VOL. XXIV, January–February, 1951

MATHEMATICS magazine

MATHEMATICS MAGAZINE

Formerly National Mathematics Magazine, founded by S. T. Sanders.

Address editorial correspondence to Glenn James, special papers to the editors of the departments for which thev are intended, and general papers to some member of the Executive Committee.

Manuscripts should be typed on 8½″ x 11″ paper, double-spaced with 1″ margins. We prefer that, in technical papers, the usual introduction be preceded by a *foreword* which states in simple terms what the paper is about.

The Mathematics Magazine is published at Pacoima, California by the managing editor, bi-monthly except July-August. Ordinary subscriptions are $3.00, sponsoring subscriptions $10.00, single copies 65¢. Reprints: 25 free to authors; 100 @ $1.25 per page with a minimum of $6.00. All reprints are bound.

Subscriptions and related correspondence should be sent to Inez James, 14068 Van Nuys Blvd., Pacoims, Calif.

Advertisers should contact Homer V. Craig, Univ. of Texas, Austin 12, Texas.

Entered as second-class matter March 23, 1948 at the Post Office, Pacoima, California under act of congress of March 8, 1876.

SPONSORING SUBSCRIBERS

Hugh M. Allen
G. A. Baker
F. F. Beckenbach
H. W. Becker
Clifford Bell
Russell M. Caldwell
H. V. Craig
Joseph W. Creely
Paul H. Daus
Alexander Ebin
Theodore M. Fdison
Henry F. Fettis
Curtis M. Fulton
Harry M. Gehman
J. Ginsburg
Merton T. Goodrich
Harold Grant
M. B. Hestenes
Robert B. Herrera
Jerome Hines
Donald H. Hyers
C. J. Jaeger
Glenn James
Robert C. James
John Kronsbein
Lillian R. Lieber
Clyde Love
A. D. Michal
L. R. Robinson
H. A. S. Sharma
C. N. Shuster
H. A. Simmons
M. E. Stark
D. Victor Steed
Paul D. Thomas
F. M. Tingley
H. S. Vandiver

CONTENTS

OUR CONTRIBUTORS

(Continued from back cover)

Jerome Hines, well known singer with the Metropolitan Opera Company, wrote his paper, appearing in this issue, while an undergraduate at the University of California at Los Angeles. While in college he majored in both chemistry and mathematics. He began his music studies at the age of sixteen and had already appeared with the philharmonic when he took his A.B. degree. While doing graduate work in the university he sang in several recitals with the New Orleans Opera. Mr. Hines won the Metropolitan $1000 Caruso award and has been with the Metropolitan since 1946-47. He has more than 30 operatic roles in his repertoire, including that of Swallow which he created at the Metropolitan premiere of "Peter Grimes". Despite the crowded life of a Metropolitan star Mr. Hines manages to continue his studies in mathematics, in which he became especially interested while in college. He is now working on the theory of general operators.

Sketches of the other authors of articles in the present number will appear in the next issue.

THE ASSOCIATION OF TEACHERS OF MATHEMATICS IN NEW ENGLAND

will sponsor

THE INSTITUTE FOR TEACHERS OF MATHEMATICS

August 23-30

at CONNECTICUT COLLEGE, NEW LONDON, CONN.

If you wish a program, please drop a card to

Charles H. Mergendahl Newton High School Newtonville, 60, Mass.	or	Katherine E. O'Brien Deering High School Portland, Maine

CONES AND THE DECOMPOSITION OF FUNCTIONALS*

R. Pierce

Grosberg and Krein [3][1] have obtained necessary and sufficient conditions for a certain decomposition of linear functionals[2] on a partially ordered, normed linear space [see 2]. The proof however depends on a lemma which seems to be accessible only to readers of Russian [1]. It is the purpose of this paper to give an independent proof of the lemma.

Let E be a normed linear space over the real numbers with zero denoted θ and norm denoted $\|x\|$. We assume that E has a subset K (a *cone*) with the following properties:

1°. $$x \epsilon K,\ y \epsilon K \Rightarrow x + y \epsilon K,$$

2°. $$x \epsilon K \text{ and } \lambda \geq 0 \Rightarrow \lambda x \epsilon K,$$

3°. $$x \epsilon K \text{ and } x \neq \theta \Rightarrow -x \bar{\epsilon} K.$$

Moreover, for the purpose of this lemma, we will assume that the following is satisfied:

4°. *$u \epsilon K$ exists satisfying $\|u\| = 1$ and such that the set of x satisfying $u - x \epsilon K$ and $u + x \epsilon K$ is contained in the closed unit sphere $\|x\| \leq 1$ and contains the open sphere $\|x\| < 1$.*

Define: $x > y$ if $x \neq y$ and $x - y \epsilon K$.

With our assumptions, we easily verify

a. $x > y$ and $y > z \Rightarrow x > z$;

b. $x > y$ and $y > x$ is impossible;

c. $x > y$ and $\lambda > 0 \Rightarrow \lambda x > \lambda y$;

d. $x > y$ and $x' > y' \Rightarrow x + x' > y + y'$;

e. Every $x \epsilon E$ has the form $x = y - z$ for some y and z in K.

Of these assertions, only the last needs proof. We note that postulate 4° can be written in the form:

$$[x\epsilon E \mid \|x\| < 1] \subseteq [x \epsilon E \mid -u \leq x \leq u] \subseteq [x \epsilon E \mid \|x\| \leq 1].$$

Now if x is arbitrary, $\left\|\frac{x}{\|2x\|}\right\| < 1$ so $-u \leq \frac{x}{\|2x\|} \leq u$ according to 4°. Thus by *c*, $-u \cdot 2\|x\| \leq x \leq u \cdot 2\|x\|$ and $x = \frac{u \cdot 2\|x\| + x}{2} - \frac{u \cdot 2\|x\| - x}{2}$.

*This article is based on material presented to Professor Michal's Seminar on Abstract Spaces at the California Institute of Technology.

1. Numbers in brackets refer to the bibliography.

2. By a linear functional, we mean an additive, homogeneous, continuous, real valued function of the elements of E. See [2].

Let E^* denote the space of real linear functionals f over E with norm denoted $\|f\|$ [see 2]. We define a subset K^* of E^* by saying that $f \epsilon K^*$ if $f(x) \geq 0$ for all $x \geq \theta$. We easily show that K^* has the properties 1^0, 2^0, 3^0 postulated above for K; hence a partial ordering of E^* can be defined precisely as it is above for E.

Following Krein, we say that a linear functional admits a canonical decomposition if it can be written in the form

$$f = g - h$$

where $g \geq 0$ and $h \geq 0$ and

$$\|f\| = \|g\| + \|h\|.$$

The result which we wish to prove can be stated as

Theorem: *If the set $K \subseteq E$ has the properties 1^0, 2^0, 3^0 and 4^0, then every linear functional in E^* admits a cononical decomposition.*

The proof of this result follows from several lemmas. Without explicitly stating it, we will always assume that conditions 1^0, 2^0, 3^0, 4^0 are satisfied.

Lemma 1: *Let f be a positive [in the sense $f(x) \geq 0$ for all $x \geq \theta$], real valued, additive, homogeneous function of degree one on E. Then f is linear and $\|f\| = f(u)$.*

Proof: $\|x\| \leq 1 \Rightarrow$ for any $\epsilon > 0$, $-u < \frac{x}{1+\epsilon} < u \Rightarrow$ for any $\epsilon > 0$,

$-f(u) < \frac{f(x)}{1+\epsilon} < f(u) \Rightarrow$ for any $\epsilon > 0$, $|f(x)| < f(u)\,[1+\epsilon] \Rightarrow |f(x)| \leq f(u)$. Since this holds for all $\|x\| \leq 1$, f is bounded on the unit sphere and hence is continuous. Also $\|f\| \leq f(u)$. But since $\|u\| = 1$, $f(u) \leq \|f\|$ and the result is established.

Now for $x \epsilon K$, we define a real valued function p by the following:

$$p(x) = \sup_{\theta \leq y \leq x} f(y), \quad \text{wherever } f \epsilon E^*.$$

Lemma 2: *If $x \epsilon K$, $p(x) \geq 0$, $p(x) \geq f(x)$.*

Lemma 3: *If $\lambda \geq 0$, $x \epsilon K$, $p(\lambda x) = \lambda p(x)$*

Proof: If $\lambda = 0$, the result is evident. Let $\lambda > 0$. Then $p(\lambda x) = \sup_{\theta \leq y \leq \lambda x} f(y) = \lambda \sup_{\theta \leq \frac{y}{\lambda} \leq x} f\left(\frac{y}{\lambda}\right) = \lambda p(x)$.

Lemma 4: *If x, $y \epsilon K$, then $p(x+y) \geq p(x) + p(y)$*

Proof: $p(x+y) = \sup_{\theta \leq z \leq x+y} f(z)$. Now if $\theta \leq x' \leq x$ and $\theta \leq y' \leq y$, $\theta \leq x' + y' \leq x + y$. Hence

$$\sup_{\theta \le z \le x+y} f(z) \ge \sup_{\theta \le x' \le x,\ \theta \le y' \le y} f(x'+y') = \sup_{\theta \le x' \le x} f(x') + \sup_{\theta \le y' \le y} f(y') = p(x) + p(y).$$

Lemma 5: $p(u) = \dfrac{\|f\| + f(u)}{2}$

Proof: $2p(u) - f(u) = 2 \sup_{\theta \le x \le u} f(x) - f(u) = \sup_{\theta \le x \le u} f(2x - u)$.

If $\theta \le x \le u$, then $-u \le 2x - u \le u$. But conversely, if $-u \le y \le u$, then $y = 2\left(\dfrac{y+u}{2}\right) - u$ where $\theta \le \dfrac{y+u}{2} \le u$. Hence

$$2p(u) - f(u) = \sup_{-u \le y \le u} f(y) = \sup_{-u \le y \le u} |f(y)| \le \sup_{\|y\| \le 1} |f(y)| = \|f\|.$$

On the other hand, if $\epsilon > 0$, $\|x\| \le 1 \Rightarrow -u \le \dfrac{x}{1+\epsilon} \le u$ so that if $\|x\| \le 1$, $\dfrac{|f(x)|}{1+\epsilon} \le \sup_{-u \le y \le u} |f(y)|$ and thus $\dfrac{\|f\|}{1+\epsilon} \le \sup_{-u \le y \le u} |f(y)|$.

Since ϵ was arbitrary, $\|f\| \le \sup_{-u \le y \le u} |f(y)|$.

Now define a linear functional g_0 on the linear subspace L_0 of E, consisting of the points of the form λu [λ real], by putting

$$g_0(\lambda u) = \lambda \frac{\|f\| + f(u)}{2}.$$

Then g_0 is clearly linear on L_0 and by lemmas 3 and 5 $g_0(x) = p(x)$ for $x \epsilon L \cap K$.

Lemma 6: *There exists an extension g of g_0 onto all of E such that*
(a) g is linear (i.e. additive and homogeneous of degree 1),
(b) $g(x) \ge p(x)$ for $x \epsilon K$.

Proof: The subspaces of E which are extensions of L_0 satisfying conditions (*a*) and (*b*) of lemma 6 evidently constitute a partially ordered set [under inclusion] where every chain has an upper bound. Hence Zorn's maximal principle is applicable and L_0 is contained in a maximal linear subspace L of E where g_0 has an extension g to L such that g is linear and $g(x) \ge p(x)$ for $x \epsilon L \cap E$. We wish to prove that $L = E$.

Suppose $L \ne E$ and let $x_0 \epsilon E - L$. Let L_1 be the linear space of all points of the form $w = x + \lambda x_0$ where λ is real and $x \epsilon L$. The representation of any w in L_1 is clearly unique. We will show that g can be extended to a linear g_1 on L_1 which also satisfies $g_1(x) \geq p(x)$ if $x \epsilon L_1 \cap K$.

Suppose y and z are points of L such that $z + x_0$ and $y - x_0$ are elements of K. Then $y + z = (y - x_0) + (z + x_0) \epsilon K$.

$$g(y) + g(z) = g(y + z) \geq p(y + z) \qquad \text{by } (b) \text{ [for } g]$$

$$p(y + z) = p[(y - x_0) + (z + x_0)] \geq p(y - x_0) + p(z + x_0)$$

by lemma 4. Hence

(1) $$p(z + x_0) - g(z) \leq -p(y - x_0) + g(y).$$

Let

(2) $$m = \sup_{z + x_0 \epsilon K;\ z \epsilon L} [p(z + x_0) - g(z)]$$

(3) $$M = \inf_{y - x_0 \epsilon K,\ y \epsilon L} [-p(y - x_0) + g(z)]$$

Then by (1), $m \leq M$. Choose r_0 with $m \leq r_0 \leq M$. Define, for $w = x + \lambda x_0$, $x \epsilon L$, λ real,

$$g_1(w) = g(x) + \lambda r_0.$$

This definition is unique since w has precisely one representation, as we showed above. Obviously g is linear and defined on all L_1. Finally, if $w \epsilon K$, that is, $x + \lambda x_0$ is in K, then either

1) $x \epsilon K$, $\quad \lambda = 0$

2) $(x/\lambda) + x_0 \epsilon K$, $\quad \lambda > 0$

3) $-(x/\lambda) - x_0 \epsilon K$, $\quad \lambda < 0$

In case 1), $g_1(w) = g(x) = g(w) \geq p(w)$.

In case 2), $p[(x/\lambda) + x_0] - g(x/\lambda) \leq m \leq r_0$ by equation (2) so that by lemma 3,

$$p(x + \lambda x_0) - g(x) \leq \lambda r_0 \qquad \text{or } p(w) \leq g_1(w).$$

In case 3), by equation (3):

$$-p[-(x/\lambda) - x_0] + g[-(x/\lambda)] \geq M \geq r_0,$$

so again by lemma 3

$$(1/\lambda)p(x + \lambda x_0) - (1/\lambda)g(x) \geq r_0, \qquad \text{or, since } \lambda < 0,$$

$$p(x + \lambda x_0) - g(x) \leq \lambda r_0.$$

Thus $p(w) \leq g_1(w)$ in this case also.

Because g_1 is an extension of g to L_1 which properly contains L, and since g_1 satisfies (a) and (b) of lemma 6, we have contradicted the fact that L was maximal. This means that $L = E$, as was to be proved.

Proof of the main theorem.

According to lemma 6, a real valued function g exists such that

i) g is defined on all E,

ii) g is linear,

iii) $g(u) = \dfrac{\|f\| + f(u)}{2}$,

iv) $g(x) \geq p(x)$ for $x \epsilon K$.

From iv) and lemma 2, it follows that $g(x) \geq 0$ for all $x \epsilon K$ and also $g(x) \geq f(x)$ for $x \epsilon K$. Then by lemma 1, g is continuous and

$$\|g\| = g(u) = \frac{\|f\| + f(u)}{2}.$$

Letting $h = g - f$, $h(x) = g(x) - f(x) \geq 0$ for $x \epsilon K$. Thus again by lemma 1,

$$\|h\| = h(u) = \frac{\|f\| - f(u)}{2}.$$

Clearly $g \epsilon K^*$, $h \epsilon K^*$,

$$f = g - h,$$

and

$$\|g\| + \|h\| = \frac{\|f\| + f(u)}{2} + \frac{\|f\| - f(u)}{2} = \|f\| .$$

This completes the proof.

As an application of this theorem, we consider the complete linear space E of all continuous real valued functions $x(t)$ defined on the closed interval $[0,1]$ with a norm defined by

$$\|x\| = \max_{0 \leq t \leq 1} |x(t)|.$$

The cone K is the set $[x \epsilon E \mid x(t) \geq 0 \text{ all } t]$. We easily verify that 1°, 2°, 3° and 4° are satisfied with the unit u being the constant function equal to 1 for all t. It can be shown that the linear functionals on E are precisely the Riemann - Stieltjes integrals

$$f_g(x) = \int_0^1 x(t)dg,$$

g being a real valued function, independent of x, having total variation on [0, 1] equal to $\|f\|$ [see 2, p. 61].

Then according to the theorem which we have just proved, we can write

$$f_g(x) = f_h(x) - f_k(x) = \int_0^1 x(t)\, dh - \int_0^1 x(t)\, dk,$$

where f_h and f_k are positive in the partial ordering of E [which means that h and k are non-decreasing functions], and

$$\underset{0 \le t \le 1}{\text{variation}}\ g(t) = \underset{0 \le t \le 1}{\text{variation}}\ h(t) + \underset{0 \le t \le 1}{\text{variation}}\ k(t)$$

By approximating the step function

$$\xi_v(t) = \begin{matrix} 1 & \text{for} & 0 \le t \le v \\ 0 & \text{for} & v < t \le 1 \end{matrix}$$

with continuous functions $x_n \epsilon E$, we can deduce the classical Jordan decomposition theorem for functions of bounded variation:

$$g(t) = h(t) - k(t).$$

Bibliography

1. N. Achyeser and M. Krein, *On Some Questions of the Theory of Moments*, (Russian), 1930.

2. S. Banach, *Théorie des Opérations Linéaires*, 1932.

3. J. Grosberg and M. Krein, *Sur la Décomposition des Fonctionelles en Composantes Positives*, C. R. (Doklady), vol. XXV (1939), pp. 723-726.

California Institute of Technology

COLLEGIATE ARTICLES

Graduate training not required for reading

ON APPROXIMATING THE ROOTS OF AN EQUATION BY ITERATION

Jerome Hines

1. Introduction: Let

$$g(x) = 0 \tag{1}$$

be the equation whose roots we wish to approximate. Assuming that x_r is the root that we seek, our nth approximation, x_n, must be a function of the $(n - 1)$th approximation, *i.e.*

$$x_n = f(x_{n-1}) \tag{2}$$

The number of forms that $f(x_{n-1})$ may have for a given equation is unlimited. We shall determine certain of these which are especially effective and which seem to be new.

In order that $\lim_{n \to \infty} x_n$ be x_r, the approximating equation (2) must be algebraically equivalent to equation (1), *i.e.* $x = f(x)$ must have the same roots as $g(x) = 0$.

In Fig. 1 we see that the intersection of $y = f(x)$ and $y = x$ is the corresponding root, x_r, of the equation $g(x) = 0$. Let the error, $x_n - x_r$, be Δx_n. The corresponding error in the ordinate, Δy_n, is Δx_{n+1}.

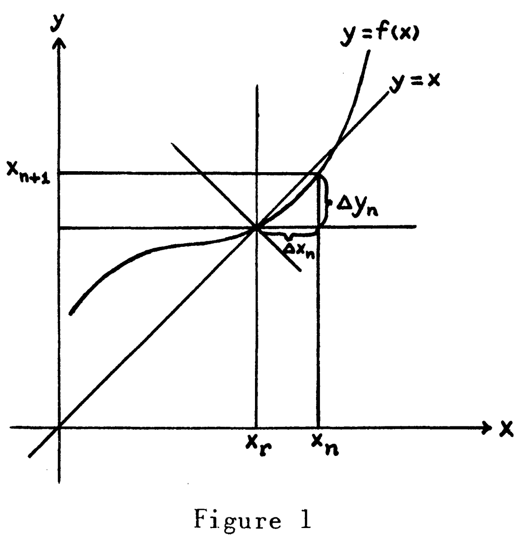

Figure 1

Sufficient conditions for convergence are

$$|\Delta x_{n+1}| < |\Delta x_n| \,, \text{ for all } n,$$

and

$$\lim_{n \to \infty} \Delta x_n = 0.$$

The first condition is equivalent to

$$\left| \frac{\Delta y_n}{\Delta x_n} \right| < 1.$$

(we assume $\Delta x_n \neq 0$ since then x_n would be the root x_r)
This is true as long as we are in a neighborhood about the root in which $y = f(x)$ crosses $y = x$ once only or touches it.

If

$$0 < \frac{\Delta y_n}{\Delta x_n} < 1,$$

then x_n is always greater than x_r in the prescribed neighborhood, thus approaching it from one side. If, however,

$$-1 < \frac{\Delta y_n}{\Delta x_n} < 0$$

we see that x_n is always opposite in sign to x_r, and x_n approaches x_r from both sides.

One simple method that suggests itself is to formally write

$$x = x + \frac{g(x)}{h(x)} \tag{3}$$

where $h(x)$ is an arbitrary function of x and $h(x) \neq 0$ in the neighborhood of the root of (1) under consideration. Numerous other forms can be written down at leisure but we will further investigate (3) since it is an interesting generalization of Newton's method.

2. Accelerated Convergence of a Sequence Approximating a Root: Referring again to Fig. 1, we see that if $y = f(x)$ were a straight line of slope zero, for any x_n, $\Delta y_n = \Delta x_{n+1} = 0$, and we would have the root. Hence we impose the condition that as many derivatives of $f(x)$ as feasible be zero at x_n thus making $y = f(x)$ lie closer to the line $y = x$ in the neighborhood of the root. For our first accelerating condition we will consider making $f'(x_n) = 0$, and will assume $g'(x)$ is different from zero in the neighborhood including the root.

Applying this condition to equation (3), with $h(x) = a$, an arbitrary constant,

$$f'(x) = 1 + \frac{g'(x)}{a}$$

When $x = x_n$, where $f'(x_n) = 0$,

$$a = -g'(x_n)$$

and

$$f(x) = x - \frac{g(x)}{g'(x_n)}$$

whence, as an approximation method,

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} \qquad (4)$$

which is *Newton's method of approximation*.

Applying the same condition to equation (3), with $h(x) = ap(x)$

$$f'(x_n) = \frac{ap(x_n)^2 + p(x_n)g'(x_n) - p'(x_n)g(x_n)}{ap(x_n)}$$

$$= 0$$

whence

$$a = \frac{p'(x_n)g(x_n) - p(x_n)g'(x_n)}{p(x_n)^2}$$

and, as an approximation formula,

$$x_{n+1} = x_n - \frac{p(x_n)g(x_n)}{p(x_n)g'(x_n) - g(x_n)p'(x_n)} \qquad (5)$$

where $p(x_n)g'(x_n) - g(x_n)p'(x_n)$ is not equal to zero, and $p(x)$ is otherwise arbitrary.

An effective form of this equation may be gotten by putting $p(x) = x^k$. This gives

$$x_{n+1} = x_n - \frac{x_n g(x_n)}{x_n g'(x_n) - kg(x_n)} \qquad (6)$$

where $x_n g'(x_n) - kg(x_n)$ is not equal to zero. It is obvious that this equation reduces to the Newtonian form for $k = 0$.

3. Further Acceleration of Convergence: Let us investigate the added condition of making $f''(x_n)$ equal to zero. We will use equation (3) with

$$h(x) = ae^{bx}$$

i.e.

$$f(x) = x + \frac{g(x)}{ae^{bx}}$$

$$f'(x_n) = 1 + \frac{g'(x_n) - bg(x_n)}{ae^{bx_n}} = 0$$

$$a = -\frac{g'(x_n) - bg(x_n)}{e^{bx_n}}$$

$$f''(x_n) = \frac{g''(x_n) - 2bg'(x_n) + bg(x_n)}{ag(x_n)} = 0$$

thus

$$b = \frac{g'(x_n) \pm \sqrt{g'(x_n)^2 - g(x_n)g''(x_n)}}{g(x_n)}$$

(7) $$x_{n+1} = x_n \pm \frac{|g(x_n)|}{\sqrt{g'(x_n) - g(x_n)g''(x_n)}}$$

If x_n is smaller than the root, then the positive sign should be taken, while if x_n is larger than the root, the negative sign should be used. While (7) is formally more complex, it seems to give more rapid convergence than (6). The following examples compare these methods:

a) $x^3 - 2x - 5 = 0,\quad x_1 = 2,\quad (x_r = 2.09455)$

Eq. (6)

$k = 0,\quad x_2 = 2.100$ (Newton's method)

$= 1,\quad = 2.095$

$= 2,\quad = 2.091$

Eq. (7)

$x_2 = 2.0945$

b) $x^4 + 4x^2 - 24x - 20 = 0,\quad x_1 = 2,\quad (x_r = 2.730)$

Eq. (6) $k = 0,\quad x_2 = 3.5$ (Newton's method)

$= 1,\quad = 2.9$

$= 2,\quad = 2.6$

Eq. (7)

$$x_2 = 2.80$$

4. Relative Accuracy of Results: Let us consider again $h(x) = ax^k$ in equation (5). Then

$$x_{n+1} = x_n - \frac{x_n g(x_n)}{x_n g'(x_n) - kg(x_n)} \tag{6}$$

where $k = 0$ gives Newton's method. In Fig. 2 we have plotted x_{n+1} versus k, holding x_n fixed and with $g(x)$, $g'(x)$, and $g''(x)$ all greater than zero in the neighborhood of the root. The resulting curve is a hyperbola asymptotic to $x_{n+1} = x_n$ and

$$k = \frac{x_n g'(x_n)}{g(x_n)}$$

Other configurations present no additional complications. We find that there is a range of values for k, $k > 0$, which all give better results than $k = 0$.

Since (6) approximates to the line, $x_{n+1} = x_r$, as we approach the root, x_r, we conclude that it would be advantageoousto take k larger and larger with successive approximations. Similarly it is seen that a large k is advantageous when dealing with a large root.

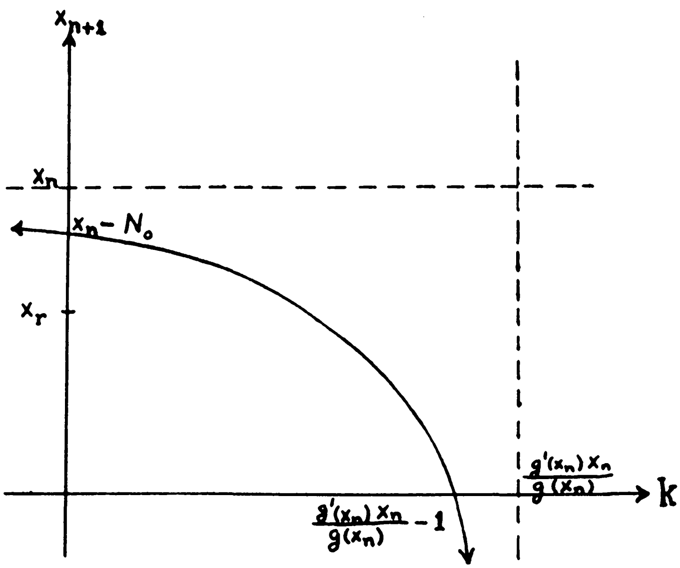

Figure 2

ON THE SUMMATION OF POWER SERIES

Eric Michalup

Recently there has been published by J. A. Pierce (1) a very interesting paper concerning some new formulae of the sum of powers, in which he refers to a paper of Ross (2) who presented formulae for the first six powers of the first n natural numbers and of the first n odd natural numbers. Niessen (3) presented some more formulae based on the calculus of finite differences.

Assuming that n and i are positive and whole numbers, so that

$$\binom{n}{i} = 0; \qquad i > n$$

we are developing according to the binomial theorem

$$(1+x)^i = 1 + \binom{i}{1} x + \binom{i}{2} x^2 + \cdots \binom{i}{i-1} x^{i-1} + \binom{i}{i} x^i .$$

Giving to x all the values from zero to n and adding these equations we get

$$\sum_{x=0}^{n} (1+x)^i = (1+n) + \binom{i}{1} \sum_{x=1}^{n} x + \binom{i}{2} \sum_{x=1}^{n} x^2 + \cdots + \binom{i}{i} \sum_{x=1}^{n} x^i$$

or

$$\sum_{1+x=1}^{n+1} x^i = \sum_{x=1}^{n} x^i + (1+n)^i = (1+n) + \binom{i}{1} \sum_{x=1}^{n} x + \cdots + \binom{i}{i} \sum_{x=1}^{n} x^i$$

then

$$(1+n)^i = (1+n) + \binom{i}{1} \sum_{x=1}^{n} x + \cdots + \binom{i}{i-1} \sum_{x=1}^{n} x^{i-1}$$

and finally

$$\sum_{x=1}^{n} x^{i-1} = \frac{1}{i} \left[(1+n)^i - (1+n) - \binom{i}{1} \sum_{x=1}^{n} x - \cdots - \binom{i}{i-2} \sum_{x=1}^{n} x^{i-2} \right]$$

the well known recursion formula for the sum of the powers of the natural numbers which is very important in connection with the use of orthogonal functions in statistics, for example. Schenker (4) presented the corresponding formulae for the first nine powers and (5) Kraitschik for the first twelve powers and Shannon (6) in a somewhat different manner. These formulae permit to obtain easily the sum of the expressions

$$(2x-1)^r; \quad (2x-n-1)^r; \quad (2x)^r \text{ and } \left(\frac{2x-n-1}{2}\right)^r .$$

We may represent the terms of an arithmetical series of the degree m by

$$a_r^{(m)} = c_m^{(m)} r^m + c_{m-1}^{(m)} r^{m-1} + \ldots + c_0^{(m)} \quad (r = 0,1,2\ldots n)$$

and the sum of the first n terms by

$$S_n^m = \sum_{r=0}^{n} a_r^m = c_m^{(m)} \sum_{r=0}^{n} r^m + \ldots\ldots + c_1^{(m)} \sum_{r=0}^{n} r + c_0^{(m)}(1 + n).$$

The coefficients $c_{m-k}^{(m)}$ are determined by the degree m of the arithmetical series and the place k, and for the sum of the powers we have to use the mentioned recursion-formula, which seems to be rather troublesome. We know that

$$a_{r+1} = a_r + \Delta a_r; \quad \Delta^i a_{r+1} = \Delta^i a_r + \Delta^{i+1} a_r$$

and easily can be proved that

$$a_r = a_o + \binom{r}{1} \Delta a_0 + \binom{r}{2} \Delta^2 a_0 + \ldots\ldots + \binom{r}{n} \Delta^n a_0$$

$$a_{r+1} = a_0 + \binom{r+1}{1} \Delta a_0 + \binom{r+1}{2} \Delta^2 a_0 + \ldots + \binom{r+1}{n} \Delta^n a_0$$

and as summation-formulae we verify without difficulties

$$s_{r-1} = \binom{r}{1} a_0 + \binom{r}{2} \Delta a_0 + \ldots + \binom{r}{n+1} \Delta^n a_0$$

formula which assumes zero the first term of the series, or the general formula

$$S_r = \binom{r}{1} a_0 + \binom{r}{2} \Delta a_0 + \ldots + \binom{r}{n+1} \Delta^n a_0.$$

Substituting the differences in this formula by the terms of the original series and remembering that $\Delta^i a_0 = a_i - \binom{i}{1} a_{i-1} + \ldots + (-1)^{i} a_0$ we get $S_r = a_0 \sum_{i=1}^{i+1} (-1)^{i+1} \binom{r}{i} \binom{i-1}{i-1} + a_1 \sum_{i=2}^{i+1} (-1)^i \binom{r}{i} \binom{i-1}{i-2} + a_2 \sum_{i=3}^{i+1} (-1)^{i-1} \binom{r}{i} \binom{i-1}{i-3} + \ldots$ which formula does not use more than the first $(n + 1)$ terms of the original series. We are not able to find a simple formula which would substitute the sums, as $\sum_{k=0}^{n} \binom{r}{n-k} \binom{s}{k} = \binom{r+s}{k}$ or the sum $\sum_{i=0}^{r} (-1)^i \binom{r}{r+i} \frac{2^i}{i+1}$ which becomes zero for r odd and is equal to $1/r$ for r even. The writer obtained this result (7) applying Lindelöf's Method of improving the convergency of power-series to the infinite series $\log_e (1 + x)$ substituting $y = x/(s + x)$ and establishing the parameter s in such a way that the coefficients of the second term of y becomes zero, and finally comparing the coefficients of the powers of the same degree, resulting by means of the development and of the well known

$$\log_e (1 + x) = 2\left[\frac{x}{2 + x} + \frac{1}{3}\left(\frac{x}{2 + x}\right)^3 + \frac{1}{5}\left(\frac{x}{2 + x}\right)^5 + \cdots \right].$$

We also may represent the sum of the powers x^r by means of Bernoulli's Polynomials.

Assuming that the signs of the terms of the power series are alternating, it is convenient to separate the positive and negative terms

$$\sum_{x=1}^{n} (-1)^x x^r = - \sum_{x=1}^{[\frac{n+1}{2}]} (2x - 1)^r + \sum_{x=1}^{[\frac{n}{2}]} (2x)^r$$

obtaining in this way the expressions already formerly considered.

So far we occupied ourselves with expressions of the class x^r and now we are going to consider two special series of the classes

$$\sum_{x=0}^{\infty} \frac{x^r}{x!} \quad \text{and} \quad \sum_{x=0}^{\infty} (-1) \frac{x^r}{x!}$$

Assuming $r = 0$, the first sum represents e and the second sum becomes equal to $1/e$. Putting in the second sum a positive and whole number for r, we get an integer multiple of $1/e$. In order to prove this we split

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx \qquad R(z) > 0$$

into the two incomplete Gamma functions

$$\Gamma(z) = P(1,z) + Q(1,z) = \int_0^1 x^{z-1} e^{-x} dx + \int_1^{\infty} x^{z-1} e^{-x} dx$$

being (8)

$$P(1,z) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!(z+n)}$$

and using ln for $\log_e$

$$Q(1,z) = \int_1^{\infty} e^{-x} e^{z \ln x} dx/x = \int_1^{\infty} \sum_{n=1} \frac{z^n}{n!} \frac{(\ln x)^n e^{-x}}{x} \frac{dx}{}$$

Introducing Hermit's substitution

$$x = e^t; \ \ln x = t; \ \frac{dx}{x} = dt$$

the limits become

x	1	∞
t	0	∞

resulting

$$Q(1,z) = \int_0^\infty \sum_{n=0}^{\infty} \frac{z^n t^n}{n!} e^{-e^t} dt = \sum_{n=0}^{\infty} \frac{z^n}{n!} \int_0^\infty t^n e^{-e^t} dt.$$

This integral can be split in the same manner as the Gamma function

$$\int_0^1 t^n e^{-e^t} dt + \int_1^\infty t^n e^{-e^t} dt.$$

Differentiating

$$\frac{d}{dt}\left[e^{-e^t}\right]_{t=0} = -e^t e^{-e^t}\Big|_{t=0} = -e^{-1}$$

we may put

$$e^{-e^t} = e^{-1} \sum_{\varkappa=0}^{\infty} \frac{c_\varkappa t^\varkappa}{\varkappa!}$$

representing $c_\varkappa$ whole numbers the values of which will be determined later on. Substituting that expression in the first integral we get an infinite series

$$e^{-1} \sum_{\varkappa=0}^{\infty} \frac{c_\varkappa}{\varkappa!} \frac{1}{n+\varkappa+1}$$

which has a pole in all the points $\varkappa = -(n+1)$, except $c_\varkappa = 0$, for instance $c_2 = 0$. Differentiating

$$\frac{d}{dt}\left[e^{-e^t}\right] = \frac{d}{dt}\left[e^{-1} \sum_{\varkappa=0}^{\infty} \frac{c_\varkappa t^\varkappa}{\varkappa!}\right]$$

results

$$-e^t e^{-e^t} = e^{-1} \sum_{\varkappa=1}^{\infty} \frac{c_\varkappa t^{\varkappa-1}}{(\varkappa-1)!} \quad ; \quad c_0 = 1$$

and on the other hand

$$-e^t e^{-e^t} = -\sum_{\mu=0}^{\infty} \frac{t^\mu}{\mu!} e^{-1} \sum_{\varkappa=0}^{\infty} \frac{c_\varkappa t^\varkappa}{\varkappa!}$$

consequently

$$-e^{-1} \sum_{\mu=0}^{\infty} \frac{t^\mu}{\mu!} \sum_{\varkappa=0}^{\infty} \frac{c_\varkappa t^\varkappa}{\varkappa!} = e^{-1} \sum_{\varkappa=0}^{\infty} \frac{c_{\varkappa+1} t^\varkappa}{\varkappa!}$$

Equalling the coefficients of $t^{\mu+\varkappa}$ results

$$-\sum \frac{c_\varkappa}{\mu!\varkappa!} = \frac{c_{\varkappa+\mu+1}}{(\varkappa+\mu)!}$$

and substituting $\varkappa + \mu = k$ the relation

$$-c_{k+1} = \sum_0^k \frac{c_\varkappa k!}{\varkappa!(k-\varkappa)!} = \sum_0^k c_\varkappa \binom{k}{\varkappa}$$

which represents a recursion formula for the determination of the coefficients $c_\varkappa$. The first ones are

$$c_1 = -1;\ c_2 = 0;\ c_3 = c_4 = 1;\ c_5 = -2;\ c_6 = c_7 = -9;\ \cdots$$

leading to the final formula

$$(A) \qquad c_\varkappa e^{-1} = \sum_{\mu=0}^{\infty} \frac{(-1)^\mu \mu^\varkappa}{\mu!}$$

For $\varkappa = 2$ we get

$$c_2 e^{-1} = -\frac{1}{1!} + \frac{4}{2!} - \frac{9}{3!} + \frac{16}{4!} - \frac{25}{5!} \cdot\cdot\cdot\cdot\cdot = 0.$$

The series at the right hand of (*A*) are consequently integer multiples of $1/e$. It may be mentioned that Fraser (9) determined the coefficients $c_\varkappa$ by means of Stirling's Numbers of second kind using the "abacus".

Bibliography

(1) "On the Summation of Progressions useful in Time Series Analysis", Journal of the American Statistical Association, September 1944, pp. 387-389; December 1944, p. 521.

(2) "Formulae for facilitating Computations in Time Series Analysis", Journal of the American Statistical Association, March 1925, pp. 75-79.

(3) "On the Summation of certain Progressions useful in Time Series Analysis", Journal of the American Statistical Association, March 1945, pp. 98-100.

(4) "Direkter Weg zur Ermittlung von Relationen zur Trendbestimmung", Journal de Statistique et Revue economique suisse, 1933, pp. 114-115.

(5) "Recherches sur la Theorie des Nombres", Paris 1924, pp. 4-5.

(6) "An alternative Method of Solution of certain fundamental Problems in the individual Theory of Risk", The Record, American Institute of Actuaries, October 1938, pp. 372-399.

(7) "Una aplicación del método de Lindelõf", presented to the Segundo Congreso Venezolano de Ingenieria, Maracaíbo, January 1945.

(8) Haskins "On the Zeros of the Function *P*(*x*), Complementary to the incomplete Gamma Function", Transactions of the American Mathematical Society, October 1915, pp. 405-412.

(9) "A Note on the Gompertz' Table", Journal of the Institute of Actuaries, London, Vol. LXXIII, Part II, No. 337, pp. 423-426.

Caracas, Venezuela

ANALYTIC FUNCTIONS RELATED TO PRIMES

R. M. Redheffer

1. *Introduction.* We show how to put certain number-theoretic problems connected with primes into analytic, or function-theoretic form. The general procedure is to construct analytic functions having special behaviour at the primes without, however, having the primes enter directly into the definition of the functions. It is thought that students knowing a little number theory as well as analysis might be interested in the ease with which the one type of problem can be converted into the other.

2. *Simple Functions.* By comparing the highest power of a prime contained in $(n - 1)!$ with the highest power contained in n, we find that n divides $(n - 1)!$ if n is any composite integer ≥ 6. Hence the function

$$p(z) = \sin \frac{\Gamma(z)}{z} \pi \tag{1}$$

has a simple zero at every composite integer ≥ 6 and no other zeros for integer values of z. In a similar manner we see, by using Wilson's theorem, that the function

$$q(z) = \sin \frac{\Gamma(z) + 1}{z} \pi \tag{2}$$

has simple zeros at the primes and no other integer zeros.[1]

In the expression

$$r_n(z) = \prod_{k=2}^{n} \sin \frac{\pi z}{k} \tag{3}$$

for $|z| \leq n$, all factors are different from zero if z is not an integer; just one factor is zero if z is a prime; and at least three factors are zero if z is composite. Hence there are simple zeros at the primes, higher order zeros at the composite integers, and no other zeros with $|z| \leq n$.

The series

$$s(z) = \sum_{k=2}^{\infty} \frac{1 - (z/k)^2}{\sin \pi z/k} \frac{e^k}{k!} \tag{4}$$

is uniformly convergent if z is bounded away from the integers, and hence represents an analytic function. Also if z is a prime, then only

[1] It has been brought to the author's attention that functions of the type (2) are actually well known; cf. Dickson's *History of Number Theory*.

one term of the sum will have a vanishing denominator, and that term will have a zero in the numerator. If z is composite, however, there is a simple pole for each term with k a factor of z. The sum of the residues for the poles of these separate terms cannot be zero, since e is transcendental. We thus conclude that $s(z)$ has simple poles at the composite integers and no other finite singularities.

3. *Primes in general.* From the above we see that the only real positive zeros of $q^2(z) + \sin^2 \pi z$ occur at the primes, and hence

$$\lim_{y \to 0} \int_{C_n} \frac{2q(z)q'(z) + \pi \sin 2\pi z}{q^2(z) + \sin^2 \pi z} dz = v(n) + 1 \tag{5}$$

if $v(n)$ is the number of primes $\leq n$. Here C_n is any simple closed contour in the right half plane containing the integers 1, 2, ... n but no others. The notation $y \to 0$ means that the contour is to contract down upon the real axis. Similarly

$$(1 - z^2) \frac{r_n(z)}{\sin \pi z} = \sum_{k=0}^{\infty} R_{nk} z^k \tag{6}$$

has its first pole at the first prime $\geq n$, so that its radius of convergence $lim\ |R_{nk}|^{-1/n}$ is equal to this function of n.

The function

$$(1 - z^2)^2 \frac{r_n(z)}{\sin^2 \pi z} \tag{7}$$

has simple poles at the primes and no others with $|z| \leq n$; the function

$$\frac{p(z)}{\sin \pi z} \tag{8}$$

has simple poles at the primes and no other singularities for $R(z) \geq 6$. On the other hand the expressions

$$\frac{q(z)}{\sin \pi z}, \quad s(z) \tag{9}$$

have simple poles at the composite integers; and the former has no other singularities in the right half plane, the latter no others in the whole plane. Problems connected with the distribution of primes, then, can be formulated in terms of the limit (5), the radius of convergence (6), or the distribution of singularities in (7) - (9).

One may modify the functions to suit the particular problem, of course; for example, if we were interested in primes of the form $m^2 + 1$ we should be led to the expression

$$\frac{p(z^2+1)}{\sin \pi z} \tag{10}$$

which has simple poles at these primes and no other singularities for $R(z) \geq 3$. The classical conjecture is thus put into analytic form, viz., to show that (10) has infinitely many poles in the right half plane. A formulation can be obtained in terms of elementary functions by the expression

$$z^2(1-z^2)\,\frac{r_n(z^2+1)}{\sin^2 \pi z}, \tag{11}$$

which has simple poles at primes $m+1$ and no other singularities in $|z| \leq \sqrt{n} - 1$. Similar analytic statements can be given for the Goldbach conjecture and for other outstanding problems in the theory of primes.

4. *Twin primes.* A question of this sort is the conjecture that there are infinitely many twin primes, that is, infinitely many prime pairs like (17,19) or (29,31) which differ by 2. From the above remarks or by inspection of the functions we see that

$$z^3\,\frac{r_n(z-1)r_n(z+1)}{\sin^3 \pi z} \tag{12}$$

has simple poles at the twin primes and no other singularities in $|z| \leq n - 1$, while

$$\frac{p(z-1)p(z+1)}{\sin \pi z} \tag{13}$$

has simple poles at the twin primes and no others for $R(z) \geq 6$. The same sort of thing can be done with $s(z-1)s(z+1)$.

To show that a function has infinitely many poles one may multiply by another function $\theta(z)$, having at most a finite number of poles, and integrate around a simple closed contour. If there are a finite number of poles the integral will eventually be constant, as the contour gets larger and larger; but if there are infinitely many poles, then there will be a $\theta(z)$ which will make the sequence of integrals diverge. Thus, a necessary and sufficient condition that there be infinitely many twin primes is that there exist an integral function $\theta(z)$ and a sequence of simple closed contours C_n free of points on the negative real axis, such that the set of numbers

$$\int_{C_n} \frac{p(z-1)p(z+1)}{\sin \pi z}\,\theta(z)\,dz$$

is unbounded.

By estimating the maximum and minimum residues, and comparing the minimum estimate for the pole having largest $|z|$ with the sum of maximum estimates for the other poles, one can choose a function $\theta(z)$ that certainly increases fast enough to give divergence if there are infinitely many poles. After carrying out these estimates of residues in (12), for example, one finds that $\phi(z) = e^{z^3}$ is good enough, with the following result: A necessary and sufficient condition that there be infinitely many twin primes is that the set of numbers

$$\int_{C_n} \frac{r_n(z - 1)r_n(z + 1)}{\sin^3 \pi z} e^{z^3} dz$$

be unbounded, if the contour C_n includes the integers, 0, ±1, ±2, ... $\pm(n - 1)$ and no others.

University of California at Los Angeles

TEACHING OF MATHEMATICS

Edited by

Joseph Seidlin and C. N. Shuster

This department is devoted to the teaching of mathematics. Thus articles on methodology, exposition, curriculum, tests and measurements, and any other topic related to teaching, are invited. Papers on any subject in which you, *as a teacher,* are interested, or questions which you would like others to discuss, should be sent to Joseph Seidlin, Alfred University, Alfred, New York.

DECOMPOSITION OF RATIONAL FRACTIONS INTO PARTIAL FRACTIONS

Alexander W. Boldyreff

Introduction.

Although the question of decomposition of rational fractions into partial fractions is both quite elementary and quite fundamentally important (e.g. in the theory of systematic integration), it has been sadly neglected in American mathematical texts.

It is true that some foreign texts devote more attention to this problem. Yet even the best of these fail to give a completely satisfactory treatment.

Therefore a brief but logically complete exposition of this subject is not out of place.

The problem presents two distinct aspects:

(1) The existence and uniqueness of the decomposition of a rational fraction into partial fractions, and

(2) The determination of the numerators of the partial fractions.

Existence and Uniqueness of Decomposition.

Let $\dfrac{f(x)}{F(x)}$ be a proper rational fraction with real coefficients and in its lowest terms.

The existence and uniqueness of decomposition of $\dfrac{f(x)}{F(x)}$ into partial fractions follows from the following two lemmas:

Lemma I. If $F(x) \equiv (x-\alpha)^n\phi(x)$, $\phi(\alpha) \neq 0$, then

$$\frac{f(x)}{F(x)} \equiv \frac{A}{(x-\alpha)^n} + \frac{\phi(x)}{(x-\alpha)^{n-1}\phi(x)}, \text{ where } A = \frac{f(\alpha)}{\psi(\alpha)} \neq 0,$$

$\psi(x)$ is unique, and $\psi(x)$ and $\phi(x)$ are relatively prime.

Proof: We have identically:

$$\frac{f(x)}{F(x)} \equiv \frac{A}{(x-\alpha)^n} + \frac{f(x) - A\phi(x)}{(x-\alpha)^n\phi(x)}$$

Let A be determined by the equality

$$f(\alpha) - A\phi(\alpha) = 0, \qquad i.e. \text{ let } A = \frac{f(\alpha)}{\phi(\alpha)}$$

This is always possible since $\phi(\alpha) \neq 0$. Also, since

$$f(\alpha) \neq 0, \qquad A \neq 0.$$

With this choice of A,

$f(x) - A\phi(x)$ must contain $x - \alpha$ as a factor:

$$f(x) - A\phi(x) \equiv (x - \alpha)\psi(x), \text{ and}$$

$$\frac{f(x)}{F(x)} \equiv \frac{A}{(x-\alpha)^n} + \frac{\psi(x)}{(x-\alpha)^{n-1}\phi(x)}$$

It may be noted that since

$$f(x) \equiv A\phi(x) + (x - \alpha)\psi(x),$$

$\psi(x)$ and $\phi(x)$ have no factor in common. Otherwise this factor would divide $f(x)$. This is impossible because $\frac{f(x)}{F(x)}$ is in its lowest terms.

To prove uniqueness, assume:

$$\frac{A_1}{(x-\alpha)^n} + \frac{\psi_1(x)}{(x-\alpha)^{n-1}\phi(x)} \equiv \frac{A_2}{(x-\alpha)^n} + \frac{\psi_2(x)}{(x-\alpha)^{n-1}\phi(x)}$$

Then

$$A_1 - A_2 \equiv \frac{\psi_2(x) - \psi_1(x)}{\phi(x)}(x - \alpha)$$

Now let

$$x = \alpha.$$

It follows that $A_1 = A_2$, and $\psi_1(x) = \psi_2(x)$.

Lemma II

If $\quad F(x) \equiv [(x-a)^2 + b^2]^m\phi(x), \quad \phi(a + bi) \neq 0, \; b \neq 0,$

then $$\frac{f(x)}{F(x)} \equiv \frac{Ax + B}{[(x-a)^2 + b^2]^m} + \frac{\psi(x)}{[(x-a)^2 + b^2]^{m-1}\phi(x)}$$

where A, B, and $\psi(x)$ are uniquely determined, A and B are not both zero, and $\psi(x)$ and $\phi(x)$ are relatively prime.

Proof:

Obviously $$\frac{f(x)}{F(x)} \equiv \frac{Ax + B}{[(x-a)^2 + b^2]^m} + \frac{f(x) - (Ax + B)\phi(x)}{[(x-a)^2 + b^2]^m\phi(x)}$$

Let A and B be so determined that

$f(a + bi) - [A(a + bi) + B]\phi(a + bi) = 0$, *i.e.* so that

$$A(a + bi) + B = \frac{f(a + bi)}{\phi(a + bi)} = \alpha + \beta i.$$

This is always possible since $\phi(a + bi) \neq 0$.

But then $A = \frac{\beta}{b}$, $B = \alpha - \frac{a\beta}{b}$

Clearly A and B cannot both be zero, because $f(a + bi) \neq 0$. With this choice of A and B

$$f(x) - (Ax + B)\phi(x) \equiv (x - a - bi)(x - a + bi)\psi(x) \equiv$$

$$\equiv [(x - a)^2 + b^2]\psi(x), \text{ and}$$

$$\frac{f(x)}{F(x)} \equiv \frac{Ax + B}{[(x - a)^2 + b^2]^m} + \frac{\psi(x)}{[(x - a)^2 + b\ ]^{m-1}\phi(x)}.$$

As in Lemma I, in this case also $\psi(x)$ and $\phi(x)$ have no common factor, and the decomposition is unique.

The Main Existence Theorem.

Let $\frac{f(x)}{F(x)}$ be a proper rational fraction with real coefficients and in its lowest terms.

If $F(x) \equiv \prod_{k=1}^{p} (x - \alpha_k)^{n_k} \cdot \prod_{r=1}^{q} [(x - a_r)^2 + b_r^2]^{m_r}$, then

$$\frac{f(x)}{F(x)} \equiv \sum_{k=1}^{p} \sum_{i=0}^{n_k - 1} \frac{A_{ki}}{(x - \alpha_k)^{n_k - i}} + \sum_{r=1}^{q} \sum_{j=0}^{m_r - 1} \frac{A_{rj}x + B_{rj}}{[(x - a_r)^2 + b_r^2]^{m_r - j}},$$

the decomposition being unique.

This theorem results from the repeated use of Lemmas I and II.

Determination of the Numerators.

It is convenient in developing the methods of determination of the numerators of the partial fractions to consider separately two cases:

Case I. The determination of the numerators of partial fractions corresponding to a real linear factor of $F(x)$;

Case II. The determination of the numerators of partial fractions corresponding to a prime quadratic factor of $F(x)$.

It is patently unnecessary to differentiate between the case when a given gactor of $F(x)$ appears raised to the first power or when it is "repeated". A so-called "distinct" factor is simply a "repeated" factor of multiplicity one.

Case I.

If $\qquad F(x) \equiv (x - \alpha)^n\phi(x), \qquad \phi(\alpha) \neq 0$, then

$$\frac{f(x)}{F(x)} \equiv \sum_{k=0}^{n-1} \frac{A_k}{(x-\alpha)^{n-k}} + \frac{\psi(x)}{\phi(x)}$$

From this

$$\frac{f(x)}{\phi(x)} \equiv \sum_{k=0}^{n-1} A_k(x-\alpha)^k + (x-\alpha)^n \frac{\psi(x)}{\phi(x)} .$$

Differentiating both sides of this identity k times and letting $x = \alpha$, we deduce

$$A_k = \frac{1}{k!}\left[\frac{d^k}{dx^k}\frac{f(x)}{\phi(x)}\right]_{x=\alpha}, \qquad k = 0, 1, \ldots, n-1.$$

Case II.

If $F(x) \equiv [(x-a)^2 + b^2]^n \phi(x), \quad \phi(a+bi) \neq 0, \quad b \neq 0,$

then $$\frac{f(x)}{F(x)} \equiv \sum_{k=0}^{n-1} \frac{A_k x + B_k}{[(x-a)^2 + b^2]^{n-k}} + \frac{\psi(x)}{\phi(x)}$$, and

$$\frac{f(x)}{\phi(x)} \equiv \sum_{k=0}^{n-1}(A_k x + B_k)[(x-a)^2 + b^2]^k + [(x-a)^2 + b^2]^n \frac{\psi(x)}{\phi(x)}$$

Let $t = (x-a)^2 + b^2$. Then

(A) $$\frac{f(x)}{\phi(x)} \equiv \sum_{k=0}^{n-1}(A_k x + B_k)t^k + t^n \frac{\psi(x)}{\phi(x)}$$

Also

$$\frac{d}{dt} \equiv \frac{1}{2(x-a)}\frac{d}{dx},$$

$$\frac{d^2}{dt^2} \equiv \frac{1}{4(x-a)^2}\frac{d^2}{dx^2} - \frac{1}{4(x-a)^3}\frac{d}{dx},$$

etc.

And by induction:

$$\frac{d^m}{dt^m} \equiv \sum_{r=1}^{m} \frac{(-1)^{r+1}\,{}_mN_r}{2^m (x-a)^{m+r-1}} \frac{d^{m-r+1}}{dx^{m-r+1}},$$ where

the numbers ${}_mN_r$ are defined by:

(1) ${}_mN_1 = 1$ for all $m \geq 1$.

(2) ${}_mN_r = 0$ for $r > m$.

(3) ${}_mN_r = {}_{m-1}N_r + (m+r-3)\,{}_{m-1}N_{r-1}$.

We now operate on both sides of (A) with $\frac{d^m}{dt^m}$, $m = 0, 1, 2, \ldots, n - 1$, and let $t = 0$, *i.e.* $x = a + bi$.

Observing that for $t = 0$,

$$\frac{d^m}{dt^m}[(A_k x + B_k)t^k] \equiv \sum_{r=0}^{m} {}_mC_r \frac{d^{m-r}}{dt^{m-r}}(A_k x + B_k) \cdot \frac{d^r}{dt^r} t^k =$$

$$= 0 \text{ for } k > m$$

$$= [A_m(a + bi) + B_m]m! \text{ for } k = m, \text{ and}$$

$$= \left[{}_mC_k \frac{d^{m-k}}{dt^{m-k}}(A_k x + B_k) \cdot k!\right]_{x=a+bi} =$$

$$= (-1)^{m-k+1} \frac{{}_mC_k \cdot {}_{m-k}N_{m-k} \cdot k!A_k}{2^{m-k}(bi)^{2(m-k)-1}}, \text{ for } k < m,$$

while

$$\frac{d^m}{dt^m}\left\{t^n \frac{\psi(x)}{\phi(x)}\right\} = 0,$$

we deduce from (A):

$$(B) \qquad \frac{f(a + bi)}{\phi(a + bi)} = A_0(a + bi) + B_0, \text{ and for } m \geq 1,$$

$$(C) \qquad \sum_{r=1}^{m} \frac{(-1)^{r+1}\,{}_mN_r}{2^m(bi)^{m+r-1}}\left[\frac{d^{m-r+1}}{dx^{m-r+1}}\frac{f(x)}{\phi(x)}\right]_{x=a+bi} =$$

$$= \sum_{k=0}^{m-1} \frac{(-1)^{m-k+1}\,{}_mC_k \cdot {}_{m-k}N_{m-k} \cdot k!A_k}{2^{m-k}(bi)^{2(m-k)-1}} + [A_m(a + bi) + B_m]m!$$

These formulas make it possible to evaluate successively A_0 and B_0, A_1 and B_1, etc. At the same time they not only represent a complete solution of the problem, but provide a practical method of evaluating the numerators of partial fractions as one can readily see by applying these formulas to particular rational fractions.

The Coefficients ${}_mN_r$.

These numbers are easily tabulated using their definition. Arranging them in a table so that ${}_mN_r$ is placed in the m^{th} row and r^{th} column we have:

1					
1	1				
1	3	3			
1	6	15	15		
1	10	45	105	105	
1	15	105	420	945	945

etc.

In practice only the first few rows of the table would be needed.

The numbers ${}_mN_r$ possess many interesting properties, resembling those of the binomial coefficients. A few of these properties are given below.

Summation by Columns

From ${}_{m-j}N_r = {}_{m-j-1}N_r + (m + r - j - 3)\ {}_{m-j-1}N_{r-1}$, letting $j = 0$, $1, \ldots, m - r$, and adding, we get

$$ {}_mN_r = \sum_{j=0}^{m-n} (m + r - j - 3)\ {}_{m-j-1}N_{r-1}\ , $$

a formula exhibiting the formation of ${}_mN_r$ from the entries of the $(r - 1)$st column of the table.

Summation by Diagonals.

From the recurrence formula we deduce readily

$$ {}_mN_r = {}_{m-1}N_r + \sum_{j=2}^{r} \left\{ \prod_{i=2}^{j} (m + r - 2i + 1) \right\}_{m-j} N_{r+1-j} $$

Summation of the Entries of any Row.

By applying the recurrence relation to each of the entries of the m^{th} row, we get

$$ \sum_{r=1}^{m} {}_mN_r = \sum_{r=1}^{m-1} (m + r - 1)\ {}_{m-1}N_r $$

The above properties are close analogues of the properties of the binomial coefficients.

Illustrative Examples.

Example 1.

$$ \frac{32}{(x - 2)(x + 2)(x^2 + 4)} = \frac{A_1}{x - 2} + \frac{A_2}{x + 2} + \frac{A_0 x + B}{x^2 + 4} $$

$$ A_1 = \frac{32}{(2 + 2)(4 + 4)} = 1. $$

$$A_2 = \frac{32}{(-2 - 2)(4 + 4)} = -1.$$

$$\frac{32}{(2i - 2)(2i + 2)} = -4 = A_0(2i) + B_0,$$

$$A_0 = 0, \qquad B_0 = -4.$$

Example 2.

$$\frac{120x + 240}{(x^2 + 1)(x^2 + 4)(x^2 + 9)} = \frac{A_1 x + B_1}{x^2 + 1} + \frac{A_2 x + B_2}{x^2 + 4} + \frac{A_3 x + B_3}{x^2 + 9}$$

$$\frac{120i + 240}{(i^2 + 4)(i^2 + 9)} = 5i + 10 = A_1 i + B_1,$$

$$A_1 = 5, \qquad B_1 = 10.$$

$$\frac{120(2i) + 240}{(4i^2 + 1)(4i^2 + 9)} = -16i - 16 = A_2(2i) + B_2,$$

$$A_2 = -8, \qquad B_2 = -16$$

$$\frac{120(3i) + 240}{(9i^2 + 1)(9i^2 + 4)} = 9i + 6 = A_3(3i) + B_3,$$

$$A_3 = 3, \qquad B_3 = 6.$$

Example 3.

$$\frac{x^5 + 2}{(x^2 + 1)^3} = \frac{A_0 x + B_0}{(x^2 + 1)^3} + \frac{A_1 x + B_1}{(x^2 + 1)^2} + \frac{A_2 x + B_2}{x^2 + 1}$$

$$i + 2 = A_0 i + B_0,$$

$$A_0 = 1, \qquad B_0 = 2\ .$$

$$\left[\frac{d}{dx}(x^5 + 2)\right]_{x=i} = 5, \qquad \left[\frac{d^2}{dx^2}(x^5 + 2)\right]_{x=i} = -20i.$$

Using (C) with $m = 1$:

$$\frac{1}{2i} \cdot 5 = \frac{1}{2i} + A_1 i + B_1,$$

$$A_1 = -2, \qquad B_1 = 0.$$

Using (C) with $m = 2$:

$$-\frac{1}{4}(-20i) + \frac{1}{4i}(5) = \frac{-1}{-4i} - \frac{4}{2i} + (A_2 i + B_2) \cdot 2,$$

$$-15 = -7 - 8A_2 + 8B_2 i,$$

$$A_2 = 1, \qquad B_2 = 0.$$

References

Osgood, W. F.: Advanced Calculus, pp. 5–18.

Edwards, Joseph: An Elementary Treatise on the Differential Calculus, pp. 72-74.
A Treatise on the Integral Calculus, pp. 143-156.

Laurent, H.: Traité d'Analyse, vol. III, pp. 1-10.

Serret, J.A.: Calcul Différentiel et Integral, Vol. I, pp 592–615.

University of New Mexico

ABSTRACT SETS, ABSTRACT SPACES AND GENERAL ANALYSIS

Maurice Fréchet

Introduction. For centuries it has been known that arithmetic deals with the number of elements of an aggregate or set. The elements of the set may be material objects or intellectual concepts. Also it is known that Euclidean geometry has to do with abstract concepts called points, straight lines, and so on, which have only a few properties in common with the concrete objects which they represent very roughly.

In modern times it has been recognized that it is possible to elaborate full mathematical theories dealing with elements of which the nature is not specified, that is, with abstract elements. A collection of these abstract elements will be called an *abstract set.* If to this set there is added some rule of association of these elements, or some relation between them, the set will be called an *abstract space.* A natural generalization of function consists in associating with any element x of an abstract set E a number $f(x)$. Functional analysis is the study of such ''functionals'' $f(x)$. More generally, *general analysis* is the theory of the transformations $y = F[x]$ of an element x of an abstract set E into an element y of another (or the same) abstract set F. It is obvious that the study of general analysis should be preceded by a discussion of abstract spaces.

It is necessary to keep in mind that these notions are *not of a metaphysical nature;* that when we speak of an abstract element we mean that the nature of this element is indifferent, but *we do not mean at all that this element is unreal.* Our theory will apply to all elements; in particular, applications of it may be made to the natural sciences. Of course, due attention must be paid to any properties which depend essentially on the nature of any special category of elements under investigation.

Abstract sets. In this article we shall restrict our study of abstract sets to a consideration of definitions and properties of what might be called the number of elements of an infinite abstract set.

Non-denumerable sets. Transfinite numbers. An extension of the notion of number arose naturally when Cantor called attention to the existence of two kinds of infinity.

It is quite easy to show that we may number by means of integers used as indices all rational numbers of arbitrary sign: in other words, the set of rational numbers is *denumerable.* Let S_n be the set of irreducible fractions $\pm p/q$ such that $p + q = n$. There is only a finite number of these fractions. We may number them a, $a + 1$, $a + 2$, ..., $a + s$, starting with a properly chosen number a. By arranging them in order, S_1, S_2, ..., S_n, ..., we may effectively set the rational numbers into one-to-one correspondence with the class of positive integers.

In an analogous manner we may number all real algebraic numbers by taking as S_n the set of the roots of the polynomials with integral coefficients (the coefficient of the highest degree term being taken as 1) for which the sum of the absolute values of the coefficients increased by the degree of the polynomial is equal to n.

It is clear that every subset of a denumerable set is denumerable and that every denumerable set of denumerable sets is denumerable.

Before Cantor it was natural to think that to represent an infinite set, that is a set including an infinity of elements, it would be sufficient to designate one of its elements as a_1, then to designate by a_2 one of the remaining elements, and so on, thus obtaining a sequence $a_1, a_2, \ldots, a_n, \ldots$. If for instance the given set was the sequence of integers $1, 2, \ldots, n, \ldots$ and if the preceding operation consisted in writing $a_{2n} = n$, then $a_1, a_3, \ldots$ would not be used. But such a difficulty could be overcome by suitable renumbering.

However, Cantor has proved that no matter how we try to count in this way all real numbers, some will always be left over: *the set of real numbers is not denumerable.* It was a great discovery, which was the source of a whole new theory; the theory of sets, which has influenced all mathematics.

One immediate result was the proof of the existence of transcendental numbers since if the set of real numbers were identical with the set of real algebraic numbers then it would be denumerable.

Another result was the creation of the theory of transfinite numbers. In defining the number of elements of a set we may suppose that its elements are arranged in a definite order, fixed in advance, in which case we have an ordinal number; otherwise, it is a question of the cardinal number. For finite sets we may for practical purposes drop the distinction between ordinal and cardinal numbers. Such is not the case for infinite sets. For this reason we shall consider these two notions of numbers separately.

In both cases it will be prudent to avoid the philosophical difficulties involved in defining in a precise way the nature of the idea of number. In order to make use of the number concept it will be sufficient to make clear, as we shall do, what is meant in saying that two sets A and B have the same number of elements or that A has a greater number of elements than B. We may go so far as to indicate two sets C and D for which the number of elements will be considered respectively as the sum and the product of the numbers of elements of A and of B.

Cardinal numbers. The practical method of comparing the number of eggs and the number of apples in two piles consists in counting separately the eggs and the apples and in comparing the numbers obtained. This is an indirect method; it consists in replacing the sets of eggs and apples by two sets of numerical symbols and in comparing the latter.

The direct method amounts to placing an apple next to an egg as long as this is possible; according to whether some eggs are left over, some apples are left over, or none are left over, we agree that the number of eggs is greater than, less than, or equal to the number of apples.

This leads to the following general definition. If there exists a one-to-one correspondence, that is element for element, between the elements of a given set E and a part of the elements of another given set F, we say that the cardinal number of elements of E is less than the cardinal number of elements of F. If the correspondence extends to all the elements of F the two cardinal numbers are said to be equal.

This definition, suggested by the case where E and F have only a finite number of elements, has a meaning even if the sets involved are infinite. We see then that:

1). The cardinal number of every finite set is less than that of every infinite set;

2). The smallest cardinal number for an infinite set is that of the sequence of natural numbers;

3). The cardinal numbers of infinite sets are not all equal.

Ordinal numbers. A set E is said to be *ordered* when a rule has been given according to which one element out of every pair of elements of E is said to precede the other. (But the rule must be such that if a precedes b and b precedes c, then a precedes c).

In every ordered finite set E every subset of E has a first element. It is not necessarily so for an infinite set. For instance, if the integers are arranged in decreasing order ... n, $n - 1$, ..., 3, 2, 1, we have an ordered set which is not well-ordered.

We have defined ordinal numbers only for well-ordered sets (which include finite sets). We say that the ordinal number of a well-ordered set E is less than or equal to the ordinal number of a well-ordered set F if it is possible to establish a one-to-one order-preserving correspondence between the elements of E and those of a subset of F or those of F itself. For a more detailed popular exposition the reader may consult our *L'Arithmétique de l'infini*, Paris, Hermann, 1935.

From Sets of Numbers to Sets of Points. The transition from the notion of number to the notion of space, or from sets of numbers to sets of points, is quite natural. A variable number u is a function of a variable number x when to each definite value of x there corresponds a definite value of u. For instance, we may take u as $x + 1$, $2x$, x^3, 10^x, etc. But u may also be a function of several variables, for instance, $x + y$, xy^2, $x^2y^4z^3$, etc., where x, y, z are independent variables. If we consider x, y, z as the coordinates of a point M, we see that u may be taken either as a function of the three variables x, y, z or as a function of the point M. The introduction of the point (rather than its coordinates) serves to simplify both notation and thought.

Extension of the Notion of Space

Euclidean space. Euclidean geometry contains a detailed study of figures drawn in space of one, two or three dimensions. One of the reasons for speaking of the dimension of these three spaces is that it is sufficient to give one, two or three numbers (coordinates) to specify the position of a point. For instance, the surface of a sphere is two-dimensional since a point on it may be determined by its two geographical coordinates, latitude and longitude. Every property of a figure in n-dimensional space, for $n = 1, 2, 3$, may be expressed by a property of systems of numbers, and conversely. It is the basic principle of Descartes' analytical geometry.

n-dimensional space. It is then natural to generalize to the case $n > 3$. This extension amounts to a definition of n-dimensional spaces for $n > 3$. It is not merely a verbal generalization. The use of geometric language has the advantage of suggesting analogies, which may be translated by properties of systems of n numbers. Henri Poincaré has said (we quote from memory): ''Mathematics is only a well made language.''

For instance, in the theory of probabilities mean values had to be calculated in the form of multiple integrals with n variables. Their evaluation baffled some of the first mathematicians who investigated these questions or at best involved long and complicated calculations. When the n variables came to be considered as coordinates of a point in n-dimensional space and the integrals were treated as volumes or masses in this space, the geometric analogies made the calculations much more intuitive.

In the study of a mixture of liquids or in the more general ergodic problem, we start out with the principle that the motion of a material system under the influence of a given field of force is known when we are given the initial positions and the initial velocities of the points of the system. In the case where at any instant the totality of these positions is determined by the value of a finite number ν of parameters, the motion will be determined by the initial values of these ν parameters and of their time derivatives. We call the collection of these $n = 2\nu$ quantities the *initial phase*. Then the motion in our three-dimensional space of the different parts of this system, which may be very complicated, will be represented exactly by the fictitious motion of a ''point'' in the auxiliary n-dimensional phase space. In the study of the ergodic problem the language, the notation and the calculations are simplified enormously by operation directly on the phase space.

Moreover, besides cases like those just cited where it was a question of convenience, the notion of space of more than three dimensions may be quite necessary. In physics, since the theory of relativity, a time coordinate must be added to the three spatial coordinates.

Functional Analysis

By taking a different path we come to other extensions of the notion of space.

The first equations studied by mathematicians were obtained by equating to zero a function of one variable, for instance, $ax^2+bx+c=0$. Also systems of equations in several variables $x_1, x_2, \ldots, x_n$ were considered. We might say that the unknown was a point of coordinates $x_1, x_2, \ldots, x_n$ in n-dimensional space.

But less simple unknowns have had to be considered. If we wish to determine the trajectory of a planet attracted by the sun, the unknown is no longer a number but a curve. To determine the orbit we have to solve a differential equation to determine the distance of the planet from the sun as a function of the angle between a fixed direction and the line joining the sun to the planet. In other words we have an equation where the unknown is not a number but a function.

Likewise in the calculus of variations problems like the following arise: to limit, by means of a rope of given length placed on a plane, a domain of maximum area. The unknown here is not a number, but the position and form of the curve to be traced by the rope (it is known that the solution is a circle).

Thus in many problems the unknown is not a number or a finite system of numbers but a curve or a function. Problems of physics or mechanics may be cited where the unknown is a surface or a function of several variables. Historically the equations considered had as their solutions numbers or systems of numbers, or, what amounts to the same thing, points of 1, 2, ..., n dimensional space; then later curves, surfaces, functions. In each case the unknown is to be taken in a given category of elements. When the unknown is a curve, surface, function, we may say it is an element of functional space.

To make use of geometric intuition, with all of its advantages, we consider a more general space whose elements are functions of like nature, each function having the role of a point of this space. To give an idea of the extent of the notions summarized here in a few lines reference may be made to the work of Volterra, the great founder of functional analysis. For a functional space to have some analogy with more familiar spaces we have to establish some way of recognizing neighboring points. This may be done if we can define a distance between two functions. For instance, the distance between $y_1(x)$ and $y_2(x)$ may be defined as the maximum of $|y_1(x) - y_2(x)|$ if $y_1(x)$, $y_2(x)$ are continuous over a common closed x interval.

General Analysis

Abstract spaces. If it is possible to study the properties of a space as complex and general as a functional space, it may be asked if there might not be something to be gained by making further generalizations. Instead of building several parallel theories of certain 1, 2, ... dimensional spaces, or of certain functional spaces, might it not be possible to include them in a single theory of spaces whose ''points'' are elements of arbitrary nature? Such a theory of abstract spaces has existed since 1904.

At first sight such an undertaking might be considered as absurd. How can we speak of a geometry in a space whose ''points'' are of an undefined nature, when we do not know if the elements are numbers, curves, surfaces, functions, series, sets, etc.? We should have to exceed the scope of this article to prove that such a theory is possible[1] and that its value, aside from purely mathematical or philosophical interest has been demonstrated by many applications in different domains.[2] However, we shall try to throw some light on this. Let us observe that, contrary to what we might think at first, the idea of reasoning mathematically on abstract elements is far from being new. The equality or inequality of two numbers, if we think of them as the numbers of elements of two sets, are two properties concerning two sets of elements whose nature does not enter at all.

Descriptive definitions. One of the best ways of reasoning on abstract elements consists in using descriptive definitions rather than constructive definitions.

Generalization of the notion of distance. We may define the distance between any two elements a, b of an abstract set as a number $(a,b) = (b,a) \geq 0$ satisfying the following conditions:

I $(a,b) = 0$ if and only if a and b are identical.

II $(a,b) \leq (a,c) + (c,b)$

This is a descriptive definition. When dealing with elements of a definite nature it is better to use a constructive definition. For the set C of continuous functions we adopt usually the definition of (y_1,y_2) previously cited, the maximum of $|y_1(x) - y_2(x)|$. It satisfies conditions I and II. For the set L_2 of functions $f(x)$, defined on the interval (α,β), whose squares are integrable on (α,β), we may call the distance of two elements $f(x)$, $g(x)$ of L_2 the quantity

$$(f,g) = \sqrt{\frac{1}{\beta-\alpha}\int_{\alpha}^{\beta}[f(x) - g(x)]^2dx}$$

For property I it is clear that $(f,g) = 0$ if $f(x) = g(x)$. In order to take care of the requirement that $(f,g) = 0$ only if $f(x)$ and $g(x)$ are identical elements, we agree to consider f and g equivalent or indistinguishable if they are equal ''almost everywhere'', that is, if the set of points x where they differ can be enclosed in a denumerable set of intervals for which the sum of their lengths may be taken as small as we wish.

In order to give some idea of the problems which may be stated and solved with respect to sets of abstract elements let us cite the following example: what condition must be satisfied by the distance defined on an abstract set E so that a correspondence may be established retaining the value of the distance between pairs of elements of E and the corresponding pairs of elements of a simple given set F, F being,

[1]M. Fréchet, *Les espaces abstraits.* Paris, Gauthier-Villars, 1928.

[2]M. Fréchet, *Mélanges mathématiques.* Internat. Congress of Math., Vol. I, Oslo, 1936.

for instance, a straight line, or a plane, or a Euclidean three-space, or the set of real numbers <2, etc.?

A whole geometrical theory has been constructed on the single basis of conditions I and II by K. Menger and his students, who have tried to retain as much as possible of the nature of Euclidean geometry. In many cases a Euclidean definition may be generalized in several ways, and it is interesting to seek the relations between the various generalizations obtained. Examples may be found in the theory of convexity, curvature, etc. A quite complete and clear exposition of this abstract metric geometry may be found in L. Blumenthal's *Distance Geometries* (University of Missouri Studies, 1938). It is one of the rather rare examples of a book which leads the reader up to the most recent discoveries without requiring extensive specialized mathematical preparation.

Topological spaces. Classical geometry is not restricted to metric considerations. It is concerned also with properties of figures which are invariant under continuous deformations: such investigations are the object of *topology*. In the latter the essential notion is not distance, but limit or neighborhood. It may even be said that an abstract space E is not clearly specified until limit or neighborhood or equivalent concepts are defined so that continuous correspondence may have a meaning.

If distance has previously been defined on such a set E, it will be natural to say that the limit of a sequence $a_1, a_2, \ldots, a_n, \ldots$ of elements of E is an element a of E (if such an element exists) for which the distance (a,a_n) tends toward zero with $1/n$, or that the ϵ-neighborhood of a is the set of elements b of E for which $(a,b) < \epsilon$.

Sometimes mathematical analysis may lead to the consideration of a space where a limit or a neighborhood is defined without the intervention of a distance concept. An example of this is furnished by the function space whose elements are functions of Baire. Hence if a "topological" space is defined (that is, if continuous transformations are defined in the space - this use of the adjective topological is more general than that accepted by some authors, who restrict its use to Hausdorff spaces), it may or may not have limit or neighborhood (or analogous concepts) associated with distance.

Let us consider the space E_ω where each element X is determined by an infinite sequence of numbers $x_1, x_2, \ldots, x_n, \ldots$, called coordinates of X, and where we consider a sequence of elements Y as converging toward X if, for each n, the coordinate y_n of Y converges toward x_n. If we write

$$(X,Y) = \sum_{n=1}^{\infty} \frac{1}{n!} \frac{|x_n - y_n|}{1 + |x_n - y_n|}$$

we see that (X,Y) satisfies conditions I and II. We note that the original natural definition of convergence on E may be expressed in terms of this distance, which is much less natural. Other distance definitions could have served equally well.

Starting with the more general concept of neighborhood a topological theory of abstract spaces may be constructed without using the distance concept.[3] A slightly more general theory starts with the idea of "closure" of a set, from which the notion of "accumulation element" of a set may be deduced. To every subset G of a topological space there corresponds a set $\bar{G}$, called the closure of G. The single requirement is that G be a subset of $\bar{G}$. An accumulation element of G is by definition an element a belonging to the closure of $G - a$. A continuous transformation of G is a transformation of each element b of G into an element b_1, such that if b belongs to the closure $\bar{g}$ of a subset g of G then b_1 belongs to the closure $\bar{g}_1$ of g_1, the transform of g.

It is now possible to generalize the notion of dimension as follows. (There are other ways in which this may be done, as shown by Brouwer, Urysohn and Menger, whose definitions are based on Poincaré's work). Let us call a transformation which with its inverse is single-valued and continuous a homeomorphism. We may say that the number of dimensions $d(G)$ of G is equal to or less than the number of dimensions $d(F)$ of F if there exists a homeomorphism between G and F or a subset of F. These two dimension numbers are equal if $d(F) \geq d(G)$ and $d(G) \geq d(F)$. Previously all spaces for which the dimension numbers were not finite were put in the same category. One advantage of this new definition is that a distinction may be made between different infinite dimension numbers. At the same time it throws light on the topological affinity which exists between the most important functional spaces considered in analysis by showing that they have the same number of dimensions.[4]

Attention should be called to another very useful generalization. Grassman and others founded an abstract vector theory from which continuity was absent. The preceding topological theory put continuity in first place. This did not prevent the magnificent use of both by Wiener and Banach to investigate normed vector spaces, and subsequent consideration of affine topological spaces. A simple and intuitive definition of the differential of a continuous abstract transformation was suggested by operations carried out on normed vector spaces.[5] It is interesting to note that the first definition of the integral of a numerical function of an abstract element was not based on topo-

[3]See pages 172-185, 224, 277-278 of *Espaces abstraits*, previously cited.

[4]See pages 30-113 of *Espaces abstraits*, already cited.

[5]M. Fréchet, La notion de différentielle dans l'analyse générale. Ann. Ec. Norm. Sup., t. XLII, 1925, pp. 293-323. M. Fréchet, Sur la notion de différentielle. Journal Math., t. XVI, 1937, pp. 233-250.

logical considerations.[6]

Some applications. Even if the reader is convinced that there are good reasons for treating abstract elements, he may still have some doubt as to the utility of such general theories. Let us eliminate immediately a possible objection by observing that, although an abstract element has been considered without specifying its nature, it may be well known and it may have a quite concrete meaning. Thus the number of elements of a set has nothing to do with the nature of the individual elements. Consequently, abstract spaces and abstract functions may be used in all branches of mathematics. Kürshak used a generalized distance in number theory. The abstract algebra introduced by Emmy Noether is based on the same type of ideas as general analysis. The recent topological theory of groups makes considerable use of the notion of ''compactness'' borrowed from abstract space theory. In the theory of ''normal'' families of analytic functions developed by Paul Montel the same situation arises. The calculus of variations existence theorems of Tonelli, Menger, Bouligand, etc., make use of compactness and generalized distance. In probability theory, integration on an abstract space and abstract integrals are indispensable in dealing with random variables. In hydrodynamics Leray and Schauder have used differentials of abstract transformations.

These examples should indicate the wide variety of applications of general analysis.

[6] M. Fréchet, Sur l'intégrale d'une fonctionnelle étendue à un ensemble abstrait. Bull. Soc. Math. Fr., t. XLII, 1915, pp. 248-265.

M. Fréchet, L'intégrale abstraite d'une fonction abstraite d'une variable abstraite. Rev. Sci. (82nd year), 1945, pp. 483-512.

University of Paris

CURRENT PAPERS AND BOOKS

Edited by

H. V. Craig

This department will present comments on papers previously published in the MATHEMATICS MAGAZINE, lists of new books, and book reviews.

In order that errors may be corrected, results extended, and interesting aspects further illuminated, comments on published papers in all departments are invited.

Communications intended for this department should be sent in duplicate to H. V. Craig, Department of Applied Mathematics, University of Texas, Austin 12, Texas.

Analytic Geometry, Revised, by Charles H. Sisam, Emeritus Professor of Mathematics, Colorado College. Cloth. Pages xvi + 304. 1948. Henry Holt and Company, 257 Fourth Avenue, New York, N.Y. Price $2.34.

A not very startling revision of a sound traditional type text, this Analytic Geometry, as did the original, reflects the years of successful classroom experience of the author. With thirteen chapters of plane geometry and three of solid geometry, there is ample material for a five hour course for a semester. Careful exposition, neatly drafted figures, an adequate number of exercises including some to challenge as well as many to drill, all combine to make a pleasant and useable textbook for teachers and students even though the small closely printed page gives an impression of crowding.

In the mind of this reviewer it is questionable if there has been a gain by bringing "together the parts of the introduction to polar coordinates into a separate chapter." This emphasizes the suspicion which many students seem to harbor that polar coordinates don't really belong - a feeling which persists into later courses in mathematics. The inclusion of four pages of tables is of questionable value since each student may be presumed to have available the more complete tables required in a course in Trigonometry.

The use of a star to mark "articles which may be omitted" should perhaps be safeguarded. Thus in Chapter One, Article 6, Intersections of Graphs, is starred. Then in Chapter Three, Article 29, Family of Lines Through the Intersection of Two Given Lines, occurs the sentence "For, the coordinates of intersection satisfy both of the given equations (Why?)." Unless Article 6 has been considered or knowledge from another course is presumed, the later sentence (and question) may be unfair.

The alert teacher will be aware of such minor points and take care of them in the daily class discussion periods.

O. H. Rechard

Analytic Geometry, by Robin Robinson, Professor of Mathematics, Dartmouth College. Cloth. Pages ix+147. 1949. McGraw-Hill Book Company, 330 W 42nd Street, New York, N.Y. Price $2.25.

The author has stated three outstanding guiding principles followed in preparing this text.

1. A course in Analytic Geometry should provide the necessary background for a later course in the calculus but should nevertheless be a course in geometry.
2. A text should not "steal the show" from the teacher, allowing the latter freedom to expand in his own way the author's brief and concise presentation.
3. A large number of problems must be included so that ample drill material is provided and original thinking is stimulated.

The third principle is admirably followed achieving at the same time emphasis on the fact that this is "a course in geometry." Numerous examples could be cited; let the four exercises – 15, 16, 17, 18 – on page 80 suffice. They all involve "ruler and compass" constructions calling for some geometric insight on the part of the student.

Principle two is likewise adhered to. In fact, presuming the author intended most of the omissions from a "standard" first course to be supplied by the instructor, the question may be asked if the principle is followed to the extreme. To list but a few of these omissions is to point up this question. The two-point form of the equation of a straight line, the problem of finding the equations of angle bisectors, parametric equations except for the line in three dimensions and surfaces of revolution, are among the topics with which the student will have no contact unless the teacher introduces them.

More serious in relation to principle one is the absence from this text of a treatment of higher plane algebraic, transcendental, and exponential equations. An analytic geometry course without these topics does not contain, in this reviewer's judgment, the minimum essentials needed for a study of elementary calculus.

The "boxed" formulas, carefully drawn figures, and generally attractive appearance of the pages make this a pleasantly readable text.

O. H. Rechard

First Year Mathematics for Colleges. By Paul R. Rider, The Macmillan Company, New York, 1949, XV + 714 pages, $5.00.

This text offers a clear exposition of the usual topics of college algebra, trigonometry, and analytic geometry taken very nearly in that order and without much intermingling or blending of the subject matter of those courses. By this arrangement, however, the duplications of separate texts are avoided. For example, a single discussion of each of the following suffices: rectangular coordinates, logarithms, logarithmic and exponential curves, functional notation, and trigonometric curves. The first eighteen chapters, with the exception of Chapter 4, present the conventional development of college algebra in the usual order of treatment. Chapter 4 has much of the material of the first chapter in many analytic geometry texts - description of rectangular coordinates, distance formula, point-of-division formula, area of a triangle formula, etc. In the next ten chapters trigonometry is developed furnishing the tools for the chapter on complex numbers which follows. Then the conventional college algebra chapter on the theory of equations is divided into two chapters entitled "Polynomials" and "Theory of Equations", the latter terminating with Horner's method but without the algebraic solutions of the cubic and quartic equations. The exposition of plane analytic geometry follows through conics, curve tracing, parametric equations, polar coordinates, and terminating with a chapter on curve fitting. Next come further topics in college algebra: permutations, combinations, probability, determinants, partial fractions, and infinite series. The book ends with three chapters on solid analytic geometry, viz.: "Rectangular Coordinates in Space", "Plane and Line", and "Surfaces and Curves", the latter dealing mostly with quadric surfaces. This is the text in broad outline. Some detailed characteristics of it follow.

The text is copiously supplied with exercises, which occur after their appropriate articles instead of being placed at the ends of the chapters. Each group of exercises is numbered, as well as the exercises within the group being numbered. This follows the practice in many newer textbooks. Answers to the odd-numbered exercises are given at the end of the book. Eleven four-place tables, which seem adequate in their number and in their size, are bound with the volume. No separate card of four-place common logarithms is supplied. Extremely few typographical errors were discovered, and no wrong answers to exercises were noted.

While the usual proof of the binomial theorem for positive, integral exponents is presented, exercises are given in expanding the binomial with both fractional and negative exponents. Upper and lower limits for real roots and Descartes' Rule of Signs are discussed before the rational roots of rational, integral equations are found. The student learns to use common logarithms before he gives consideration to logarithms with other bases. The rule for finding the characteristic of a common logarithm is stated in terms of the number of places from the standard position to the decimal point. The so-called Computor's Rule for favoring the even digit is given for rounding off a single five. There is an excellent chapter on calculation with approximate numbers. In discussing determinants of the nth order, double subscripts are used instead of changing the letter representing the element and using one subscript. The method of determining the signs of the products in the expansion of a determinant is unusual. The hyperharmonic series is called the p series. In connection with the chapter on probability, the newer Commissioners 1941 Standard Ordinary Mortality Table might have been used instead of the older American Experience Table of Mortality.

Trigonometric functions are first introduced for acute angles before they are defined for the general angle. The tangent law and the half-angle formulas are obtained geometrically, thus making it possible to introduce the solution of oblique triangles early before formal work on trigonometric identities. This seems to be in accordance with the historical order of development. Exercises are given in the use of both four and five-place logarithms for the solution of triangles. Illustrative examples, however, use four-place logarithms, tables of which are found at the back of the book. Angle data for use with five-place logarithms are given in tenths of a minute instead of seconds. Mollweide's equations are used to check the first three cases in the solution of oblique triangles. A summary is given of the correspondence between the number of figures in the sides of triangles and the accuracy of their angles.

The ellipse and the hyperbola are defined first in terms of their focal radii. Later the conic is defined in terms of its eccentricity. There is no proof that conics may be obtained by cutting a cone. Polar coordinates are allocated to one chapter instead of being scattered through several chapters. The limacon is defined by adding a constant to the radius vector in the polar equation of a circle passing through the pole. The intersection of polar curves is carefully discussed. The chapter on curve fitting includes a brief discussion of the mean, the standard deviation, and the coefficient of correlation. The normal equations for the least squares line are derived while the normal equations for polynomial, exponential, and logarithmic curves are used without being derived. The use of logarithmic and semilogarithmic graph paper is described. The sigma notation for summation is used in this chapter. In the analytic geometry of space, attention is called to the plane as a cylinder whose directrix is a straight line. No formulas for the rotation of axes in space are given.

This book seems suitable for students with a year each of secondary algebra and plane geometry although it should be readily adaptable to those with more extensive preparation. The exposition is concise but lucid, and it leaves the impression of thoroughness and care. This book is a very satisfactory contribution to the textbook literature of elementary mathematics.

University of Arizona R. F. Graesser

THE PERSONAL SIDE OF MATHEMATICS

Articles intended for this Department should be sent to the Mathematics Magazine, 14068 Van Nuys Blvd., Pacoima, California.

WHAT MATHEMATICS HAS MEANT TO ME

F. T. Bell

The Fditor has asked for about 400 words on "what mathematics has meant to me." Notice the 'me' – not somebody else. This will account for all the 'I', 'me' in what follows, for which I apologise. I am as embarrassed as if I had inadvertently stood up in church to tell the congregation how and why I had been saved. You may be even more embarrassed in witnessing my testimony.

My interest in mathematics began with two school prizes, one in Greek, the other for physical laboratory, both richly bound in full calf. The Greek prize was Clerk Maxwell's classic on electricity and magnetism, the other, Homer's Odyssey. My cousin got the prize for Greek, I got the other. He read mine, I tried, and failed, to read his. The integral signs were particularly baffling to one who had not gone beyond the binomial theorem for a positive integral exponent. The calculus was not a school subject at the time, so my mother paid for private lessons from a man – the late E. M. Langley – who was the best teacher I ever had. From him I learned what dy/dx and $\int ydx$ mean. The rest was comparatively easy, and I found myself in possession of a key that unlocks a hundred doors. Although I have never done anything in mathematical physics, I have been able to read some of the great classics which, without the calculus, would have been incomprehensible. This has been one thing that has made life interesting. How some philosophers of science and others have the audacity to write on relativity and the quantum theory without a reading knowledge of the calculus is the wonder of the ages.

Another thing I got from mathematics has meant more to me than I can say. No man who has not a decently skeptical mind can claim to be civilized. Euclid taught me that without assumptions there is no proof. Therefore, in any argument, examine the assumptions. Then, in the alleged proof, be alert for inexplicit assumptions. Euclid's notorious oversights drove this lesson home. Thanks to him, I am (I hope!) immune to all propaganda, including that of mathematics itself. Mathematical 'truth' is no 'truer' than any other, and Pilate's question is still meaningless. There are no absolutes, even in mathematics.

California Institute of Technology

(A series of articles on "What Mathematics Means to Me," written by other eminent mathematicians and people in various professions, will appear in subsequent issues of the Mathematics Magazine. Editor.)

A REALISTIC VIEW OF DIFFERENTIAL CALCULUS

Sister Helen Sullivan

A modern mathematician has made two significant statements regarding the calculus. First, it is one of the great achievements of the human mind: second, it occupies a place between the natural and the humanistic sciences and hence should be a fruitful medium of higher education.

In the attempt to prove the foregoing assertions while at the same time establishing a case that will appeal to the average non-technical reader, this paper will investigate what is calculus and what is its scope or field of influence. In so doing, the approach will be from common experience with a minimum of mathematical formulas or technicalities.

If there is one fact in man's experience that stands out more than all others, it is the fact of motion. This fact so impressed the late Mr. Gilbreth of Cheaper-By-the-Dozen fame, that he spent his life studying the daily motions made by individuals in order to ascertain how to reduce them to a minimum in given cases and thus speed up efficiency. Motion is the most nöte-worthy feature of the sense world and is so obvious as to defy proof. Although careful philosophical distinctions may be drawn to show the difference between motion and change, in a general way they may be used synonymously. Each day witnesses the birth of something new and the decay of something old. Seedlings develop into fullgrown plants. Warm days are followed by cold ones. Clear skies succeed rainy ones. Nature is never still. Neither is man in a state of rest, as the five o'clock traffic rush daily shows. Philosophers have said that all the motion in the world can be reduced to six types - generation, corruption, local movement, alteration, augmentation or diminution. Water is generated by the chemist who combines oxygen and hydrogen in proper porportions and under suitable conditions - there has been a change in substances. The student who rises from his desk to close his window displays local motion. He has changed the position of his body and likewise the position of the window glass. The sapling planted on the college campus by the class president manifests to the returning alumni augmentation or change due to growth. The average Alumnus likewise is aware of a diminution in the hairs of his head as he approaches middle age. The art of alteration, or a change in accessories, is the explanation why the well-dressed woman appears so often in different attire. So rapid and continuous are the changes of everything in the universe that some philosophers have asked - Is there anything but change?

Associated with the fact of motion or change is the equally incontestable fact of relationship. There is a relationship between productivity of the soil and the amount of rainfall. Good spring rains - correctly spaced and a bumper corn crop results (other essential factors being present). Normally speaking, a girl's popularity is dependent on her sincere friendliness and congeniality. A student's semester mark is largely determined by his native intelligence and academic industry. Mental efficiency is conditioned by one's physical health and habits of thought. Examples could be multiplied to show the fact of relationship between various things. Actually it is no overstatement to say that all beings are inter-related and inter-dependent because there is unity, harmony, plan, order, and hierarchy in the universe as designed by the Divine Architect.

In the language of the mathematician, anything that changes is termed a *variable* and is accordingly designated by an "x" or "y" . Anything that bears a relationship to something else is said by the mathematician to be a function of that other; it is dependent on it in order to exercise its privilege of assuming different values. To state that "y" is a function of "x" means that we are dealing with two varying entities that are so related that one depends on the other for its operation - just as popularity depends on sociability.

Of all the myriads of things in the universe which vary, undergo changes, or suffer mutations and alterations, very few come under the scope of the calculus simply because they are not amenable to *exact* mensuration which is the ticket of admittance for mathematicians whose major concern is with quantity (which in turn comprises both magnitude and multitude). Sociability is a personal human trait incapable of exact measurement. The same holds for popularity - its related variable.

At this point in our discussion, the formal definition of differential calculus becomes meaningful. It is that branch of mathematics which studies (1) with precision (2) the rate of change of (3) related variables. But in order to accomplish this it employs the notion of limits. A limit is a boundary towards which the changing entity tends in such a manner that the difference between the ultimate goal or limit and the value of the variable quantity at any moment can be rendered arbitrarily negligible or nearly so. It is somewhat analagous to the situation involved in the following example. A teen-ager sets out to determine how much speed he can produce from his antiquated car. The mechanic tells him that ninety miles per hour is the upper bound. The daring chap attempts to attain that limit. He may approach it at instantaneous intervals; on one occasion he may reach eighty-four miles per hour; at another eighty-seven miles per hour. He is rendering the difference between the speed attained by him and the upper limit set by the mechanic (who knows the workings of the car) more nearly negligible. Any analogy limps and a mechanical illustration introduces even more errors than one would wish to reckon with.

In the geometrical world it is not difficult to see that the area of an n-sided polygon approaches the area of its circumscribed circle as n grows larger. One can reduce the difference between the two areas to an amount as small as one pleases by letting n grow very large. From this relation certain inferences concerning the limit may be drawn and one gains an insight into the type of analysis employed by the calculus.

To say that the method of calculus is characterized by *precision* is to say that it aims to express its findings in numerically exact terms. If an entity is neither mensurable nor numerable it can not be handled with precision by a mathematician. This rules out all purely qualitative beings whose quantitative aspect is indeterminable. Thus, justice, honesty and other realities of like kind defy mensuration. This is not to say that they do not vary: rather that they are of a different order.

It has been said that the calculus studies "the rate of change of related variables" and this is to focus attention on its essential nature. It is concerned with beings which suffer *change*. It is also concerned with mutable beings which are related to other mutable beings and its concern is primarily quantitative for it asks such questions as - if the volume of a cylindrical container depends on the radius, what change will a two-inch increase of radius produce in the container, or again: what is the rate at which a man's shadow lengthens when said man (whose height is known) moves away from a lamp (whose elevation is known) at a definite rate of speed?

It should be pointed out that while calculus studies change or motion, it does so by stopping the motion or preventing the change for the moment of investigation. It is roughly analagous to the slowing down of the movie projector in order to view the individual still picture which is the element of the series which presents the motion-picture. Thus, the calculus permits one variable to take on an increase, its related variable increases correspondingly.It sets up a ratio of the two increases (called increments in the technical calculus) for the sake of comparison. Then it permits the increase of the independent variable to approach zero and thus, in the limit, a notion of the changing quantity at that instant is obtained. The rate of change at each instant is thereby observed. It will be noticed that it is not actual motion that is being studied but the limiting condition of the ratio of the small increases at each time interval. The calculus can show how a curve changes its direction at every consecutive point - it does not examine the reality of motion itself - it is powerless to do so. In this paradoxical procedure of studying motion by preventing it, or better by conceiving of it as motion-less, calculus is not unlike the natural sciences which interfere with and control the operations of the objects studied by them. The biologist studies life by destroying it - the experiments are performed on the lifeless animal as dissection is thought to provide greater knowledge of life. The physicist studies forces by considering bodies *at rest* or in a state of equilibrium.

Conceived in the strict sense there is no such thing as motion in mathematics. Points do not generate lines in the sense that there is *motion* involved. Mathematics is static, incapable of dealing with the reality of motion. It can describe the position of bodies before and after the motion has occurred and, by reducing the interval to a minimum, hopes the reader will be satisfied.

A return now to the initial paragraph of this paper will render the statements presented there more meaningful. Calculus is truly one of the great achievements of the human mind for man is able to begin his investigations in the *actual*, physical order and then by a power, peculiar to his intellect and denied to any other corporeal creature, he is able to pass instantaneously to the realm of the *possible* and conceive what could happen in the limit. Calculus lays claim to greatness also because it is the mathematician's explanation of something universally known to all - *the fact of change*. This likewise accounts for its being located between the natural and the humanistic sciences. It occupies itself not only with the inanimate order of *nature* but also with things which bear a close relationship to man and his *human* concerns.

Differential Calculus is the first step in basic analysis and proceeds to examine the phenomena of motion in the only way open to it, i.e. by analyzing it in terms of mathematical ultimates and resorting to mathematical imagination.

Mount St. Scholastica College

PROBLEMS AND QUESTIONS

Edited by

C. W. Trigg, Los Angeles City College

Readers of this department are invited to submit for solution problems believed to be new and subject-matter questions that may arise in study, in research, or in extra-academic situations. Proposals should be accompanied by solutions, when available, and by such information as will assist the editor. Ordinarily, problems in well-known text-books should not be submitted.

In order to facilitate their consideration, solutions should be submitted on separate, signed sheets within three months after publication of the problems. Readers are invited to offer heuristic discussions in addition to formal solutions. Manuscripts typewritten on 8½″ by 11″ paper, double-spaced and with margins at least one inch wide are preferred. Figures should be drawn in india ink and in exact size for reproduction.

Send all communications for this department to *C. W. Trigg, Los Angeles City College, 855 N. Vermont Ave., Los Angeles 29, Calif.*

PROPOSALS

84. *Proposed by B. F. Crow, Roxbury, Mass.*

In a game which some of my friends play, one man holds a $1.00 bill which has an eight-digit number imprinted twice on its face. Another man calls three digits. If these digits are in the imprinted number, he wins. For example, if B 27588607 F appears on the face of the bill and the second man calls 277, he wins. What is the probability of winning?

85. *Proposed by Victor Thébault, Tennie, Sarthe, France.*

Find the three smallest consecutive integers each of which is the sum of two squares (zero excepted).

86. *Proposed by Dewey Duncan, East Los Angeles Junior College.*

We define a heterosquare as a square array of the first n^2 positive integers, so arranged that no two of the rows, columns, and diagonals (broken, as well as straight) have the same sum. (*a*) Show that no heterosquare of order 2 exists. (*b*) Find a heterosquare of order 3.

87. *Proposed by Leo Moser, Texas Technological College.*

A right circular cone is cut by a plane. The intersection of course is a conic. Find the equation of the curve that this conic goes into if the cone is unrolled on to a plane. In particular, if the cone is a cylinder and the plane cuts the axis of the cylinder at 45°, then the ellipse formed will unroll into a sine curve.

88. *Proposed by O.E. Stanaitis, St. Olaf College, Northfield, Minnesota.*

Establish the convergence or divergence of

a) $1 - \frac{1}{2} + \frac{1}{3\sqrt{3}} - \frac{1}{4} + \frac{1}{5\sqrt{5}} - \frac{1}{8} + \frac{1}{7\sqrt{7}} - \frac{1}{16} + \cdots;$

b) $1 - \frac{1}{\sqrt{2}} + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} - \frac{1}{\sqrt{6}} + \frac{1}{7} - \frac{1}{2\sqrt{2}} \quad \cdots .$

89. *Proposed by H. T. R. Aude, Colgate University.*

If the graph of the quartic $y = f(x) = x^4 + px^2 + qx + s$ has points of inflection, then there exist uniquely three pairs of parallel lines which are tangent to the quartic. Find the equations of the six lines, and note that the sum of their three slopes is $3q$.

90. *Proposed by D. L. MacKay, Manchester Depot, Vt.*

Triangle ABC is divided into two parts, triangle DBE and quadrilateral $ADEC$, by the line DE. Construct a line which will bisect each of these parts.

SOLUTIONS

Bisectors of the Area of a Triangle

58. [March 1950] *Proposed by W. B. Clarke, San Jose, California.*

Through a point P in the plane of a given triangle lines are drawn bisecting the area of the triangle. Discuss the location of points P for which there are one, two, or three bisecting lines.

Solution by C. C. Oursler, Indiana University, Gary Center. Any line bisecting the area will form at least one triangle having one vertex in common with the original triangle. Such a triangle must have an area equal to one half of the area of the given triangle.

Consider the following problem: Given two intersecting lines, a third line is required to intersect these two lines and thereby form a triangle of given area. What positions may the third line have? It must be tangent to one of a pair of conjugate hyperbolas whose asymptotes are the two given lines and such that the product of the semi-transverse axis and the semi-conjugate axis is the required area. Let us first verify that any tangent to such a hyperbola does form with the two given lines a triangle meeting the required conditions. Without loss of generality we may assume that equations of the given lines are $y = bx/a$ and $y = -bx/a$ and the hyperbola is $x^2/a^2 - y^2/b^2 = 1$. Any tangent to this hyperbola at a point (x_1, y_1) on the hyperbola is $x_1 x/a^2 - y_1 y/b^2 = 1$. The intercepts on the asymptotes are $ab(bx_1 \pm ay_1)^{-1}\sqrt{a^2 + b^2}$. Since (x_1, y_1) is on the hyperbola, the product of the intercepts is $a^2 + b^2$. The area of the triangle determined is one half the product of the intercepts times the sine of the angle θ between them. This angle is twice arc $\tan(b/a)$. Therefore, $\sin\theta = 2ab(a^2 + b^2)^{-1}$.

Hence the area of the triangle is ab. No other line having the same slope as the tangent will form a triangle of the same area because then there would be two triangles similar and equal in area but not equal in dimensions, which is a contradiction.

In the original problem, we can utilize only that branch of the hyperbola which cuts the side opposite the vertex being considered. Indeed, only those tangents which cut both adjacent sides can be considered. The medians of the original triangle meet the requirements of the problem and must obviously be limiting positions of acceptable tangents. An elementary theorem says that that part of a tangent to a hyperbola between two asymptotes is bisected by the point of tangency.

Construct two medians of the original triangle. Construct the arc of a hyperbola between the midpoints of these two medians such that the hyperbola has as asymptotes the two sides of the triangle intersecting in the opposite vertex and such that the hyperbola is tangent to the two given medians at their midpoints. Similarly construct arcs connecting the midpoint of the other median. We now have a curvilinear triangle within the original triangle. To construct a bisecting line through any point we construct a line through the point tangent to one of these arcs. For the vertices and for any point outside the curvilinear triangle, only one such tangent can be constructed. Through a point on the sides of the curvilinear triangle (excluding vertices) two such tangents can be constructed. Through any point in the interior of the curvilinear triangle three such tangents can be constructed.

Also solved by P. N. Nagara, College of Agriculture, Bangkok, Thailand; and the proposer.

The envelope of the lines bisecting the area of a triangle is discussed in *American Mathematical Monthly*, **42**, 455, (1935). References to other problems dealing with these area-bisectors are:

Shortest bisector - *A.M.M.*, **24**, 129, (1917); **46**, 171, (1939); *School Science and Mathematics*, **33**, 781, (1933); **39**, 581, (1939).

Bisector perpendicular to a side -*S.S.M.*, **7**, 414, (1907); **42**, 687, (1942); **45**, 776, (1945).

Bisector equal to circumradius - *S.S.M.*, **12**, 234, (1912).

Bisector passing through a given point - *S.S.M.*, **13**, 613, (1913); **18**, 556, (1918); **22**, 878, (1922); **23**, 77, (1923); **36**, 217, (1936).

Area-bisector also bisecting perimeter - *A.M.M.*, **49**, 64, (1942).

Some Almost Regular Polyhedrons

59. [March 1950] *Proposed by D. L. MacKay, Manchester Depot, Vermont.*

The definition of regular polyhedrons gives three requirements: (a) faces regular polygons, (b) faces congruent, (c) polyhedral angles congruent. Give illustrations of polyhedrons possessing each pair of these requirements but not the third.

Solution by Michael Goldberg, Washington, D.C. (a) and (b) only.

Triangular and pentagonal dipyramids formed by joining the bases of regular pyramids whose lateral faces are equilateral triangles.
(*a*) and (*c*) only. The Archimedean solids which have two kinds of regular faces. (For those having three kinds of faces, the polyhedral angles are symmetric.)
(*b*) and (*c*) only. The irregular tetrahedra known as disphenoids. They are formed by folding an acute-angled triangle along the lines joining the midpoints of the sides.

An Undenary Square

63. [May 1950] *Proposed by Victor Thébault, Tennie, Sarthe, France.*

In the system of numeration having base 11, find a six-digit square of the form *abcabc*.

I. *Solution by H. M. Gehman, University of Buffalo.* We shall write all numbers to the base 11. Since $N = a\,b\,c\,a\,b\,c = (1001)(abc) = (2^2)(3^2)(34)(abc)$, the condition that N be a square is that abc be of the form $34t^2$. Letting $t = 2, 3$ we find the only two values of N: 125 125 ($= 374^2$) and 283 283 ($= 556^2$).

If $t = 1, 4, 5$ we have 034034 ($= 192^2$), 499 499 ($= 738^2$), 771771 ($= 91X^2$). These are not admissible solutions since the statement of the problem implies that $a \neq 0$, and that a, b, c are distinct.

II. *Heuristic Discussion by Monte Dernham, San Francisco.* The ensuing discussion reflects the actual process by which the writer happened to *find* the solution to this problem. It makes no attempt at formal proof in the accepted sense.

Given that there is a number of the form $abcabc_{11} = N^2$, we are to find a, b, and c; that is, we have to solve in non-negative integers

$$11^5a + 11^4b + 11^3c + 11^2a + 11b + c = N^2$$

where a, b and c each < 11, $a \neq 0$ and $a \neq b \neq c$. Collecting terms and expanding, we have

$$161172a + 14652b + 1332c = N^2.$$

How now shall we proceed to find values for a, b and c such that the left side will be a perfect square? Could it be that a factorization of the coefficients would furnish a clue? On factoring, we discover that 1332 ($= 6^2 \cdot 37$) happens to divide each coefficient, so we write

$$6^2 \cdot 37(121a + 11b + c) = N^2.$$

It is now clear that the left side is a square if, and only if, $121a + 11b + c = 37m^2$. Then if we succeed in finding admissible values for m we have solved the problem. What is the upper bound for m? Since a, b, c each < 11, $37m^2 \leqq 1330$, whence $m \leqq 5$. Testing for $m = 1,2,3,4,5$, successively, we obtain

m	$37m^2$	$abcabc$	*Radix* 11	*Radix* 10	*Radix* 10	
1	37	034034	$= 192^2$	$= 222^2$	$= 49284$	Trivial case since $a = 0$
2	148	125125	$= 374^2$	$= 444^2$	$= 197136$	Admissible result
3	333	283283	$= 556^2$	$= 666^2$	$= 443556$	Admissible result
4	592	499499	$= 738^2$	$= 888^2$	$= 788544$	Inadmissible since $b = c$
5	925	771771	$= 91X^2$	$= 1110^2$	$= 1232100$	Inadmissible since $a = b$

It follows that there are no other solutions.

Retrospective Observations. This process was exceptionally devoid of heuristic difficulties; in plain language, it was an easy problem. On looking back, this appears to have been due in large measure to the seemingly fortuitous circumstance that the three coefficients "happen" to be respectively 11^2, 11 and 1 times a common factor, $6^2 \cdot 37$. However, upon closer examination we discover this relation is not accidental, it arises inevitably from the identity:

$$(11^5 + 11^2)a + (11^4 + 11)b + (11^3 + 1)c \equiv (11^3 + 1)(11^2 a + 11b + c).$$

This the writer was not sharp enough to notice on his first approach. Therefore in constructing a "formal proof" it is unnecessary to expand the coefficients; in fact, better not. Rather, simply set forth the foregoing identity and note that $11^3 + 1 = 6^2 \cdot 37$. It is also clear, since every integer may be expressed uniquely in the scale of 11, that every positive integral value of m within its upper bound yields a distinct solution, though not necessarily one consistent with all the restrictions imposed upon a, b and c. Again, on reviewing the foregoing tabulation, we observe a conspicuous uniformity amongst the numbers in the first column headed "Radix 10". Each entry $= (222m)^2$. This we find, is not accidental. The explanation when discovered suggests a slightly different way of completing the solution to the problem, by arriving at once at the possible values of N_{10}, thus: If

$$N_{10}^2 = 6^2 \cdot 37(11^2 a + 11b + c) = 6^2 \cdot 37(37m^2),$$

then

$$N_{10} = 6 \cdot 37m = 222m,$$

from which we obtain N_{10}^2, thence N_{11}^2, $= abcabc$.

Can the foregoing method be used "for some other problem"—for example, if base 11 be replaced by another base? Yes. We have merely to replace 11 by r, the radix. We obtain

$$(r^3 + 1)(ar^2 + br + c) = N^2.$$

Now, if $r^3 + 1 = kp^2$, where p^2 denotes the greatest square dividing $r^3 + 1$, then for reasons similar to those already suggested

$$ar^2 + br + c = km^2, \qquad N_{10} = kpm,$$

where, disregarding for the moment the restrictions $a \neq 0$, $a \neq b \neq c$, m ranges through all positive integral values from unity through

$$\left[\sqrt{(r^2 + r + 1)(r - 1)/k}\right] = \left[\sqrt{(r^3 - 1)/k}\right] = \left[\sqrt{(kp^2 - 2)/k}\right] =$$

$$\left[\sqrt{p^2 - 2/k}\right] = p - 1,$$

where, as usual, $[x]$ denotes the greatest integer in x. Thus, still disregarding the restrictions just mentioned, there are always exactly $p - 1$ solutions. Also, there is no solution for any radix where $r^3 + 1$ is not divisible by a square number, for example in systems of numeration having base 4, 6, 9, 10, 12, 13, 16, 18, 21, 22 or 25. Now, it can readily be verified that $a > 0$ if, and only if, $km^2 \geqq r^2$; also, that $km^2 = r^2$ if, and only if, $a = 1$, $b = 0$, $c = 0$. It follows that for each radix there are in all

$$S = p - [r/\sqrt{k}] - 1$$

solutions restricted to integers composed of six significant digits, excluding the solution 100100 when it occurs, as it does, for example, when $r = 2$. If the restriction that a, b and c represent distinct digits be restored, all we can say is that there are then *at most* S solutions.

In solving for a particular radix, we find it convenient to use the formula $km^2_{10} = x_r$, where x denotes the number formed by the last three digits of the desired square. Thus, for radix 7, we have $7^3 + 1 = 344 = 86 \cdot 2^2$. Here $p = 2$, and we know at once that the solution, if there be one, is unique, and that m can assume only one value, unity. We then write $86_{10} = 152_7$, giving $abcabc = 152152_7 = 334^2_7 = 172^2_{10}$. One other example will suffice. For $r = 23$, we have $23^3 + 1 = 12168 = 2 \cdot 78^2$, which gives in all 77 solutions, 61 of which have six significant digits, which, however, in a number of instances are not distinct. Selecting as a sample $m = 17$, we find that $2 \cdot 17^2_{10} = 123_{23}$, and that $123123_{23} = 507^2_{23} = 2652^2_{10}$.

These observations have been confined to the pattern *abcabc*. It may well be that other patterns lend themselves to similar treatment.

Also solved by M. P. de Regt, Walnut Creek, Calif.; Dewey Duncan, East Los Angeles Junior College; A. L. Epstein, Cambridge Research Laboratories, Mass.; P. N. Nagara, College of Agriculture, Bangkok, Thailand; L. A. Ringenberg, Eastern Illinois State College; E. D. Schell, Arlington, Va.; and W. R. Talbot, Jefferson City, Mo.

Construction of a Trapezoid

64. [May 1950] *Proposed by D. L. MacKay, Manchester Depot, Vt.*

Construct a trapezoid $ABCD$ given its diagonals and its non-parallel sides.

I. *Solution by Dewey Duncan, East Los Angeles Junior College.* If the two given sides are equal, or if two given diagonals are equal, the trapezoid is necessarily isosceles, and *indeterminate,* for such data (two equal non-parallel sides or two equal diagonals) immediately yield an isosceles trapezoid inscriptible within any circle whose diameter exceeds the length of a given diagonal. Accordingly, a given pair of equal non-parallel sides and a given pair of unequal diagonals, or vice versa, are incompatible data.

Denote the non-parallel sides AD by a, BC by d, the parallel base AB by x, the diagonal AC by c, and the diagonal BD by b. The triangles ABD and ABC, having common base and altitude, have equivalent areas. Accordingly, Heron's formula yields the identity

$$(a+b+x)(a+b-x)(a-b+x)(-a+b+x) = (c+d+x)(c+d-x)(c-d+x)(-c+d+x),$$

from which one obtains

$$x = \sqrt{(a^2+d^2-b^2-c^2)(a^2+c^2-b^2-d^2)/2(a^2+b^2-c^2-d^2)}$$

$$= \sqrt{(e^2-f^2)(g^2-h^2)/2(k^2-q^2)} = mn/p,$$

where

$$\begin{array}{lll} e^2 = a^2+d^2 & h^2 = b^2+d^2 & m^2 = |e^2-f^2| \\ f^2 = b^2+c^2 & k^2 = a^2+b^2 & n^2 = |g^2-h^2| \\ g^2 = a^2+c^2 & q^2 = c^2+d^2 & p^2 = 2|k^2-q^2| \end{array}$$

Hence m, n, and p may be obtained by use of right triangle constructions, and x may be constructed as the fourth proportional of p, m, and n, if and only if an odd number of these relations hold: $e > f$, $g > h$, $k > q$. The construction of the trapezoid follows immediately.

II. *Solution by the Proposer.* If $ABCD$ is the required trapezoid, $AB \| CD$ and $AB < CD$, translate AD to BG and AC to BF and draw the circles $C_1 = B(BG)$, $C_2 = B(BC)$, $C_3 = B(BD)$, $C_4 = B(BF)$. The circle C_1 cuts CD in G and H, C_2 cuts it in K and C, C_3 cuts it in D and J, C_4 cuts it in E and F. Thus on DC we have the point order $EDKGHCJF$.

By the translation, $CJ + JF = AB = DK + KG$, and since the segments intercepted on any secant by two concentric circles are equal, $CJ = DK$. Hence $JF = KG = ED = HC$.

Let t_1 and t_2 be the tangents from E and D to circles C_1 and C_2. Then $t_1^2 = (EG)(EH)$, $t_2^2 = (DK)(DC)$. Since $EH = DC$,

$$EG : DK :: t_1^2 : t_2^2 \text{ and } ED : EK :: (t_1^2 - t_2^2) : (t_1^2 + t_2^2).$$

Since $t_1^2 = AC^2 - AD^2$ and $t_2^2 = BD^2 - BC^2$, t_1 and t_2 can be constructed and hence a right triangle can be constructed with $(t_1^2 - t_2^2)^{\frac{1}{2}}$,and $(t_1^2 + t_2^2)^{\frac{1}{2}}$ as legs. If m and n are the projections of these legs on the hypotenuse we know the locus of a point K such that $ED : EK :: m : n$, where E is any point on circle C_4.

Draw $KB' \| DB$ meeting EB prolonged at B'. Then $EB : EB' :: ED : EK :: m : n :: DB : KB'$. Thus, since $EB' = (n/m)EB$ and $KB' = (n/m)DB$, both constants, point B' is fixed, and the locus of K is the circle $B'\{(n/m)DB]$. Point K being determined, the trapezoid can be completed.

Also solved by Howard Eves, Oregon State College; H. E. Fettis, Dayton, Ohio; L. M. Kelly, Michigan State College; and W. I. Thompson, Los Angeles City College.

N. A. Court observes that method II is essentially the same as the one given by Julius Peterson, *Geometric Construction*, Copenhagen (1866), articles 296 and 142. The solution may be found again in Ivan Aleksandrov, *Geometric Constructions*, Moscow (1934), 112. G. Fontene discussed this solution in *Bulletin de Sciences Mathématiques et Physiques*, **8**, No. 11, 164-6, (1902-3). A detailed study of the trapezoid was given by L. Vautré in *Journal de Mathématiques élémentaires*, series 4, **3**, 99-107, (1894).

QUICKIES

From time to time as space permits 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.

Q 21. Four snails start at the vertices of a unit square and move directly towards one another in cyclic order, at unit rate. How far will they travel before they meet? [*Submitted by Leo Moser.*]

Q 22. Which of the following is a square number in the decimal scale?

(A)	13841287208	(E)	54875873526
(B)	27680640645	(F)	61919364224
(C)	34296447247	(G)	78364964096
(D)	42180533641	(H)	90458382179

Q 23. If $f(x,y)$ is not identically zero, and if $f(x,y) = (k)\ f(y,x)$ for all values of x and y, what are the possible values of k? [*Submitted by Leo Moser.*]

Q 24. Two factories *A* and *B* are located 1 and 4 miles respectively from the same side of a river with a straight bank. On the bank the nearest points, A_1 and B_1, from the factories are 12 miles apart. Where can a common loading dock be located so that the shortest railroad track may serve both factories? How long will the track be?

Q 25. If p_n denotes the *n*th prime, show that $p_1 p_2 \dots p_n + 1$ is not a perfect square. [*Submitted by Leo Moser.*]

Q 26. In the expansion of $(x^2 + y^3)^{15}$ find the sum of the coefficients of the alternate terms beginning with the second.

Q 27. Which has greater area, an isosceles triangle 13 by 13 by 10, or one 13 by 13 by 24? [*Submitted by Leo Moser.*]

ANSWERS

A 21. By symmetry, the snails will always be at the vertices of some square. Snail 1 is headed toward snail 2 at unit rate, while snail 2 has no component of velocity toward snail 1. Hence the square is decreasing at unit rate and will become a point in unit time. By this time the snails will each have traveled unit distance.

A 22. The terminal digit of a square number can only be 0, 1, 4, 5, 6 or 9 (this eliminates A and C). The penultimate digit of a square is 2 if the number terminates in 5 (this eliminates B), and is even unless the terminal digit is 6 in which case it is odd (this eliminates E and H). Any number is congruent to the sum of its digits, and a square is congruent to 0, 1, 4 or 7 modulo 9 (this eliminates F and G). That D = $(205379)^2$ may be confirmed in the conventional manner.

A 23. Since $f(x,y) = (k)\, f(y,x)$ for all x and y, we have $f(y,x) = (k)\, f(x,y)$ so that $f(x,y) = (k)^2 f(x,y)$ and since $f(x,y)$ is not identically zero this gives $k^2 = 1$ or $k = \pm 1$. That $k = -1$ is actually possible may be seen from $f(x,y) = x - y$.

A 24. Reflect *A* about the river bank into A'. Let AA' cut A_1B_1 at *C*. Then $ACB = AA' = \sqrt{(1 + 4)^2 + 12^2} = 13$ miles, the required shortest distance. Furthermore, $(A_1C) : (A'A_1) :: (CB_1) : (BB_1)$ or $A_1C/1 = (12 - A_1C)/4$, so $A_1C = 2.4$ miles.

A 25. Since $p_1 p_2 \dots p_n$ is divisible by 2 but not by 4, $p_1 p_2 \dots p_n + 1$ will leave a remainder of 3 on division by 4, while any square leaves a remainder of 0 or 1.

A 26. The set of 16 coefficients is palindromic, so each set of alternate coefficients has the same sum. Let $x = y = 1$, then the desired sum is $\frac{1}{2}(1 + 1)^{15}$ or 2^{14} or 16384.

A 27. They both have the same area since both can be made by placing two right angled triangles 5 by 12 by 13 back to back.

OUR CONTRIBUTORS

Raymond M. Redheffer was born in Chicago in 1921, and has lived there, in Florida, and in Massachusetts. He attended the Coburn School, the Asheville School, and M.I.T., where he obtained the doctorate in Mathematics in 1946. In 1942-1948 he was a member of the M.I.T. Radiation Laboratory, and subsequently he taught at M.I.T., Harvard University, and now at U.C.L.A.

René Maurice Fréchet was born in Maligny, France in 1878. After receiving the doctorate at the University of Paris ('06) he was appointed Maitre des Conferences, Rennes ('08), and Professor of Mechanics at Poitiers ('10). In 1919 he became Professor of Higher Analysis at Strasbourg. Dr. Fréchet has been Professor at the Institute Henri Poincaré, University of Paris since 1928.

Internationally famous among mathematicians for his pioneering work in general analysis and his many researches in analysis, probability and statistics, Professor Fréchet is the author of a number of books on various phases of advanced mathematics, perhaps the best known being "Les Èspaces Abstraits".

Eric Temple Bell, Professor of Mathematics, California Institute of Technology, was born in Aberdeen, Scotland, in 1883. Educated at London, Stanford (A.B., '04), Washington (A.M., '08) and Columbia (Ph.D., '12), he joined Washington's faculty and remained there until 1927 when he was appointed to his present position. A prominent figure in mathematical circles for many years, Dr. Bell has received many honors, including the Bocher prize ('24) and election to the National Academy of Science, for his researches in number theory. He has served as President, A.M.A. and Vice President, A.M.S. and of Section A, A.A.A.S. Widely known as an author, he has written over twenty books, ranging from adventure stories to "Algebraic Arithmetic".

Sister Helen Sullivan, O.S.B., attended the Catholic University of America (M.A., '31; Ph.D., '34). Sister Helen is Professor of Mathematics and chairman of the department at Mount Saint Scholastica College, Atchison, Kansas, and has held this position since 1934. She has also served as visiting professor at Loyola University of the South in New Orleans. Sister Helen is the author of a book for college students entitled "The Christian Approach to Science".

(Continued on Page 128)