1 with $\|p_1\| = \|p\|$, it follows that the preceding inequality holds when A_{i1} is replaced by $\overline{A}_{i1}$, $1 \leqq i \leqq r$. Now, let $p_1^- = \{\overline{A}_{i1} \mid \inf_{x \in \overline{A}_{i1}} f^-(x) > 0\}$ and set $A_1^- = \bigcup_{\overline{A}_{i1} \in p_1^-} \overline{A}_{i1}$, and $A_1^+ = \bigcup_{\overline{A}_{i1} \notin p_1^-} \overline{A}_{i1}$. The sets A_1^-, A_1^+ are clearly compact measurable subsets of A_1 with $A_1^- \cup A_1^+ = A_1$. In addition, the sets A_1^-, A_1^+ are nonoverlapping since

$$\overline{\overline{A_1^-}} \cap \overline{\overline{A_1^+}} = \overline{A_1^-} \cap \overline{A_1^+} = \bigcup_{\overline{A}_{i1} \in p_1^-} \bigcup_{\overline{A}_{j1} \notin p_1^-} (\overline{A}_{i1} \cap \overline{A}_{j1}),$$

whence $0 \leqq m(\overline{A_1^-} \cap \overline{A_1^+}) \leqq \sum_i \sum_j m(\overline{A}_{i1} \cap \overline{A}_{j1}) = 0$ on using (2.5). Moreover, by the first of (2.7) and the above

$$\int_{A_1^-} f^- \geqq \sum_{\overline{A}_{i1} \in p_1^-} \{\inf_{x \in \overline{A}_{i1}} f^-(x)\} m(\overline{A}_{i1}) > \int_{A_1} f^- - \frac{1}{2} \cdot$$

In addition, $f^+(x) = 0$ for $x \in A_1^-$. For otherwise, $f^+(x) > 0$ for some $x \in \overline{A}_{i1}$ with $\overline{A}_{i1} \in p_1^-$, and hence $f^-(x) = \max[-f(x), 0] = 0$, which is not the case. Hence

$$\int_{A_1} f^+ = \int_{A_1^+} f^+ + \int_{A_1^-} f^+ = \int_{A_1^+} f^+,$$

completing the proof that (4.1) and (4.2) are satisfied for $n=1$.

Suppose we have succeeded in constructing such pairs of sets A_k^-, A_k^+ for $k=1, 2, \cdots, n$. Then corresponding to $k=n+1$ and the compact measurable set $\overline{A_{n+1}-A_n}$ we may construct—by precisely the method just used for the case $n=1$—nonoverlapping compact measurable sets B_{n+1}^- B_{n+1}^+ such that $\overline{A_{n+1}-A_n} = B_{n+1}^- \cup B_{n+1}^+$ and

$$\int_{B_{n+1}^-} f^- > \int_{\overline{A_{n+1}-A_n}} f^- - 2^{-n-1} = \int_{A_{n+1}-A_n} f^- - 2^{-n-1},$$

$$\int_{B_{n+1}^+} f^+ = \int_{\overline{A_{n+1}-A_n}} f^+ = \int_{A_{n+1}-A_n} f^+;$$

here we used the property (2.13) for the last equality in each case. Now, set $A_{n+1}^- = A_n^- \cup B_{n+1}^-$ and $A_{n+1}^+ = A_n^+ \cup B_{n+1}^+$. Then $A_n^- \subset A_{n+1}^-$, $A_n^+ \subset A_{n+1}^+$ and the sets A_{n+1}^-, A_{n+1}^+ are nonoverlapping compact measurable sets with

$$A_{n+1}^- \cup A_{n+1}^+ = (A_n^- \cup A_n^+) \cup (B_{n+1}^- \cup B_{n+1}^+) = A_n \cup (\overline{A_{n+1} - A_n}) = A_{n+1}.$$

To see that A_{n+1}^-, A_{n+1}^+ are nonoverlapping, we note that both of these sets are closed, and that

$$A_{n+1}^- \cap A_{n+1}^+ = (A_n^- \cap A_n^+) \cup (A_n^- \cap B_{n+1}^+) \cup (B_{n+1}^- \cap A_n^+) \cup (B_{n+1}^- \cap B_{n+1}^+).$$

By construction, the first and the last of the four sets of this union have zero content, so by (2.5),

$$0 \leqq m(A_{n+1}^- \cap A_{n+1}^+) \leqq m(A_n^- \cap B_{n+1}^+) + m(B_{n+1}^- \cap A_n^+).$$

Both of the terms on the right side of the last inequality have the value zero. For example, since $A_n^- \subset A_n$ and $B_{n+1}^+ \subset (\overline{A_{n+1}-A_n}) \subset A_{n+1} \cap \overline{C(A_n)}$, it follows that

$$A_n^- \cap B_{n+1}^+ \subset A_n \cap A_{n+1} \cap \overline{C(A_n)} = \overline{A}_n \cap \overline{C(A_n)} = \partial A_n,$$

whence $0 \leqq m(A_n^- \cap B_{n+1}^+) \leqq m(\partial A_n) = 0$ by (2.5) and property (iv) of Jordan content. Hence A_{n+1}^- and A_{n+1}^+ are nonoverlapping, as asserted. Similarly, since A_n^-, B_{n+1}^- are nonoverlapping, it follows from (2.17) that

$$\int_{A_{n+1}^-} f^- = \int_{A_n^-} f^- + \int_{B_{n+1}^-} f^- > \left(\int_{A_n} f^- - \sum_{i=1}^{n} 2^{-i}\right) + \left(\int_{A_{n+1}-A_n} f^- - 2^{-n-1}\right)$$
$$= \int_{A_{n+1}} f^- - \sum_{i=1}^{n+1} 2^{-i},$$

and similarly

$$\int_{A_{n+1}^+} f^+ = \int_{A_n^+} f^+ + \int_{B_{n+1}^+} f^+ = \int_{A_n} f^+ + \int_{A_{n+1}-A_n} f^+ = \int_{A_{n+1}} f^+.$$

This completes the induction and the theorem is proved.

Thus the adoption of Definition 1 for the existence of an improper integral, $\int_A f$, eliminates the possibility of conditionally convergent improper integrals. Nevertheless, as the following example shows, in case $n \geqq 2$ Definition 1 (or a similar definition) seems to be both sensible and reasonable.

Example 1. Let $A = \{(x, y) \mid 0 \leqq x < \infty, 0 \leqq y < \infty\}$ and $f(x, y) = \sin(x+y)$. Obviously $f \in \mathcal{R}(A')$ for every compact measurable subset $A' \subset A$. Moreover, if $\{a_n\}$, $\{b_n\}$ are any two strictly increasing sequences of positive real numbers, then $A_n = \{(x, y) \mid 0 \leqq x \leqq a_n, 0 \leqq y \leqq b_n\}$ is clearly an admissible sequence for A. Now,

$$\int_{A_n} f = \int_0^{a_n} \int_0^{b_n} \sin(x+y)\,dy\,dx = 4 \sin(a_n/2) \sin(b_n/2) \sin((a_n+b_n)/2).$$

Obviously the value of $\lim \int_{A_n} f$ (when it exists) depends on the choice of the admissible sequence $\{A_n\}$. For example, if $a_n = b_n = n\pi$, then $\lim \int_{A_n} f = 0$, if $a_n = b_n = (4n+1)\pi/2$, $\lim \int_{A_n} f = 2$, while if $a_n = n$, $b_n = 3n$, $\lim \int_{A_n} f$ does not exist. $\int_A f$ is thus divergent according to Definition 1. Moreover, since any one of the sequences $\{A_n\}$ appears to be as natural as the others, it would seem that there is no reason why one should restrict the class of admissible sequences in the definition in order to ensure the existence of $\int_A f$ (according to the new definition).

We conclude by considering the following well-known example in the light of Definition 1.

Example 2. Let $A = \{x \mid 0 \leqq x < \infty\}$, and let $f(x) = \sin x/x$ for $x > 0$ and $f(0) = 1$. As in Example 1, if $\{a_n\}$ is any strictly increasing sequence of positive real numbers, then the sequence $\{A_n\}$ with $A_n = \{x \mid 0 \leqq x \leqq a_n\}$ is admissible for A. Now

$$\int_{A_n} f = \int_0^{k_n\pi} f + \int_{k_n\pi}^{k_n\pi + r_n} f = I_n + J_n,$$

where k_n is the unique nonnegative integer such that $a_n = k_n\pi + r_n$, and $0 \leqq r_n < \pi$. We have

$$|J_n| < \int_{k_n\pi}^{(k_n+1)\pi} (1/k_n\pi) = k_n^{-1},$$

so that $\lim J_n = 0$. Also, for all n such that $k_n \geqq 1$,

$$I_n = \sum_{j=1}^{k_n} \int_{(j-1)\pi}^{j\pi} (\sin x/x)\,dx = \sum_{j=1}^{k_n} b_j.$$

The sequence $\{\sum_{j=1}^{k_n} b_j\}$ is a subsequence of the sequence of partial sums of the infinite series $\sum_1^\infty b_j$ (possibly diluted). The convergence of this series follows easily from the alternating series test. Hence

$$\lim \int_{A_n} f = \lim \int_0^{a_n} (\sin x/x)dx = \sum_1^\infty b_j, \tag{4.4}$$

and unlike the situation in Example 1, this limit is independent of the increasing sequence $\{a_n\}$. Nevertheless, $\int_A f$ still diverges *according to Definition* 1. To see this, consider the admissible sequence $\{A_n'\}$ for A defined by

$$A_n' = [0, (2n-1)\pi] \cup \bigcup_{j=n}^{2n} [2j\pi, (2j+1)\pi].$$

Here,

$$\begin{aligned}\int_{A_n'} f = \int_0^{(2n-1)\pi} (\sin x/x)dx + \sum_{j=n}^{2n} \int_{2j\pi}^{(2j+1)\pi} (\sin x/x)dx \\ > I_n + \sum_{j=n}^{2n} \int_{2j\pi}^{(2j+1)\pi} \{\sin x/(2j+1)\pi\} dx = I_n + \sum_{j=n}^{2n} \{2/(2j+1)\pi\},\end{aligned}$$

where I_n is defined as before, with $k_n = (2n-1)\pi$ in this case. However,

$$\sum_{j=n}^{2n} \frac{2}{(2j+1)\pi} > \frac{1}{\pi} \sum_{j=n}^{2n} \frac{1}{j+1} > \frac{1}{\pi} \int_{n+1}^{2n+2} \frac{1}{x}\, dx = \frac{1}{\pi} \log 2$$

holds for each $n \geqq 1$, so that if $\lim \int_{A_n'} f$ exists, we have

$$\lim \int_{A_n'} f \geqq \sum_1^\infty b_j + \frac{1}{\pi} \log 2. \tag{4.5}$$

Comparing (4.4) and (4.5), we see that $\int_A f$ diverges according to Definition 1.

On the other hand, in this example, the sequences $\{A_n\}$ with $A_n = [0, a_n]$ appear to be more natural choices rather than any sequence of the second type. It is because this natural restriction on admissible sequences is actually made that we are led to conditionally convergent improper (single) integrals. In Example 2, the improper integral $\int_0^\infty (\sin x/x)dx$ is such a conditionally convergent integral, and its value is given by (4.4).

References

1. T. M. Apostol, Mathematical Analysis, Addison-Wesley, Reading, 1957.
2. R. C. Buck, Advanced Calculus, 2nd ed., McGraw-Hill, New York, 1965.
3. W. Fulks, Advanced Calculus, Wiley, New York, 1961.
4. C. Goffman, Calculus of Several Variables, Harper and Row, New York, 1965.
5. E. Goursat (tr. E. R. Hedrick), A Course in Mathematical Analysis, vol. 1, repr. Dover, New York, 1959.
6. W. Kaplan, Advanced Calculus, Addison-Wesley, Cambridge, Mass., 1952.
7. W. W. Rogosinski, Volume and Integral, 2nd ed., Oliver and Boyd (Interscience), New York, 1962.
8. R. E. Williamson, R. H. Crowell, H. F. Trotter, Calculus of Vector Functions, 2nd ed., Prentice-Hall, Englewood Cliffs, 1968.

ECONOMIC TRAVERSAL OF LABYRINTHS

A. S. FRAENKEL, The Weizmann Institute of Science, Rehovot, Israel and
Bar Ilan University, Ramat Gan, Israel

1. Introduction. A labyrinth is a system of enclosed (normally underground) alleys which are interconnected in various ways. Any labyrinth can be described by a finite connected graph, where edges correspond to alleys, and vertices to alley intersections and terminal points of blind (dead end) alleys. Henceforth we shall use the terms "edges" and "alleys" interchangeably; similarly for "vertices" and "intersections". An example of a labyrinth is presented in Figure 1. Its associated graph is given in Figure 2. An alternative form of the graph is given in Figure 3.

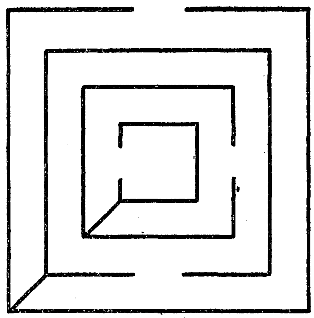

FIG. 1.

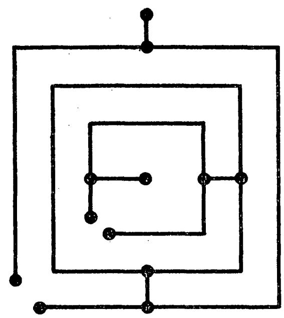

FIG. 2.

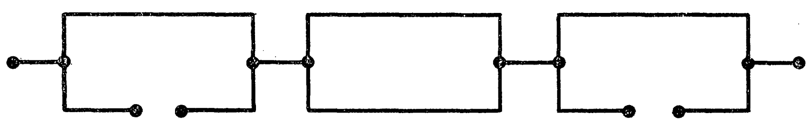

FIG. 3.

A typical problem concerning labyrinths is to go from an entrance to the "center", where a treasure is hidden or a monstrous Minotaur waits to be killed by the hero Theseus (who is to be subsequently extricated from the labyrinth by rewinding a thread whose end is held by Ariadne); or to find one's way to an exit if one is lost somewhere in the labyrinth. Since it is always assumed that the plan of the labyrinth is not known to the person touring it, and, in particular, the location of a "center" is not known, a solution to the following more general problem is normally sought: starting from any given vertex, traverse *every* edge of the graph.

There exist a number of algorithms for tracing a labyrinth such that every edge is traversed. Each of these algorithms is associated with a set of assumptions about the information available at any intersection about the path which was already traversed. For example, an algorithm of Wiener insures that every edge is traversed at least twice; the algorithms of Trémaux and Tarry insure that every edge is traversed exactly twice, once in each direction [**2, 3**]. The same result is obtained by the "right-hand-on-the-wall" rule, provided the graph representing the labyrinth is a tree [1]. Trakhtenbrot [4] considered also nonconnected graphs (labyrinths). For the special case of a connected graph, his algorithm also yields a double traversal of each edge. Trakhtenbrot concluded with the comment " . . . it is doubtful that an algorithm simpler than the one we have given can be constructed."

It is natural to seek an algorithm which permits an excursion such that every edge is traversed at least once and *at most* twice. The purpose of this note is to formulate and prove validity of such an algorithm for a general labyrinth (Section 3). For example, if the labyrinth represents an Euler graph, the excursion, if "lucky", may result in a *single* traversal of every alley.

Since our algorithm is based on that of Tarry, we start, for the sake of completeness, by proving the latter.

2. Tarry's algorithm. Tarry's assumptions are that upon arrival at any intersection v, two things are known: (i) the subset of those edges incident to v along which we have previously *left* v, i.e., those that were traversed in the direction pointing away from v; (ii) the *entrance edge*, i.e., the edge via which we first arrived at v.

Under these assumptions, Tarry's algorithm states: arriving at an intersection v, continue via an edge (v, v') which was not yet traversed in the direction of v to v'; but choose the entrance edge only as a last resort.

THEOREM I (Tarry). *Suppose that a labyrinth is toured according to Tarry's algorithm, starting from an initial vertex v_0. The excursion will terminate at v_0 and every edge will then have been traversed exactly twice, once in each direction.*

Proof. (i) Since the graph is finite, the excursion terminates. In other words, a vertex v is reached such that every edge incident to v was already traversed in the direction pointing away from v. Now it is clear that no edge of the graph was traversed more than once in each direction. If $v \neq v_0$, the vertex v has been entered once more than it has been left. Hence there exists an edge incident to v along which v has not yet been traversed in the direction pointing away from v, which is a contradiction. Hence $v = v_0$.

(ii) Suppose that the excursion of the labyrinth amounts to the traversal of the vertices

$$v_0, v_1, v_2, \cdots, v_0 \tag{1}$$

in this order. (The sequence may contain repetitions besides v_0.) At the conclusion of the excursion, all edges incident to v_0 were traversed once in the direction pointing away from v_0, as otherwise the excursion could be continued. Hence all edges incident to v_0 were also traversed once in the direction pointing towards v_0.

We shall now show that the same holds for all other vertices in the sequence (1). Suppose that this is not the case. Let v_n be the first vertex in (1) such that not all edges incident to it were traversed. Since also at v_n the number of entrances and exits must be the same, there exists, in particular, an edge incident to v_n along which v_n was not traversed in the direction pointing away from v_n.

Suppose that (v_k, v_n) is the entrance edge of v_n. Then v_k appears in (1) before v_n. Because of the minimality of v_n, the edge (v_k, v_n) was traversed in the direction from v_n to v_k. But this contradicts the rule stating that the entrance edge is to be used as an exit only when all other possibilities have been exhausted.

(iii) It remains only to show that all the vertices of the graph are contained in (1). Let W be a vertex which was not traversed. Since the graph is connected, there exists a path

$$v_0 = W_0, W_1, \cdots, W_{m-1}, W_m = W$$

connecting v_0 with W. Let W_i be the first vertex in this sequence not contained in (1). Then W_i was not traversed even once. Since however W_{i-1} is in (1), the edge (W_{i-1}, W_i) was traversed. This contradiction establishes the result.

3. An economic algorithm. The following assumptions are made: upon arrival at a vertex v, its entrance edge is known, as well as the edges incident to v which have been traversed previously, and the direction of their traversal. In addition, the person touring the labyrinth has a *counter* (e.g., in the form of pencil and paper).

Let $\rho(v)$ be the *valence* of v, that is, the number of edges incident to v, and let v_0 be the initial vertex. Without loss of generality we may assume $\rho(v_0) = 1$. Because if this is not the case, we can adjoin a short alley leading to the labyrinth's entrance. The valence of the initial terminal of this alley is 1.

The algorithm:

(1) Start out from v_0 with the counter containing zero. Whenever we arrive at a vertex not traversed before, the counter is increased by 1.

(2) If we arrive at a vertex v such that before entering it there was at least one edge incident to it which was not yet traversed, and upon arrival at v there remains at most one such edge, decrease the counter by 1.

(3) As long as the counter is positive, the tour is conducted according to Tarry's algorithm, but, whenever possible, an edge not traversed at all before, is used in preference over an edge already traversed before.

(4) As soon as the counter contains zero, leave all vertices via their entrance edges.

THEOREM II: *The tour terminates at v_0 such that every edge was traversed at least once and at most once in each direction.*

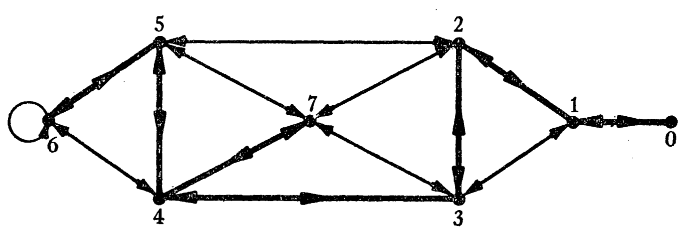

FIG. 4.

Examples. 1. Consider the labyrinth represented by the graph given in Figure 4, whose vertices have been labeled by the numbers 0–6, the initial vertex being 0. An excursion according to the algorithm is given by the sequence of vertex traversals:

$$0, 1, 2, 3, 1, 3, 4, 5, 6, 4, 7, 2, 5, 7, 3, 7, 5,$$
$$2, 7, 4, 6, 6, 5, 4, 3, 2, 1, 0.$$

As the reader may easily check, the last occurrence of vertex 6 represents the point at which the counter becomes zero. From there on, the journey amounts to entrance edge backtracks. All entrance edges in Figure 4 and in subsequent figures are represented by heavy lines. Note that the only saving of this excursion over an excursion according to Trémaux, Tarry or Trakhtenbrot, is that the loop at vertex 6 is traversed only once.

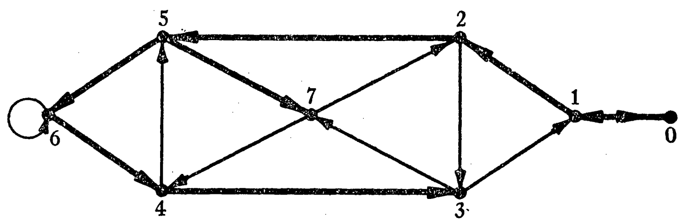

FIG. 5.

2. Another journey of the same labyrinth, also consistent with the algorithm, is given by the sequence

$$0, 1, 2, 5, 6, 6, 4, 5, 7, 4, 3, 7, 2, 3, 1, 0.$$

(See Figure 5.) It is seen that this is a most economical traversal, as all edges incident to vertices of even valence are traversed exactly once, and only the edge joining vertices 0 and 1, which are of odd valence, is traversed twice.

Proof. (i) We show first that the counter will eventually contain zero. Suppose this is not the case. Then the labyrinth is traversed according to Tarry's algorithm only. If its vertices are $v_0, v_1, \cdots, v_n$, n units are added into the counter and n units are subtracted from it during the excursion, so that the counter will eventually contain zero after all.

(ii) Next we show that when the counter contains zero, all edges have been traversed at least once. By renaming vertices if necessary, we may suppose that the counter first became zero after a traversal of the vertices $v_0, v_1, \cdots, v_k$. Clearly *all* edges incident to v_i $(0 \leqq i \leqq k)$ must have been traversed. In other words, there is no vertex such that only part of the edges incident to it were traversed. Suppose that W is a vertex not traversed even once. There exists a simple path $v_0 = W_0, W_1, \cdots, W_{m-1}, W_m = W$ connecting v_0 with W. Let W_i be the first vertex in the sequence which was never traversed. Then W_{i-1} was traversed, and, in particular, the edge (W_{i-1}, W_i) was traversed; a contradiction.

(iii) No edge was traversed more than once in each direction. This is clear for the part of the excursion performed with positive counter, which is a subtour of Tarry. It remains only to show that if p is the intersection at the arrival of which the counter became zero, the rest of the excursion is consistent with the claim. We proceed by induction. By Tarry's algorithm, the entrance edge of p was not yet traversed in the direction pointing away from p, and we can do so now. Suppose that during the course of the excursion we leave a vertex q via its entrance edge (q, r), which was not traversed before in the direction from q to r. If this is our first visit at r with the counter containing zero, Tarry's algorithm insures that the entrance edge of r was not yet used as an exit. Suppose that we already visited at r after the counter became zero. Then there exists a circuit of entrance edges

$$r = W_0, W_1, \cdots, W_{m-1}, W_m = W_0.$$

Suppose that W_i is the first among them that was entered during the excursion with positive counter. Then it was entered via the edge (W_{i-1}, W_i). This means that W_{i-1} was entered before, a contradiction.

(iv) After the counter contains zero, each traversal of an edge must thus bring us to a new vertex. Because the graph is finite, we must return to v_0, completing the proof.

The following remarks can be easily verified. Proofs are omitted.

(1) If the object of the labyrinth's exploration is to hunt after a treasure or Minotaur, Tarry's algorithm can be abandoned and the backtrack entrance edge traversal commenced just as soon as the object hunted after has been located. This will normally yield a faster return to the entrance.

(2) If a counter is used in which the unit additions are recorded and then *crossed out* rather than *erased*, the final number of crossed out entries equals the number of vertices of the labyrinth (not counting v_0, but counting all other intersections, including terminal ends of blind alleys).

(3) If the intersections are numbered as the excursion proceeds and their sequence of traversal S is recorded, the traveler will know, as soon as the counter becomes zero, the number N of edges he still has to traverse before reaching the entrance. The number N can be read off from S as follows. Let v_k be the vertex at which the counter became zero. Let v_j be the vertex immediately preceding the first appearance of v_k in S; let v_i be the vertex immediately preceding the first appearance of v_j in S, and so on. Then the sequence $v_k, v_j, v_i, \cdots, v_0$ represents the backtrack journey, and the number of vertices in the sequence is $N+1$.

The sequence S also enables one to construct a complete plan of the labyrinth after the first traversal.

Examples. 3. Consider the labyrinth represented in Figure 4 and its traversal represented in the figure and by the sequence S given in Example 1. The counter became zero at the last appearance of 6 in the sequence. The number immediately preceding the first appearance of 6 is 5, and the complete backtrack journey is represented by the sequence 6, 5, 4, 3, 2, 1, 0. Thus $N=6$.

4. Another traversal of the same labyrinth, consistent with the algorithm, is represented in Figure 6, and by the sequence S

$$0, 1, 2, 5, 6, 6, 4, 3, 1, 3, 2, 7, 5, 4, 7, 3, 4, 6,$$
$$5, 2, 1, 0.$$

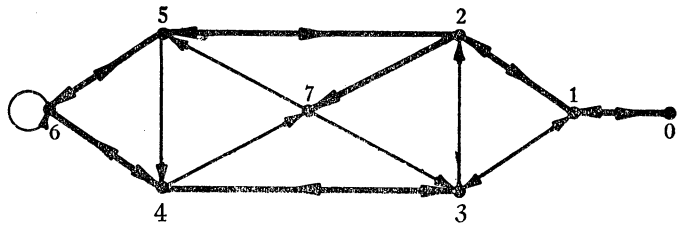

FIG. 6.

The counter becomes zero upon arrival at 3 from 7. Hence the backtrack journey is given by 3, 4, 6, 5, 2, 1, 0, and $N=6$.

The recording of S even permits us sometimes to shorten the route leading to 0 from the vertex at which the counter became zero. In Example 4, suppose that we leave vertex 3 by an edge other than its entrance edge. We then arrive at vertex 7, 2 or 1. The backtrack journey from these vertices includes only 3, 2 or 1 edges respectively. Thus the route from 3 back to 0 includes only 4, 3 or 2 edges respectively.

If in addition to numbering intersections and recording their sequence of traversal also edges are labeled on the labyrinth's corridors—one label at each of its two ends—additional advantages result. For example, a shortest route back to 0 can be pursued after the counter became zero. Such a route can be planned with the aid of the sequence S alone. The labels on the corridors enable one to realize the plan. In Example 4, it enables the traveler to go from vertex 3 at which the counter became zero, directly to 0 via 1.

The labeling of edges also clearly enables one to carry out subsequent excursions of the labyrinth in a most economical fashion.

This work was sponsored, in part, by the U.S. National Bureau of Standards. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

References

1. W. W. R. Ball, Mathematical Recreations and Essays, rev. by H. S. M. Coxeter, Macmillan, New York, 1947.

2. D. König, Theorie der endlichen und unendlichen Graphen, Chelsea, New York, 1950.

3. O. Ore, Theory of graphs, Amer. Math. Soc. Coll. Publ., 38 (1962).

4. B. A. Trakhtenbrot, Algorithms and Automatic Computing Machines, (transl.), Heath, Boston, 1966.

A REPRESENTATION PROBLEM

R. G. STANTON, University of Manitoba

1. Introduction. The purpose of this paper is to give an exposition of an interesting problem in number theory in such a way that it can be used as an enrichment topic in secondary schools. It is also hoped that students will be stimulated to search out and prove some elementary theorems on their own.

Suppose that n is a nonnegative integer and that x and y are integers. Then

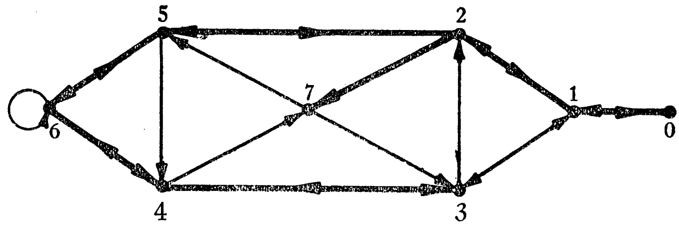

FIG. 6.

The counter becomes zero upon arrival at 3 from 7. Hence the backtrack journey is given by 3, 4, 6, 5, 2, 1, 0, and $N=6$.

The recording of S even permits us sometimes to shorten the route leading to 0 from the vertex at which the counter became zero. In Example 4, suppose that we leave vertex 3 by an edge other than its entrance edge. We then arrive at vertex 7, 2 or 1. The backtrack journey from these vertices includes only 3, 2 or 1 edges respectively. Thus the route from 3 back to 0 includes only 4, 3 or 2 edges respectively.

If in addition to numbering intersections and recording their sequence of traversal also edges are labeled on the labyrinth's corridors—one label at each of its two ends—additional advantages result. For example, a shortest route back to 0 can be pursued after the counter became zero. Such a route can be planned with the aid of the sequence S alone. The labels on the corridors enable one to realize the plan. In Example 4, it enables the traveler to go from vertex 3 at which the counter became zero, directly to 0 via 1.

The labeling of edges also clearly enables one to carry out subsequent excursions of the labyrinth in a most economical fashion.

This work was sponsored, in part, by the U.S. National Bureau of Standards. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

References

1. W. W. R. Ball, Mathematical Recreations and Essays, rev. by H. S. M. Coxeter, Macmillan, New York, 1947.
2. D. König, Theorie der endlichen und unendlichen Graphen, Chelsea, New York, 1950.
3. O. Ore, Theory of graphs, Amer. Math. Soc. Coll. Publ., 38 (1962).
4. B. A. Trakhtenbrot, Algorithms and Automatic Computing Machines, (transl.), Heath, Boston, 1966.

A REPRESENTATION PROBLEM

R. G. STANTON, University of Manitoba

1. Introduction. The purpose of this paper is to give an exposition of an interesting problem in number theory in such a way that it can be used as an enrichment topic in secondary schools. It is also hoped that students will be stimulated to search out and prove some elementary theorems on their own.

Suppose that n is a nonnegative integer and that x and y are integers. Then we ask: in how many ways can n be represented as the sum of two squares, that is, in how many ways can we write $n=x^2+y^2$? If we denote this frequency function by $f(n)$, it can easily be found, for n small, by trial and error. Thus

$$2 = 1^2 + 1^2 = (-1)^2 + (-1)^2 = (-1)^2 + 1^2 = 1^2 + (-1)^2,$$

and we see that $f(2)=4$ (note that we consider x^2+y^2 and y^2+x^2 as different representations of n, for $x\neq y$).

2. Tabulation of $f(n)$. It is always useful in any piece of research to get some experimental facts, and $f(n)$ is easily found by an algorithm which works down from the larger value of x^2 or y^2. For example, if $n=1025$, we can make the following table, in which we assign squared values to x^2 and mark with an asterisk suitable y^2 values.

x^2	$y^2 = n - x^2$
32^2	1*
31^2	64*
30^2	125
29^2	184
28^2	241
27^2	296
26^2	349
25^2	400*
24^2	449
23^2	496

The solutions (32, 1), (31, 8), and (25, 20) give, by interchange and sign alteration, a total of 24 solutions; thus $f(1025)=24$.

We record, in Table I, the values of $f(n)$ for $n\leqq 51$.

3. Some theorems concerning $f(n)$. From Table I, we immediately note that $f(n)$ behaves very erratically, and it would appear very difficult to obtain an explicit formula for $f(n)$. In such a case, it often pays us to rephrase the question, and one obvious change is to ask when $f(n)$ is equal to zero. Even a cursory search through Table I suggests:

THEOREM 1. *If $n=4k+3$, then $f(n)=0$.*

Proof. We suppose that $x^2+y^2=4k+3$, and discuss three cases.

Case 1. x and y even; then $(2a)^2+(2b)^2=4k+3$, $4(a^2+b^2-k)=3$, which is impossible.

Case 2. x and y odd; then $(2a+1)^2+(2b+1)^2=4k+3$, $4(a^2+a+b^2+b-k)=1$, which is impossible.

Case 3. One of x or y even, the other odd; then $(2a)^2+(2b+1)^2=4k+3$, $4(a^2+b^2+b-k)=2$, which is impossible.

Since no case is possible, the theorem is proved.

TABLE I. Table of $f(n)$ where $n=x^2+y^2$.

n	(x, y)	$f(n)$	n	(x, y)	$f(n)$
0	(0, 0)	1	26	(±5, ±1), (±1, ±5)	8
1	(1, 0,), (−1, 0), (0, 1), (0, −1)	4	27		0
2	(1, 1), (1, −1), (−1, 1), (−1, −1)	4	28		0
3		0	29	(±5, ±2), (±2, ±5)	8
4	(±2, 0), (0, ±2)	4	30		0
5	(±2, ±1), (±1, ±2)	8	31		0
6		0	32	(±4, ±4)	4
7		0	33		0
8	(±2, ±2)	4	34	(±5, ±3), (±3, ±5)	8
9	(±3, 0), (0, ±3)	4	35		0
10	(±3, ±1), (±1, ±3)	8	36	(±6, 0), (0, ±6)	4
11		0	37	(±6, ±1), (±1, ±6)	8
12		0	38		0
13	(±2, ±3), (±3, ±2)	8	39		0
14		0	40	(±6, ±2), (±2, ±6)	8
15		0	41	(±5, ±4), (±4, ±5)	8
16	(±4, 0), (0, ±4)	4	42		0
17	(±4, ±1), (±1, ±4)	8	43		0
18	(±3, ±3)	4	44		0
19		0	45	(±6, ±3), (±3, ±6)	8
20	(±4, ±2), (±2, ±4)	8	46		0
21		0	47		0
22		0	48		0
23		0	49	(±7, 0), (0, ±7)	4
24		0	50	(±7, ±1), (±1, ±7), (±5, ±5)	12
25	(±5, 0), (0, ±5), (±3, ±4), (±4, ±3)	12	51		0

Many students will realize that the proof of Theorem 1 can be shortened using congruence notation, but we avoid it in order to stress the fact that all our results can be obtained by even more elementary methods.

As an immediate development from Theorem 1, we obtain two other results.

THEOREM 2. *If n is a multiple of* 4, *then x and y must both be even.*

Proof. We use cases, as in Theorem 1.

THEOREM 3. *If $n=16k+12$, then $f(n)=0$.*

Proof. Let $x^2+y^2=16k+12$. By Theorem 2, $x=2a$, $y=2b$. Hence $a^2+b^2=4k+3$. By Theorem 1, this is impossible. Observation of Table I now suggests:

THEOREM 4. *If $n=8k+6$, then $f(n)=0$.*

Proof. Trying $x=2a$, $y=2b$, leads to the result that 4 divides 6. Trying $x=2a+1$, $y=2b+1$, leads to the result $a(a+1)+b(b+1)=2k+1$; this is impossible since $a(a+1)$ and $b(b+1)$ are both even numbers. So Theorem 4 is proved.

Now we consider some theorems dependent upon the value of n modulo 3. The most easily derived is:

THEOREM 5. *If $n=12k+9$ and 3 does not divide k, then $f(n)=0$.*

Proof. If $x^2+y^2=12k+9$, we note that $x^2=(3a)^2=9a^2$ or $x^2=(3a\pm 1)^2=3A+1$, and similarly for y^2. Any combination of these forms leads to a contradiction.

As a generalization of Theorem 3, we prove a result which is quite similar.

THEOREM 6. *If $f(n_1)=0$, then $f(4n_1)=0$.*

Proof. Let $x^2+y^2=4n_1$. Then x and y must both be even, say $x=2a$, $y=2b$. This requires that $a^2+b^2=n_1$, which is impossible, by hypothesis.

Students might like to try their hand on the case not of $4n_1$ but of $9n_1$. A result related to Theorem 6 appears in:

THEOREM 7. *If 3 does not divide k, then $f(6k)=0$.*

Proof. Let $x^2+y^2=6k$; then clearly x^2 and y^2 have the form $9a^2$ or $3A+1$. Any combination of these forms leads to a contradiction.

The sort of arguments presented so far are very direct, and interested students can derive more theorems of this kind. Naturally, the ideal thing would be to determine all values of n for which $f(n)=0$; this is difficult, but we have illustrated how to make considerable progress on finding some classes of n's for which $f(n)=0$. Perhaps it would be valuable to include two more difficult theorems as a conclusion for this section.

THEOREM 8. *If $f(n_1)>0$ and $f(n_2)>0$, then $f(n_1n_2)>0$.*

Proof. The hypothesis shows that there exist numbers a, b, c, d, such that $a^2+b^2=n_1$, $c^2+d^2=n_2$. Then

$$\begin{aligned} n_1n_2 &= (a^2+b^2)(c^2+d^2) \\ &= a^2c^2+a^2d^2+b^2c^2+b^2d^2 \\ &= (ac+bd)^2+(ad-bc)^2. \end{aligned}$$

Thus $f(n_1n_2)>0$.

THEOREM 9. *If $n=p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}$ where the p_i are distinct primes and the a_i are even, then $f(n)>0$.*

Proof. The result follows from the fact that, if $a_i=2b_i$, then $p_i^{a_i}=p_i^{2b_i}=(p_i^{b_i})^2+0^2$. Thus $f(p_i^{a_i})>0$, and, by Theorem 8, $f(n)>0$.

By more advanced methods, it is possible to extend Theorem 9 to a result which we shall not prove, namely,

THEOREM 10. *If $n=2^a p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}$, then $f(n)>0$ if and only if every prime p_i of the form $4r+3$ appears to an even power a_i.*

4. The average value of $f(n)$. Sometimes, when a function behaves as erratically as $f(n)$ does, the average value of the function is a more stable quantity. In Table II, we tabulate

$$\operatorname{ave} f(n) = \frac{1}{n+1}\sum_{i=0}^{n} f(i)$$

TABLE II. Average value of $f(n)$.

n	ave $f(n)$		n	ave $f(n)$		n	ave $f(n)$		n	ave $f(n)$	
0	1/1	1.00	13	45/14	3.21	26	89/27	3.30	39	121/40	3.03
1	5/2	2.50	14	45/15	3.00	27	89/28	3.18	40	129/41	3.15
2	9/3	3.00	15	45/16	2.81	28	89/29	3.09	41	137/42	3.26
3	9/4	2.50	16	49/17	2.88	29	97/30	3.23	42	137/43	3.19
4	13/5	2.60	17	57/18	3.17	30	97/31	3.13	43	137/44	3.11
5	21/6	3.50	18	61/19	3.21	31	97/32	3.03	44	137/45	3.04
6	21/7	3.00	19	61/20	3.05	32	101/33	3.06	45	145/46	3.15
7	21/8	2.63	20	69/21	3.29	33	101/34	2.97	46	145/47	3.08
8	25/9	2.78	21	69/22	3.14	34	109/35	3.11	47	145/48	3.02
9	29/10	2.90	22	69/23	3.00	35	109/36	3.03	48	145/49	2.96
10	37/11	3.36	23	69/24	2.88	36	113/37	3.06	49	149/50	2.98
11	37/12	3.08	24	69/25	2.76	37	121/38	3.18	50	161/51	3.16
12	37/13	2.85	25	81/26	3.12	38	121/39	3.10	51	161/52	3.10

for the same values of n as in Table I.

Table II certainly suggests that

$$\lim_{n\to\infty} \text{ave } f(n)$$

exists, and is a number slightly greater than 3. In the next section, we shall verify this empirical result.

5. A geometric approach to ave $f(n)$. If we draw a circle of radius $\sqrt{n}$ in the Cartesian plane (this is done in Figure 1 for $n=10$), we see that all the lattice

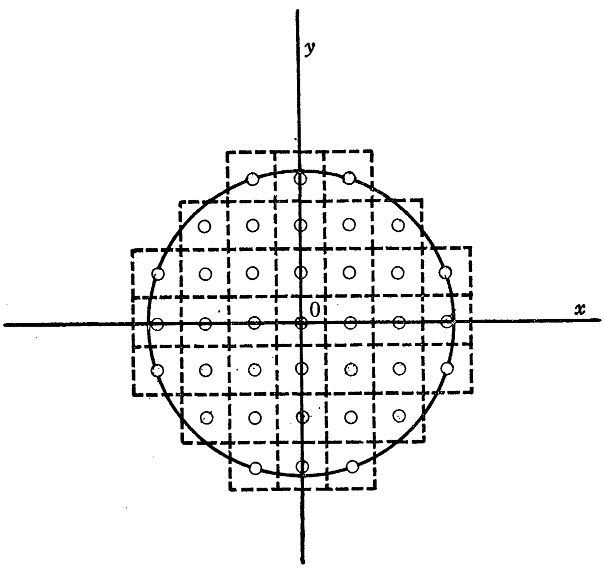

FIG. 1. Value of cum $f(n)$ for $n=10$.

points (that is, points with integer coordinates) on this circle are just those number pairs (x, y) such that $x^2+y^2=n$. Also, the lattice points inside the circle are exactly those number pairs (x, y) such that $x^2+y^2=i$, where $i<n$. Thus, we have the result stated in:

THEOREM 11. *The expression* cum $f(n)$ *is equal to the number of lattice points in or on the circle* $x^2+y^2=n$, *where* cum $f(n)=\sum_{i=0}^{n} f(i)$.

Suppose we now draw a grid of unit squares with centers at the lattice points illustrated in Figure 1. These squares are dotted in Figure 1. Clearly the number of these squares is equal to their area, and so we have:

THEOREM 12. *The expression* cum $f(n)$ *is equal to the area of the figure F formed by taking all unit squares whose centers lie in or on the circle* $x^2+y^2=n$.

We cannot exactly evaluate the area described in Theorem 12. However, we can note that any unit square can jut at most $1/\sqrt{2}$ units past the circle $x^2+y^2=n$ (this is half the diagonal of a unit square). Similarly, unused squares can jut at most $1/\sqrt{2}$ units into the circle. So we have:

THEOREM 13. *The figure F described in Theorem* 12 *completely contains the circle* $x^2+y^2=(\sqrt{n}-1/\sqrt{2})^2$, *and is completely contained within the circle* $x^2+y^2=(\sqrt{n}+1/\sqrt{2})^2$.

We can now produce upper and lower bounds on

$$\begin{aligned} \operatorname{ave} f(n) &= \frac{1}{n+1} \operatorname{cum} f(n) \\ &= \frac{1}{n+1} (\text{area of } F). \end{aligned}$$

We use Theorem 13 to give the result:

$$\frac{1}{n+1}\,\pi\left[\sqrt{n}-\frac{1}{\sqrt{2}}\right]^2 < \operatorname{ave} f(n) < \frac{1}{n+1}\,\pi\left[\sqrt{n}+\frac{1}{\sqrt{2}}\right]^2.$$

This result can be simplified to:

THEOREM 14.

$$\pi\left[1-\frac{\sqrt{2n}+1/2}{n+1}\right] < \operatorname{ave} f(n) < \pi\left[1+\frac{\sqrt{2n}-1/2}{n+1}\right].$$

The bounds given in Theorem 14 are quite good; for example, suppose we let $n=5000$. Then Theorem 14 tells us that

$$\pi\left[1-\frac{100.5}{5001}\right] < \operatorname{ave} f(5000) < \pi\left[1+\frac{99.5}{5001}\right],$$

a result which simplifies to

$$\frac{4900.5}{5001}\pi < \text{ave } f(50) < \frac{5100.5}{5001}\pi ,$$

or, simply, $3.08 < \text{ave } f(50) < 3.20$.

It is of course clear from Theorem 14 that, as $n \to \infty$, the upper and lower bounds both approach π, and so we have:

THEOREM 15. $\lim_{n\to\infty} \text{ave } f(n) = \pi$.

This result is well illustrated in Table II.

6. Extensions to more than two squares. Exactly similar considerations can be brought into play if $x^2+y^2+z^2=n$. Using cases, it is easy to prove:

THEOREM 16. *If $f_3(n)$ is the number of solutions of $x^2+y^2+z^2=n$, and if $n=8k+7$, then $f_3(n)=0$.*

While more difficult to prove, a converse of Theorem 16 is also true, that is, $f_3(n)=0$ only if n has the form $8k+7$.

Using the fact that the length of the diagonal of a unit cube is $\sqrt{3}$ and that the volume of a sphere is $(4/3)\pi r^3$, it is easy to obtain the three-dimensional analogue of Theorem 14, namely,

THEOREM 17.

$$\frac{4\pi}{3(n+1)}\left[\sqrt{n} - \frac{\sqrt{3}}{2}\right]^3 < \text{ave } f_3(n) < \frac{4\pi}{3(n+1)}\left[\sqrt{n} + \frac{\sqrt{3}}{2}\right]^3.$$

Theorem 17 shows that ave $f_3(n)$ does not approach a limit as $n \to \infty$; however, it is easy to deduce:

THEOREM 18.

$$\lim_{n\to\infty} \frac{\text{ave } f_3(n)}{\sqrt{n}} = \frac{4\pi}{3}.$$

If more than 3 squares are permitted, that is, if we define $f_t(n)$ to be the number of ways of writing n as the sum of t squares, then it is possible to prove:

THEOREM 19. *If $t \geqq 4$, then $f_t(n) > 0$.*

Also, by using the formula for the volume of a sphere in t-dimensional space, one can get the generalized analogue of Theorems 15 and 18 as:

THEOREM 20.

$$\lim_{n\to\infty} \frac{\text{ave } f_t(n)}{n^{(t-2)/2}} = \frac{\pi^m}{m!}$$

for $t = 2m$, and

$$\lim_{n\to\infty} \frac{\text{ave } f_t(n)}{n^{(t-2)/2}} = \frac{\pi^m m!\, 2^{2m+1}}{(2m+1)!}$$

for $t = 2m + 1$.

7. Conclusion. We have attempted in this exposition to illustrate a few general ideas which can be useful to students carrying out an investigation; perhaps it might help to summarize these ideas.

1. It is very helpful to obtain experimental data as in Table I.

2. Computer tabulation can be useful; this would be especially noticeable if we wanted to extend Table I or to make a table of ave $f_3(n)$, as defined in Section 6.

3. If one can guess theorems from data, then it is often relatively easy to prove them. Almost all the results of Section 3 were guessed from Table I.

4. A geometric representation, as in Section 5, often makes a problem more concrete, easier to grasp, and easier to solve.

5. If one cannot solve a problem, one should try to formulate other problems that one can solve. We did not obtain a formula for $f(n)$, but we did manage to find many classes of values for which $f(n)=0$, and we also managed to determine the limiting value of ave $f(n)$.

6. From almost any problem, one can begin thinking of generalizations, as in Section 6.

EXTENSIONS OF THE WEIERSTRASS PRODUCT INEQUALITIES

M. S. KLAMKIN, Ford Scientific Laboratory and
D. J. NEWMAN, Yeshiva University

If $A_1, A_2, \cdots, A_n$ are numbers in $[0, 1]$ whose sum is denoted by S_1, then the Weierstrass inequalities [1] are given by

$$1 - S_1 \leqq (1 - A_1)(1 - A_2) \cdots (1 - A_n) \leqq [1 + S_1]^{-1} \tag{1}$$

$$1 + S_1 \leqq (1 + A_1)(1 + A_2) \cdots (1 + A_n) \leqq [1 - S_1]^{-1} \tag{2}$$

where, in the last inequality, it is supposed that $S_1 \leqq 1$. These inequalities are useful in treating the convergence of infinite products [1].

We extend the above inequalities by first showing that

$$1 - S_1 + S_2 - \cdots + S_{2p} \geqq \prod_{i=1}^{n} (1 - A_i), \qquad p = 0, 1, \cdots, [n/2], \qquad n = 1, 2, \cdots; \tag{3}$$

$$\prod_{i=1}^{n} (1 - A_i) \geqq 1 - S_1 + S_2 - \cdots - S_{2q-1}, \qquad q = 0, 1, \cdots, [(n+1)/2], \qquad n = 1, 2, \cdots; \tag{4}$$

$$1 - S_1 + S_2 - \cdots + (-1)^r S_r \geqq 0, \qquad r = 0, 1, \cdots, n \quad \text{and} \quad S_1 \leqq 1. \tag{5}$$

Here the S_i's are the elementary symmetric functions of the A_i's, i.e.,

$$\prod_{i=1}^{n} (x + A_i) = x^n + S_1 x^{n-1} + S_2 x^{n-2} + \cdots + S_n.$$

7. Conclusion. We have attempted in this exposition to illustrate a few general ideas which can be useful to students carrying out an investigation; perhaps it might help to summarize these ideas.

1. It is very helpful to obtain experimental data as in Table I.
2. Computer tabulation can be useful; this would be especially noticeable if we wanted to extend Table I or to make a table of ave $f_3(n)$, as defined in Section 6.
3. If one can guess theorems from data, then it is often relatively easy to prove them. Almost all the results of Section 3 were guessed from Table I.
4. A geometric representation, as in Section 5, often makes a problem more concrete, easier to grasp, and easier to solve.
5. If one cannot solve a problem, one should try to formulate other problems that one can solve. We did not obtain a formula for $f(n)$, but we did manage to find many classes of values for which $f(n)=0$, and we also managed to determine the limiting value of ave $f(n)$.
6. From almost any problem, one can begin thinking of generalizations, as in Section 6.

EXTENSIONS OF THE WEIERSTRASS PRODUCT INEQUALITIES

M. S. KLAMKIN, Ford Scientific Laboratory and
D. J. NEWMAN, Yeshiva University

If $A_1, A_2, \cdots, A_n$ are numbers in $[0, 1]$ whose sum is denoted by S_1, then the Weierstrass inequalities [1] are given by

$$1 - S_1 \leqq (1 - A_1)(1 - A_2) \cdots (1 - A_n) \leqq [1 + S_1]^{-1} \tag{1}$$

$$1 + S_1 \leqq (1 + A_1)(1 + A_2) \cdots (1 + A_n) \leqq [1 - S_1]^{-1} \tag{2}$$

where, in the last inequality, it is supposed that $S_1 \leqq 1$. These inequalities are useful in treating the convergence of infinite products [1].

We extend the above inequalities by first showing that

$$1 - S_1 + S_2 - \cdots + S_{2p} \geqq \prod_{i=1}^{n} (1 - A_i), \qquad p = 0, 1, \cdots, [n/2], \quad n = 1, 2, \cdots; \tag{3}$$

$$\prod_{i=1}^{n} (1 - A_i) \geqq 1 - S_1 + S_2 - \cdots - S_{2q-1}, \qquad q = 0, 1, \cdots, [(n+1)/2], \quad n = 1, 2, \cdots; \tag{4}$$

$$1 - S_1 + S_2 - \cdots + (-1)^r S_r \geqq 0, \qquad r = 0, 1, \cdots, n \quad \text{and} \quad S_1 \leqq 1. \tag{5}$$

Here the S_i's are the elementary symmetric functions of the A_i's, i.e.,

$$\prod_{i=1}^{n} (x + A_i) = x^n + S_1 x^{n-1} + S_2 x^{n-2} + \cdots + S_n.$$

The proof is by induction. Assume that (3) is valid for $n=1, 2, \cdots, m$. Then by multiplying (3) (for $n=m$) by $1-A_{m+1}$, we get

$$1-(S_1+A_{m+1})+(S_2+S_1A_{m+1})-\cdots+(S_{2p}+S_{2p-1}A_{m+1})$$
$$\geqq \prod_{i=1}^{m+1}(1-A_i)+S_{2p}A_{m+1}.$$

The new symmetric functions S_k' (of the $n+1$ terms $A_1, A_2, \cdots, A_{m+1}$) can be expressed in terms of the old symmetric functions (of $A_1, A_2, \cdots, A_m$) by

$$S_k' = S_k + S_{k-1}A_{m+1} \qquad (k = 1, 2, \cdots, m+1; S_0 = 1).$$

Whence,

$$1 - S_1' + S_2' - \cdots + S_{2p}' \geqq \prod_{i=1}^{m+1}(1-A_i).$$

This does not include the case when $n=m+1$ (m odd) and $2p=m+1$. However, this case is valid since then the l.h.s. identically equals the r.h.s. Since (1) is valid for $n=1$, it is now valid for all n.

The proof of (4) is virtually identical to that of (3).

In order to establish (5), it suffices to show that $S_k \geqq S_{k+1}$. Since $S_1 \leqq 1$,

$$S_k \geqq S_1S_k \geqq (k+1)S_{k+1}.$$

We now show that the bounds in (3) and (4) are better than the corresponding bounds in (1) for the case when $S_1 \leqq 1$. Here

$$1 - S_1 + S_2 \leqq \frac{1}{1+S_1}$$

since this is equivalent to

$$(1+S_1)S_2 \leqq S_1^2 = 2S_2 + \sum A_i^2.$$

The rest follows from (5).

The lower bound of (2) can be trivially bettered to $(1+S_1+S_2+\cdots+S_p)$ for $p=0, 1, 2, \cdots, n$. For an improved upper bound, we start with:

$$\prod_{i=1}^{n}(1-A_i^2) = \prod_{i=1}^{n}(1+A_i)(1-A_i) \leqq 1, \tag{6}$$

to give

$$\prod_{i=1}^{n}(1+A_i) \leqq \prod_{i=1}^{n}(1-A_i)^{-1} \leqq \{1-S_1+S_2-\cdots-S_{2q-1}\}^{-1}, \tag{7}$$
$$q = 0, 1, \cdots, [(n+1)/2].$$

Another set of bounds for the products in (1) and (2) can be obtained by extending the following known inequalities,

$$8xyz \leqq (1-x)(1-y)(1-z) \leqq \frac{8}{27}, \tag{8}$$

where $x, y, z \geqq 0$, $x+y+z=1$ [2, p. 226];

$$(a+1)(b+1)(c+1)(d+1) \leqq 8(abcd+1), \tag{9}$$

where $a, b, c, d \geqq 1$ [2, p. 225];

$$\prod_{i=1}^{n} (1+A_i) \geqq \frac{2^n}{n+1}(1+S_1), \qquad A_i \geqq 1 \quad [3, \text{p. } 99]; \tag{10}$$

$$\prod_{i=1}^{n} (1-A_i) \geqq \left\{\frac{n-S_1}{S_1}\right\}^n S_n, \qquad 1/2 \geqq A_i > 0 \quad [4, \text{p. } 5]. \tag{11}$$

Another related inequality (for which we do not have an extension) is

$$\prod_{i=1}^{n} (1+A_i) \geqq \left\{1+S_n^{1/n}\right\}^n, \qquad A_i \geqq 0 \quad [3, \text{p. } 130]. \tag{12}$$

The extensions are given by,

$$\prod_{i=1}^{n} (1+A_i) \leqq \left\{\frac{n+S_1}{n}\right\}^n < e^{S_1}, \qquad A_i \geqq 0; \tag{13}$$

$$\prod_{i=1}^{n} (1-A_i) \leqq \left\{\frac{n-S_1}{n}\right\}^n < e^{-S_1}, \qquad 1 \geqq A_i \geqq 0; \tag{14}$$

$$\prod_{i=1}^{n} (1+A_i) \geqq (n+1)^n S_n \tag{15}$$

where $A_i \geqq 0$ and $S_1 = 1$;

$$\prod_{i=1}^{n} (1-A_i) \geqq (n-1)^n S_n \tag{16}$$

where $1 \geqq A_i \geqq 0$ and $S_1 = 1$;

$$\prod_{i=1}^{n} (1+A_i) \leqq 2^{n-1}(1+S_n) \tag{17}$$

where $A_i \geqq 1$ (or equivalently $1 \geqq A_i \geqq 0$);

$$\prod_{i=1}^{n} (1+A_i) \geqq \frac{(1+\lambda)^n}{1+n\lambda}(1+S_1), \qquad A_i \geqq \lambda \geqq 0; \tag{18}$$

$$\prod_{i=1}^{n} (1+A_i) \leqq \frac{(1+\lambda)^n}{1+n\lambda}(1+S_1), \qquad \lambda \geqq A_i \geqq 0; \tag{19}$$

$$\prod_{i=1}^{n} (1-A_i) \geqq \left\{\frac{1-S_1/n}{(S_1/n)^{a^2}}\right\}^n \prod_{i=1}^{n} A_i^{a^2}, \qquad \frac{a}{a+1} \geqq A_i > 0. \tag{20}$$

Proofs. (13) and (14) follow from the arithmetic-geometric mean inequality (A.M.-G.M.) (or from the convexity of $\pm\log(1\pm x)$).

$$\sum_{i=1}^{n} \frac{1 \pm A_i}{n} = \frac{n \pm S_1}{n} \geqq \prod_{i=1}^{n} (1 \pm A_i)^{1/n}.$$

Also, $(1\pm S_1/n)^n$ increases monotonically with n and $\to \exp(\pm S_1)$.

(15) and (16) also follow from the A.M.-G.M.

$$\frac{1 \pm A_i}{n \pm 1} = \frac{S_1 \pm A_i}{n \pm 1} \geqq \left\{ S_n A_i^{\pm 1} \right\}^{1/(n\pm 1)}.$$

Thus

$$\prod_{i=1}^{n} (1 \pm A_i) \geqq (n \pm 1)^n S_n.$$

For (17), the proof is by induction. Assume (17) is valid for $n=k-1$. Then

$$\prod_{i=1}^{k} (1 + A_i) \leqq 2^{k-2}(1 + S_{k-1})(1 + A_k).$$

(17) will also be valid for $n=k$, if

$$2^{k-2}(1 + S_{k-1})(1 + A_k) \leqq 2^{k-1}(1 + S_{k-1}A_k).$$

But the latter is valid since it is equivalent to $(S_{k-1}-1)(A_k-1)\geqq 0$.

Since (17) is valid for $n=1$, it is valid for all n.

The proofs of (18) and (19) are similar to that of (17). An alternate form for (18) is

$$\prod_{i=1}^{n} (1 + A_i) \geqq \frac{(1 + \lambda)^n}{1 + n\lambda} (S_n + S_{n-1}), \qquad 1/\lambda \geqq A_i > 0.$$

In [4, p. 5], inequality (11) is ascribed to an unpublished result of Ky Fan and it is noted that it can be established by forward and backward induction. Here, we establish the extension (20) by a convexity argument. Since

$$D^2\left\{\log(1 - x) - a^2 \log x\right\} = \frac{a^2}{x^2} - \frac{1}{(1 - x)^2} \geqq 0$$

for $0\leqq x\leqq a/(a+1)$, we then have

$$\frac{1}{n} \sum_{i=1}^{n} \left\{\log(1 - A_i) - a^2 \log A_i\right\} \geqq \log(1 - S_1/n) - a^2 \log S_1/n$$

which is equivalent to (20).

The proof of (12) follows immediately from the extension of Hölder's inequality [4, p. 20]

$$\prod_{i=1}^{n} (A_i + B_i) \geqq \left(\prod_{i=1}^{n} A_i^{1/n} + \prod_{i=1}^{n} B_i^{1/n} \right)^n$$

where $A_i, B_i \geqq 0$.

References

1. T. J. I'A. Bromwich, An Introduction to the Theory of Infinite Series, Macmillan, London, 1965, pp. 104–106.
2. S. Barnard and J. M. Child, Higher Algebra, Macmillan, London, 1949.
3. D. S. Mitrinović, Elementary Inequalities, Noordhoff, Groningen, 1964.
4. E. F. Beckenbach and R. Bellman, Inequalities, Springer-Verlag, New York, 1965.

DEFINING THE INTERCHANGE OF A LOOP

B. L. SCHWARTZ, The Mitre Corporation

1. Introduction. In graph theory, an interesting concept is that of the interchange of a graph. However, the definition of the interchange graph of a loop is not provided in a consistent manner by existing references. The purpose of this note is to suggest a defintion of interchange under which loops will be included in a natural manner. We also provide some heuristic justification, showing why the proposal we here put forth is a reasonable and consistent generalization of the interchange concept, in preference to other alternatives.

2. Existing conventions. In [1], the interchange $I(G)$ of a graph G is defined as the graph whose vertices are the edges of G; and two vertices of $I(G)$ are adjacent iff the corresponding edges of G have a common vertex. This definition is restricted to loop-free graphs without parallel edges; and the phraseology used does not admit to relaxing these limitations.

In a generalization to s-graphs (parallel edges admitted), Menon [2] proposed that any pair of vertices e_1 and e_2 in $I(G)$ be joined by 0, 1, or 2 edges respectively, accordingly as the corresponding edges in G had 0, 1, or 2 vertices in common. Although it is not explicitly stated in that definition, it is clear from the examples that the definition applies only when $e_1 \neq e_2$. It is therefore implicitly limited to loop-free s-graphs. Such a limitation is certainly desirable. The reader can quickly verify that use of the definition when $e_1 = e_2$ would lead to unreasonably extensive proliferation of loops in interchange graphs.

For graphs with loops, the only references dealing with the interchange operation have treated the loops as special cases, to be accommodated by a separate convention. In [3], the present writer has proposed that the interchange of a loop be another loop. But in [4], Menon has implied that it should be two loops on a single vertex. The two papers reach apparently contradicting results about the same problem. The need is therefore underscored to have a common understanding on this concept.

3. A proposed new convention. We propose to fill this definitional gap by extending Ore's concept in a consistent manner. Ore's discussion stated ([1], p. 20), where A_i, $B_i \geqq 0$.

References

1. T. J. I'A. Bromwich, An Introduction to the Theory of Infinite Series, Macmillan, London, 1965, pp. 104–106.
2. S. Barnard and J. M. Child, Higher Algebra, Macmillan, London, 1949.
3. D. S. Mitrinović, Elementary Inequalities, Noordhoff, Groningen, 1964.
4. E. F. Beckenbach and R. Bellman, Inequalities, Springer-Verlag, New York, 1965.

DEFINING THE INTERCHANGE OF A LOOP

B. L. SCHWARTZ, The Mitre Corporation

1. Introduction. In graph theory, an interesting concept is that of the interchange of a graph. However, the definition of the interchange graph of a loop is not provided in a consistent manner by existing references. The purpose of this note is to suggest a defintion of interchange under which loops will be included in a natural manner. We also provide some heuristic justification, showing why the proposal we here put forth is a reasonable and consistent generalization of the interchange concept, in preference to other alternatives.

2. Existing conventions. In [1], the interchange $I(G)$ of a graph G is defined as the graph whose vertices are the edges of G; and two vertices of $I(G)$ are adjacent iff the corresponding edges of G have a common vertex. This definition is restricted to loop-free graphs without parallel edges; and the phraseology used does not admit to relaxing these limitations.

In a generalization to s-graphs (parallel edges admitted), Menon [2] proposed that any pair of vertices e_1 and e_2 in $I(G)$ be joined by 0, 1, or 2 edges respectively, accordingly as the corresponding edges in G had 0, 1, or 2 vertices in common. Although it is not explicitly stated in that definition, it is clear from the examples that the definition applies only when $e_1 \neq e_2$. It is therefore implicitly limited to loop-free s-graphs. Such a limitation is certainly desirable. The reader can quickly verify that use of the definition when $e_1 = e_2$ would lead to unreasonably extensive proliferation of loops in interchange graphs.

For graphs with loops, the only references dealing with the interchange operation have treated the loops as special cases, to be accommodated by a separate convention. In [3], the present writer has proposed that the interchange of a loop be another loop. But in [4], Menon has implied that it should be two loops on a single vertex. The two papers reach apparently contradicting results about the same problem. The need is therefore underscored to have a common understanding on this concept.

3. A proposed new convention. We propose to fill this definitional gap by extending Ore's concept in a consistent manner. Ore's discussion stated ([1], p. 20), "The actual construction of the interchange graph $I(G)$ from the diagram of G is simple. On each edge E one selects a fixed point e_E; for instance, the midpoint of E. Then one connects a pair of such vertices (e, e') by a new edge belonging to $I(G)$ if and only if the corresponding edges E and E' have a vertex in common in G."

We use this as a starting point, making precise certain necessary details.

Let e_1 and e_2 be edges of G with generic interior points C_1 and C_2 respectively. Then C_1 and C_2 are vertices of $I(G)$. And the number of edges in $I(G)$ joining C_1 and C_2 is the number of distinct ways in which a vertex α in G can be chosen so that it is possible that, starting at C_1, a point can move along the edges of G to C_2 via α without retracing and without meeting any other vertex of G.

Not only will this absorb the previous definitions for loop-free graphs and s-graphs, and extend them to s-graphs with loops, but it will also permit extension to directed graphs, with or without loops and parallel arcs. The result of that extension is consistent with usage in [5].

4. Examples. In Figure 1 below, a number of graphs are shown with their interchanges, using the above convention. Since the interchange operation was already defined in a generally acceptable manner for loop-free s-graphs, all the

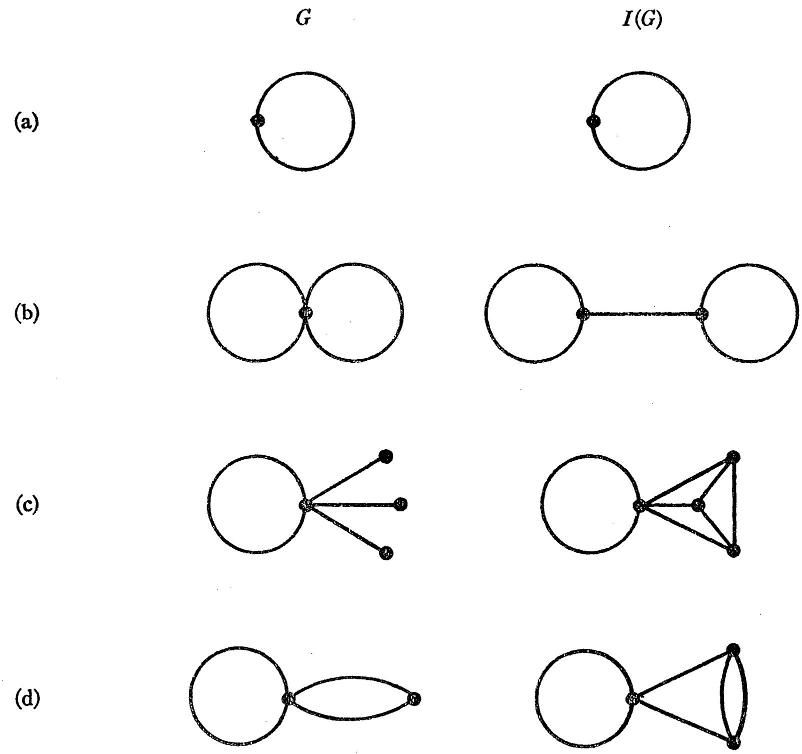

FIG. 1.

examples given are graphs with loops. We have tried to include several typical cases of one or more loops in conjunction with nonloop edges.

References

1. Oystein Ore, Theory of Graphs, Amer. Math. Soc. Colloq. Publ., 38, 1962.
2. V. Menon, The isomorphism between graphs and their adjoint graphs, Canad. Math. Bull., 8 (1965) 7–15.
3. B. L. Schwartz, On interchange graphs, Pacific J. Math., 27(1968) 393–396.
4. V. Menon, On repeated interchange graphs, Amer. Math. Monthly, 13 (1966) 986–989.
5. L. Muracchini and A. Ghirlanda, Sul grafo commutato e sul grafo opposto di un grafo orientato, Atti. Sem. Mat. Fis. Univ. Modena, 14 (1965) 87–97.

NOTE ON A NONLINEAR RECURRENCE RELATED TO $\sqrt{2}$

R. L. GRAHAM and H. O. POLLAK, Bell Telephone Laboratories

In a recent study of sorting algorithms for a partially-sorted set, F. K. Hwang and S. Lin [2] introduced the following sequence:

$$a_1 = 1, \qquad a_{n+1} = \left[\sqrt{2a_n(a_n+1)}\right], \qquad n \geqq 1,$$

where $[\,\cdot\,]$ denotes the greatest integer function. Thus, the sequence begins

n	1	2	3	4	5	6	7	8	9	10	11	12	13	
a_n	1	2	3	4	6	9	13	19	27	38	54	77	109	$\cdots$.

One notes that $a_{2n+1}-a_{2n}=2^{n-1}$ for $1\leqq n\leqq 6$ and it might be conjectured that this holds in general.

In this note we investigate the sequence $\{a_n\}$. We obtain an explicit expression for a_n from which the conjecture follows as well as the following curious result: $a_{2n+1}-2a_{2n-1}$ is just the nth digit in the binary expansion of $\sqrt{2}$.

We begin by making the preliminary observation that if $S(\alpha)$ denotes the set of integers $\{[\alpha], [2\alpha], [3\alpha], \cdots\}$ then every positive integer occurs in exactly one of the two sets $S(1+1/\sqrt{2})$ and $S(1+\sqrt{2})$. This follows from well-known results for $S(\alpha)$ (cf. [1]) together with the fact that $(1+1/\sqrt{2})^{-1}+(1+\sqrt{2})^{-1}=1$ and $1+\sqrt{2}$ is irrational. Thus, for any positive integer m there exists a unique integer t such that either $m=[t(1+1/\sqrt{2})]$ or $m=[t(1+\sqrt{2})]$.

THEOREM. *Let* $a_1=m$, $a_{n+1}=[\sqrt{2a_n(a_n+1)}]$, $n\geqq 1$. *Then*

$$a_n = \begin{cases} [t(2^{(n-1)/2}+2^{(n-2)/2})] & \text{if } m=\left[t\left(1+\dfrac{1}{\sqrt{2}}\right)\right] \\ [t(2^{n/2}+2^{(n-1)/2})] & \text{if } m=[t(1+\sqrt{2})]. \end{cases}$$

Proof. First note that since no integral square lies between

$$2t^2+2t \quad \text{and} \quad 2t^2+2t+\tfrac{1}{2}=2(t+\tfrac{1}{2})^2$$

examples given are graphs with loops. We have tried to include several typical cases of one or more loops in conjunction with nonloop edges.

References

1. Oystein Ore, Theory of Graphs, Amer. Math. Soc. Colloq. Publ., 38, 1962.
2. V. Menon, The isomorphism between graphs and their adjoint graphs, Canad. Math. Bull., 8 (1965) 7–15.
3. B. L. Schwartz, On interchange graphs, Pacific J. Math., 27(1968) 393–396.
4. V. Menon, On repeated interchange graphs, Amer. Math. Monthly, 13 (1966) 986–989.
5. L. Muracchini and A. Ghirlanda, Sul grafo commutato e sul grafo opposto di un grafo orientato, Atti. Sem. Mat. Fis. Univ. Modena, 14 (1965) 87–97.

NOTE ON A NONLINEAR RECURRENCE RELATED TO $\sqrt{2}$

R. L. GRAHAM and H. O. POLLAK, Bell Telephone Laboratories

In a recent study of sorting algorithms for a partially-sorted set, F. K. Hwang and S. Lin [2] introduced the following sequence:

$$a_1 = 1, \qquad a_{n+1} = \left[\sqrt{2a_n(a_n+1)}\right], \qquad n \geqq 1,$$

where $[\,\cdot\,]$ denotes the greatest integer function. Thus, the sequence begins

n	1	2	3	4	5	6	7	8	9	10	11	12	13	
a_n	1	2	3	4	6	9	13	19	27	38	54	77	109	$\cdots$.

One notes that $a_{2n+1}-a_{2n}=2^{n-1}$ for $1 \leqq n \leqq 6$ and it might be conjectured that this holds in general.

In this note we investigate the sequence $\{a_n\}$. We obtain an explicit expression for a_n from which the conjecture follows as well as the following curious result: $a_{2n+1}-2a_{2n-1}$ is just the nth digit in the binary expansion of $\sqrt{2}$.

We begin by making the preliminary observation that if $S(\alpha)$ denotes the set of integers $\{[\alpha], [2\alpha], [3\alpha], \cdots\}$ then every positive integer occurs in exactly one of the two sets $S(1+1/\sqrt{2})$ and $S(1+\sqrt{2})$. This follows from well-known results for $S(\alpha)$ (cf. [1]) together with the fact that $(1+1/\sqrt{2})^{-1}+(1+\sqrt{2})^{-1}=1$ and $1+\sqrt{2}$ is irrational. Thus, for any positive integer m there exists a unique integer t such that either $m=[t(1+1/\sqrt{2})]$ or $m=[t(1+\sqrt{2})]$.

THEOREM. *Let $a_1=m$, $a_{n+1}=[\sqrt{2a_n(a_n+1)}]$, $n\geqq 1$. Then*

$$a_n = \begin{cases} [t(2^{(n-1)/2}+2^{(n-2)/2})] & \text{if } m = \left[t\left(1+\dfrac{1}{\sqrt{2}}\right)\right] \\ [t(2^{n/2}+2^{(n-1)/2})] & \text{if } m = [t(1+\sqrt{2})]. \end{cases}$$

Proof. First note that since no integral square lies between

$$2t^2+2t \quad \text{and} \quad 2t^2+2t+\tfrac{1}{2} = 2(t+\tfrac{1}{2})^2$$

then

$$[\sqrt{2a_n(a_n+1)}] = [\sqrt{2}(a_n+\tfrac{1}{2})].$$

Thus we can assume

$$a_{n+1} = [\sqrt{2}(a_n+\tfrac{1}{2})], \qquad n \geqq 1.$$

Suppose $x=t(1+1/\sqrt{2})$ for some positive integer t. Then

(1) $$[\sqrt{2}([x]+\tfrac{1}{2})] = [\sqrt{2}\,x]$$

Proof of (1). Let

$$\beta = \frac{t}{\sqrt{2}} - \left[\frac{t}{\sqrt{2}}\right].$$

Then $x=[x]+\beta$. Also

$$\sqrt{2}\,x = t(1+\sqrt{2}), \qquad \sqrt{2}\,t - [\sqrt{2}\,t] \equiv \beta'$$

and

$$2\beta = \beta' + \alpha_1, \qquad \alpha_1 = 0 \quad \text{or } 1,$$

(i.e., $\beta = \cdot\alpha_1\alpha_2 \cdots \quad \beta' = \cdot\alpha_2\alpha_3 \cdots$ expressed base 2).

$$\therefore \ (1) \text{ iff } [t(\sqrt{2}+1)] \left[= \sqrt{2}\left(t\left(1+\frac{1}{\sqrt{2}}\right) - \beta + \tfrac{1}{2}\right)\right]$$

$$\text{iff } t + [t\sqrt{2}] = \left[t + [t\sqrt{2}] + \beta' - \sqrt{2}\,\beta + \frac{1}{\sqrt{2}}\right]$$

$$\text{iff } 0 = \left[\beta' - \sqrt{2}\,\beta + \frac{1}{\sqrt{2}}\right] = \left[\beta' - \frac{\sqrt{2}}{2}(\beta'+\alpha_1) + \frac{1}{\sqrt{2}}\right]$$

$$\text{iff } 0 = \left[\beta'\left(1 - \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}(1-\alpha_1)\right].$$

The expression inside is $\geqq 0$. Also $\beta' < 1$ so that the expression inside is $< 1 - 1/\sqrt{2} + 1/\sqrt{2}\cdot 1 = 1$. $\therefore$ (1) holds.

Next, suppose $x=t(1+\sqrt{2})$, for some positive integer t Then

(1′) $$[\sqrt{2}([x]+\tfrac{1}{2})] = [\sqrt{2}\,x].$$

Proof of (1′). As before, let $\beta' = \sqrt{2}\ t - [x\sqrt{2}]$. Then $x\sqrt{2} = t(2+\sqrt{2})$ and $x\sqrt{2} - [x\sqrt{2}] = t\sqrt{2} - [t\sqrt{2}] = \beta'$.

$$\therefore \ (1') \text{ iff } [2t + t\sqrt{2}] = \left[\sqrt{2}\,t(1+\sqrt{2}) - \sqrt{2}\,\beta + \frac{1}{\sqrt{2}}\right]$$

$$\text{iff } 2t + [t\sqrt{2}] = \left[2t + [\sqrt{2}\,t] + \beta - \sqrt{2}\,\beta + \frac{1}{\sqrt{2}}\right]$$

$$\text{iff } 0 = \left[\beta(1-\sqrt{2}) + \frac{1}{\sqrt{2}}\right].$$

But

$$0 < \frac{2-\sqrt{}}{2} = 1 - \sqrt{2} + \frac{1}{\sqrt{2}} < \beta(1-\sqrt{2}) + \frac{1}{\sqrt{2}} \leqq \frac{1}{\sqrt{2}} < 1;$$

$\therefore$ (1′) holds. Hence,

$$\text{if} \quad m = \left[t\left(1 + \frac{1}{\sqrt{2}}\right)\right] = a_n \quad \text{then} \quad a_{n+1} = [t(\sqrt{2}+1)];$$

$$\text{if} \quad m = [t(\sqrt{2}+1)] = a_n \quad \text{then} \quad a_{n+1} = \left[2t\left(1 + \frac{1}{\sqrt{2}}\right)\right].$$

A minor induction argument on n now proves the theorem.

We point out that it is possible to express the conclusion of the theorem in a somewhat more concise form:

If $a_1 = m$ and $a_{n+1} = [\sqrt{2a_n(a_n+1)}]$ then

$$a_n = [\tau(2^{(n-1)/2} + 2^{(n-2)/2})], \qquad n > 1,$$

where τ is the mth smallest real number in the set $\{1, 2, 3. \cdots\} \cup \{\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}, \cdots\}$. The first few values are:

m	1	2	3	4	5	6	7	8	9	10	11	12	
τ	1	$\sqrt{2}$	2	$2\sqrt{2}$	3	4	$3\sqrt{2}$	5	$4\sqrt{2}$	6	7	$5\sqrt{2}$	$\cdots$.

The fact that for $m=1$, $a_{2n+1} - 2a_{2n-1}$ is the nth digit in the binary expansion of $\sqrt{2}$ is now immediate (as is the fact $a_{2n+1} - a_{2n} = 2^{n-1}$).

It would be interesting to know if similar results hold for sequences defined by

$$a_{n+1} = [\sqrt{3a_n(a_n+1)}], \qquad a_{n+1} = [\sqrt[3]{2a_n(a_n+1)(a_n+2)}], \text{ etc.}$$

References

1. I. Niven, Diophantine Approximations, Wiley, New York, 1963.

2. F. K. Hwang and S. Lin, An Analysis of Ford and Johnson's Sorting Algorithm, to appear in Proc. 3rd Annual Princeton Conference on Information Sciences and Systems.

GENERAL SUBTRACTION

JAMES W. PETTICREW, Indiana State University

There seems to be a reluctance in abstract algebra to consider partial operations, although in elementary mathematics, subtraction and division in the natural numbers are considered very early. It is true that some recent work in universal algebra has been extended to include partial operations, for example, see Pierce [1]. However, very few examples of algebraic systems with partial operations are given. A set equipped with subtraction is an elementary example of such a system and is understandable to students in a first course in abstract

But

$$0 < \frac{2-\sqrt{}}{2} = 1 - \sqrt{2} + \frac{1}{\sqrt{2}} < \beta(1-\sqrt{2}) + \frac{1}{\sqrt{2}} \leqq \frac{1}{\sqrt{2}} < 1;$$

$\therefore$ (1′) holds. Hence,

$$\text{if} \quad m = \left[t\left(1+\frac{1}{\sqrt{2}}\right)\right] = a_n \quad \text{then} \quad a_{n+1} = [t(\sqrt{2}+1)];$$

$$\text{if} \quad m = [t(\sqrt{2}+1)] = a_n \quad \text{then} \quad a_{n+1} = \left[2t\left(1+\frac{1}{\sqrt{2}}\right)\right].$$

A minor induction argument on n now proves the theorem.

We point out that it is possible to express the conclusion of the theorem in a somewhat more concise form:

If $a_1 = m$ and $a_{n+1} = [\sqrt{2a_n(a_n+1)}]$ then

$$a_n = [\tau(2^{(n-1)/2} + 2^{(n-2)/2})], \qquad n > 1,$$

where τ is the mth smallest real number in the set $\{1, 2, 3. \cdots\} \cup \{\sqrt{2}, 2\sqrt{2}, 3\sqrt{2}, \cdots\}$. The first few values are:

m	1	2	3	4	5	6	7	8	9	10	11	12
τ	1	$\sqrt{2}$	2	$2\sqrt{2}$	3	4	$3\sqrt{2}$	5	$4\sqrt{2}$	6	7	$5\sqrt{2}$

$\cdots$.

The fact that for $m=1$, $a_{2n+1} - 2a_{2n-1}$ is the nth digit in the binary expansion of $\sqrt{2}$ is now immediate (as is the fact $a_{2n+1} - a_{2n} = 2^{n-1}$).

It would be interesting to know if similar results hold for sequences defined by

$$a_{n+1} = [\sqrt{3a_n(a_n+1)}], \qquad a_{n+1} = [\sqrt[3]{2a_n(a_n+1)(a_n+2)}], \text{ etc.}$$

References

1. I. Niven, Diophantine Approximations, Wiley, New York, 1963.

2. F. K. Hwang and S. Lin, An Analysis of Ford and Johnson's Sorting Algorithm, to appear in Proc. 3rd Annual Princeton Conference on Information Sciences and Systems.

GENERAL SUBTRACTION

JAMES W. PETTICREW, Indiana State University

There seems to be a reluctance in abstract algebra to consider partial operations, although in elementary mathematics, subtraction and division in the natural numbers are considered very early. It is true that some recent work in universal algebra has been extended to include partial operations, for example, see Pierce [1]. However, very few examples of algebraic systems with partial operations are given. A set equipped with subtraction is an elementary example of such a system and is understandable to students in a first course in abstract algebra, as the axioms and the proofs are motivated by the properties of subtraction in the natural numbers. In this note we give a characterization of abelian semigroups with cancellation and identity in terms of a partial operation.

Groups have previously been characterized in terms of subtraction in [2]; however, in a group subtraction is a full operation. As a corollary we can characterize abelian groups by requiring that subtraction be an operation.

Let us suppose that $(A, +)$ is an abelian semigroup with cancellation and identity.

DEFINTION 1. *If* $a, b, c \in A$ *such that* $a+b=c$, *then define* $c-a=b$. *Note that the cancellation law implies that* $c-a$ *is well defined, if it exists.*

The following properties are easily verified in $(A, +)$.

1. If $a \in A$, then $a-a$ exists.
2. If $a, b, c \in A$ and $a-b=c$, then $a-c=b$.
3. If $a, b \in A$, then there exists $c \in A$ such that $c-a=b$.
4. If $a, b, c \in A$, $b-c$ exists, $a-(b-c)$ exists, and $b-a$ exists, then $a-(b-c) = c-(b-a)$.
5. If $a, b, c \in A$ and $a-c=b-c$, then $b-a$ exists.
6. If $a, b, c \in A$, $a-b$ exists, and $b-c$ exists; then $a-(b-c)$ exists.

Let us note that in the case of an abelian group conditions 1, 5, and 6 may be replaced by the stronger condition:

1′. "$-$" is a binary operation on A.

We now assume that we have a nonempty set A and a partial binary operation $-$ on A satisfying conditions 1 through 6.

THEOREM 1. *If* $a, b, c \in A$ *and* $a-c=a-b$, *then* $c=b$.

This is immediate from condition 2.

THEOREM 2. *If* $a, b \in A$, *then* $a-(b-b)=a$.

We first note that condition 2 implies that if $a, b \in A$ and $a-b$ exists, then $a-(a-b)=b$. By condition 3 we may assume the existence of $c \in A$ such that $c-a=b$ and by condition 2 $c-b=a$. Thus we have $c-(c-a)=a$. An application of conditions 1 and 4 now yields that $a-(c-c)=a$ and that $b-(c-c)=b$; from condition 2 we obtain that $a-a=c-c=b-b$.

DEFINITION 2. *If* $a \in A$ *define* $a-a=0$. *By the proof of Theorem* 2 *we see that* 0 *is well defined and that* $a-0=a$.

COROLLARY. *If* $a, b \in A$ *with* $a-b=0$ *then* $a=b$.

THEOREM 3. *If* $a, b \in A$, *then there exists a unique* $c \in A$ *such that* $c-a=b$.

The existence of such a c is given by condition 3. Uniqueness clearly will follow if we show that $a-c=b-c$ implies that $a=b$. By conditions 2, 5 and 4 we have that $c=a-(b-c)=c-(b-a)$. Thus by the remarks following Definition 2 and the corollary we have that $b-a=0$ and that $a=b$.

DEFINITION 3. *If* $a, b \in A$ *define* $a+b=c$ *such that* $c-a=b$.

The following properties are readily verified:
If a, $b \in A$, then $a+b=b+a$.
If $a \in A$, then $a+0=a$.
If a, b, $c \in A$ and $a+b=c+b$, then $a=c$.
If a, $b \in A$, then $(a+b)=b-a$.

THEOREM 4. *If* a, b, $c \in A$ *and* $(a-b)-c$ *exists, then* $(a-b)-c=(a-c)-b$.

We first note from the proof of Theorem 2 that $a-(a-b)=b$ and that $a-(a-c)=c$ if both $a-b$ and $a-c$ exist. From this we see that if $a-b$ and $a-c$ exist, then $(a-b)-c=(a-b)-(a-(a-c))=(a-c)-b$ by condition 4. Now $a-b$ exists by hypothesis, whence it remains to show that $a-c$ exists. Let $(a-b)-c=e$, then $(a-b)-e=c$ by condition 2; however $a-((a-b)-e)$ exists by condition 6, when ce $a-c$ exists.

THEOREM 5. *If* a, b, $c \in A$ *then* $a+(b+c)=(a+b)+c$.

Let $(a+b)+c=d$, then by Definition 3 and the commutative law $d-c=a+b$. However, $(a+b)-a=b$, hence $(d-c)-a=b$. From Theorem 4 we obtain $b=(d-c)-a=(d-a)-c$. A double application of Definition 3 now yields that $d-a=b+c$ and that $d=(b+c)+a$.

From the remarks following Definition 3 and from Theorem 5, we see that the system $(A, +, 0)$ derived from the system $(A, -)$ is an abelian semigroup with cancellation and identity. It is clear that the procedure of going from $+$ to $-$ to $+'$ or vice versa yields the original. It is also readily verified that 1′, 2, 3, and 4 are necessary and sufficient for $(A, -)$ to lead to an abelian group. Merely note that since $a+(b-a)=b$, $b-a$ is a solution to $a+x=b$.

The author wishes to acknowledge the advice and encouragement of Dr. L. R. Wilcox in the preparation of this paper.

References

1. R. S. Pierce, Introduction to the Theory of Abstract Algebra, Holt, Rinehart and Winston, 1968.
2. B. A. Bernstein, Groups and abelian groups in terms of negative addition and subtraction, Duke Math. J., (1939) 871–874.

A NOTE ON GENERALIZED SEMILINEAR TRANSFORMATIONS

W. G. DOTSON, JR., North Carolina State University

By a *linear system* we mean an ordered pair (E, F) where E is a vector space over the field F. If each of (E, F) and (E', F') is a linear system, then a *generalized semilinear transformation* from (E, F) to (E', F') is an ordered pair (T, S) such that $T: E \to E'$ and $S: F \to F'$ are mappings which satisfy

$$T(ax + by) = Sa \cdot Tx + Sb \cdot Ty$$

for all a, b in F and all x, y in E.

The following properties are readily verified:
If $a, b \in A$, then $a+b=b+a$.
If $a \in A$, then $a+0=a$.
If $a, b, c \in A$ and $a+b=c+b$, then $a=c$.
If $a, b \in A$, then $(a+b)=b-a$.

THEOREM 4. *If $a, b, c \in A$ and $(a-b)-c$ exists, then $(a-b)-c=(a-c)-b$.*

We first note from the proof of Theorem 2 that $a-(a-b)=b$ and that $a-(a-c)=c$ if both $a-b$ and $a-c$ exist. From this we see that if $a-b$ and $a-c$ exist, then $(a-b)-c=(a-b)-(a-(a-c))=(a-c)-b$ by condition 4. Now $a-b$ exists by hypothesis, whence it remains to show that $a-c$ exists. Let $(a-b)-c=e$, then $(a-b)-e=c$ by condition 2; however $a-((a-b)-e)$ exists by condition 6, when ce $a-c$ exists.

THEOREM 5. *If $a, b, c \in A$ then $a+(b+c)=(a+b)+c$.*

Let $(a+b)+c=d$, then by Definition 3 and the commutative law $d-c=a+b$. However, $(a+b)-a=b$, hence $(d-c)-a=b$. From Theorem 4 we obtain $b=(d-c)-a=(d-a)-c$. A double application of Definition 3 now yields that $d-a=b+c$ and that $d=(b+c)+a$.

From the remarks following Definition 3 and from Theorem 5, we see that the system $(A, +, 0)$ derived from the system $(A, -)$ is an abelian semigroup with cancellation and identity. It is clear that the procedure of going from $+$ to $-$ to $+'$ or vice versa yields the original. It is also readily verified that 1′, 2, 3, and 4 are necessary and sufficient for $(A, -)$ to lead to an abelian group. Merely note that since $a+(b-a)=b$, $b-a$ is a solution to $a+x=b$.

The author wishes to acknowledge the advice and encouragement of Dr. L. R. Wilcox in the preparation of this paper.

References

1. R. S. Pierce, Introduction to the Theory of Abstract Algebra, Holt, Rinehart and Winston, 1968.
2. B. A. Bernstein, Groups and abelian groups in terms of negative addition and subtraction, Duke Math. J., (1939) 871–874.

A NOTE ON GENERALIZED SEMILINEAR TRANSFORMATIONS

W. G. DOTSON, JR., North Carolina State University

By a *linear system* we mean an ordered pair (E, F) where E is a vector space over the field F. If each of (E, F) and (E', F') is a linear system, then a *generalized semilinear transformation* from (E, F) to (E', F') is an ordered pair (T, S) such that $T: E \to E'$ and $S: F \to F'$ are mappings which satisfy

$$T(ax + by) = Sa \cdot Tx + Sb \cdot Ty$$

for all a, b in F and all x, y in E.

In case $E'=E$, $F'=F$, and S is an automorphism of F, the generalized semilinear transformation (T, S) reduces to the *semilinear transformation* T *with automorphism* S, which was first studied by C. Segre (see Jacobson [1] p. 26).

THEOREM 1. *Suppose* (T, S) *is a generalized semilinear transformation from the linear system* (E, F) *to the linear system* (E', F'). *Then either*: (1) $Tx=\theta$ *for all* $x\in E$, *and* $S: F\to F'$ *is arbitrary, or* (2) *there is a* $v\in E'(v\neq\theta)$ *such that* $Tx=v$ *for all* $x\in E$, *and* $S: F\to F'$ *is given by* $Sa=1/2$ *for all* $a\in F$ (*implying that* F' *does not have characteristic* $=2$), *or* (3) $T: E\to E'$ *is a group homomorphism, and* $S: F\to F'$ *is a field monomorphism.*

Proof. We have $ax=ax+0y=ax+b\theta$ for all a, b in F and all x, y in E. Hence

$$T(ax) = S(a)Tx + S(0)Ty = S(a)Tx + S(b)T\theta,$$

and so $S(0)Ty=S(b)T\theta$ for all $b\in F$ and all $y\in E$. If $S(0)\neq 0$, then (with $b=0$) we get $Ty=T\theta=v\in E'$ for all $y\in E$. Of course, if $v=T\theta=\theta$ then (1) holds. If $v\neq\theta$ then

$$v = T(ay + ay) = S(a)Ty + S(a)Ty = 2S(a)v$$

implies $S(a)=1/2$ for all $a\in F$, so that (2) holds. Therefore, if neither (1) nor (2) holds then we must have $S(0)=0$; and so

$$T(ax) = T(ax + 0y) = S(a)Tx$$

for all $a\in F$ and all $x\in E$. Choose $x\in E$ such that $Tx\neq\theta$ (still under the assumption that neither (1) nor (2) holds). Then $ab\cdot x=a\cdot(bx)$ implies $S(ab)=SaSb$ [i.e., $S(ab)Tx=T(ab\cdot x)=T(a\cdot(bx))=SaT(bx)=(SaSb)Tx$, and $Tx\neq\theta$]. Similarly, $(a+b)x=ax+bx$ implies $S(a+b)=Sa+Sb$. Hence S is a field homomorphism. It follows that $S1=1$, and so

$$T(x + y) = T(1x + 1y) = 1Tx + 1Ty = Tx + Ty$$

for all x, y in E. Thus T is a group homomorphism. To see that S must actually be a monomorphism (still under the assumption that neither (1) nor (2) holds), observe that if $Sa=0$ with $a\neq 0$ then $Tx=T(a^{-1}ax)=(Sa^{-1})(0)Tx=\theta$ for all $x\in E$.

In the following example, $T: E\to E'$ is a group isomorphism, $S: F\to F'$ is a field isomorphism, but (T, S) is *not* a generalized semilinear transformation from (E, F) to (E', F'). Let $E=E'=$ the real numbers with ordinary addition, and let $T: E\to E'$ be the identity mapping on E. Let $F=F'=\{a+b\sqrt{2}: a, b \text{ rational real numbers}\}$ with ordinary addition and multiplication. The scalar product function on $F\times E$ to E, and on $F'\times E'$ to E', is given by $(a+b\sqrt{2})\cdot x=(a+b\sqrt{2})x$ —ordinary multiplication of real numbers. Let $S: F\to F'$ be given by $S(a+b\sqrt{2})=a-b\sqrt{2}$. If (T, S) were a generalized semilinear transformation, then we would have $T((a+b\sqrt{2})x)=S(a+b\sqrt{2})Tx=(a-b\sqrt{2})x$ for all rational a, b and real x; and so with $a=b=x=1$ we would get $T(1+\sqrt{2})=1-\sqrt{2}$, which would contradict $Tx=x$ for all real x. Nevertheless, one can obtain a sort of partial converse of Theorem 1.

THEOREM 2. *Suppose that* (E, F) *and* (E', F') *are linear systems, and let* R *denote the prime subfield of* F. *If* $T:E\to E'$ *is a group homomorphism and* $S: F\to F'$ *is a field monomorphism, then* $(T, S|R)$ *is a generalized semilinear transformation from* (E, R) *to* (E', F').

Proof. We have $T\theta=\theta$, $S(0)=0$, $S(1)=1$. Hence for $x\in E$, $T(0x)=T\theta=\theta=0\cdot Tx=S(0)Tx$ and $T(1x)=Tx=1\cdot Tx=S(1)Tx$. For $n=1+1+\cdots+1$ (n times) we have $S(n)=nS(1)=n$ and so

$$T(nx) = T(x+x+\cdots+x) = nTx = S(n)Tx.$$

Also, $T((-n)x)=T(-(nx))=-T(nx)=-S(n)Tx=S(-n)Tx$. If n is an integer in R, $n\neq 0$, and $x\in E$ then $Tx=T(1x)=T(nn^{-1}x)=S(n)T(n^{-1}x)$; whence $T(n^{-1}x)=S(n)^{-1}Tx=S(n^{-1})Tx$ [note that $n\neq 0$ implies $S(n)\neq 0$ since S is a monomorphism]. If $p=mn^{-1}$ is any element of R and $x\in E$ then

$$T(px) = T(mn^{-1}x) = S(m)T(n^{-1}x) = S(m)S(n^{-1})Tx = S(p)Tx.$$

Finally, if $p, q\in R$ and $x, y\in E$ then we have

$$T(px+qy) = T(px)+T(qy) = S(p)Tx+S(q)Ty.$$

Reference

1. N. Jacobson, The theory of rings, Math. Surveys No. 2, Amer. Math. Soc., New York, 1943.

RELATION BETWEEN CORRELATION AND ECCENTRICITY

W. E. BLEICK, Naval Postgraduate School, Monterey, California

The nondimensional bivariate normal frequency function is

$$f(x, y) = \frac{(1-\rho^2)^{-1/2}}{2\pi}\exp[-(x^2-2\rho xy+y^2)/2(1-\rho^2)] \tag{1}$$

(see [1]). The x and y variables in (1) are the nondimensional deviations from the mean measured in units of the corresponding σ_x and σ_y standard deviations, and ρ is the correlation coefficient between x and y. The double integral of $f(x, y)$ over the $0xy$ plane is unity. It is of pedagogical interest to relate ρ to the eccentricity e of the elliptical cross sections of the $f(x, y)$ surface. If the ellipse

$$x^2-2\rho xy+y^2 = 2(1-\rho^2)k^2 \tag{2}$$

becomes

$$(\xi/\alpha)^2+(\eta/\beta)^2 = k^2 \tag{3}$$

when referred to its principal axes, the invariants are

$$(1/\alpha)^2+(1/\beta)^2 = 1/(1-\rho^2) \tag{4}$$

and

$$1/\alpha^2\beta^2 = 1/4(1-\rho^2). \tag{5}$$

THEOREM 2. *Suppose that* (E, F) *and* (E', F') *are linear systems, and let* R *denote the prime subfield of* F. *If* $T: E \to E'$ *is a group homomorphism and* $S: F \to F'$ *is a field monomorphism, then* $(T, S|R)$ *is a generalized semilinear transformation from* (E, R) *to* (E', F').

Proof. We have $T\theta = \theta$, $S(0) = 0$, $S(1) = 1$. Hence for $x \in E$, $T(0x) = T\theta = \theta = 0 \cdot Tx = S(0)Tx$ and $T(1x) = Tx = 1 \cdot Tx = S(1)Tx$. For $n = 1 + 1 + \cdots + 1$ (n times) we have $S(n) = nS(1) = n$ and so

$$T(nx) = T(x + x + \cdots + x) = nTx = S(n)Tx.$$

Also, $T((-n)x) = T(-(nx)) = -T(nx) = -S(n)Tx = S(-n)Tx$. If n is an integer in R, $n \neq 0$, and $x \in E$ then $Tx = T(1x) = T(nn^{-1}x) = S(n)T(n^{-1}x)$; whence $T(n^{-1}x) = S(n)^{-1}Tx = S(n^{-1})Tx$ [note that $n \neq 0$ implies $S(n) \neq 0$ since S is a monomorphism]. If $p = mn^{-1}$ is any element of R and $x \in E$ then

$$T(px) = T(mn^{-1}x) = S(m)T(n^{-1}x) = S(m)S(n^{-1})Tx = S(p)Tx.$$

Finally, if $p, q \in R$ and $x, y \in E$ then we have

$$T(px + qy) = T(px) + T(qy) = S(p)Tx + S(q)Ty.$$

Reference

1. N. Jacobson, The theory of rings, Math. Surveys No. 2, Amer. Math. Soc., New York, 1943.

RELATION BETWEEN CORRELATION AND ECCENTRICITY

W. E. BLEICK, Naval Postgraduate School, Monterey, California

The nondimensional bivariate normal frequency function is

$$f(x, y) = \frac{(1 - \rho^2)^{-1/2}}{2\pi} \exp[-(x^2 - 2\rho xy + y^2)/2(1 - \rho^2)] \tag{1}$$

(see [1]). The x and y variables in (1) are the nondimensional deviations from the mean measured in units of the corresponding σ_x and σ_y standard deviations, and ρ is the correlation coefficient between x and y. The double integral of $f(x, y)$ over the $0xy$ plane is unity. It is of pedagogical interest to relate ρ to the eccentricity e of the elliptical cross sections of the $f(x, y)$ surface. If the ellipse

$$x^2 - 2\rho xy + y^2 = 2(1 - \rho^2)k^2 \tag{2}$$

becomes

$$(\xi/\alpha)^2 + (\eta/\beta)^2 = k^2 \tag{3}$$

when referred to its principal axes, the invariants are

$$(1/\alpha)^2 + (1/\beta)^2 = 1/(1 - \rho^2) \tag{4}$$

and

$$1/\alpha^2\beta^2 = 1/4(1 - \rho^2). \tag{5}$$

The reciprocal square root of (5) multiplied by (4) yields

$$(\beta/\alpha) + (\alpha/\beta) = 2/(1 - \rho^2)^{1/2}. \tag{6}$$

The squaring of (6) yields

$$\rho = \pm e^2/(2 - e^2) \tag{7}$$

where use is made of $\beta^2/\alpha^2 = 1 - e^2$ for $\beta \leqq \alpha$. Equation (7) may help a student to understand the roughly elliptical shape of the swarm of points in a non-dimensional x, y scatter diagram in terms of correlation between x and y.

Reference

1. P. G. Hoel, Introduction to Mathematical Statistics, 3rd ed., Wiley, New York, 1962, p. 198.

A NOTE ON THE VECTOR PRODUCT $a \times (b \times c)$

C. A. GRIMM, South Dakota School of Mines and Technology

The product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ is at best tedious to derive. The intent of this note is to make the computation reasonably palatable. The usual attack is to argue that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is parallel to the plane of $\mathbf{b}$ and $\mathbf{c}$ and hence

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = M\mathbf{b} + N\mathbf{c}. \tag{1}$$

One dot multiplies (1) by $\mathbf{a}$ to get

$$0 = M\mathbf{a} \cdot \mathbf{b} + N\mathbf{a} \cdot \mathbf{c}. \tag{2}$$

From (2) it is easy to argue that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \lambda[(\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}] \tag{3}$$

in which λ may depend on $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. With $\mathbf{b} = b_1 i + b_2 j + b_3 k$ and $\mathbf{c} = c_1 i + c_2 j + c_3 k$ we have

$$\mathbf{b} \times \mathbf{c} = (b_2 c_3 - b_3 c_2) i + (b_3 c_1 - b_1 c_3) j + (b_1 c_2 - b_2 c_1) k. \tag{4}$$

From (3) we have

$$\begin{aligned} i \times (\mathbf{b} \times \mathbf{c}) &= \lambda_1[(i \cdot \mathbf{c})\mathbf{b} - (i \cdot \mathbf{b})\mathbf{c}] \\ j \times (\mathbf{b} \times \mathbf{c}) &= \lambda_2[(j \cdot \mathbf{c})\mathbf{b} - (j \cdot \mathbf{b})\mathbf{c}] \\ k \times (\mathbf{b} \times \mathbf{c}) &= \lambda_3[(k \cdot \mathbf{c})\mathbf{b} - (k \cdot \mathbf{b})\mathbf{c}]. \end{aligned} \tag{5}$$

We dot multiply these three equations on the left by k, i and j respectively, interchange the dot and first cross, to obtain

$$\begin{aligned} j \cdot (\mathbf{b} \times \mathbf{c}) &= \lambda_1[c_1 b_3 - b_1 c_3] \\ k \cdot (\mathbf{b} \times \mathbf{c}) &= \lambda_2[c_2 b_1 - b_2 c_1] \\ i \cdot (\mathbf{b} \times \mathbf{c}) &= \lambda_3[c_3 b_2 - b_3 c_2], \end{aligned}$$

from which by (4) $\lambda_1 = \lambda_2 = \lambda_3 = 1$. We now return to (5) with all λ's equal to 1 and with $\mathbf{a} = a_1 i + a_2 j + a_3 k$. Multiply the three equations in (5) by a_1, a_2 and a_3 respectively and add to obtain $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

The reciprocal square root of (5) multiplied by (4) yields

$$(\beta/\alpha) + (\alpha/\beta) = 2/(1 - \rho^2)^{1/2}. \tag{6}$$

The squaring of (6) yields

$$\rho = \pm e^2/(2 - e^2) \tag{7}$$

where use is made of $\beta^2/\alpha^2 = 1 - e^2$ for $\beta \leqq \alpha$. Equation (7) may help a student to understand the roughly elliptical shape of the swarm of points in a non-dimensional x, y scatter diagram in terms of correlation between x and y.

Reference

1. P. G. Hoel, Introduction to Mathematical Statistics, 3rd ed., Wiley, New York, 1962, p. 198.

A NOTE ON THE VECTOR PRODUCT $a\times(b\times c)$

C. A. GRIMM, South Dakota School of Mines and Technology

The product $a\times(b\times c) = (a\cdot c)b - (a\cdot b)c$ is at best tedious to derive. The intent of this note is to make the computation reasonably palatable. The usual attack is to argue that $a\times(b\times c)$ is parallel to the plane of b and c and hence

$$a \times (b \times c) = Mb + Nc. \tag{1}$$

One dot multiplies (1) by a to get

$$0 = Ma\cdot b + Na\cdot c. \tag{2}$$

From (2) it is easy to argue that

$$a \times (b \times c) = \lambda[(a\cdot c)b - (a\cdot b)c] \tag{3}$$

in which λ may depend on a, b and c. With $b = b_1i + b_2j + b_3k$ and $c = c_1i + c_2j + c_3k$ we have

$$b \times c = (b_2c_3 - b_3c_2)i + (b_3c_1 - b_1c_3)j + (b_1c_2 - b_2c_1)k. \tag{4}$$

From (3) we have

$$\begin{aligned} i \times (b \times c) &= \lambda_1[(i\cdot c)b - (i\cdot b)c] \\ j \times (b \times c) &= \lambda_2[(j\cdot c)b - (j\cdot b)c] \\ k \times (b \times c) &= \lambda_3[(k\cdot c)b - (k\cdot b)c]. \end{aligned} \tag{5}$$

We dot multiply these three equations on the left by k, i and j respectively, interchange the dot and first cross, to obtain

$$\begin{aligned} j\cdot(b \times c) &= \lambda_1[c_1b_3 - b_1c_3] \\ k\cdot(b \times c) &= \lambda_2[c_2b_1 - b_2c_1] \\ i\cdot(b \times c) &= \lambda_3[c_3b_2 - b_3c_2], \end{aligned}$$

from which by (4) $\lambda_1 = \lambda_2 = \lambda_3 = 1$. We now return to (5) with all λ's equal to 1 and with $a = a_1i + a_2j + a_3k$. Multiply the three equations in (5) by a_1, a_2 and a_3 respectively and add to obtain $a\times(b\times c) = (a\cdot c)b - (a\cdot b)c$.

ON A LAPLACE INTEGRAL

HIROSHI HARUKI, University of Waterloo, Canada

We consider the Laplace integral

$$(1) \qquad f(a) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{\cos ax}{x^2+1}\, dx (a \geqq 0).$$

The usual methods of evaluating this integral ($f(a) = \exp(-a)$) are the following three:

(A) The method of deducing the differential equation $f''(a) = f(a)$ and using initial conditions $f(0) = 1, f'(0) = -1$. (See [6].)

(B) The method of using analytic function theory. (See [5].)

(C) The method of using a Fourier's integral theorem for $\exp(-|x|)$. (See [3].)

Now we shall prove that $f(a) = \exp(-a)$ from the standpoint of functional equation theory.

The function $\exp(kx)$ where k is a real constant and x is a real variable can be characterized by means of the functional equation in two real variables:

$$(2) \qquad \phi(x+y) = \phi(x)\phi(y),$$

where $\phi(\neq 0)$ is a real-valued function of a real variable x on $-\infty < x < +\infty$ and is continuous at one point (or can be majorized by a measurable function on a set of positive measure). (See [1].)

In this note, instead of using (2), we shall use the following functional equation in a single real variable:

$$(3) \qquad \phi(2x) = \phi(x)^2.$$

We shall apply the following lemma:

LEMMA. *The function* $\phi(x) = \exp(kx)$ *where* k *is a given constant is the unique function which is of class* C^1 *on* $0 \leqq x < +\infty$, *satisfies* (3) *on* $0 \leqq x < +\infty$ *and fulfils the conditions* $\phi(0) = 1, \phi'(0) = k$.

Proof. See [4, 2].

We may now prove that $f(a) = \exp(-a)$. It is clear that f is of class C^1 on $0 \leqq a < +\infty$ and $f(0) = 1, f'(0) = -1$. (See [6].) By (1) we have

$$(4) \qquad f(a)^2 = \frac{1}{\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{\cos ax \cos ay}{(x^2+1)(y^2+1)}\, dxdy,$$

where a is an arbitrary nonnegative real number. By (4) we have

$$(5) \qquad \begin{aligned} f(a)^2 &= \frac{1}{2\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{\cos a(x+y)}{(x^2+1)(y^2+1)}\, dxdy \\ &+ \frac{1}{2\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{\cos a(x-y)}{(x^2+1)(y^2+1)}\, dxdy \\ &= \frac{1}{2\pi^2}(I_1 + I_2). \end{aligned}$$

Putting $u=x+y$, $v=y$ in I_1, we have (the Jacobian $=1$)

$$I_1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\cos au}{((u-v)^2+1)(v^2+1)}\,dudv \tag{6}$$
$$=\int_{-\infty}^{+\infty}\cos au\left(\int_{-\infty}^{+\infty}\frac{1}{((u-v)^2+1)(v^2+1)}\,dv\right)du.$$

Using the method of partial fractions, we have

$$\int_{-\infty}^{+\infty}\frac{1}{((u-v)^2+1)(v^2+1)}\,dv=\frac{2}{u^2+4}\pi. \tag{7}$$

Putting $u=x-y$, $v=y$ in I_2, we have (the Jacobian $=1$)

$$I_2=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{\cos au}{((u+v)^2+1)(v^2+1)}\,dudv \tag{8}$$
$$=\int_{-\infty}^{+\infty}\cos au\left(\int_{-\infty}^{+\infty}\frac{1}{((u+v)^2+1)(v^2+1)}\,dv\right)du.$$

Putting $w=-v$ in (8) and using (7), we have

$$\int_{-\infty}^{+\infty}\frac{1}{((u+v)^2+1)(v^2+1)}\,dv=\frac{2}{u^2+4}\pi. \tag{9}$$

By (5), (6), (7), (8), (9) we have

$$f(a)^2=\frac{2}{\pi}\int_{-\infty}^{+\infty}\frac{\cos au}{u^2+4}\,du. \tag{10}$$

By (1) we have

$$f(2a)=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{\cos 2ax}{x^2+1}\,dx. \tag{11}$$

Putting $x=u/2$ in (11), we have

$$f(2a)=\frac{2}{\pi}\int_{-\infty}^{+\infty}\frac{\cos au}{u^2+4}\,du. \tag{12}$$

By (10), (12) we have

$$f(2a)=f(a)^2 \qquad (a\geqq 0). \tag{13}$$

Since f is of class C^1 on $0\leqq a<+\infty$, satisfies (13) on $0\leqq a<+\infty$ and fulfils the conditions $f(0)=1$, $f'(0)=-1$, the above lemma gives ($k=-1$)

$$f(a)=\exp(-a)\ (a\geqq 0).$$

References

1. J. Aczél, Lectures on Functional Equations and Their Applications, Academic Press, New York and London, 1966, pp. 38–39.

2. J. Aczél and H. Kiesewetter, Über die Reduktion der Stufe bei einer Klasse von Funktionalgleichungen, Publ. Math. Debrecen, 8 (1958) 348–363.
3. J. Indritz, Methods in Analysis, Macmillan, New York, 1963, pp. 446–447, p. 451.
4. M. Kuczma, A characterization of the exponential and logarithmic functions by functional equations, Fund. Math., 52 (1963) 283–288.
5. Z. Nehari, Introduction to Complex Analysis, Allyn and Bacon, Boston, 1961, p. 129.
6. J. M. H. Olmsted, Advanced Calculus, Appleton-Century-Crofts, New York, 1961, p. 479.

AN EXERCISE IN VECTOR IDENTITIES

CLARK GIVENS, Michigan Technological University

The problem of finding the distance between two skew lines, which is usually discussed in a course in calculus or vector analysis and is solved by elementary vector methods, is solved here as a problem in the minimization of a function of two variables and thereby leads to an interesting exercise in vector identities. Let L_1 and L_2 be lines in space whose vector equations are

$$L_1\colon\ R_1(\alpha) = U_1 + \alpha V$$
$$L_2\colon\ R_2(\beta) = U_2 + \beta W$$

with V and W not parallel so that L_1 and L_2 are in fact skew lines. Then the vector $D = D(\alpha, \beta)$ between points on L_1 and L_2 corresponding to the parameter values α and β is given by

$$D = R_1 - R_2 = \alpha V - \beta W + U_1 - U_2 = \alpha V - \beta W + U$$

and the square of the distance between any two such points is

$$\rho^2 = (D, D) = (V, V)\alpha^2 + (W, W)\beta^2 - 2(V, W)\alpha\beta + 2(U, V)\alpha - 2(U, W)\beta + (U, U).$$

Setting the first partial derivatives of ρ^2 with respect to α and β equal to zero, as the necessary condition for a minimum, we obtain

$$(V, V)\alpha - (V, W)\beta = -(U, V)$$
$$-(V, W)\alpha + (W, W)\beta = (U, W).$$

Solutions of this system exist for all right hand sides, that is, for all possible positions and orientations of the *skew* lines L_1 and L_2 only if the determinant of the system, Δ, is nonvanishing. But

$$\Delta = \begin{vmatrix} (V, V) & -(V, W) \\ -(V, W) & (W, W) \end{vmatrix} = (V, V)(W, W) - (V, W)^2 = (V \times W, V \times W).$$

Since V and W are not parallel by hypothesis, we have that $V \times W \neq 0$ and thus $\Delta \neq 0$. Solution for α and β then yields

$$\alpha_{\min} = [(V, W)(U, W) - (U, V)(W, W)]/\Delta = (V \times W, W \times U)/\Delta = (N, W \times U)/\sqrt{\Delta}$$
$$\beta_{\min} = [(V, V)(U, W) - (V, U)(V, W)]/\Delta = (V \times W, V \times U)/\Delta = (N, V \times U)/\sqrt{\Delta}$$

with $N = V \times W/\sqrt{\Delta}$. Therefore,

2. J. Aczél and H. Kiesewetter, Über die Reduktion der Stufe bei einer Klasse von Funktionalgleichungen, Publ. Math. Debrecen, 8 (1958) 348–363.
3. J. Indritz, Methods in Analysis, Macmillan, New York, 1963, pp. 446–447, p. 451.
4. M. Kuczma, A characterization of the exponential and logarithmic functions by functional equations, Fund. Math., 52 (1963) 283–288.
5. Z. Nehari, Introduction to Complex Analysis, Allyn and Bacon, Boston, 1961, p. 129.
6. J. M. H. Olmsted, Advanced Calculus, Appleton-Century-Crofts, New York, 1961, p. 479.

AN EXERCISE IN VECTOR IDENTITIES

CLARK GIVENS, Michigan Technological University

The problem of finding the distance between two skew lines, which is usually discussed in a course in calculus or vector analysis and is solved by elementary vector methods, is solved here as a problem in the minimization of a function of two variables and thereby leads to an interesting exercise in vector identities. Let L_1 and L_2 be lines in space whose vector equations are

$$L_1:\ R_1(\alpha) = U_1 + \alpha V$$
$$L_2:\ R_2(\beta) = U_2 + \beta W$$

with V and W not parallel so that L_1 and L_2 are in fact skew lines. Then the vector $D = D(\alpha, \beta)$ between points on L_1 and L_2 corresponding to the parameter values α and β is given by

$$D = R_1 - R_2 = \alpha V - \beta W + U_1 - U_2 = \alpha V - \beta W + U$$

and the square of the distance between any two such points is

$$\rho^2 = (D, D) = (V, V)\alpha^2 + (W, W)\beta^2 - 2(V, W)\alpha\beta + 2(U, V)\alpha - 2(U, W)\beta + (U, U).$$

Setting the first partial derivatives of ρ^2 with respect to α and β equal to zero, as the necessary condition for a minimum, we obtain

$$(V, V)\alpha - (V, W)\beta = -(U, V)$$
$$-(V, W)\alpha + (W, W)\beta = (U, W).$$

Solutions of this system exist for all right hand sides, that is, for all possible positions and orientations of the *skew* lines L_1 and L_2 only if the determinant of the system, Δ, is nonvanishing. But

$$\Delta = \begin{vmatrix} (V, V) & -(V, W) \\ -(V, W) & (W, W) \end{vmatrix} = (V, V)(W, W) - (V, W)^2 = (V \times W, V \times W).$$

Since V and W are not parallel by hypothesis, we have that $V \times W \neq 0$ and thus $\Delta \neq 0$. Solution for α and β then yields

$$\alpha_{\min} = [(V, W)(U, W) - (U, V)(W, W)]/\Delta = (V \times W, W \times U)/\Delta = (N, W \times U)/\sqrt{\Delta}$$
$$\beta_{\min} = [(V, V)(U, W) - (V, U)(V, W)]/\Delta = (V \times W, V \times U)/\Delta = (N, V \times U)/\sqrt{\Delta}$$

with $N = V \times W/\sqrt{\Delta}$. Therefore,

$$\begin{aligned} D_{\min} &= [(N, W \times U)/\sqrt{\Delta}]V - [(N, V \times U)/\sqrt{\Delta}]W + U \\ &= [(U \times N, W)/\sqrt{\Delta}]V - [(U \times N, V)/\sqrt{\Delta}]W + U \\ &= [(U \times N) \times (V \times W)]/\sqrt{\Delta} + U \\ &= N \times (N \times U) + U = (N, U)N - (N, N)U + U = (N, U)N. \end{aligned}$$

But,

$$U = D(\alpha, \beta) - \alpha V - \beta W \qquad \text{for any values of } \alpha, \beta$$

and therefore

$$D_{\min} = (D, N)N$$

Thus we obtain the well-known result that the minimum distance between two skew lines is $\rho_{\min} = |(D, N)|$, the absolute value of the dot product of any vector connecting the two lines with their common unit normal. Letting a, c, b denote the second partial derivatives of ρ^2 with respect to α, β, α and β, respectively, we obtain

$$a = 2(V, V) \qquad b = -2(V, W) \qquad c = 2(W, W).$$

Thus,

$$b^2 - ac = 4(V, W)^2 - 4(V, V)(W, W) = -4(V \times W, V \times W)$$

and the criteria for a mimimum, namely

$$a > 0 \quad \text{and} \quad b^2 - ac < 0$$

are satisfied.

CHOREOGRAPHIC PROOF OF A THEOREM ON PERMUTATIONS

F. CUNNINGHAM, JR., Bryn Mawr College

Let S be a finite set. A permutation $\tau: S \to S$ is called a *transposition* if there are two elements $s_1 \neq s_2$ of S which are interchanged by τ while all other elements of S stay still. In other words, $\tau(s_1) = s_2$, $\tau(s_2) = s_1$, and $\tau(s) = s$ for all $s \neq s_1$, s_2. Every permutation π of S can be factored as a finite succession of transpositions, but by no means uniquely. An important fact, which arises for instance in the theory of determinants, is that the parity (oddness or evenness) of the number of transpositions in any such factorization of π is always the same, depending only on π.

Applying this fact with π the identity permutation gives in particular that any succession of transpositions which in the end returns every s in S to its starting point must be of even length. This is because one such factorization of the identity is the trivial one of length 0, and 0 is even. Conversely, this special case implies the theorem. Indeed, suppose that $\sigma_1 \cdots \sigma_m = \pi = \tau_1 \cdots \tau_n$, where $\sigma_1, \cdots, \sigma_m$ and $\tau_1, \cdots, \tau_n$ are transpositions. Then $\sigma_1 \cdots \sigma_m \tau_n \cdots \tau_1$ is the

$$\begin{aligned} D_{\min} &= [(N, W \times U)/\sqrt{\Delta}]V - [(N, V \times U)/\sqrt{\Delta}]W + U \\ &= [(U \times N, W)/\sqrt{\Delta}]V - [(U \times N, V)/\sqrt{\Delta}]W + U \\ &= [(U \times N) \times (V \times W)]/\sqrt{\Delta} + U \\ &= N \times (N \times U) + U = (N, U)N - (N, N)U + U = (N, U)N. \end{aligned}$$

But,

$$U = D(\alpha, \beta) - \alpha V - \beta W \qquad \text{for any values of } \alpha, \beta$$

and therefore

$$D_{\min} = (D, N)N$$

Thus we obtain the well-known result that the minimum distance between two skew lines is $\rho_{\min} = |(D, N)|$, the absolute value of the dot product of any vector connecting the two lines with their common unit normal. Letting a, c, b denote the second partial derivatives of ρ^2 with respect to α, β, α and β, respectively, we obtain

$$a = 2(V, V) \qquad b = -2(V, W) \qquad c = 2(W, W).$$

Thus,

$$b^2 - ac = 4(V, W)^2 - 4(V, V)(W, W) = -4(V \times W, V \times W)$$

and the criteria for a mimimum, namely

$$a > 0 \quad \text{and} \quad b^2 - ac < 0$$

are satisfied.

CHOREOGRAPHIC PROOF OF A THEOREM ON PERMUTATIONS

F. CUNNINGHAM, JR., Bryn Mawr College

Let S be a finite set. A permutation $\tau: S \to S$ is called a *transposition* if there are two elements $s_1 \neq s_2$ of S which are interchanged by τ while all other elements of S stay still. In other words, $\tau(s_1) = s_2$, $\tau(s_2) = s_1$, and $\tau(s) = s$ for all $s \neq s_1, s_2$. Every permutation π of S can be factored as a finite succession of transpositions, but by no means uniquely. An important fact, which arises for instance in the theory of determinants, is that the parity (oddness or evenness) of the number of transpositions in any such factorization of π is always the same, depending only on π.

Applying this fact with π the identity permutation gives in particular that any succession of transpositions which in the end returns every s in S to its starting point must be of even length. This is because one such factorization of the identity is the trivial one of length 0, and 0 is even. Conversely, this special case implies the theorem. Indeed, suppose that $\sigma_1 \cdots \sigma_m = \pi = \tau_1 \cdots \tau_n$, where $\sigma_1, \cdots, \sigma_m$ and $\tau_1, \cdots, \tau_n$ are transpositions. Then $\sigma_1 \cdots \sigma_m \tau_n \cdots \tau_1$ is the identity (transpositions being their own inverses), so that $m+n$ is even. This implies that m and n are both odd or both even.

This theorem is usually proved by arranging the elements of S in some order and then counting the number of times the order is reversed by π. This attaches to π a number which by its definition does not depend on any factorization. One then shows that this number changes parity when π is changed by a transposition.

I offer here a different proof whose main claim to interest is that it does not call for giving names to the elements of S. In fact it is most clearly expressed not using mathematical notation at all. While the specific content of the choreographic model used is clearly irrelevant to the matter in hand, the reader is challenged to present the proof clearly without using some such model.

As noted above, we need only prove the special case where π is the identity. We do this by induction on the size of S. If S has only one or two elements, the conclusion is certainly clear, so assume the theorem true for sets of n elements, and consider permutations of a set with $n+1$ elements.

Imagine a dance club whose membership consists of n married couples and one bachelor, thus $n+1$ men and n women. The group gets together to dance in a "set" of n couples, with one man left over to play the fiddle. This means that at any time during the dance each couple occupies a certain position in the set, with two couples occasionally changing places as called for by the figures of the dance. As is customary, the dance ends with all the couples back in their orginal positions. Husbands, as much as possible, dance with their own wives as partners, but since the bachelor wants to dance too, the men take turns playing the fiddle, while the bachelor, if not playing the fiddle himself, dances with the wife of the fiddle player. At the beginning and the end of the dance the bachelor is playing the fiddle. Each change of fiddler is accomplished by the fiddler changing places with one of the other men. Notice that if both of these men happen to be married men, their wives must simultaneously change places, so that the old fiddler gets his wife back as partner, while the bachelor gets the new fiddler's wife as partner. This completes the description of what happens.

Now let's count. Any sequence of transpositions of the $n+1$ men can be accomplished in this way, provided all the men return to their original positions. Separate these transpositions into two classes: type I is those transpositions in which the bachelor gives up the fiddle or takes the fiddle; type II is all other transpositions of men. Since the bachelor plays the fiddle at the beginning and at the end, it is clear that the number of transpositions of type I is even. Now consider the transpositions of women. The women do not move when the men do a transposition of type I. They experience a transposition whenever the men do a transposition of type II. (This is obvious if the transposition is part of the dance, and it has already been explained in the case of a change of fiddlers.) Since the women also return in the end to their original positions, we know from the inductive hypothesis that the number of transpositions of type II must be even. Adding the numbers of transpositions of types I and II, both even, we see that the total number of transpositions experienced by the men must be even, and the induction is finished.

ON THE LEAF CURVES

A. A. AUCOIN, University of Houston

The equation $\rho^n = a^n \cos p\theta$, where p is a positive integer, (the results remain valid if we replace $\cos p\theta$ by $\sin p\theta$) n is a positive odd integer, ρ is the principal root of ρ^n, are generalizations of the equation of rose curves to which they reduce if $n=1$. We wish to consider the area enclosed by curves of this type. If p is odd there are p leaves and if p is even there are $2p$ leaves.

THEOREM 1. *For every k the area enclosed by* $\rho^n = a^n \cos (2k+1)\theta$ is

$$\frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{n+2}{2n}\right)}{\Gamma\left(\frac{n+1}{n}\right)} a^2.$$

Proof. If A is the area enclosed

$$A = 2(2k+1)\int_0^{\pi/2(2k+1)} \frac{\rho^2}{2} d\theta$$

$$= (2k+1)a^2 \int_0^{\pi/2(2k+1)} \cos^{2/n}(2k+1)d\theta.$$

If now we let $(2k+1)\theta = \alpha$,

$$A = a^2 \int_0^{\pi/2} \cos^{2/n} \alpha d\alpha$$

$$= \frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{n+2}{2n}\right)}{\Gamma\left(\frac{n+1}{n}\right)} a^2.$$

COROLLARY. *If $n=1$ the area enclosed is $\pi a^2/4$.*

THEOREM 2. *For every k the area enclosed by* $\rho^n = a^n \cos 2k\theta$ *is*

$$\sqrt{\pi} \frac{\Gamma\left(\frac{n+2}{2n}\right)}{\Gamma\left(\frac{2n+1}{2n}\right)} a^2.$$

Proof. The proof is similar to that of Theorem 1.

COROLLARY. *If $n=1$ the area enclosed is $\pi a^2/2$.*

The above curves contain either an odd number or twice an even number of leaves. The curve $\rho^{2n} = a^{2n} \cos 2(2k+1)\theta$ contains $2(2k+1)$ leaves. If $n=1$, $k=0$ the equation reduces to the equation of the leminscate.

THEOREM 3. *For every k the area enclosed by $\rho^{2n}=a^{2n}\cos 2(2k+1)\theta$ is*

$$\frac{\sqrt{\pi}}{2}\,\frac{\Gamma\left(\dfrac{n+1}{2n}\right)}{\Gamma\left(\dfrac{2n+1}{2n}\right)}\,a^2.$$

Proof. The proof is similar to that of Theorem 1.

COROLLARY. *If $n=1$ the area enclosed is a^2.*

A NEW NECESSARY AND SUFFICIENT CONDITION FOR LINEAR DEPENDENCE OF VECTORS IN A FINITE DIMENSIONAL VECTOR SPACE

MICHAEL LAIDACKER, Lamar State College of Technology

This note contains a new necessary and sufficient condition for linear dependence of vectors in a complex n dimensional vector space which will be denoted throughout this paper as the vector space V or simply V. For vectors in the vector space V, any m vectors are linearly dependent for $m>n$. It is well known that a necessary and sufficient condition for linear dependence of n vectors in the vector space V is given by the following lemma.

LEMMA 1. *If $\bar{a}_1=(a_{11},\cdots,a_{1n}),\cdots,\bar{a}_{n-1}=(a_{n-1,1},\cdots,a_{n-1,n})$ and $\bar{a}_n=(a_{n1},\cdots,a_{nn})$ belong to the vector space V, then a necessary and sufficient condition that the vectors $\bar{a}_1,\cdots,\bar{a}_n$ are linearly dependent is*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0.$$

We introduce the notational change whereby the determinant equation of Lemma 1 becomes $|\bar{a}_1,\cdots,\bar{a}_n|=0$ in the new notational scheme. With this scheme in mind, we write the following theorem which gives a new necessary and sufficient condition for linear dependence of vectors in the vector space V.

THEOREM 1. *A necessary and sufficient condition that $\bar{a}_1,\cdots,\bar{a}_m$ of V, $m<n$, be linearly dependent is*

$$\sum_{1\le i_1<\cdots<i_{n-m}} |\bar{a}_1,\cdots,\bar{a}_m,\bar{b}_{i_1},\cdots,\bar{b}_{i_{n-m}}|\,\overline{|\bar{a}_1,\cdots,\bar{a}_m,\bar{b}_{i_1},\cdots,\bar{b}_{i_{n-m}}|}=0$$

where the second determinant is the conjugate of the first and $\bar{b}_1,\cdots,\bar{b}_n$ of V are linearly independent.

THEOREM 3. *For every k the area enclosed by $\rho^{2n} = a^{2n} \cos 2(2k+1)\theta$ is*

$$\frac{\sqrt{\pi}}{2} \frac{\Gamma\left(\frac{n+1}{2n}\right)}{\Gamma\left(\frac{2n+1}{2n}\right)} a^2.$$

Proof. The proof is similar to that of Theorem 1.

COROLLARY. *If $n=1$ the area enclosed is a^2.*

A NEW NECESSARY AND SUFFICIENT CONDITION FOR LINEAR DEPENDENCE OF VECTORS IN A FINITE DIMENSIONAL VECTOR SPACE

MICHAEL LAIDACKER, Lamar State College of Technology

This note contains a new necessary and sufficient condition for linear dependence of vectors in a complex n dimensional vector space which will be denoted throughout this paper as the vector space V or simply V. For vectors in the vector space V, any m vectors are linearly dependent for $m>n$. It is well known that a necessary and sufficient condition for linear dependence of n vectors in the vector space V is given by the following lemma.

LEMMA 1. *If $\bar{a}_1 = (a_{11}, \cdots, a_{1n}), \cdots, \bar{a}_{n-1} = (a_{n-1,1}, \cdots, a_{n-1,n})$ and $\bar{a}_n = (a_{n1}, \cdots, a_{nn})$ belong to the vector space V, then a necessary and sufficient condition that the vectors $\bar{a}_1, \cdots, \bar{a}_n$ are linearly dependent is*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0.$$

We introduce the notational change whereby the determinant equation of Lemma 1 becomes $|\bar{a}_1, \cdots, \bar{a}_n| = 0$ in the new notational scheme. With this scheme in mind, we write the following theorem which gives a new necessary and sufficient condition for linear dependence of vectors in the vector space V.

THEOREM 1. *A necessary and sufficient condition that $\bar{a}_1, \cdots, \bar{a}_m$ of V, $m<n$, be linearly dependent is*

$$\sum_{1 \le i_1 < \cdots < i_{n-m}} |\bar{a}_1, \cdots, \bar{a}_m, \bar{b}_{i_1}, \cdots, \bar{b}_{i_{n-m}}| \, \overline{|\bar{a}_1, \cdots, \bar{a}_m, \bar{b}_{i_1}, \cdots, \bar{b}_{i_{n-m}}|} = 0$$

where the second determinant is the conjugate of the first and $\bar{b}_1, \cdots, \bar{b}_n$ of V are linearly independent.

The proof of Theorem 1 follows easily from Lemma 1 and from the following two lemmas whose proofs can be found in any reliable book on linear vector spaces.

LEMMA 2. *If* $\bar{a}_1, \cdots, \bar{a}_m$ *of* V *are linearly independent,* $m<n$, *and* $\bar{b}_1, \cdots, \bar{b}_n$ *of* V *are linearly independent, then there are vectors* $\bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *from the set* $\bar{b}_1, \cdots, \bar{b}_n$ *such that* $\bar{a}_1, \cdots, \bar{a}_m, \bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *are linearly independent.*

LEMMA 3. *If* $\bar{a}_1, \cdots, \bar{a}_m$ *of* V *are linearly dependent,* $m<n$, *and* $\bar{b}_1, \cdots, \bar{b}_n$ *of* V *are linearly independent, then there are no vectors* $\bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *from the set* $\bar{b}_1, \cdots, \bar{b}_n$ *such that* $\bar{a}_1, \cdots, \bar{a}_m, \bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *are linearly independent.*

Any basis of V could serve as our set of b's in Theorem 1; however the standard basis of V in most cases would be more convenient.

I wish to thank Joseph Baj for his helpful comments.

Reference

E. D. Nering, Linear Algebra and Matrix Theory, Wiley, New York, 1967.

TWO-DIMENSIONAL POWER-ASSOCIATIVE ALGEBRAS

E. W. WALLACE, University of Leeds, England

1. Introduction. Many of the familiar nonassociative algebras are *power-associative*, in the sense that each element generates an associative subalgebra. Thus, associative algebras, Lie algebras, Jordan algebras and alternative algebras are all power-associative.

It has proved useful to have a complete list available of all complex Lie algebras of dimension $\leqq 4$ in a simple canonical form, one use of such a list being to test conjectures (Wallace [4]). The purpose of this note is to obtain, in an elementary way, canonical forms for the two-dimensional power-associative algebras over the real and complex fields, and to provide invariants which distinguish them. These invariants also enable us to locate the canonical form corresponding to a given algebra. At the end of the note we point out whether each of the canonical forms is one of the types mentioned earlier.

If R is an arbitrary nonassociative algebra, and $x \in R$, define $x^1 = x$, $x^{i+1} = xx^i$ for all $i \geqq 1$. A straightforward proof then gives the standard result that R is power-associative if and only if $x^i x^j = x^{i+j}$ for all $x \in R$ and for all i, j. This is the form in which we shall use power-associativity.

Throughout the note, the base field is denoted by F, and this can usually be taken to be any field of characteristic 0. From Section 3 onwards, F will be restricted to be the real field, although reference will be made to the complex field.

The algebras are obtained from a discussion of two distinct cases, and an algebra arising from one cannot be isomorphic to an algebra arising from the other.

The proof of Theorem 1 follows easily from Lemma 1 and from the following two lemmas whose proofs can be found in any reliable book on linear vector spaces.

LEMMA 2. *If* $\bar{a}_1, \cdots, \bar{a}_m$ *of* V *are linearly independent,* $m<n$, *and* $\bar{b}_1, \cdots, \bar{b}_n$ *of* V *are linearly independent, then there are vectors* $\bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *from the set* $\bar{b}_1, \cdots, \bar{b}_n$ *such that* $\bar{a}_1, \cdots, \bar{a}_m, \bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *are linearly independent.*

LEMMA 3. *If* $\bar{a}_1, \cdots, \bar{a}_m$ *of* V *are linearly dependent,* $m<n$, *and* $\bar{b}_1, \cdots, \bar{b}_n$ *of* V *are linearly independent, then there are no vectors* $\bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *from the set* $\bar{b}_1, \cdots, \bar{b}_n$ *such that* $\bar{a}_1, \cdots, \bar{a}_m, \bar{b}'_{m+1}, \cdots, \bar{b}'_n$ *are linearly independent.*

Any basis of V could serve as our set of b's in Theorem 1; however the standard basis of V in most cases would be more convenient.

I wish to thank Joseph Baj for his helpful comments.

Reference

E. D. Nering, Linear Algebra and Matrix Theory, Wiley, New York, 1967.

TWO-DIMENSIONAL POWER-ASSOCIATIVE ALGEBRAS

E. W. WALLACE, University of Leeds, England

1. Introduction. Many of the familiar nonassociative algebras are *power-associative*, in the sense that each element generates an associative subalgebra. Thus, associative algebras, Lie algebras, Jordan algebras and alternative algebras are all power-associative.

It has proved useful to have a complete list available of all complex Lie algebras of dimension $\leqq 4$ in a simple canonical form, one use of such a list being to test conjectures (Wallace [4]). The purpose of this note is to obtain, in an elementary way, canonical forms for the two-dimensional power-associative algebras over the real and complex fields, and to provide invariants which distinguish them. These invariants also enable us to locate the canonical form corresponding to a given algebra. At the end of the note we point out whether each of the canonical forms is one of the types mentioned earlier.

If R is an arbitrary nonassociative algebra, and $x \in R$, define $x^1 = x$, $x^{i+1} = xx^i$ for all $i \geqq 1$. A straightforward proof then gives the standard result that R is power-associative if and only if $x^i x^j = x^{i+j}$ for all $x \in R$ and for all i, j. This is the form in which we shall use power-associativity.

Throughout the note, the base field is denoted by F, and this can usually be taken to be any field of characteristic 0. From Section 3 onwards, F will be restricted to be the real field, although reference will be made to the complex field.

The algebras are obtained from a discussion of two distinct cases, and an algebra arising from one cannot be isomorphic to an algebra arising from the other.

2. The first case. In this section a typical algebra is denoted by A, and has genus 0. These algebras have been completely determined (Patterson [2]), and multiplication tables are given by him in the form:

(i) $e_ie_j = 0$.

(ii) $e_ie_j = \lambda e_i + \mu e_j$ ($\lambda, \mu \in F$, not both zero).

(iii) $e_ie_j = \delta_{1j}e_i + \delta_{2i}e_j$, where δ_{ij} is the Kronecker delta.

We shall denote by A_1 the algebra of (i).

The algebras (ii): If $\lambda = 0$ we can take $\mu = 1$, and a new basis $f_1 = e_1, f_2 = e_1 - e_2$ gives the algebra

$$A_2: f_1^2 = f_1, \quad f_1f_2 = f_2.$$

If $\lambda \neq 0$ and $\mu = 0$ we can take $\lambda = 1$. The same change of basis gives the algebra

$$A_3: f_1^2 = f_1, f_2f_1 = f_2.$$

Finally, if $\lambda \neq 0$ and $\mu \neq 0$, put $\sigma = \lambda/\mu$ and choose a new basis $f_1 = (1/\mu)e_1, f_2 = (1/\lambda)(e_1 - e_2)$, giving the family of algebras:

$$A(\sigma): f_1^2 = (1 + \sigma)f_1, \quad f_1f_2 = f_2, \quad f_2f_1 = \sigma f_2.$$

Interchanging e_1, e_2 in the algebra (iii) gives

$$A_4: f_1^2 = f_1, \quad f_1f_2 = f_1 + f_2, \quad f_2^2 = f_2.$$

3. The second case. From now on, F is assumed to be the real field, unless otherwise stated, and a typical algebra for this section is denoted by B. Each algebra has genus 1 and a basis e_1, e_2 such that,

$$e_1^2 = e_2, \quad e_2^2 = \alpha e_1 + \beta e_2, \quad e_1e_2 = \gamma e_1 + \delta e_2, \quad e_2e_1 = \lambda e_1 + \mu e_2$$

for some scalars $\alpha, \beta, \gamma, \delta, \lambda, \mu$. We intend to obtain the nonisomorphic algebras directly.

Since B is power-associative it is necessary that $e_1e_1^2 = e_1^2e_1$ and so B is commutative. Moreover, $e_1e_1^3 = e_1^2e_1^2$, which yields $\alpha = \gamma\delta$, $\beta = \gamma + \delta^2$. Thus

$$e_1^2 = e_2, \ e_2^2 = \gamma\delta e_1 + (\gamma + \delta^2)e_2, \ e_1e_2 = e_2e_1 = \gamma e_1 + \delta e_2.$$

Conversely, algebras defined by these products are associative, and hence power-associative.

a) $\gamma = 0$. Since $e_1(pe_1 + qe_2) = (p + q\delta)e_2$, for all p, q, none of the algebras has an identity.

If $\delta = 0$ we immediately have the algebra:

$$B_1: e_1^2 = e_2.$$

If $\delta \neq 0$, the new basis $f_1 = (1/\delta)e_1, f_2 = (1/\delta^2)e_2$ gives

$$B_2: f_1^2 = f_2, \quad f_2^2 = f_2, \quad f_1f_2 = f_2f_1 = f_2.$$

b) $\gamma \neq 0$. For these algebras, $(-\delta/\gamma)e_1 + (1/\gamma)e_2$ is always an identity, and may be taken as a new basis vector f_2. If f_1 is another basis vector, we certainly have $f_2^2=f_2$, $f_1f_2=f_2f_1=f_1$. Now the radical (Albert [1]) is zero or nonzero according as $\delta^2+4\gamma\neq 0$ or $\delta^2+4\gamma=0$.

b$_1$) $\gamma\neq 0$ and $\delta^2+4\gamma>0$. If $t=(\delta^2+4\gamma)^{1/2}$, the new basis

$$f_1=-(1/\gamma t)[(2\gamma+\delta^2)e_1-\delta e_2], f_2=-(1/\gamma)(\delta e_1-e_2)$$

yields

$$B_3: f_1^2=f_2,\quad f_2^2=f_2,\quad f_1f_2=f_2f_1=f_1.$$

b$_2$) $\gamma\neq 0$ and $\delta^2+4\gamma<0$. Now put $s=[-(\delta^2+4\gamma)]^{1/2}$, and

$$f_1=(1/\gamma s)[(2\gamma+\delta^2)e_1-\delta e_2], f_2=-(1/\gamma)\ (\delta e_1-e_2),$$

giving

$$B_4: f_1^2=-f_2,\quad f_2^2=f_2,\quad f_1f_2=f_2f_1=f_1.$$

(The algebras B_3, B_4 are not isomorphic over the real field, but they are isomorphic over the complex field. In fact, start with the products for B_3 and let $g_1=if_1$, $g_2=f_2$ where $i=\sqrt{(-1)}$. Then we obtain the products for B_4.)

b$_3$) $\gamma\neq 0$ and $\delta^2+4\gamma=0$. If $f_1=\delta e_1-2e_2$, then $f_1^2=0$ and so $(f_1x)^2=0$ for all x. Hence f_1 is in the radical. Put $f_1=\delta e_1-2e_2$, $f_2=(4/\delta^2)\ (\delta e_1-e_2)$. Then

$$B_5: f_2^2=f_2,\quad f_1f_2=f_2f_1=f_1.$$

4. Summary. We have now determined all the real two-dimensional power-associative algebras, and they are listed here for convenience. The basis is always denoted by e_1, e_2 and, apart from A_1, only nonzero products are given.

$$A_1: e_ie_j=0 \quad \text{for all } i,j, =1,2;$$

$$A_2: e_1^2=e_1,\quad e_1e_2=e_2;$$

$$A_3: e_1^2=e_1,\quad e_2e_1=e_2;$$

$$A_4: e_1^2=e_1,\quad e_2^2=e_2,\quad e_1e_2=e_1+e_2;$$

$$A(\sigma): e_1^2=(1+\sigma)e_1,\quad e_1e_2=e_2,\quad e_2e_1=\sigma e_2;$$

$$B_1: e_1^2=e_2;$$

$$B_2: e_1^2=e_2,\quad e_2^2=e_2,\quad e_1e_2=e_2e_1=e_2;$$

$$B_3: e_1^2=e_2,\quad e_2^2=e_2,\quad e_1e_2=e_2e_1=e_1;$$

$$B_4: e_1^2=-e_2,\quad e_2^2=e_2,\quad e_1e_2=e_2e_1=e_1;$$

$$B_5: e_2^2=e_2,\quad e_1e_2=e_2e_1=e_1.$$

5. Invariants. The family $A(\sigma)$ will be considered in detail in the next section. Invariants, or properties, of the algebras will now be given which distinguish the members of $A(\sigma)$ from the other algebras, but not from each other. In particular, $A_2(=A(0))$ is not distinguished from the other members of $A(\sigma)$.

Twelve invariants and properties $Q_1 - Q_{12}$ are given in the table below. Only $Q_1 - Q_6$ are needed to distinguish the algebras, and the others are given for information.

Note that among $Q_1 - Q_6$, if one is omitted, the remaining five are the same for at least two of the algebras.

The reader is referred to Schafer [3] for definitions of any unfamiliar terms.

Q_1: The algebra has a left identity. If an algebra has this property put $\surd$. If not, put X.

Q_2: The algebra is simple. [This can be verified by assuming there is a one-dimensional ideal N spanned by $x \neq 0$, and showing that $xy \notin N$ or $yx \notin N$ for some y.]

Q_3: Dimension of the nucleus.

Q_4: Dimension of the center.

Q_5: The index r. (For a general algebra H, define $H^1 = H$, $H^{i+1} = H^i H^i$ for all $i \geqq 1$. Then $H^r = H^{r+1}$, $H^r \neq H^{r-1}$. If $H \neq (0)$ and $H^2 = H$, define $r = 1$.)

Q_6: Dimension of the nilradical (= maximum nil ideal).

Q_7: Genus of the algebra.

Q_8: The algebra is associative.

Q_9: The algebra is commutative.

Q_{10}: The algebra has a right identity.

Q_{11}: The algebra has a nonzero idempotent.

Q_{12}: Dimension of the maximum solvable ideal.

Algebra	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}	Q_{11}	Q_{12}
A_1	X	X	2	2	2	2	0	$\surd$	$\surd$	X	X	2
$A_2 = A(0)$	$\surd$	X	2	0	1	1	0	$\surd$	X	X	$\surd$	1
A_3	X	X	2	0	1	1	0	$\surd$	X	$\surd$	$\surd$	1
A_4	X	$\surd$	0	0	1	0	0	X	X	X	$\surd$	0
$A(\sigma): \sigma \neq 0, \pm 1$	X	X	0	0	1	1	0	X	X	X	$\surd$	1
$A(1)$	X	X	0	0	1	1	0	X	$\surd$	X	$\surd$	1
$A(-1)$	X	X	0	0	3	2	0	X	X	X	X	2
B_1	X	X	2	2	3	2	1	$\surd$	$\surd$	X	X	2
B_2	X	X	2	2	2	1	1	$\surd$	$\surd$	X	$\surd$	1
B_3	$\surd$	X	2	2	1	0	1	$\surd$	$\surd$	$\surd$	$\surd$	0
B_4	$\surd$	$\surd$	2	2	1	0	1	$\surd$	$\surd$	$\surd$	$\surd$	0
B_5	$\surd$	X	2	2	1	1	1	$\surd$	$\surd$	$\surd$	$\surd$	1

6. The family $A(\sigma)$. If $x = x^i e_i$ is an element of some member $A(\sigma)$ of the family, then the matrices C, D which represent the left and right multiplications L_x, R_x are respectively

$$C = \begin{pmatrix} (1+\sigma)x^1 & \sigma x^2 \\ 0 & x^1 \end{pmatrix} \qquad D = \begin{pmatrix} (1+\sigma)x^1 & x^2 \\ 0 & \sigma x^1 \end{pmatrix}.$$

Thus $trC : trD = 2+\sigma : 1+2\sigma$.

This ratio is independent of x and of the basis, and

$$2+\sigma_1 : 1+2\sigma_1 = 2+\sigma_2 : 1+2\sigma_2 \Leftrightarrow \sigma_1 = \sigma_2.$$

Hence $A(\sigma_1) \approx A(\sigma_2) \Leftrightarrow \sigma_1 = \sigma_2$.

Thus, knowing that a given algebra is a member of the family $A(\sigma)$, (by using $Q_1 - Q_6$), we can determine its parameter σ by finding the ratio $trC : trD$.

7. Special types of algebras. We now list all the real two dimensional algebras of certain special types. Recall that B_3, B_4 are isomorphic over the complex field.

(a) *Associative*: $A_1, A_2, A_3, B_1, B_2, B_3, B_4, B_5$.
(b) *Lie*: $A_1, A(-1)$.
(c) *Jordan*: $A_1, A(1), B_1, B_2, B_3, B_4, B_5$.
(d) (*Noncommutative*) *Jordan*: $A_1, A_2, A_3, A(\sigma)$ (all σ), B_1, B_2, B_3, B_4, B_5.
(e) *Alternative* (*left, right or two sided*): $A_1, A_2, A_3, B_1, B_2, B_3, B_4, B_5$.
(f) *Flexible*: $A_1, A_2, A_3, A(\sigma)$ (all σ), B_1, B_2, B_3, B_4, B_5.

References

1. A. A. Albert, Structure of Algebras, Colloquium Publications, vol. 24, Amer. Math. Soc., 1964.
2. E. M. Patterson, Linear algebras of genus zero, J. London Math. Soc., 31 (1956) 326–331.
3. R. D. Schafer, An Introduction to Nonassociative Algebras, Academic Press, New York, 1966.
4. E. W. Wallace, Complex four-dimensional Lie algebras, Proc. Roy. Soc. Edinburgh, 65 (1958) 72–83.

A NEW PROOF OF A COMBINATORIAL IDENTITY

DAVID C. SHIPMAN, University of Oregon

The following seems to be a new proof of an important combinatorial identity used by Marshall Hall Jr. in his proof of the Principle of Inclusion and Exclusion. This proof seems to make the identity conceptually clearer than the proof offered by Hall on p. 4 of his book, *Combinatorial Theory*. Here $\binom{k}{j}$ denotes the binomial coefficient $k!/j!(k-j)!$.

PROPOSITION. $\sum_{k=j}^{i} \binom{k}{j}\binom{i}{k}(-1)^{j+k} = 0$ *for* $i > j$.

Proof. Let F be the vector space of all polynomials of degree $\leqq p-1$ over the

6. The family $A(\sigma)$. If $x=x^ie_i$ is an element of some member $A(\sigma)$ of the family, then the matrices C, D which represent the left and right multiplications L_x, R_x are respectively

$$C=\begin{pmatrix}(1+\sigma)x^1 & \sigma x^2\\ 0 & x^1\end{pmatrix}\qquad D=\begin{pmatrix}(1+\sigma)x^1 & x^2\\ 0 & \sigma x^1\end{pmatrix}.$$

Thus $trC: trD=2+\sigma:1+2\sigma$.

This ratio is independent of x and of the basis, and

$$2+\sigma_1:1+2\sigma_1=2+\sigma_2:1+2\sigma_2\Leftrightarrow\sigma_1=\sigma_2.$$

Hence $A(\sigma_1)\approx A(\sigma_2)\Leftrightarrow\sigma_1=\sigma_2$.

Thus, knowing that a given algebra is a member of the family $A(\sigma)$, (by using Q_1-Q_6), we can determine its parameter σ by finding the ratio trC: trD.

7. Special types of algebras. We now list all the real two dimensional algebras of certain special types. Recall that B_3, B_4 are isomorphic over the complex field.

(a) *Associative*: $A_1, A_2, A_3, B_1, B_2, B_3, B_4, B_5$.
(b) *Lie*: $A_1, A(-1)$.
(c) *Jordan*: $A_1, A(1), B_1, B_2, B_3, B_4, B_5$.
(d) (*Noncommutative*) *Jordan*: $A_1, A_2, A_3, A(\sigma)$ (all σ), B_1, B_2, B_3, B_4, B_5.
(e) *Alternative* (*left, right or two sided*): $A_1, A_2, A_3, B_1, B_2, B_3, B_4, B_5$.
(f) *Flexible*: $A_1, A_2, A_3, A(\sigma)$ (all σ), B_1, B_2, B_3, B_4, B_5.

References

1. A. A. Albert, Structure of Algebras, Colloquium Publications, vol. 24, Amer. Math. Soc., 1964.
2. E. M. Patterson, Linear algebras of genus zero, J. London Math. Soc., 31 (1956) 326–331.
3. R. D. Schafer, An Introduction to Nonassociative Algebras, Academic Press, New York, 1966.
4. E. W. Wallace, Complex four-dimensional Lie algebras, Proc. Roy. Soc. Edinburgh, 65 (1958) 72–83.

A NEW PROOF OF A COMBINATORIAL IDENTITY

DAVID C. SHIPMAN, University of Oregon

The following seems to be a new proof of an important combinatorial identity used by Marshall Hall Jr. in his proof of the Principle of Inclusion and Exclusion. This proof seems to make the identity conceptually clearer than the proof offered by Hall on p. 4 of his book, *Combinatorial Theory*. Here $\binom{k}{j}$ denotes the binomial coefficient $k!/j!(k-j)!$.

PROPOSITION. $\sum_{k=j}^{i}\binom{k}{j}\binom{i}{k}(-1)^{j+k}=0$ *for* $i>j$.

Proof. Let F be the vector space of all polynomials of degree $\leqq p-1$ over the real field. Then a basis for F is $(1, x, x^2, \cdots, x^{p-1})$. Now let $y=1-x$, so $(1, y, y^2, \cdots, y^{p-1})$ is also a basis for F. Since

$$y^s = \sum_{r=0}^{s} \binom{s}{r}(-x)^r + 0x^{s+1} + \cdots + 0x^{p-1} \quad \text{for} \quad 0 \leqq s \leqq p-1$$

the change of basis matrix $A=(a_{ij})$ from $(1, x, \cdots, x^{p-1})$ to $(1, y, \cdots, y^{p-1})$ is given by

$$a_{ij} = \begin{cases} \binom{i}{j}(-1)^{j+1} & \text{if } i \geqq j \\ 0 & \text{if } i < j. \end{cases}$$

But similarly, since $x=1-y$, we get $A^{-1}=A$, so $I=AA=(b_{ij})$, where

$$b_{ij} = \sum_{k=1}^{p} a_{ik}a_{kj} = \begin{cases} \sum_{k=j}^{i} a_{ik}a_{kj} & \text{if } i \geqq j \\ 0 & \text{if } i < j. \end{cases}$$

So for

$$i > j, 0 = b_{ij} = \sum_{k=j}^{i} a_{ik}a_{kj} = \sum_{k=j}^{i} \binom{k}{j}\binom{i}{k}(-1)^{j+k+2},$$

and thus

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

Reference

1. Marshall Hall, Jr., Combinatorial Theory, Blaisdell, Waltham, 1967.

BOOK REVIEWS

Edited by D. Elizabeth Kennedy, University of Victoria

Materials intended for review should be sent to: Professor D. Elizabeth Kennedy, Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada.

Reviews of texts at the freshman-sophomore level based upon classroom experience will be welcomed by the Book Review Editor.

A boldface capital **C** *in the margin indicates a classroom review.*

C *Modern Mathematics for Business Students.* By R. E. Wheeler and W. D. Peeples, Jr. Brooks/Cole, Belmont, Calif., 1969. xii+589 pp. \$9.95.

At present we are using this book at the University of Colorado in our three semester-hour course for social science and business students. We have two real field. Then a basis for F is $(1, x, x^2, \cdots, x^{p-1})$. Now let $y=1-x$, so $(1, y, y^2, \cdots, y^{p-1})$ is also a basis for F. Since

$$y^s = \sum_{r=0}^{s} \binom{s}{r}(-x)^r + 0x^{s+1} + \cdots + 0x^{p-1} \quad \text{for} \quad 0 \leqq s \leqq p-1$$

the change of basis matrix $A=(a_{ij})$ from $(1, x, \cdots, x^{p-1})$ to $(1, y, \cdots, y^{p-1})$ is given by

$$a_{ij} = \begin{cases} \binom{i}{j}(-1)^{j+1} & \text{if } i \geqq j \\ 0 & \text{if } i < j. \end{cases}$$

But similarly, since $x=1-y$, we get $A^{-1}=A$, so $I=AA=(b_{ij})$, where

$$b_{ij} = \sum_{k=1}^{p} a_{ik}a_{kj} = \begin{cases} \sum_{k=j}^{i} a_{ik}a_{kj} & \text{if } i \geqq j \\ 0 & \text{if } i < j. \end{cases}$$

So for

$$i > j, 0 = b_{ij} = \sum_{k=j}^{i} a_{ik}a_{kj} = \sum_{k=j}^{i} \binom{k}{j}\binom{i}{k}(-1)^{j+k+2},$$

and thus

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

Reference

1. Marshall Hall, Jr., Combinatorial Theory, Blaisdell, Waltham, 1967.

BOOK REVIEWS

EDITED BY D. ELIZABETH KENNEDY, University of Victoria

Materials intended for review should be sent to: Professor D. Elizabeth Kennedy, Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada.

Reviews of texts at the freshman-sophomore level based upon classroom experience will be welcomed by the Book Review Editor.

A boldface capital **C** *in the margin indicates a classroom review.*

C *Modern Mathematics for Business Students.* By R. E. Wheeler and W. D. Peeples, Jr. Brooks/Cole, Belmont, Calif., 1969. xii+589 pp. $9.95.

At present we are using this book at the University of Colorado in our three semester-hour course for social science and business students. We have two lecture sections of about 150 students each taught by full time members of the department and 9 small sections of about 35 each taught by teaching assistants. For the big sections there are two lectures a week and six recitation sections, each meeting twice a week, in charge of teaching assistants. The syllabus calls for a minimum of Chapters 2, 3, 4, 6 and a little of Chapter 8, that is: sets, relations and functions, systems of linear equations and inequalities, vectors and matrices, probability and, in Chapter 8, linear programming. In some of the small sections they will cover somewhat more material.

I asked for some students' comments. There were a number like the following: "I think the book is ok for a math book. . . . It is very easy to read, not like most math books." Of course there were comments of all kinds beginning with those who felt it was hard to read to those who felt it contained a lot of "useless reading." But I think in general the students find the book worth reading and understandable; they appreciate the examples worked out. One student had the suggestion that the answers in the back of the book would be more useful on occasion if there were also given the steps leading to the answer. Incidentally, there are quite a number of annoying mistakes in the answers which are given.

The "nonmathematical" approach of the book which appeals to many of the students, of course, is not so attractive to the instructor. Here of course the teacher must keep in mind for whom the book was written. But there are spots where something is lacking from both points of view. There are places where the authors make a bow to a subject or concept and do not either develop it or use it. One of these is linear dependence. The definition is given in the two-dimensional case and illustrated with examples and exercises but it is not mentioned for higher dimensions nor, as far as I could see, is it used elsewhere. Another like case is the concept of a convex set, which is defined in six lines with a picture in the beginning of the chapter on linear programming, but I could find no other place where it is mentioned, much less used.

In places, the authors cut corners too much. For instance, in the definition of rank of a matrix, having worked in detail with row operations, they suddenly use column operations for their definition. This is not only confusing to students as well as to the instructor, but also needless, since rank could be defined only in terms of the results of row operations. Another example of cutting corners is in the explanation of the simplex method in linear programming. There one is left completely without aid if he wants to find out why the process works. Furthermore, if he contrasts the simplex method given with that in which one merely evaluates the function to be maximized at the critical points, the latter would certainly seem simpler in the examples shown.

For students of this level of mathematical training, it is very hard to find applications which have at least some appearance of being genuine. I think the authors have done quite well in this regard but I do find myself wishing there were some better substitute for the various mixture problems. For instance, I think mixtures of antifreeze would be more meaningful than mixtures of nuts.

Sometimes the treatment seems roundabout to students as well as to the teacher. Here, perhaps, the authors should explain that what they are trying to do is to give a kind of preview, that is, to give a feeling for what is coming. It is good to do this, but the reader should be warned.

I have mentioned chiefly the points in the book where I think improvement is needed. Its best virtue is that it is apparently readable and understandable to most of the students. There is also a wide choice of subject matter in the book, which we could not touch on: mathematics of finance, random variables and distribution functions, theory of games and Markov analysis, an introduction to differential and integral calculus and a chapter entitled "from probability to statistics." There are suggested course outlines. It is a very useable text as it stands and the choice of material and examples is good. With some revision it could become an excellent text.

B. W. JONES

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

ASSOCIATE EDITOR, MURRAY S. KLAMKIN, Ford Scientific Laboratory, Dearborn, Michigan

Readers of this department are invited to submit for solution problems believed to be new that may arise in study, in research, or in extra-academic situations. Problems may be submitted from any branch of mathematics and ranging in subject content from that accessible to the talented high school student to problems challenging to the professional mathematician. Proposals should be accompanied by solutions, when available, and by any information that will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted.

The asterisk (*) will be placed by the problem number to indicate that the proposer did not supply a solution. Readers' solutions are solicited for all problems proposed. Proposers' solutions may not be "best possible" and solutions by others will be given preference.

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and exactly the size desired for reproduction.

Send all communications for this department to Robert E. Horton, Los Angeles Valley College, 5800 Fulton Avenue, Van Nuys, California 91401.

To be considered for publication, solutions should be mailed before October 1, 1970.

PROPOSALS

761. *Proposed by John Hudson Tiner, High Ridge, Missouri.*

Solve the cryptarithm

$$\begin{array}{r} C\ H\ U\ C\ K \\ T\ R\ I\ G\ G \\ T\ U\ R\ N\ S \\ \hline T\ R\ I\ C\ K\ S \end{array}$$

do is to give a kind of preview, that is, to give a feeling for what is coming. It is good to do this, but the reader should be warned.

I have mentioned chiefly the points in the book where I think improvement is needed. Its best virtue is that it is apparently readable and understandable to most of the students. There is also a wide choice of subject matter in the book, which we could not touch on: mathematics of finance, random variables and distribution functions, theory of games and Markov analysis, an introduction to differential and integral calculus and a chapter entitled "from probability to statistics." There are suggested course outlines. It is a very useable text as it stands and the choice of material and examples is good. With some revision it could become an excellent text.

B. W. JONES

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles Valley College

ASSOCIATE EDITOR, MURRAY S. KLAMKIN, Ford Scientific Laboratory, Dearborn, Michigan

Readers of this department are invited to submit for solution problems believed to be new that may arise in study, in research, or in extra-academic situations. Problems may be submitted from any branch of mathematics and ranging in subject content from that accessible to the talented high school student to problems challenging to the professional mathematician. Proposals should be accompanied by solutions, when available, and by any information that will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted.

The asterisk (*) will be placed by the problem number to indicate that the proposer did not supply a solution. Readers' solutions are solicited for all problems proposed. Proposers' solutions may not be "best possible" and solutions by others will be given preference.

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and exactly the size desired for reproduction.

Send all communications for this department to Robert E. Horton, Los Angeles Valley College, 5800 Fulton Avenue, Van Nuys, California 91401.

To be considered for publication, solutions should be mailed before October 1, 1970.

PROPOSALS

761. *Proposed by John Hudson Tiner, High Ridge, Missouri.*

Solve the cryptarithm

$$\begin{array}{r} C\ H\ U\ C\ K \\ T\ R\ I\ G\ G \\ T\ U\ R\ N\ S \\ \hline T\ R\ I\ C\ K\ S \end{array}$$

762. *Proposed by Arthur Marshall, Madison, Wisconsin.*

Let n be a natural number greater than three. Prove that there exist two odd primes p_1 and p_2 such that

$$2n \equiv p_1 + p_2 \pmod{p_2}.$$

763. *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey.*

Prove:

$$\left(1+\frac{1}{3^{10}}+\frac{1}{5^{10}}+\cdots\right)=\left(1+\frac{1}{3^4}+\frac{1}{5^4}\cdots\right)\left(1-\frac{1}{2^6}+\frac{1}{3^6}-\frac{1}{4^6}+\cdots\right).$$

764. *Proposed by F. D. Parker, St. Lawrence University.*

Let $A=[a_{ij}]$ be a nonsingular square matrix, and denote its determinant by $d(A)$. If the same nonzero number x is added to each element of A to produce the matrix $A+x=[a_{ij}+x]$ then $d(A+x)=d(A)$ if and only if the sum of the elements of A^{-1} is zero.

765. *Proposed by Stanley Rabinowitz, Far Rockaway, New York.*

Let ABC be an isosceles triangle with right angle at C. Let $P_0=A$, $P_1=$ the midpoint of BC, $P_{2k}=$ the midpoint of AP_{2k-1}, and $P_{2k+1}=$ the midpoint of BP_{2k} for $k=1, 2, 3, \ldots$ Show that the cluster points of the sequence $\{P_n\}$ trisect the hypotenuse.

766. *Proposed by Warren Page, New York City Community College.*

Let N_1 and N_2 be two $n\times n$ nilpotent matrices over the field F. If N_1 and N_2 have the same nullity k and the same minimal polynomial p, what is the largest value of n for which N_1 and N_2 must be similar?

767. *Proposed by Harry Sitomer, C. W. Post College, New York.*

Given triangle ABC. Let A', B' and C' be in open segments BC, CA and AB, respectively, such that AA', BB' and CC' are concurrent at G and $AG/GA'=BG/GB'=CG/GC'$.

Prove that AA', BB' and CC' are medians of ABC. How does the conclusion change if "open segments" is replaced by "lines"?

SOLUTIONS

Late Solutions

Gladwin E. Bartel, Washington State University: 736; *E. M. Clarke, Madison College, Virginia*: 738, 739; *Renaldo E. Giudici, Universidad Santa Maria, Valparaiso, Chile*: 733; *Shiv Kumar, Panjab University, India, and Miss Nirmal, Government Girls' High School, Panipat, India (jointly)*: 725, 733, 735, 737, 738 *and* 739; *Mark Witnell, Kansas State University*: 722.

An Old Chestnut

740. [November, 1969] *Proposed by Dewey C. Duncan, Los Angeles, California.*

Can the decimal natural number consisting of $6k-1$ ones be a prime number?

Solution by Kenneth M. Wilke, Topeka, Kansas.

The answer is yes. The number consisting of 23 ones, $(10^{23}-1)/9$ is cited as a prime on Page 84 of Albert Beiler's *Recreations in the Theory of Numbers.*

The number consisting of $6k-1$ ones may be written in the form $(10^q-1)/9$ where $q=6k-1$. Since $6k-1\not\equiv 0 \pmod 9$, all prime factors of $(10^q-1)/9$ may be found by considering the congruence $10^q\equiv 1 \pmod p$ where p divides $(10^q-1)/9$. It is well known that if q is composite so is $(10^q-1)/9$. Hence if $(10^q-1)/9$ is prime, q must be prime. Furthermore, Fermat's theorem shows that any prime divisor p of $(10^q-1)/9$ must be of the form $2\,kq+1$ where k is an integer.

Furthermore, if $p=2q+1$ is prime and q is prime, and $(10/p)=1$, then $p=2q+1$ divides $(10^q-1)/9$, a modification of Euler's result in Dickson's *History of the Theory of Numbers*, Vol. 1, Page 160.

Solutions or references also submitted by Joel Brenner, University of Arizona; William F. Fox, Moberly Junior College, Missouri; Michael Goldberg, Washington, D.C.; Joseph J. Heed, Norwich University, Vermont; C. C. Oursler, Southern Illinois University at Edwardsville; Henry J. Ricardo, Yeshiva University; E. F. Schmeichel, Itasca, Illinois; Charles W. Trigg, San Diego, California; and Samuel Yates, Moorestown, New Jersey. One incorrect solution was received.

Inscribed Polygons

741. [November, 1969] *Proposed by Leon Bankoff, Los Angeles, California.*

A square and a triangle of equivalent area are inscribed in a semicircle, one side of the triangle forming the diameter of the semicircle. Show that the incenter of the triangle lies on one of the sides of the square.

I. *Solution by Paul D. Thomas, Naval Research Laboratory, Washington, D.C.*

Let the diameter of the semicircle be $2R$, with midpoint O. The side of the inscribed square is $2(5^{1/2})R/5$. If A,B are the vertices of the square on the semicircle, then the perpendicular from A upon OB is $u=4R/5$. If h is the altitude from the vertex angle, $90°$, upon the hypotenuse $2R$, then the required equality of areas implies $h=u$. Now, if the incenter is on a side of the square and d is the distance between the incenter and circumcenter, then $d^2=r^2+R^2/5=R^2-2Rr$, where r is the inradius. Hence $r=R(3\cdot 5^{1/2}/5-1)$. But in a right triangle, $h=2r+r^2/R$, whence $h=4R/5=u$ as was to be shown.

II. *Solution by Lawrence A. Ringenberg, Eastern Illinois University.*

Set up an xy-coordinate system so that the semicircle is given by $x^2+y^2=1$, $y\geqq 0$. Then $(\pm 1/\sqrt{5},\ 2/\sqrt{5})$ are vertices and $4/5$ is the area of the inscribed square. Let $A=(-1,0)$ and $B=(1,0)$. The inscribed triangle is ABC where $C=(3/5,4/5)$ or $(-3/5,4/5)$. Take $C=(3/5,4/5)$. Then $\sin A=\cos B=1/\sqrt{5}$, $\cos A=\sin B=2/\sqrt{5}$, $\tan A/2=\sqrt{5}-2$, and $\tan B/2=(\sqrt{5}-1)/2$. The intersection

$(1/\sqrt{5},\ (3-\sqrt{5})/5)$ of the lines $y=\tan\frac{1}{2}A\ (x+1)$ and $y=-\tan\frac{1}{2}B\ (x-1)$ is the incenter of the triangle and a point on the right side of the square. Similarly if $C=(-3/5,\ 4/5)$, it may be shown that the incenter of the triangle is on the left side of the square.

III. *Solution by the proposer.*

In any right triangle ABC ($C=90°$), $r=(s-c)$. Hence

$$rs = \sqrt{s(s-a)(s-b)(s-c)} = (s-a)(s-b).$$

In other words, the area of a right triangle is equal to the product of the segments into which the incircle contact divides the hypotenuse.

Here, each of the two vertices of the inscribed square lying on the hypotenuse divides it symmetrically into two segments whose product is equal to the area of the square.

Since the areas of the square and the triangle are equal, one of the vertices of the square coincides with the incircle contact on the hypotenuse. Thus the incenter lies on one of the sides of the square.

Also solved by Patrick J. Boyle, San Jose, California; Brother Brendan Kneale, St. Mary's College, California; Martin J. Brown, Northern Community College, Kentucky; Huseyin Demir, Middle East Technical University, Ankara, Turkey; Gerald C. Dodds, HRB-Singer, Inc., State College Pennsylvania; Clayton W. Dodge, University of Maine; Herta T. Freitag, Hollins College, Virginia; Michael Goldberg, Washington, D.C.; George Gruber, Brooklyn, New York; Philip Haverstick, Ft. Belvoir, Virginia; Lew Kowarski, Morgan State College, Maryland; Thomas Hun-tak Lee, Swarthmore College; R. J. Mahoney, Ottawa, Canada; John Riegsecker, Chicago, Illinois; M. Rodeen, Balboa High School, San Francisco, California; Steve Rohde, Lehigh University; E. F. Schmeichel, Itasca, Illinois; E. P. Starke, Plainfield, New Jersey; A. W. Walker, Toronto, Canada; Charles W. Trigg, San Diego, California; P. Weygang, Levack High School, Ontario, Canada; and Edward J. Zoll, Newark State College, New Jersey.

Euler's Function

742. [November, 1969] *Proposed by S. Srinivasan, Panjab University, Chandigarh, India.*

Let $F(n)=(1/n)\sum_{q=1}^{n}\phi(q)/q$, $n\geqq 1$, ϕ the Euler phi function. Show that:

1) $F(n)>\frac{1}{2}$

2) $\lim_{n\to\infty} F(n)=6/\pi^2$.

Solution by Michael Goldberg, Washington, D.C.

It is shown in *An Introduction to the Theory of Numbers*, G. H. Hardy and E. M. Wright, Page 268, that

$$\sum_{n=1}^{\infty}\phi(n)$$

is $3n^2/\pi^2$, and that the average order of $\phi(n)$ is $6n/\pi^2$. Hence $\phi(n)/n$ is asymptotic to $6/\pi^2$ and $\lim_{n\to\infty} F(n)=6/\pi^2$.

References to earlier solutions of this problem, given in *History of the Theory of Numbers*, L. E. Dickson, Vol. 1, Page 130, are E. Cesaro, Comptes Rendus, Paris, 106 (1888), 1651–1654, 107 (1888), 81–82, 426–427; Annali di Mat., (2) 16 (1888–1889), 178–179.

Also solved by E. F. Schmeichel, Itasca, Illinois; and the proposer.

Tetrahedral Inequality

743. [November, 1969] *Proposed by Huseyin Demir, Middle East Technical University, Ankara, Turkey.*

Let P be an interior point of a regular tetrahedron, $T\equiv A_1A_2A_3A_4$, with $p_i=PA_i$, and let x_{ij} denote the distance of P from the edge A_iA_j. Then prove

$$\sum_{i=1}^{4} p_i \geqq 2\sqrt{3}/3 \sum_{i<j} x_{ij},$$

equality holding if and only if P is at the center O of T.

Solution by Michael Goldberg, Washington, D.C.

Given a base of fixed length and the sum of two other lengths, the triangle of greatest height is obtained when the triangle is isosceles. Similarly, given the same base, and the sum of three other lengths to form three triangles by using the three pairs of these three sides, the sum of the heights is maximized when the triangles are congruent isosceles triangles. This can be generalized to n triangles. Hence, the sum of the distances of P from the edges of a regular tetrahedron divided by the sum of the distances of P from the vertices is maximized when P is at the center of the tetrahedron.

If e is the length of the edge of the tetrahedron then the distance h between opposite edges is given by

$$h^2 + (e/2)^2 + (e/2)^2 = e^2, \quad \text{or} \quad h = e/\sqrt{2}.$$

The distance R from the center to a vertex is given by

$$R^2 = (e/2)^2 + (h/2)^2 = e^2/4 + e^2/8, \quad \text{or} \quad R = \sqrt{3}e/2\sqrt{2}.$$

Hence, when P is at the center, the ratio of the sums is

$$4R/3h = (2e\sqrt{3}/\sqrt{2})/(3e/\sqrt{2}) = 2\sqrt{3}/3.$$

A similar extremal is obtained for each of the regular polyhedra, and for each of the regular polygons. Of course, the ratio of the two sums depends upon the figure considered.

Also solved by the proposer. One incorrect solution was received.

An Alternating Series

744. [November, 1969] *Proposed by B. J. Cerimele, North Carolina State University.*

Show that the alternating series

$$\sum_{i=1}^{\infty} (-1)^{i+1} \ln(1 + 1/i)$$

is conditionally convergent and determine its sum.

Solution by J. L. Brown, Jr., Pennsylvania State University.

We consider the partial sum,

$$S_{2n} = \sum_{i=1}^{2n} (-1)^{i+1} \ln\left(\frac{i+1}{i}\right)$$

$$= \ln\left[\left(\frac{2}{1}\cdot\frac{4}{3}\cdots\frac{2n}{2n-1}\right)\Big/\left(\frac{3}{2}\cdot\frac{5}{4}\cdots\frac{2n+1}{2n}\right)\right].$$

Denoting the bracketed argument of the logarithm by σ_{2n}, we have

$$\sigma_{2n} = \frac{(2\cdot 4\cdot 6\cdots 2n)^2}{[1\cdot 3\cdot 5\cdots(2n-1)]^2(2n+1)}$$

$$= \frac{(2\cdot 4\cdot 6\cdots 2n)^2(2n+1)}{[1\cdot 3\cdot 5\cdots(2n+1)]^2} = \left[\frac{(2^n n!)^2}{(2n+1)!}\right]^2\cdot(2n+1).$$

Using Stirling's asymptotic form for the factorials, we find

$$\lim_{n\to\infty} \sigma 2n = \lim_{n\to\infty}\left[\frac{\pi}{2(2n+2)}\right]\cdot(2n+1) = \pi/2,$$

so that $\lim_{n\to\infty} = \ln \pi/2$ by continuity of the logarithm. Similarly, $S_{2n+1} = (2n+2/2n+1)\cdot S_{2n}$, implying that $\lim_{n\to\infty} S_{2n+1} = \ln \pi/2$, and thus the given series converges to $\ln \pi/2$.

Also solved by Donald Batman, APO, San Francisco, California; Martin J. Brown, Northern Community College, Kentucky; L. Carlitz, Duke University; C. Robert Clement, Phillips Exeter Academy; Clayton W. Dodge, University of Maine; Michael Goldberg, Washington, D.C.; Ray Haertel, Central Oregon Community College; Philip Haverstick, Ft. Belvoir, Virginia; Walt Hillman, California State Polytechnic College, Pomona; Frank C. Kost and David E. Manes, State University College, Oneonta, New York (jointly); Lew Kowarski, Morgan State College, Maryland; J. F. Leetch, Bowling Green State College, Ohio; Peter A. Lindstrom, Genesee Community College, New York; George A. Novacky, Jr., St. Anselm High School, Swissdale, Pa.; C. P. A. Peck, State College, Pennsylvania; B. E. Rhoades, Indiana University; Steve Rohde, Lehigh University; E. F. Schmeichel, Itasca, Illinois; Benjamin L. Schwartz, McLean, Virginia; Donna J. Seaman, Olympic College, Washington; E. P. Starke, Plainfield, New Jersey; Raymond E. Whitney, Lockhaven State College, Pennsylvania; Robert L. Young, Cape Cod Community College, Massachusetts; and the proposer.

A Nine Point Problem

745. [November, 1969] *Proposed by Sidney H. L. Kung, Jacksonville University, Florida.*

Given any nine points in a unit square, show that among all the triangles having vertices on the given points there exists at least one triangle whose area does not exceed 1/8. Generalize this result.

Solution by Benjamin L. Schwartz, McLean, Virginia.

It is easy to see that any three points in a rectangle bound a triangle whose area does not exceed half of the rectangle. (The rectangle can be shrunk, if necessary, until it circumscribes the triangle. Then it can be subdivided into smaller rectangles so that the given triangle is seen to cover half, or less, of each such subrectangle.)

Next: Divide the unit square into quarters by bisecting it horizontally and vertically. At least one of the four quadrants (including boundaries) must contain 3 of the given nine points. From the lemma, the desired result now follows.

Generalization: Among a set of $2pq+1$ points in the unit square, there must be three points which define a triangle of area not exceeding $1/(2pq)$. Proof: Divide the unit square into a pq congruent rectangle by $p-1$ equally spaced horizontal lines and $q-1$ equally spaced vertical lines. Apply the lemma.

Also solved by Gary L. Britton, University of Wisconsin; Frank M. Eccles, Phillips Academy, Andover, Massachusetts; Michael Goldberg, Washington, D.C.; E. F. Schmeichel, College of Wooster, Ohio; Tej N. Tiwari, University of California, San Diego; Charles W. Trigg, San Diego, California; and the proposer.

Sum of Two Means

746. [November, 1969] *Proposed by Murray S. Klamkin, Ford Scientific Laboratory, and Morris Morduchow, Polytechnic Institute of Brooklyn.*

Determine the extreme values of $S_1/r+S_2/(n-r)$ where n is a fixed integer, $S_1=p_1+p_2+\cdots+p_r$,

$$S_1+S_2=\sum_{i=0}^{n-1} i,$$

and the p's are distinct integers in the interval $[0, n-1]$.

Solution by L. Carlitz, Duke University.

We may assume without loss of generality that $r\leqq n/2$. We shall show that

$$(*)\qquad \frac{1}{2}(n+2r-2)\leqq A\leqq\frac{1}{2}(3n-2r-2)\qquad\left(1\leqq r\leqq\frac{n}{2}\right),$$

where $A=S_1/r+S_2/(n-r)$.

Proof. If

$$\begin{aligned}S_1&=0+1+\cdots+(r-1)=\tfrac{1}{2}r(r-1),\\ S_2&=r+(r+1)+\cdots+(n-1)=\tfrac{1}{2}n(n-1)-\tfrac{1}{2}r(r-1)\\ &=\tfrac{1}{2}(n-r)(n+r-1),\end{aligned}$$

then

$$(1)\qquad A=\tfrac{1}{2}(n+2r-2).$$

If

$$\begin{aligned}S_1&=(n-r)+(n-r+1)+\cdots+(n-1)\\ &=\tfrac{1}{2}n(n-1)-\tfrac{1}{2}(n-r)(n-r-1)\\ &=\tfrac{1}{2}r(n-2r-1),\\ S_2&=0+1+\cdots+(n-r-1)=\tfrac{1}{2}(n-r)(n-r-1),\end{aligned}$$

then

$$(2) \qquad A = \tfrac{1}{2}(3n - 2r - 2).$$

Now let

$$(3) \qquad S_1 = p_1 + \cdots + p_r \ (p_1 < p_2 < \cdots < p_r), S_1 + S_2 = \sum_{k=0}^{n-1} k,$$

where $p_1, p_2, \cdots p_r$ are any r distinct numbers in $[0, 1, \cdots n-1]$. Assume that the corresponding value of A satisfies (*). Let $a \in S_1$, $b \in S_2$ and put

$$S_1' = S_1 - a + b, \qquad S_2' = S_2 + a - b.$$

Then

$$A' = \frac{1}{r}(S_1 - a + b) + \frac{1}{n-r}(S_2 + a - b) = A + \frac{n-2r}{r(n-r)}(b-a).$$

Hence if $b > a$ it follows that $A' \geqq A$ (with strict inequality except when $n = 2r$)· Thus setting out with

$$S_1 = 0 + 1 + \cdots + (r-1), \quad S_2 = r + (r+1) + \cdots + (n-1),$$

it is clear that after a number of interchanges $a \leftrightarrow b$ we get S_1, S_2 as in (3) and that the corresponding A satisfies

$$A \geqq \tfrac{1}{2}(n + 2r - 2).$$

Similarly, starting with

$$S_1 = (n-r) + (n-r+1) + \cdots + (n-1),$$
$$S_2 = 0 + 1 + \cdots + (n-r-1),$$

then again after a number of interchanges $a \leftrightarrow b$ we get S_1, S_2 as in (3) and that the corresponding A satisfies

$$A \leqq \tfrac{1}{2}(3n - 2r - 2).$$

Also solved by Michael Goldberg, Washington, D.C.; Thomas Hun-tak Lee, Swarthmore College; E. F. Schmeichel, Itasca, Illinois; Benjamin L. Schwartz, McLean, Virginia; and Kenneth M. Wilke, Topeka, Kansas.

Comment on Problem 699

699. [September, 1968, and March, 1969] *Proposed by James G. Seiler, San Diego City College.*

Find an oblique Heronic triangle with sides of three nonzero digits each, such that the nine digits involved are distinct. A Heronic triangle is defined as one that has integers for the lengths of the sides and also an integer representing the area.

Comment by Charles W. Trigg, San Diego, California.

Two Pythagorean triangles with a leg of one equal to a leg of the other may be combined to form a Heron triangle. There are 828 Pythagorean triangles, primitive and nonprimitive, with three-digit hypotenuses. The hypotenuses of 611 of these have three distinct digits. Upon pairing the triangles with equal legs and discarding those where duplicate digits appear, the 18 tabulated Heron triangles with distinct perimeter digits are obtained. They are listed according to the digits missing from the perimeter. Five of them have 0 missing (as required in the proposal) and two involve all ten digits. a is the common leg of the two Pythagorean triangles and the altitude of the Heron triangle; c and c' are the hypotenuses become sides of the Heron triangle. The third sides of the Pythagorean triangles, b and b', sum to the third side, d, of the Heron triangle which has area A.

Missing Digit	a	c	c'	b +	b' =	d	A
1	480	492	730	108	550	658	157920
2	168	357	410	315	374	689	57876
3	195	507	291	468	216	684	66690
3	420	609	427	441	77	518	108780
3	252	615	420	561	336	897	113022
3	420	609	541	441	341	782	164220
6	120	218	409	182	391	573	34380
9	63	287	105	280	84	364	11466
9	144	306	145	270	17	287	20664
9	260	701	325	651	195	846	109980
9	720	801	725	351	85	436	156960
0	120	725	169	715	119	834	50040
0	540	612	549	288	99	387	104490
0	264	814	265	770	23	793	104676
0	660	715	692	275	208	483	159390
0	756	819	765	315	117	432	163296
—	232	493	857	435	825	1260	146160
—	540	675	829	405	629	1034	279180

Comment on Q459

Q459. [September, 1969] Find the fallacy in the following:

(a)
$$x^2 = (x)(x) = x + x + \cdots + x$$
(a total of x addends)

(b)
$$\frac{d(x^2)}{dx} = \frac{d(x + x + \cdots + x)}{dx}$$
(a total of x addends)

(c)
$$2x = 1 + 1 + \cdots + 1$$
(a total of x addends)

(d) $2x = x$

(e) $2 = 1.$

[*Submitted by Steven R. Conrad*]

Comment by Sidney Spital, California State College at Hayward.

Answer A459, based on the nondifferentiability of the right hand side, is incomplete since it ignores its variation with number of addends. This is brought out by an analogous fallacy which does not require differentiability:

$$n^2 = n + n + \cdots + n, \qquad \text{integer } n \geqq 0$$
$$(n \text{ addends})$$
$$\Delta n^2 = \Delta(n + n + \cdots + n)$$
$$2n + 1 = 1 + 1 + \cdots + 1 = n$$
$$1 = 0 \quad \text{by letting } n = 0.$$

QUICKIES

From time to time this department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solution and the source, if known.

Q477. Polar coordinates (r, θ) of a point with rectangular coordinates (x, y) are often determined incorrectly by the formulas $r = \sqrt{x^2 + y^2}$ and $\theta = \arctan y/x$. Find a correct formula for θ assuming that $r > 0$.

[*Submitted by Leroy F. Meyers*]

Q478. Can $\sqrt{\sin\theta}$ and $\sqrt{\cos\theta}$ both be rational for some θ in the open interval $(0, \pi/2)$?

[*Submitted by Norman Schaumberger*]

Q479. The probability of winning a dollar in a game is p. A player who has n dollars decides to wager one dollar n times. What is the expected value of his winnings?

[*Submitted by John Howell*]

Q480. Prove that $8^n \pm 1$ is composite for every integer $n > 1$.

[*Submitted by E. F. Schmeichel*]

Q481. Let T be a normal subgroup of a finite group G. If g is any element of G and k is the smallest nonnegative integer such that g^k is contained in T, prove that k divides the order of g.

[*Submitted by Erwin Just*]

(Answers on p. 124)

ANSWERS

A477. Since

$$\tan \theta/2 = \frac{\sin \theta}{1 + \cos \theta} = \frac{y/r}{1 + (x/r)} = \frac{y}{r + x}$$

and

$$\tan \theta/2 = \frac{1 - \cos}{\sin \theta} = \frac{1 - (x/r)}{y/r} = \frac{r - x}{y}$$

we obtain

$$\theta = 2 \arctan \frac{y}{r + x} = 2 \arctan \frac{r - x}{y}$$

where the values of θ differ, correctly, by integral multiples of 2π, not π.

A478. No. Suppose $\sqrt{\sin \theta} = a/b$ and $\sqrt{\cos \theta} = c/d$ where a, b, c and d are positive integers then

$$\sin^2 \theta + \cos^2 \theta = a^4/b^4 + c^4/d^4$$

and using

$$\sin^2 \theta + \cos^2 \theta = 1$$

we have $(bd)^4 = (ad)^4 + (bc)^4$. The last equation is false because there are no positive integral solutions of $x^4 + y^4 = z^4$.

A479. The expected value is $n(2p-1)$ because

$$\sum_{x=0}^{n} (2x - n) \binom{n}{x} p^n q^{n-x} = 2pn - n$$
$$= n(2p - 1).$$

A480. If $n > 1$ we have $2^n \pm 1$ divides

$$2^{3n} \pm 1 = 8^n \pm 1$$

and

$$1 < 2^n \pm 1 < 8^n \pm 1.$$

A481. Let $n = o(g)$. There exist integers m and r, $0 \leqq r < k$, such that $n = mk + r$. Therefore, $e = g^n = g^{mk+r} = (g^k)^m g^r$, which yields $g^r = (g^k)^{-m}$. Since $g^r \in T$ it follows that $g \in T$. However, the minimality condition on k requires that $r = 0$. Thus $n = mk$.

(Quickies on p. 174)

THE CARUS MATHEMATICAL MONOGRAPHS

The Monographs are a series of expository books intended to make topics in pure and applied mathematics accessible to teachers and students of mathematics and also to non-specialists and scientific workers in other fields.

These numbers are currently available:

1. *Calculus of Variations,* by G. A. Bliss.
2. *Analytic Functions of a Complex Variable,* by D. R. Curtiss.
3. *Mathematical Statistics,* by H. L. Rietz.
4. *Projective Geometry,* by J. W. Young.
6. *Fourier Series and Orthogonal Polynomials,* by Dunham Jackson.
7. *Vectors and Matrices,* by C. C. MacDuffee.
8. *Rings and Ideals,* by N. H. McCoy.
9. *The Theory of Algebraic Numbers,* by Harry Pollard.
10. *The Arithmetic Theory of Quadratic Forms,* by B. W. Jones.
11. *Irrational Numbers,* by Ivan Niven.
12. *Statistical Independence in Probability, Analysis and Number Theory,* by Mark Kac.
13. *A Primer of Real Functions,* by Ralph P. Boas, Jr.
14. *Combinatorial Mathematics,* by H. J. Ryser.
15. *Noncommutative Rings,* by I. N. Herstein.

One copy of each Carus Monograph may be purchased by individual members of the Association for $3.00 each.

Orders with remittance should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

Additional copies and copies for nonmembers of Monographs 1-4, 6-8 are priced at $6.00 each and are available from the Open Court Publishing Company, La Salle, Illinois 61301.

Similarly, Monographs 9-15 are available at $6.00 each from John Wiley and Sons, 605 Third Avenue, New York, N. Y. 10016.