MATHEMATICS

magazine

Address general editorial correspondence to the managing editor, Glenn James, special papers to the editors of the departments for which they 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. Because of the wide variety of papers published in this magazine, it is highly desirable that each introduction should state what the paper accomplishes and that each paper should be accompanied by a brief statement of what mathematical training is required to read it. Expository papers which can be read by men who are not specializing in the field in which a given paper is written, are especially desired.

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 cents.

Send *subscriptions* and related correspondence to Inez James, 14068 Van Nuys Blvd., Pacoima, Calif.

Concerning everything pertaining to *advertizing* in this magazine, contact Homer V. Craig, University of Texas, Austin 12, Texas.

At present we can supply copies of Vol. XXI, numbers 4 and 5 and would like to buy copies of numbers 1, 2 and 3. We, also, can supply a few copies of the following articles: "The Generalized Weierstrass Approximation Theorem" by Marshall H. Stone; "Orthogonal Polynomials and Polynomials in General" by the late J. A. Shohat, and other special papers appearing in this and subsequent issues.

A list of sponsor subscribers for Vol. XXII will appear in the Nov.-Dec. Issue.

CONTENTS

ON TWO NEW CHAPTERS IN THE THEORY OF PROBABILITY

Maurice Fréchet

Introduction

First Chapter: The probabilities associated with a system of events. A number of scattered memoirs, apparently unrelated, and written by authors unfamiliar with each other's work, have followed two fundamental innovations (apparently also independent), expressed by an inequality of Boole and an equality of Poincaré. The trait common to all these works is that they study the probabilities concerning a system of events, and their novelty consists of not restricting the discussion to the simple cases where the events are supposed mutually exclusive or independent. In a work in two parts* published in 1939 and in 1943, we have brought together the ideas contained in these papers in a systematic theory, completed them at a number of points, studied particular cases, and given applications. Thus, we believe, a new chapter in the calculus of probabilities has been inaugurated, which will continue to develop, and will find its place in the future treatises on probability.

If we may mention it here, the two booklets, appearing just before and during the war, have necessarily remained unknown to many of those who are interested in probability theory.

Second Chapter: Theory of random elements of any nature whatever. In discussing the first chapter, it was possible to restrict ourselves to the mention of the title of the new chapter and to refer to a work already published for the details. On the contrary, the subject of the second chapter (abstract random elements) is not yet in any book and, moreover, is still in the period of formation. We know of nothing on this subject except for the two memoirs we have written, of which the first has already appeared** and the second*** is still in press. Even if the third memoir which we are preparing is added together with the works on statistics in which we have considered the related concrete examples, the subject, far from being exhausted, still presents many unsolved problems. For this reason it would seem useful to summarize here the results of the second memoir, in order to draw the attention of scholars to a new domain which may prove to be very fruitful.

1. *Abstract random elements.* The theory of probability, after being almost exclusively devoted to the study of random numbers, has been extended to random points, random vectors, and more recently to random series and finally to random functions. But in the study of Nature and in the "Social" Sciences, as well as in various technical applications, one encounters various other random elements: random curves, random surfaces, etc. Is it necessary

---

to study each one of these categories separately and successively? Without
ignoring the differences in their properties, can one discover their common
traits?

We are going to show that some definitions, some theorems of the theory of
random numbers, can be extended in a convenient form to the study of random
elements of *any nature*. There will result a technical simplification and an
overall philosophical view analogous to that furnished by vector analysis
preceding the particular study of forces, of velocities, of rotations, etc.—
that is, of the various kinds of vectors.

But how can one discuss random elements of indeterminate nature? We will
proceed as in the theory of abstract spaces, by employing descriptive rather
than constructive definitions. It will suffice here—although this weak restric-
tion may be further weakened—to suppose that the random elements considered
are chosen at random in a *metric* space. The elements (or points) of this space
are unspecified, except that we suppose that with each pair of elements x, y of
the space is associated a number $(x,y) \geq 0$ such that

(i) $(x,y) = 0$ if, and only if, $x = y$.

(ii) $(x,y) = (y,x)$

(iii)   the "triangular" inequality

   (1) $(x,y) \leq (x,z) + (z,y)$

is satisfied for every three points x, y, z in the space.

Among all the definitions of the theory of random numbers which can be
extended to the theory of abstract random elements, we retain here only these:
the typical values (mean, etc. . . . .), the various measures of the dispersion,
and the various kinds of stochastic convergence. We will give here some of
their principal properties and refer to the memoir in press cited above for
the demonstrations.

We think it useful to insist on the *prodigious extension* which the simple
introduction of the notion of distance gives to the domain of validity of the
theory of random numbers.

Part I. Dispersion

2. *Mean deviation.* If X,Y,Z are random elements chosen by chance from
among the elements of a metric space $\mathfrak{D}$ whose elements are arbitrary in nature,
and if X,Y,Z are determined simultaneously by each trial of the same category
$\mathfrak{C}$ of trials, it can be shown that

   (2) $\sqrt[k]{\mathfrak{M}(X,Y)^k} \leq \sqrt[k]{\mathfrak{M}(X,Z)^k} + \sqrt[k]{\mathfrak{M}(Y,Z)^k}$

for each number $k \geq 1$, where $\mathfrak{M}U$ denotes the mean value* of the random
number U.**

We will say that a random element X of $\mathfrak{D}$ is *bounded in the mean of order* k

---

*It is, of course, assumed that the laws of probability of the random numbers
$(X,Y)^k$, etc., are known. If U is a random number, then $\mathfrak{M}U = \int_{-\infty}^{\infty} u dF(u)$, where $F(u)$
Prob $[U<u]$. (Ed. Note)

**When $k \leq 1$, (2) may be replaced by: $\mathfrak{M}(X,Y)^k \leq \mathfrak{M}(X,Z)^k + \mathfrak{M}(Z,Y)^k$.

if there exists at least one "certain" element a of $\mathfrak{D}$ such that the mean value $\mathfrak{M}(X,a)^k$ is finite.

It follows from (2) that if X is bounded in the mean of order $k \geqq 1$, then $\mathfrak{M}(X,b)^k$ is finite for each "certain" element b of $\mathfrak{D}$. We will call the quantity $\sqrt[k]{\mathfrak{M}(X,Y)^k}$ the mean deviation of X and Y of order k. We will call the lower bound of the mean deviation of order k of X with a "certain" element a of $\mathfrak{D}$ when a varies arbitrarily on $\mathfrak{D}$, the mean deviation of X of order k. One can consider this mean deviation as giving one of the possible methods* of evaluating the *dispersion* of X. One may say that X is bounded in the mean of order k when its mean deviation of order k is finite.

3. *Typical positions.* When X is bounded in the mean of order k, every definite element $\gamma^{(k)}$, if such exist, such that the mean deviation of order k of X is equal to the mean deviation of order k of X and $\gamma^{(k)}$, will be called a *typical position of order* k of X. In the case where X is a random number and $(x,y) = |x-y|$, the typical positions of orders one and two are none other than the median and the mean of X.

In general, the application of the preceding definitions will vary not only with the set of points of the space $\mathfrak{D}$ but with the definition of the distance adopted for $\mathfrak{D}$, as we will see in examples.

4. *Case of infinite dispersion.* When X is not bounded in the mean of order k, one can define a typical position of order k of X in a manner analogous to the preceding, by substituting for $\mathfrak{M}(X,a)^k$, which is infinite, a sort of "finite part" of this quantity. To do this we choose arbitrarily a fixed point b of $\mathfrak{D}$ and designate by $X_n(b)$, or more briefly, by X_n, a random element identical with X when $(X,b) \leqq n$, and identical with b when $(X,b) > n$. We designate by m_n the lower bound ($\geqq 0$) of $\mathfrak{M}(X_n,a)^k$, for a in $\mathfrak{D}$. Finally, consider the difference

$$\phi_n(a) = \mathfrak{M}(X_n,a)^k - m_n.$$

It is a non-negative function of a. Let $h(a)$ be the lower limit of $\phi_n(a)$ when $n \to \infty$, i.e.

$$h(a) = \lim_{n\to\infty} \inf \phi_n(a)$$

We shall say that X has a typical position of order k when

1. $h(a)$ is finite for at least one position of a. One may then call $h(a)$ the finite part of $\mathfrak{M}(X,a)^k$.

2. $h(a)$ attains its lower bound (necessarily finite and non-negative) for at least one location of a independent of b. This point will be called a typical position of order k of X.

It can be shown that this definition, which also applies when X is bounded in the mean of order k, is equivalent in this case to the original definition (sec. 3).

Remark. It might appear more natural to define the typical position $\overline{X}$ as the limit of $\overline{X}_n$ as $n \to \infty$. But even when X_n is bounded, $\overline{X}_n$ does not necessarily

*We shall see others below.

exist. However, one may show that if X is bounded in the mean of order k, if X_n has a typical position $\overline{X}_n$ of order k, and if $\overline{X}_n$ tends to a limit, then this limit is a typical position of order k of X.

5. *Particular Cases.*

I. When X is a random number, and $(X,a) = |X-a|$, and if X has a mean value $\overline{X}$ in the classical sense, that is, if the integral

$$\overline{X} = \int_{-\infty}^{\infty} x \; d \text{ Prob } [X<x]$$

is absolutely convergent, X has a typical position of order 2 in the above sense, namely $\overline{X}$. But the converse is not true. We refer for more details to our article, "Nouvelles définitions de la valeur moyenne et des valeurs équiprobables d'un nombre aléatoire," Ann. Univ. Lyon, 3rd Series, Section A (1946), pp. 5-26.

II. Consider the case where the random element X is a numerical function $X(t)$ of one real variable t, the function being chosen at random in a family $\mathfrak{F}$ of such functions defined, for example, on the fixed line segment $S: a \leqq t \leqq b$. One may define the distance in various ways. The mean $\overline{X(t)}$ of $X(t)$, that is, the typical position of order two, will vary according to the definition adopted for the distance between two functions of $\mathfrak{F}$. It is necessary to distinguish between $\overline{X(t)}$ and the function $\overline{X}(t)$ which is equal, for each value of t, to the mean value of the random number $X(t)$. These functions may be identical or *distinct* depending on the particular case.

(i) If $\mathfrak{F}$ is the family of functions whose squares are integrable on S and if

$$(f,g) = \sqrt{\frac{1}{b-a} \int_a^b [f(t) - g(t)]^2 dt} \; ,$$

as in the method of least squares, then one has

(3) $\overline{X(t)} = \overline{X}(t)$

in very general cases, including in particular the simple case where $X(t)$ has only a finite number of determinations $x_1(t), x_2(t), \cdots x_n(t)$, each having an integrable square on S.

(ii) On the contrary, if $\mathfrak{F}$ is the family of all continuous functions on S, and if we put

$$(f,g) = \max_{a \leqq t \leqq b} |f(t) - g(t)| \; ,$$

then (3) is not necessarily true. As an example, consider the case where $X(t)$ has only two determinations, $x_1(t)$ with probability p and $x_2(t)$ with probability q, where $x_1(t) \leqq x_2(t)$, and where

$$q \leqq m^2 / M^2 \; ,$$

in which m and M denote the minimum and maximum of $x_2(t) - x_1(t)$, respectively. Then one may show that

$$\overline{X(t)} = x_1(t) + qM,$$

whereas

$$\overline{\overline{X}}(t) = x_1(t) + q[x_2(t) - x_1(t)] ,$$

so that the two functions $\overline{X}(t)$ and $\overline{\overline{X}}(t)$ are distinct, except when $x_2(t) - x_1(t)$ is constant, or when $p = 1$. We refer for more details and for the study of further examples to our memoir cited above (p. 4).

6. *Global elements and their distances.* The inequality (2) resembles the triangular inequality (1), and this fact leads to the conception of what we shall call *global elements.* If each trial of a certain category $\mathfrak{C}$ of trials determines simultaneously the random elements $X, Y, Z, \cdots$, *one may consider these elements as transformations*

$$X = \mathfrak{F}(R),$$
$$Y = \mathfrak{U}(R),$$
$$Z = \mathfrak{B}(R),$$
$$\cdots\cdots\cdots$$

of the result R of each trial into a corresponding position of each of the elements X, Y, Z.

If, for example, $X, Y, Z, \cdots$, are taken at random in a metric space $\mathfrak{D}$ and are each bounded in the mean of order $k \geqq 1$, it follows from (2) that each of the means $\mathfrak{M}(X,Y)^k$, $\mathfrak{M}(Y,Z)^k$, $\cdots$, is finite. Now they are determined by the transformations (or functions) $\mathfrak{F}, \mathfrak{U}, \mathfrak{B}, \cdots$ of R. One may then consider

$$\sqrt[k]{\mathfrak{M}(X,Y)^k}$$

as a distance between the functions $\mathfrak{F}, \mathfrak{U}$ taken in a certain function space $\mathfrak{B}_k$. Remembering that $\mathfrak{F}$ or $\mathfrak{U}$ generate the *entire set* of determinations of X or Y, we may speak of this distance as the distance between the "global elements $[X], [Y]$." This expression and this notation emphasize the fact that we are interested here in the distance between the entire or "global" set of determinations of X and the set of determinations of Y rather than the distance between the determinations of X and Y from one and the same trial. One can say that the distance of these global elements is the measure of a sort of dispersion of the pair X, Y determined by each trial. This distance may be represented by the notation

$$(4) \qquad ((\,[X]\,, \,[Y]\,)) = \sqrt[k]{(X,Y)^k}$$

It is clear that this expression defines a distance satisfying conditions (*i*), (*ii*), (*iii*), (sec. 1), providing that we agree to consider two global elements the same, if the corresponding random elements coincide at each trial, except perhaps in a case of probability zero.

But this introduction of a space of global elements—where the distance is defined by the formula (4) in which (X,Y) is the random distance in $\mathfrak{D}$ of X,Y, in the same trial—may be carried out in a natural way with the same global elements but with other non-equivalent expressions of the distance. We shall see an example later. This will lead us to a more general concept. Considering again the transformations $\mathfrak{F}, \mathfrak{U}, \mathfrak{B}, \cdots$, associating with the result R of each

trial the elements X, Y, Z, $\cdots$ of the space $\mathfrak{E}$ (which does not even have to be supposed metric in advance), we may consider these transformations as belonging to a function space Δ.

We may make correspond to each transformation $\mathfrak{J}$ a "global element" $[X]$ representing the set of determinations of X (but essentially related to the set of results R by the transformation $\mathfrak{J}$) and consider Δ as a space of global elements defined for the same category $\mathfrak{E}$ of trials.

Thus (as shown by the example of B_k, above), one may consider Δ as a metric space by associating with each pair of global elements $[X]$, $[Y]$ of Δ a distance $([X], [Y])$ satisfying the conditions (i) (ii) (iii), (sec. 1). But if we wish to take account of the fact that the result R is obtained by chance, then the distance $([X], [Y])$ should be a sort of dispersion of the pair X,Y, that is, the contribution to this distance of pairs corresponding to a set of results R of small probability whould be small. This is the case, for example, in the expression (4).

When the space $\mathfrak{E}$ from which the elements X,Y are selected at random is a metric space, we see that the distance (X,Y) is a random number, whereas the distance $([X], [Y])$ is a "certain" number determined by the transformations $\mathfrak{J}$, $\mathfrak{U}$ and the law of probability of R.

7. *Generalization of the dispersion and typical positions.* Whether the space $\mathfrak{E}$ in which X,Y, $\cdots$ are chosen at random is metric or not, we suppose that a distance has been defined for the space Δ of global elements $[X]$, $[Y]$, $\cdots$ Let a denote a "certain" element of $\mathfrak{E}$ and consider the distance $\phi(a) = ([X], [a])$ in Δ. If $\phi(a)$ is finite for at least one position of a, one can consider the lower bound of $\phi(a)$ as a measure (associated with Δ) of the dispersion of X. And if, in addition, this bound is attained for at least one determination $\widehat{X}$ of X, so that $\widehat{X}$ is a "certain" element of which the distance in Δ to the global element $[X]$ is the smallest, and therefore may be considered as the "certain" element most representative of the global element, one may consider $\widehat{X}$ as a typical position of X (relative to the distance defined in Δ).

When the distance $([X], [a])$ is infinite, one can try to extend the definition of the typical position in a manner analogous to that already presented.

Part II. *Stochastic Convergence*

8. *Definitions of various types of stochastic convergence.* One says that a sequence y_n in a metric space D tends to x in D if the distance (y_n, x) tends to zero. When in each trial of the category C, the random elements X, X_1, X_2, $\cdots$, chosen by chance in the same metric space D, are determined simultaneously, it may happen that in each trial, $X_n \to X$ when $n \to \infty$. Besides this classical type of convergence of $[X_n]$ towards $[X]$, the calculus of probabilities leads one to consider various types of *stochastic* convergence, in which a given property of the classical convergence is realized *except* perhaps with a small probability when n becomes large. Among these various types, we consider at first the following and indicate other more general types later.

$X_n \to X$ *almost certainly* as $n \to \infty$; when the probability that $X_n \to X$, in a given trial, as $n \to \infty$ is equal to one.

$X_n \to X$ *in the mean of order* k $(k \geq 1)$ as $n \to \infty$ if $\mathfrak{M}(X_n, X)^k \to 0$ with $1/n$.

$X_n \to X$ *in probability* as $n \to \infty$ if, for each $\epsilon > 0$, one has

$$\lim_{n \to \infty} \{\text{Prob. } [(X_n, X) \leq \epsilon]\} = 1$$

$X_n \to X$ *lawfully* as $n \to \infty$ if the law of probability of X_n tends* toward that of X when $n \to \infty$.

One sees here again how the simple introduction of the notion of distance permits an immediate and far reaching extension of the domain of validity of the definitions for stochastic convergence of random numbers.

It can be shown that if X and X_n are points of a Cartesian space $\Re_s$ of s dimensions, where

$$(X, \ X_n) = \sqrt{\sum_{j=1}^{s} \ (X^{(j)} - X_n^{(j)})^2} \ ,$$

the necessary and sufficient condition that $X_n \to X$ almost certainly, or in mean of order $k > 1$, or in probability is that each coordinate X_n^j of X_n converge almost certainly, or in mean of order k, or in probability, respectively, towards the corresponding coordinate $X^{(j)}$ of X. One can also demonstrate that either almost certain convergence or convergence in the mean of a given order k implies convergence in probability. Moreover, if $X_n \to X$ in probability when $n \to \infty$, there exists a subsequence $Y_r = X_{n(r)}$ of the sequence X_n such that $Y_r \to X$ almost certainly.

In certain very general cases, convergence in probability is actually equivalent to convergence in the mean of order k. In order to state sufficient conditions of this equivalence, we define the random elements Y of a given family $\mathfrak{F}$ of elements to be *uniformly summable of order k*, if for at least one "certain" element a of $\mathfrak{D}$, the numbers

$$U_n = \begin{cases} (a, \ Y) & \text{for } (a, \ Y) \geqq n \\ 0 & \text{otherwise} \end{cases}$$

converge in the mean of order k to zero, uniformly on $\mathfrak{F}$. (That is, for every $\epsilon > 0$, there exists an integer N, independent of Y, such that $\mathfrak{M} U_n^k < \epsilon$ for $n > N$ and for every Y in $\mathfrak{F}$.)

Moreover, if this property holds for an element a, it will hold for every "certain" element b of $\mathfrak{D}$. A simple particular case where the Y of $\mathfrak{F}$ are uniformly summable of order k for every $k \geqq 1$ is that in which the Y are uniformly bounded in the sense that there exists a "certain" element a of $\mathfrak{D}$ and a finite number A such that $(a, \ Y) < A$ in every trial and for every element Y of $\mathfrak{F}$. (Again if this holds, then for every "certain" element b of $\mathfrak{D}$, there exists a number B such that $(b, Y) < B$ in every trial and for every Y of $\mathfrak{F}$.)

Now in order for convergence in probability of X_n to X to be equivalent to the convergence in the mean of order $k \geqq 1$ of X_n to X, it is sufficient that the elements X_n and X be uniformly summable of order k.

9. *Stochastic convergence deduced from a distance.* We have seen in sec. 6 that the expression $\sqrt[k]{\mathfrak{M}(X,Y)^k}$ can be considered as a distance $((\ [X] , \ [Y] \ ))$ of the global elements $[X]$, $[Y]$. Thus convergence in the mean of order k of X_n towards X can be defined, as in classical analysis, by means of a distance: to say that X_n converges in the mean of order k to X is equivalent to saying that the distance $((\ [X_n] , \ [X] \ ))$ between the global elements $[X_n]$ and $[X]$ tends to zero.

It is interesting to show that convergence in probability may also be

*We will make this definition precise further on (sections 13 and 16).

expressed by means of a distance. For this purpose we generalize the expressions that we have given in the case of random numbers. One may take for a new distance between the global elements $[X]$, $[Y]$ the expression

(5) $(((\ X \ , \ Y \ ))) = \inf_{\epsilon > 0} \{ \ \epsilon + \text{Prob} \ [(X,Y) \geqq \epsilon] \ \}$.

It can be shown that: (1) this expression is a distance satisfying the conditions (i), (ii), (iii) of sec. 1 (providing that two random elements are again considered equal when they are identical at each trial except for a case of zero probability), (2) to say that X_n converges to X in probability is to say that the distance between the global elements $[X_n]$ and $[X]$ tends to zero, the distance being defined as in Eq. (5).

10. *Complete spaces.* We now introduce the notion of a *complete* metric space. It is clear that in any metric space $\mathfrak{D}$, if the sequence of elements x_1, x_2, x_3, $\cdots$ is convergent then one has

(6) $\lim \ (x_n, \ x_n) = 0$, as $\frac{1}{n} + \frac{1}{m} \to 0$.

But the converse is not always true. We say that the metric space $\mathfrak{D}$ is *complete* if the Cauchy criterion (6) is not only necessary but also sufficient for the convergence of the sequence x_n.

Now let us suppose that the random elements X, Y, $\cdots$ are chosen by chance from a complete metric space $\mathfrak{D}$. Let $\mathfrak{B}_k$ and $\mathfrak{P}$ denote the spaces consisting of the same global elements $[X]$, $[Y]$, $\cdots$ but with the distances defined by formulas (4) and (5), respectively. Then one can show that the *two spaces $\mathfrak{B}_k$ and $\mathfrak{P}$ are complete.*

As we have seen, convergence in probability can be expressed in terms of a metric, such as that given by (5). However, our expression has the inconvenience of requiring the addition of unlike quantities. For example, (X,Y) and ϵ might represent the measure of a length, and depend on the units chosen, whereas the probability term occurring in formula (5) is, of course, a pure number. To avoid this inconvenience, Ky Fan has proposed a different expression (for the case of random numbers) which applies immediately to the actual case. Namely, one may also express convergence in probability by means of another distance

$((((([X],[Y])))) = \inf \{\epsilon > 0; \ \epsilon^{-1} \ \text{Prob} \ [(X,Y) > \epsilon] < 1\}$

Finally, one could use for the same purpose any of the infinite number of metrics representable by the expression $\mathfrak{M} f((X,Y))$, where $f(\lambda)$ is a non-negative continuous, increasing function, defined for $\lambda \geqq 0$, zero for $\lambda = 0$ and such that $f(\lambda+\mu) \leqq f(\lambda) + f(\lambda)$ for all λ, $\mu \geqq 0$.

11. *New types of stochastic convergence.* Let $X, Y, \cdots$ be chosen at random from a space $\mathfrak{E}$ (metric *or not*). Suppose that a distance $([X],[Y])$ is defined in any manner in the space Δ of global elements. By means of this distance one may define a corresponding type of stochastic convergence: a sequence X_n will be said to converge stochastically to X (in the sense of this new definition) if the distance $([X_n],[X])$ tends to zero.

12. *Limits of dispersions and typical positions.* Beginning with a given definition of the distance $([X],[Y])$ of two global elements $[X]$ and $[Y]$, it can be shown that if X_n converges stochastically·towards X (in the sense

corresponding to the given definition of distance), then the dispersion D_n of X_n converges to the dispersion D of X, where these dispersions are those associated with the above definition of distance.

In the case in which each X_n has a typical position $\widehat{X}_n$ (also associated with this distance), one has

$$D = \lim_{n \to \infty} ([X], [\widehat{X}_n]).$$

And if, in addition, one can find a subsequence of the X_n which converges to a "certain" point $\widehat{a}$, then: (1) $\widehat{X}$ also has at least one typical position, (2) the limit a of $\widehat{X}_n$ is one of the typical positions of X.

13. *Law of probability of a random element.* We consider an element X chosen at random in a space $\mathfrak{G}$(metric or not). The law of probability is determined if we know the probability that X satisfies an arbitrary condition γ. This condition will be satisfied for some points of $\mathfrak{G}$ belonging to a set E. Thus, to know the law of probability of X is equivalent to knowing the values of the set function $p(E)$, which represents the probability that X belongs to the set of points E. It is always necessary to restrict to some extent the arbitrariness of E. Thus in the case where $\mathfrak{G}$ represents a straight line segment of length one, and X is a point selected at random from this segment [with uniform distribution of the probabilities, we know that the probability $p(E)$ is the measure of the set E of points of the line, and we must suppose that E ranges over *measurable* sets. We do not know how to assign a probability to non-measurable sets.

Returning to the general case where $\mathfrak{G}$ is any space, we shall suppose then that $p(E)$ is defined only for certain sets E called *"probabilisable"* (relative to the random element X). In order for the theorem on total probabilities to apply, we must assume that if two disjoint sets E, F (i.e., having no elements in common) are probabilisable, then so is their sum (or union), and if E is a part of G, and E and G are probabilisable, then so is their difference $G - E$. Finally, $p(\mathfrak{G}) = 1$. Moreover, we take $p(E)$ to be an additive* set function defined on the additive family W of probabilisable sets. Moreover, if we wish to extend the theorem on total probabilities to the sum of a denumerable number of incompatible events, W and p must be completely additive.

14. *Distribution functions.* We shall call the set function

$$p(E) = \text{Prob } [X \in E]$$

the *distribution function* of X. One might object to this definition, since when X is a number it is usual to determine the law of probability by an "apportioning function" of the form

$$F(x) = \text{Prob } [X < x].$$

That is, in this case it is sufficient to know $p(E)$ for the special sets $X < x$, i.e., for the half-lines. But this is because it is possible to deduce the value of $p(E)$ for any probabilisable set from the values of $F(x)$. It is very clear that it is $p(E)$ which determines completely the law of probability and

---

*A family of sets is said to be additive if the sum of any finite number of sets of the family is also in the family. A set function $f(E)$ is said to be additive if $f(E_1 + \ldots + E_n) = f(E_1) + \ldots + f(E_n)$ for disjoint sets E_j. If these statements can be extended to a denumerable number of sets, the family and the function are each called completely additive. (Ed. note)

that the knowledge of $F(x)$ is just a simplification, which solves the problem indirectly. Even in the most practical applications it is necessary to find the probability that a random point X belongs to one or to several segments, and it is necessary to *deduce* from $F(x)$ the value of this probability. However, the function $F(x)$ is certainly useful, and in the general case where X is an abstract random element, one might also like to be able to deduce the values of the set function $p(E)$ from a function of one or more numerical variables. There is a case, theoretically rather special, but quite general enough for applications, in which this can actually be done.

15. *Separable Metric Spaces*. Let $\mathfrak{G}$ be a separable metric space, that is a metric space S containing a denumerable set N: a_0, a_1, a_2, $\cdots$, such that every element x of S either belongs to N or is a point of accumulation of N (that is, S is the closure of N). Then to each element x of S corresponds the sequence of numbers, which one may call the coordinates of x:

$$x_1 = (x, a_1) - (a_0, a_1), \cdots, \quad x_n = (x, a_n) - (a_0, a_n), \cdots$$

Evidently we have $|x_n| \leqq (x, a_0)$, so that the coordinates of x form, for each x, a bounded set of real numbers. One can prove that $(x, y) =$ least upper bound of $|x_n - y_n|$, for $n = 1, 2, 3, \cdots$ In particular, it follows that distinct points have distinct sets of coordinates.

By means of these coordinates we can generalize the notion of the "apportioning function" to the case of a random variable X taken from a separable metric space by defining

(7) $F(x_1, x_2, \cdots, x_n, \cdots) = \text{Prob.}\ [X_1 < x_1, X_2 < x_2, \cdots, X_n < x_n, \cdots]$,

where $x_1, x_2, \cdots, x_n \cdots$ are arbitrary real numbers.* If one prefers, one may substitute for the function F of an infinite number of variables, an infinite sequence of functions F_n of a finite number of variables defined by

$$F_n(x_1, x_2, \cdots, x_n) = \text{Prob}\ [X_1 < x_1, X_2 < x_2, \cdots X_n < x_n].$$

Knowing the function F is equivalent to knowing the sequence of functions F_n, since one has $F_n(x_1, \cdots, x_n) = F(x_1, x_2, \cdots x_n, +\infty, +\infty, \cdots)$ and $F(x_1, x_2, \cdots x_p, \cdots) = \lim_{n \to \infty} F_n(x_1, \cdots, x_n)$.

From the function F or functions F_n for X, we may deduce the expression

$$G(x_1, x_2, \cdots, y_1, y_2 \cdots) = \text{Prob}\ [x_1 \leqq X_1 < y_1, x_2 \leqq X_2 < y_2, \cdots]$$

$$= \lim_{n \to \infty} \text{Prob}\ [x_1 \leqq X_1 < y_1, \cdots, x_n \leqq X_n < y_n]$$

$$= \lim \triangle_n F_n(x_1, \cdots, x_n),$$

where $\triangle_n$ denotes the nth difference of $F_n(x_1, \cdots, x_n)$, when we give to x_k the increment $y_k - x_k$, $1 \leqq k \leqq n$. It follows that if we know the apportioning function F for X, we can find the distribution function $p(E)$ for the simple sets

(8) $x_1 \leqq X_1 < y_1, \cdots, x_n \leqq X_n < y_n, \cdots$

which might be called semi-open parallelopipeds. By means of the theorem of total probabilities one can then find the value of $p(E)$ for all the sets that can be expressed by a denumerable sequence of additions or subtractions of

---

*We suppose that the sets such as those written in the brackets are probabilisable. This is the generalization of the hypothesis made implicitly in the case where X is a random number.

parallelopipeds of the form (8).

In our memoir cited above (p.4), we show how the coordinates of X may be given a simple significance.

16. *Lawful convergence.* Let $P_n(E)$, $P(E)$ denote the distribution functions of X_n, X, respectively, where X_n, X are elements chosen at random in a metric space $\mathfrak{D}$. It can be shown that, if X_n converges in probability to X, $P_n(E)$ tends to $P(E)$ for every set E such that the set function $P(h)$ is continuous for $h = E$. It is necessary to make the meaning of this last restriction precise. It is sufficient to suppose that if f is the boundary* of E, then $P(f) = 0$, or what is the same thing, that $P(i) = P(E) = P(j)$, where i denotes the interior of E and j the complement of the interior of the complementary set $\mathfrak{D} - E$. This definition not only suffices to make the above theorem true, but this condition $P(f) = 0$ is necessary in order that the theorem be correct, that is, no less strict definition of continuity would do. We are then led to say in general that Y_n converges lawfully to Y when the distribution function $\pi_n(E)$ of Y_n converges to the distribution function $\pi(E)$ of Y for every set E *such that $\pi(h)$ is continuous for $h = E$.*

We have just seen that convergence in probability implies lawful convergence. However, as is known in the case in which $\mathfrak{D}$ is a line, the converse is not true.

One can obtain a necessary and sufficient condition by generalizing a proposition obtained by Kozakiewicz in the case of random numbers. We have shown that if X and the X_n are chosen at random in a metric space $\mathfrak{D}$ which is separable and complete, then putting

$$p_n(E,F) = \text{Prob} \; [X_n \in E, \; X \in F]$$
$$p_{n,m}(E,F) = \text{Prob} \; [X_n \in E, \; X_m \in F],$$

the necessary and sufficient condition that X_n converge in probability to X is that $p_n(E,E)$ converges to $P(E)$ for every set E for which $P(h)$ is continuous for $h = E$. (In the demonstration, it is assumed that every sphere is probabilisable for X, for X_n and for the pairs X, X_n. The condition that $\mathfrak{D}$ be separable and complete doesn't come in except in the proof of sufficiency).

It is interesting to give a condition for the convergence of X_n in probability not involving the knowledge of the limit X. In order for the sequence X_n of elements chosen at random from a separable, complete metric space to converge in probability, it is necessary and sufficient (1) that $p_{n,m}(E)$ converge as $1/n + 1/m \to 0$ when E belongs to a certain family $\bar{\Phi}$ of sets, (2) that the limit $\pi(E)$ thus obtained has the properties of a distribution function, that is that $\pi(E)$ is a non-negative completely additive set function with $\pi(\mathfrak{D}) = 1$, (3) that the family $\bar{\Phi}$ contain all the sets E such that $\pi(h)$ be continuous for $h = E$. If these conditions are satisfied then, if X is the limit of X_n in probability, one can take $\pi(E)$ as the distribution function of X.

17. *Almost certain convergence.* Let $X, X_1, X_2, \cdot \cdot \cdot$ be chosen at random in a metric space $\mathfrak{D}$, let C designate the event "convergence of the X_n in one trial, and let Γ designate "convergence of X_n to X in one trial." One can extend the formula of Kolmogoroff established for the case of random numbers to this more general case, as follows:

$$(9) \qquad \text{Prob} \; \Gamma = \lim_{\epsilon \to 0} \{ \; \lim_{n \to \infty} (\lim_{m \to \infty} \text{Prob} \; [\Gamma_{n,m}(\epsilon)]) \},$$

---

*The interior i of E is the set of points X of E such that each x is the center of a sphere: $(x,y) \leqq r$ lying entirely within E. The *boundary* f of E is the set of points of $\mathfrak{D}$ which belong neither to i nor to the interior $\mathfrak{D} - j$ of $\mathfrak{D} - E$.

where $\Gamma_{n,\mathbf{m}}(\epsilon)$ is the event consisting in the simultaneous realization of the relations:

$$(X, X_{n+1}) < \epsilon, \ (X, X_{n+2}) < \epsilon, \cdot \cdot \cdot, \ (X, X_{\mathbf{m}}) < \epsilon,$$

and similarly, if we suppose that the space $\mathfrak{D}$ is *complete*:

$$(10) \qquad \text{Prob } C = \lim_{\epsilon \to 0} \left\{ \lim_{n \to \infty} \left(\lim_{\mathbf{m} \to \infty} \text{Prob } [H_{n,\mathbf{m}}(\epsilon)] \right) \right\},$$

where $H_{n,\mathbf{m}}(\epsilon)$ is the event consisting in the simultaneous realization of the relations:

$$(X_n, X_{n+1}) < \epsilon, \ (X_n, X_{n+2}) < \epsilon, \cdot \cdot \cdot, \ (X_n, X_{\mathbf{m}}) < \epsilon.$$

From relations (9) and (10) we can immediately obtain conditions for almost certain convergence involving a triple limit. By generalizing a remark of Kozakiewicz, however, we can restrict ourselves to double limits and state: when the X_n and X are chosen at random in a metric space, (1) the necessary and sufficient condition for X_n to converge almost certainly to X is that

$$\lim_{n \to \infty} \left(\lim_{\mathbf{m} \to \infty} \text{Prob } \Gamma_{n,\mathbf{m}}(\epsilon) \right) = 0, \quad \text{for every } \epsilon > 0;$$

(2) for X_n to converge almost certainly (when $\mathfrak{D}$ is complete), it is necessary and sufficient that

$$\lim_{n \to \infty} \left(\lim_{\mathbf{m} \to \infty} \text{Prob } H_{n,\mathbf{m}}(\epsilon) \right) = 0$$

for every $\epsilon > 0$. The condition that $\mathfrak{D}$ be complete is not used in the proof of the necessary condition.

One can also formulate the conditions for almost certain convergence by means of functions analogous to distribution functions without using the distance (explicitly) in the definitions. When X and X_n are chosen at random in a separable metric space, the necessary and sufficient condition for X_n to converge almost certainly to X is that

$$p(E) = \text{Prob } [X \epsilon E] = \lim_{n \to \infty} \lim_{\mathbf{m} \to \infty} \text{Prob } [X \epsilon E, \ X_n \epsilon E, \ X_{n+1} \epsilon E, \cdot \cdot \cdot X_{\mathbf{m}} \epsilon E]$$

for every set E such that $p(h)$ is continuous for $h = E$. The hypothesis that $\mathfrak{D}$ is separable is used only in the proof of sufficiency. When we know only the X_n's, we can say that if they are chosen at random in a separable and *complete* metric space $\mathfrak{D}$, then in order that X_n converge almost certainly, it is necessary and sufficient that if in the expression

$$\text{Prob } [X_n \epsilon E, \ X_{n+1} \epsilon E, \cdot \cdot \cdot, \ X_{\mathbf{m}} \epsilon E]$$

we let m and then n tend to infinity:

(1) the expression tends to a limit when E belongs to a certain family ψ of sets, (2) that this limit $\pi(E)$ be a distribution function, (3) that the family ψ include all the sets E such that $\pi(h)$ is continuous for $h = E$. When those conditions are satisfied, we may take $\pi(E)$ as the distribution function of the "almost certain" limit of X_n. Once again the hypothesis that $\mathfrak{D}$ is separable and complete is used only for the sufficient condition.

18. *Stochastic continuity.* When to each value of a real variable t there corresponds an element $X(t)$ chosen at random in a metric space $\mathfrak{D}$, the set of values of $X(t)$ defines a random function (not numerically valued) of t. It is clear that to each type of stochastic convergence there corresponds a definition of stochastic continuity of $X(t)$. We refer to our memoir cited p.   , for the study of these various types of stochastic continuity.

Institut Henri Poincaré, Paris

ON A GROUP OF CONTACT TRANSFORMATIONS

Nilos Sakellariou

1.   It is known that an element of contact in the Euclidian three dimensional space is the set of a point and a plane which passes through this point. Such an element is defined by its five coordinates; these are the coordinates of a unit vector perpendicular to the plane [1].

In the present paper we consider a family or a multiplicity of elements of contact depending on a parameter. An element of this multiplicity is defined by the vector of its point X, that is by $\overline{OX} = \overline{x}$ (x_1, x_2, x_3) = $\overline{x}$ (x_i), i = 1, 2, 3, and the unit vector $\overline{XT} = \overline{t}$ (t_1, t_2, t_3) = $\overline{t}$ (t_i), which is perpendicular to the plane of the element and lies on the plane-direction (P) of the element, whereby $\overline{XB} = \overline{b}$ (b_1, b_2, b_3) = $\overline{b}$ (b_i) is the unit vector perpendicular to (P). The space is referred to a set of dextro rotatory rectangular axes $ox_1\ x_2\ x_3$. We consider the straight line which is determined by $\overline{t}$ as tangent to the edge of regression of the developable surface, which is tangent or circumscribed to the mentioned multiplicity. If the element is transferred on the direction of the straight line of $\overline{t}$ to a distance c, X^* is the new place of X and $0x_1^*\ x_2^*\ x_3^*$ is another set of dextro rotatory rectangular axes, the vector $\overline{OX^*} = \overline{x}^*$ (x_i^*) is referred to the x_i^*-system, then we say that we submit our multiplicity to a group of transformations

$$(1.1) \quad \begin{cases} \overline{x} = \overline{a}_j x_j + \overline{a}_j t_j \cdot c, \ j = 1, 2, 3 \\ \overline{t} = \overline{a}_j t_j \end{cases}$$

where $\overline{a}_j = \overline{OA}_j$ are unit vectors of the x_i-axes and a_{ji} their coordinates with respect to the x_i^*-system. C and a_{ji} are (real) constants. These a_{ji} are the coefficients of an orthogonal transformation with determinant

$$(1.2) \quad |a_{ji}| = \delta = +1, \quad \text{and they satisfy the relations}$$

$$(1.3) \quad \left.\begin{array}{l} \Sigma_i\ a_{ji} a_{ki} \\ \Sigma_i\ a_{ij} a_{ik} \end{array}\right\} = \ 0 \text{ or } 1 \text{ according as } j \neq k \text{ or } j = k.$$

We determine our multiplicity as one-parameter family, the elements depending on an invariant paramater ω, which is the angle of $\overline{t}$ with a fixed direction. We will find invariant expressions (2,3,4) concerning our multiplicity submitted to (1.1) and some properties and relations (5) existing between the edges of regression of the tangent, the polar and the rectifying surface of the multiplicity respectively to (1.1) and between the locus of the multiplicity and and the edge of regression of this tangent surface (6).

2.   Suppose (S) is the developable tangent surface of our multiplicity and (γ) the edge of regression of (S). Let s denote the length of the arc of (γ) and ρ, τ their radii of the curvature and the torsion, $\overline{b}$ and $\overline{t}$ are the unit vectors of the binormal and the tangent to (γ) referred to the x_i-system, and if $\overline{n}$ (n_i) is the unit vector of the principal normal of (γ), we will have [2]

$$(2.1) \quad d\omega = \sqrt{(d\overline{t})^2} = \sqrt{(dt_1)^2\ (dt_2)^2\ (dt_3)^2}$$
$$(ds^*)^2 = (ds)^2 + 2 \cdot d\overline{t} \cdot d\overline{x} + (d\omega)^2,$$

where s^* is the corresponding expression of s with respect to the x^*-system after the transformation (1.1), and

(2.2) $\quad \dfrac{ds}{\sqrt{(dt)^2}} = \dfrac{ds}{dw} = \rho.$

According to Frenet's formulas we have

$$\frac{d\bar{t}}{dw} = \frac{d\bar{t}}{ds} \cdot \frac{ds}{dw} = \frac{\bar{\eta}}{\rho} \cdot \rho$$

$$\cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot$$
$$\cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot$$

that is,

(2.3) $\quad \begin{cases} \dfrac{d\bar{t}}{dw} = \bar{\eta} \\[2mm] \dfrac{d\bar{\eta}}{dw} = -\bar{t} + \dfrac{\rho}{\tau}\,\bar{b} \\[2mm] \dfrac{d\bar{b}}{dw} = -\dfrac{\rho}{\tau}\,\bar{\eta}. \end{cases}$

We know that $|\bar{t}\ \bar{\eta}\ \bar{b}| = 1$ and therefore $\dfrac{\tau}{\rho}\Big|\bar{t}\ \dfrac{d\bar{t}}{dw}\ \dfrac{d\bar{\eta}}{dw}\Big| = 1$ or $\Big|\bar{t}\ \dfrac{d\bar{t}}{dw}\ \dfrac{d^2\bar{t}}{(dw)^2}\Big| \cdot \dfrac{\tau}{\rho} = -1$ and

(2.4) $\quad \dfrac{\rho}{\tau} = -\Big|\bar{t}\ \dfrac{d\bar{t}}{dw}\ \dfrac{d^2\bar{t}}{(dw)^2}\Big|.$

Let ρ^*, τ^*, $\bar{t}^*$, $\bar{n}^*$, $\bar{b}^*$, denote the correspondents of ρ, τ, $\bar{t}$, $\bar{n}$, $\bar{b}$, with respect to the x^*-system. Then we find by (1.2) and (1.3) that

(2.5) $\quad (d\bar{t}^*)^2 = (d\bar{t})^2$

and $\quad \Big|\bar{t}^*\ \dfrac{d\bar{t}^*}{dw}\ \dfrac{d^2\bar{t}^*}{(dw)^2}\Big| = \Big|\bar{t}\ \dfrac{d\bar{t}}{dw}\ \dfrac{d^2\bar{t}}{(dw)^2}\Big|$ or

(2.6) $\quad \dfrac{\rho^*}{\tau^*} = \dfrac{\rho}{\tau}.$

In addition we have $\dfrac{ds}{\sqrt{(d\bar{b})^2}} = \tau, \quad \dfrac{ds^*}{\sqrt{(d\bar{b}^*)^2}} = \tau^*, \quad \dfrac{\rho}{\tau} = \dfrac{\sqrt{(d\bar{b})^2}}{\sqrt{(d\bar{t})^2}}, \quad \dfrac{\rho^*}{\tau^*} = \dfrac{\sqrt{(d\bar{b}^*)^2}}{\sqrt{(d\bar{t}^*)^2}}.$

Combining this result with (2.5) we obtain $\sqrt{(d\bar{b}^*)^2} = \sqrt{(d\bar{b})^2}$. Therefore $\dfrac{\rho}{\tau}$, $\sqrt{(d\bar{t})^2}$, $\sqrt{(d\bar{b})^2}$ of the curve (γ) are invariant for the transformation (1.1).

3. We now ask to find the distance of the point X from the intersection of the normal planes at X and X' $(\bar{x} + d\bar{x})$ of (γ). The equations of these planes are [3]

(3.1) $\quad (\bar{X} - \bar{x}) \cdot \bar{t} = 0$

(3.2) $\quad (\bar{X} - \bar{x}) \cdot \dfrac{d\bar{t}}{dw} - \bar{t} \cdot \dfrac{d\bar{x}}{dw} = 0,$

and the distance p_t of X from the intersection of (3.1) and (3.2) is equal to the distance of X from (3.2), for $\bar{t}\,d\bar{t} = 0$, and then we have by (2.1),

$P_t = \dfrac{-\bar{t}\,dx}{\sqrt{(d\bar{t})^2}} = -\dfrac{\bar{t}\cdot d\bar{x}}{dw}.$ Likewise we find that the distance p_n of X from the intersection of the rectifying planes to (γ) at the points X and X' is, by (2.3), $P_n = \dfrac{-\bar{\eta}\ d\bar{x}}{dw\sqrt{1 + \dfrac{\rho^2}{\tau^2}}};$ The distance p_b of X from the intersection of the osculating planes of (γ) at X and X' is given by $P_b = \dfrac{-\tau\bar{b}d\bar{x}}{\rho\,dw}.$ Let p_t^*, p_η^*, p_b^*,

be the corresponding expressions of P_t^*, P_η^*, P_b^*. We will then have by (2.5), (2.6)

$P_t^* = \dfrac{-\bar{t}^* d\bar{x}^*}{dw}$, $P_\eta^* = -\dfrac{\bar{\eta}^* d\bar{x}^*}{dw} : \sqrt{1 + \dfrac{\rho^2}{\tau^2}}$, $P_b^* = -\dfrac{\tau}{\rho} \cdot \bar{b}^* \cdot \dfrac{d\bar{x}^*}{dw}$.   But by (1.1), (1.2)

and (1.3) we find that $\bar{t}^* \dfrac{d\bar{x}^*}{dw} = \bar{t}\dfrac{d\bar{x}}{dw}$, $\bar{\eta}^* \dfrac{d\bar{x}^*}{dw} = \bar{\eta}\dfrac{d\bar{x}}{dw} + c$, $\bar{b}^* \dfrac{d\bar{x}^*}{dw} = \bar{b}\dfrac{d\bar{x}}{dw}$, and

therefore $P_t^* = P_t$, $P_b^* = P_b$, $P_\eta^* = P_\eta + \dfrac{c}{\sqrt{1 + \dfrac{\rho^2}{\tau^2}}}$.   We put

(3.3) $\bar{t}\dfrac{d\bar{x}}{dw} = T$, $\bar{\eta}\dfrac{d\bar{x}}{dw} = N$, $\bar{b}\dfrac{d\bar{x}}{dw} = B$,

and we find

(3.4) $T^* = T$, $B^* = B$, $N^* = N + c$,

That is, T and B are invariant for the transformation (1.1) but not N.

4.   By use of (3.3), (3.4) and (2.3) we find the following expressions

(4.1) $\dfrac{dT^*}{dw} = \dfrac{dT}{dw} = \dfrac{d\bar{t}}{dw} \cdot \dfrac{d\bar{x}}{dw} + \bar{t} \cdot \dfrac{d^2\bar{x}}{(dw)^2}$

(4.2) $\dfrac{dN^*}{dw} = \dfrac{dN}{dw}$

(4.3) $\dfrac{dB^*}{dw} = \dfrac{dB}{dw}$

(4.4) $\bar{t}\dfrac{d^2\bar{x}}{(dw)^2} = \dfrac{dT}{dw} - N$ $\bar{\eta}\dfrac{d^2\bar{x}}{(dw)^2} = \dfrac{dN}{dw} + T - \dfrac{\rho}{\tau} \cdot B$.

If we put $\dfrac{d\bar{t}}{dw} \cdot \dfrac{d^2\bar{x}}{(dw)^2} = N_1$, we will have $\dfrac{dN}{dw} = N_1 - T + \dfrac{\rho}{\tau} \cdot B$

(4:5) $\bar{b}\dfrac{d^2\bar{x}}{(dw)^2} = \dfrac{dB}{dw} + \dfrac{\rho}{\tau} \cdot N$.

All these expressions except (4.4) and (4.5) are invariant.

5.   We will now seek to find how the edge of regression of the polar surface of (γ) will be transformed. In order to find it we make use of

(5.1) $(\bar{X} - \bar{x})\bar{t} = 0$

and we have [4] $(\bar{X} - \bar{x})\dfrac{d\bar{t}}{dw} - \bar{t}\dfrac{d\bar{x}}{dw} = 0$.   By (2.3) and (3.3)

(5.2) $(\bar{X} - \bar{x}) \cdot \bar{\eta} = T$

and furthermore $(\bar{X} - \bar{x})\dfrac{d\bar{\eta}}{dw} - \bar{\eta}\dfrac{d\bar{x}}{dw} = \dfrac{dT}{dw}$ or by (2.3), (3.3) and (5.1) we get

(5.3) $(\bar{X} - \bar{x})\bar{b} = \dfrac{\tau}{\rho}(N + \dfrac{dT}{dw})$.

We put $\bar{X} - \bar{x} = \lambda_1\bar{t} + \lambda_2\bar{\eta} + \lambda_3\bar{b}$, λ_i are independent from $\bar{t}$, $\bar{n}$, $\bar{b}$, and we

find $(\bar{X} - \bar{x})\bar{t} = \lambda_1 = 0$, by (5.1)

          $(\bar{X} - \bar{x})\bar{\eta} = \lambda_2 = T$, by (5.2)

          $(\bar{X} - \bar{x})\bar{b} = \lambda_3 = \dfrac{\tau}{\rho}(N + \dfrac{dT}{dw})$, by (5.3)

and finally $\bar{X} = \bar{x} + T\bar{\eta} + \dfrac{\tau}{\rho}(N + \dfrac{dT}{dw})\bar{b}$.

For the transformation $\bar{x}^* = \bar{x} + c\bar{t}$, because $\bar{n}$, $\bar{b}$ satisfy the system (2.3) with the same initial values, we will have

$$\bar{X}^* = \bar{X} + (\bar{t} + \frac{\tau}{\rho}\bar{b})\, c,$$

or

$$\bar{X}^* = \bar{X} + \frac{c\sqrt{\rho^2 + \tau^2}}{\rho} \cdot \frac{\rho\bar{t} + \tau\bar{b}}{\sqrt{\rho^2 + \tau^2}}$$

where $\dfrac{\rho\bar{t} + \tau\bar{b}}{\sqrt{\rho^2 + \tau^2}}$ is the unit vector of the rectifying straight line of (γ).

Thus we see that: *the point X of the locus of the centers of the osculating spheres of (γ) will be transferred on the rectifying straight line of (γ) a distance $\frac{c}{\rho}\sqrt{\rho^2 + \tau^2}$* .

For the edge of regression of (S) starting [5] from the equation

$$(\bar{X} - \bar{x})\bar{b} = 0$$

we likewise find that for the transformation $\bar{x}^* = \bar{x} + c\bar{t}$,

$$\bar{X}^* = \bar{X} \qquad \text{and for } (1.1)$$

$$\bar{X}^* = \bar{x}^* + \left[\frac{d(B\frac{\tau}{\rho})}{dw} - (N + C)\right]\bar{t} - \frac{\tau}{\rho}B\,\bar{\eta}.$$

If we use the transformation $\bar{x}^* = \bar{x} + c\bar{b}$ and consider as invariant parameter the angle of $\bar{b}$ with a fixed direction, then we will get the conclusion of A. Haimovici [6]. Finally for the edge of regression of the rectifying surface of (γ) we find

$$\bar{X} = \bar{x} + \left[\frac{B\frac{\tau}{\rho} - T + \frac{dN}{dw}}{\frac{d}{dw}(\frac{\rho}{\tau})}\right] \cdot (\frac{\rho}{\tau}\bar{t} + \bar{b}) - N\bar{t},$$

and for $\bar{x}^* = \bar{x} + c\bar{t}$, $\bar{X}^* = \bar{X}$.

6.   We now put $\dfrac{d\bar{x}}{dw} = \mu_1\bar{t} + \mu_2\bar{\eta} + \mu_3\bar{b}$, μ_i are independent from $\bar{t}$, $\bar{n}$, $\bar{b}$, and

we obtain $\bar{t}\dfrac{d\bar{x}}{dw} = \mu_1 = T$, $\bar{\eta}\dfrac{d\bar{x}}{dw} = \mu_2 = N$, $\bar{b}\dfrac{d\bar{x}}{dw} = \mu_3 = B$,   and consequently

$$(6.1) \quad \frac{d\bar{x}}{dw} = T\bar{t} + N\bar{\eta} + B\bar{b}.$$

We suppose that $T = \bar{t}\dfrac{d\bar{x}}{dw} = t_1\dfrac{dx_1}{dw} + t_2\dfrac{dx_2}{dw} + t_3\dfrac{dx_3}{dw} = 0$.

In this case the tangent straight line to the locus of the elements of our multiplicity, let it be denoted by (γ_*), is perpendicular to $\bar{t}$ of the edge of regression of (S). If we furthermore suppose

$$B = \bar{b}\frac{d\bar{x}}{dw} = 0, \text{ we will have}$$

$$\frac{d\bar{x}}{dw} = N\bar{\eta}.$$

If s_* is the length of the arc of (γ_*), we will have

$$(d\bar{x})^2 = N^2(dw)^2, \quad ds_* = N dw.$$

With $\bar{t}_*$, $\bar{n}_*$, $\bar{b}_*$ we denote the unit vectors of the tangent, the principal normal and the binormal to (γ_*). Thus we have

$$\bar{t}_* = \frac{d\bar{x}}{ds_*} = \bar{\eta},$$

$$\frac{d\bar{t}_*}{ds_*} = \frac{d\bar{\eta}}{Ndw} = \frac{\bar{\eta}_*}{\rho_*}$$

ρ_* is the radius of curvature of (γ_*), and

$$\frac{\bar{\eta}_*}{\rho_*} = (-\bar{t} + \frac{\rho}{\tau}\,\bar{b}) \cdot \frac{1}{N} \qquad \rho_* = \frac{N}{\sqrt{1 + \frac{\rho^2}{\tau^2}}}$$

On the other hand we have

$$\bar{b}_* = \bar{t}_* \times \bar{\eta}_* = (\frac{\rho}{\tau}\,\bar{t} + \bar{b}) : \sqrt{1 + \frac{\rho^2}{\tau^2}}$$

and

$$(6.2) \qquad \frac{d\bar{b}_*}{ds_*} = \frac{(\bar{t} - \frac{\rho}{\tau}\,\bar{b})d(\frac{\rho}{\tau})}{N \cdot \left(1 + \frac{\rho^2}{\tau^2}\right)^{\frac{3}{2}} \cdot dw} = -\frac{\bar{\eta}_*}{\tau_*},$$

where τ_* is the radius of the torsion of (γ_*). From (6.2) we get

$$\frac{1}{\tau_*} = \frac{d(\frac{\rho}{\tau})}{N \cdot \left(1 + \frac{\rho^2}{\tau^2}\right) dw} \qquad \text{and} \qquad \frac{\rho_*}{\tau_*} = \frac{d(\frac{\rho}{\tau})}{\left(1 + \frac{\rho^2}{\tau^2}\right)^{\frac{3}{2}} \cdot dw}$$

and therefore $\frac{\rho_*}{\tau_*}$ is invariant for the transformation (1.1). If $\frac{\rho}{\tau}$ is constant, then we will have $\frac{\rho_*}{\tau_*} = 0$, that is, *if the edge of regression of (S) is a curve with constant inclination, e.g. a cylindrical curve, then the orthogonal trajectories to their tangents are plane curves.*

Differentiating the expression $N = \bar{\eta}\,\frac{d\bar{x}}{dw}$ we obtain by (2.3), (3.3)

$$(6.3) \qquad \frac{dN}{dw} = -T + \frac{\rho}{\tau}\,B + \frac{d\bar{t}}{dw} \cdot \frac{d^2\bar{x}}{(dw)^2}\ .$$

If we take $T = 0$, $B = 0$, $N_1 = 0$, we will have N = constant, and putting $N = 0$, we have $dx_i = 0$, $x_i = c_i$. For $c_i = 0$, we have $x_i = 0$, that is, *in this case all the planes (P) of the considered multiplicity pass through the same fixed point.*

7. We now suppose that T, B, N, ρ/τ are given as functions of ω and we consider the system (2.3) and the equation (2.4) with respect to $\bar{t}$. The solution of (2.4) determines $\bar{t}$ as function of ω nearly to a rotation. We can consider the system (2.3) as one of Frenet's for a skew curve with curvature 1 and torsion ρ/τ. The set $\bar{t}$, $\bar{n}$, $\bar{b}$ is a particular solution of (2.3) and the general solution of it, is

$$(7.1) \qquad \bar{t}^* = \bar{a}_j\,t_j, \qquad \bar{n}^* = \bar{a}_j\,n_j, \qquad \bar{b}^* = \bar{a}_j\,b_j,$$

wherby a_{ji} are the coefficients of an orthogonal transformation with determinant equal to 1. By (6.3) we see that N can be determined by T, B, N and an arbitrary constant. If N is a particular integral of (6.3) its general integral will be $N^* = N + c$, c = constant.

On the other hand we have (6.1)

$$(7.2) \qquad \frac{d\bar{x}}{dw} = T\bar{t} + N\bar{\eta} + B\bar{b},$$

and $\bar{x}$ is a solution of this equation corresponding to the integrals of $\bar{t}$ and N of (2.3) and (6.3). Thus the general integral Z^* of (7.2) corresponding to the general solutions of (2.3) and (6.3) will be given by

$$\frac{d\bar{x}^*}{dw} = T^*\bar{t}^* + N^*\bar{\eta}^* + B^*\bar{b}^* \qquad \text{or by}$$

$$\frac{d\bar{x}^*}{dw} = T\bar{t}^* + (N + c)\,\bar{\eta}^* + B\bar{b}^* \qquad \text{that is, by}$$

$$\frac{d\bar{x}^*}{dw} = T\bar{a}_j\,t_j + (N + c)\,\bar{a}_j\eta_j + {+}B\bar{a}_j\,b_j \qquad \text{or}$$

$$\frac{d\bar{x}^*}{dw} = (\bar{t}\,\frac{d\bar{x}}{dw})t_j\bar{a}_j + (\bar{\eta}\,\frac{d\bar{x}}{dw} + c)\eta_j\bar{a}_j + (\bar{b}\,\frac{d\bar{x}}{dw})b_j\bar{a}_j$$

(7.3) $\quad \dfrac{d\bar{x}^*}{dw} = \bar{a}_j\,\dfrac{dx_j}{dw} + c\cdot\bar{a}_j\,\dfrac{dt_j}{dw}$

and finally $\quad \bar{x}^* = \bar{a}_j \cdot x_j + c\bar{a}_j\,t_j + \bar{C}_o, \quad \bar{C}_o$ is a constant vector. Reciprocally the equations (7.1) and (7.3) determine an one parameter multiplicity of elements of contact, satisfying the relations

$$\bar{t}^* \cdot \frac{d\bar{x}^*}{dw} = \bar{t} \cdot \frac{d\bar{x}}{dw} = T$$

$$\bar{\eta}^* \cdot \frac{d^2\bar{x}^*}{(dw)^2} = \bar{\eta} \cdot \frac{d^2\bar{x}}{(dw)^2} = N_1$$

$$\bar{b}^* \cdot \frac{d\bar{x}^*}{dw} = \bar{b} \cdot \frac{d\bar{x}}{dw} = B$$

$$\left| \bar{t}^* \; \frac{d\bar{t}^*}{dw} \; \frac{d^2\bar{t}}{(dw)^2} \right| = -\frac{\varrho}{\tau}.$$

References

1. E. Vessiot: Leçons de la géométrie supérieure (1919, pp. 137, 293.)
2. N. Sakellariou: Contribution à la théorie des surfaces, Bull. de la Soc. Math. de Gréce, t. I, 1, p. 126.
3. N. Sakellariou: Lectures on vector's calculus (in Greek), 1947, pp. 167-8.
4. N. Sakellariou: Lectures on vector's calculus (in Greek), 1947, p. 166.
5. N. Sakellariou: Lectures on vector's calculus (in Greek), 1947, p. 169.
6. A. Haïmovici: Sur la géométrie d'un group de contact, Annales Sc. de l'Univ. de Jasey, t. XXIX, fasc. 1-2, p. 123.

University of Athens, Greece.

THE COMPLETE QUADRILATERAL

Henry E. Fettis

(1) A *complete quadrilateral* may be defined as the configuration of four lines in general position, and the six points which they determine. The four lines will be referred to as the *sides,* and designated as (1), (2), (3), and (4). The intersection of any two sides determines a *vertex,* and these will be designated by the letter "A" with a double subscript indicating the two sides by which the vertex is determined. In particular, the intersection of sides (1) and (2) determines the vertex A_{12}, etc., Fig. 1.
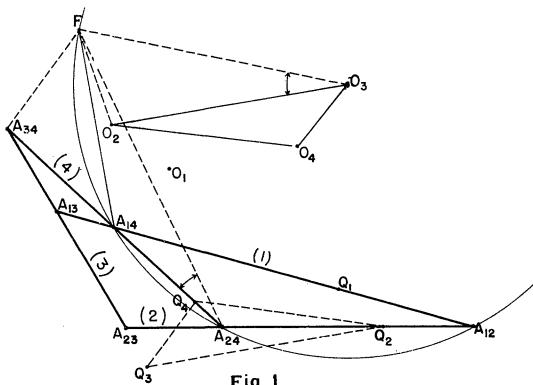
Fig. 1

Points which are definitely associated with respective sides are also designated by the proper subscript; in particular 0_1 is the circumcenter of the triangle which would be formed if the side (1) were omitted, H_1 is the orthocenter, and N_1 the nine-point center of this triangle. Collectively these points may be designated as 0_i, H_i, N_i.

The following properties of the complete quadrilateral are well known, and so are recalled without formal proof:

(a) The four circumcircles are on a common point, F, and the four circumcenters are on a circle through F. F is commonly called the *focus,* and the circle on the circumcenters, the *centric circle* of the quadrilateral.

(b) The four orthocenters are on a line which also contains the images of the point F in the four sides of the quadrilateral. This is the *directrix,* or *line of orthocenters.*

In order to prove Theorems (1) and (2), the following two lemmas will be needed:

Lemma 1. *The four triangles of the quadrangle $0_1 0_2 0_3 0_4$ are directly similar to the respective four triangles of the quadrilateral, and the center of similitude in each case is F.*

To demonstrate this, we note that $\angle F0_3 0_2$ (Fig. 1) equals half the arc FA_{14} of circle 0_3 as also does $\angle FA_{24}A_{34}$; the two angles are therefore equal. Similarly, $\angle F0_2 0_3$ equals $\angle FA_{34}A_{24}$, whence triangle $F0_2 0_3$ is directly similar to triangle $FA_{24}A_{34}$, with F as center of similitude. Applying the same reasoning to triangles $F0_2 0_4$ and $FA_{34}A_{24}$; $F0_3 0_4$ and $FA_{24}A_{23}$, it is seen that F is the center of
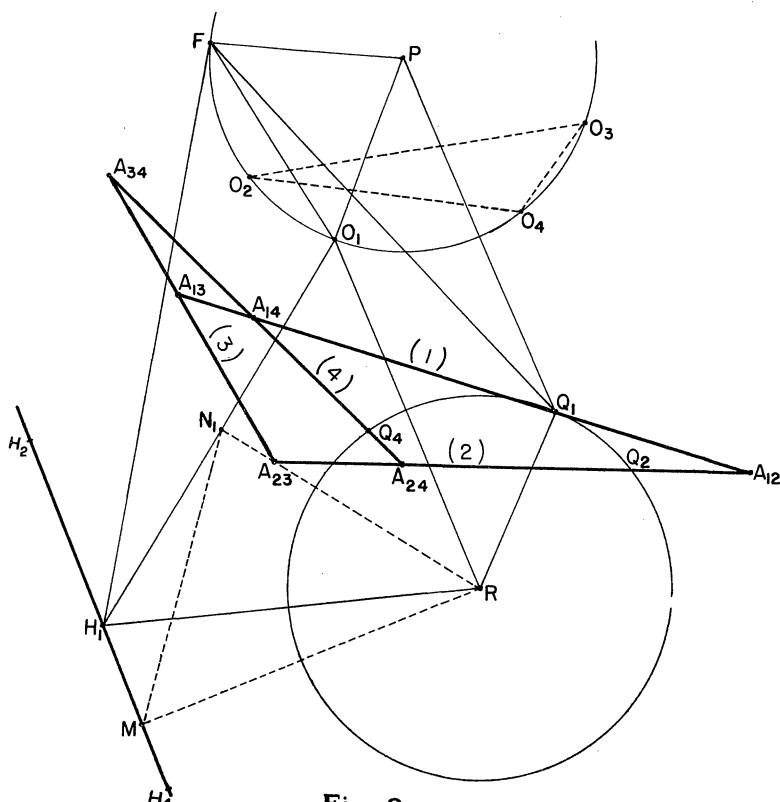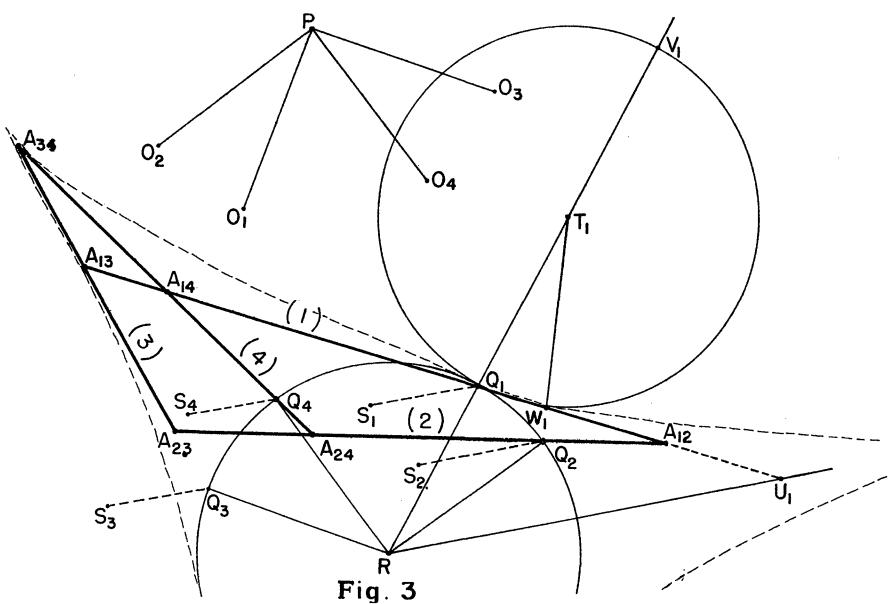
Fig. 2

Fig. 3

similitude of triangles $O_2O_3O_4$ and $A_{34}A_{24}A_{23}$, and similarly for the remaining three triangles of the quadrilateral.

Lemma 2. *The orthocenters of the four triangles of the quadrangle $O_1O_2O_3O_4$ lie on the respective sides of the given quadrilateral.*

This may be easily seen from the fact that O_2O_3 is the perpendicular bisector of FA_{14}, etc., so that side (1) contains the images of the point F in the sides of the triangle $O_2O_3O_4$, and is therefore on the orthocenter, Q_1, of of this triangle (Cf. Ref. 1, Art. 108).

Making use of the fact that the orthocenters of the four triangles of any cyclic quadrangle are vertices of a quadrangle whose sides are equal and parallel to the given one, the following result may be obtained:

"If lines are drawn through the circumcenters of the four triangles of a complete quadrilateral parallel to the respective sides, the resulting quadrilateral is equal to the given one, and the circumcenters of the four triangles of this quadrilateral are on the sides of the given one."

Theorem 1. *The perpendiculars to the Euler lines of the four triangles of a complete quadrilateral at the nine-point centers are concurrent.*

Let P be the circumcenter of $O_1O_2O_3O_4$, and R the circumcenter of $Q_1Q_2Q_3Q_4$ (Fig. 2). Then in the directly similar triangles $O_2O_3O_4$ and $A_{34}A_{24}A_{23}$, P corresponds to O_1, and Q_1 corresponds to H_1. Thus triangle FPQ_1 is directly similar to FO_1H_1, whence $FP:FO_1 = PQ_1:O_1H_1$ and, also, the angle between PQ_1 and O_1H_1 is the angle between FP and FO_1. Furthermore, since PO_1RQ_1 is a parallelogram, PQ_1 is equal and parallel to O_1R. Therefore, the angle between FP and FO_1 is the angle between O_1R and O_1H_1, and $FP:FO_1 = O_1R:O_1H_1$ so that triangles FPO_1 and O_1RH_1 are directly similar. But FPO_1 is isosceles by $\S(1)a$, whence triangle O_1RH_1 *is isosceles, or, the perpendicular bisector of* O_1H_1 passes through R.

Repeating the arguments for the other three triangles, it is clear that the four perpendicular bisectors are concurrent at R.

Theorem 2. *The perpendiculars from the nine-point centers of the four triangles to the respective sides are concurrent.*

Since the reflections of F in the four sides of the quadrilateral lie on the line of orthocenters, and the reflections of F in the sides of $O_2O_3O_4$ lie in the side (1), the line of orthocenters and the side (1) are corresponding parts related to the directly similar triangles $A_{34}A_{24}A_{23}$ and $O_2O_3O_4$, whence the angle between these lines is the angle between any two other corresponding parts. Thus if perpendiculars from N_1 and R to (1) and the line of orthocenters, respectively, intersect at M, the angle N_1MR equals the angle N_1H_1R so that the quadrangle N_1RMH_1 is cyclic, whence H_1MR is a right angle, and M is on the line of orthocenters.

The four perpendiculars from the nine-point centers to the sides are therefore concurrent at M, the foot of the perpendicular from R to the line of orthocenters.

Lemma 3. *The lines Q_iS_i, which trisect the angles between the sides of the quadrilateral and the lines Q_iR, in the sense* $\angle(i), Q_iS_i = 1/3 \angle(i), Q_iR$, *are parallel.*

For we have (Fig. 3) $\angle(i), Q_iR = \angle Q_jRQ_i + \angle(i), Q_jR$ Also

$\angle Q_jRQ_i = \angle O_jPO_i = 2\angle(i), (j)$ and $\angle(i), Q_jR = \angle(j), Q_jR + \angle(i), (j)$

whence $\angle(i), Q_iR = 3\angle(i), (j) + \angle(j), Q_jR$ or $\angle(i), Q_iS_i = \angle(i), (j) +$

$1/3\angle(j), Q_jR = \angle(i), (j) + \angle(j), Q_jS_j$. But $\angle(i), (j) + \angle(j), Q_jS_j = \angle(i), Q_jS_j$

so that $\angle(i), Q_iS_i = \angle(i), Q_jS_j$ or, Q_iS_i is parallel to Q_jS_j.

Theorem 3. *The sides of the quadrilateral are tangent to a deltoid which is circumscribed to the circle on the points Q_i.*

Let R be the circle on the points Q_i, and let the circle T_i be described equal to the circle R, and tangent to it at the point Q_i, intersecting side (i) at W_i, and RT_i produced at V_i. Also, let side (i) intersect the lines through R, parallel to Q_iS_i, at U_i. Then $\angle (i), Q_iR = \angle Q_iU_iR + \angle U_iRQ_i =$

$\frac{1}{2}\angle U_iRQ_i + \angle U_iRQ_i = 3/2\angle U_iRQ_i$ and $V_iT_iW_i = 2\angle (i), Q_iR = 3\angle U_iRQ_i$

which is the necessary and sufficient condition that the side (i) touch a deltoid circumscribed to the circle R.[1] Since the direction of Q_iS_i is the same for all four sides, each side of the quadrilateral is tangent to this same deltoid.

The Theorem of Steiner, that the pedal lines of a triangle are tangent to a deltoid circumscribed to the nine-point circle, is a special case of the above, as may be seen by considering the quadrilateral formed by a triangle and the pedal line of any point, relative to the triangle. In this case, the focus F of the quadrilateral is the point itself, and the center of the centric circle is the midpoint of OF, where O is the circumcenter of the triangle. The center of the deltoid is thus the midpoint of OH, and the circle R is the nine-point circle. Furthermore, the position of the deltoid is given by the direction of the trisectors of the lines joining R to the points Q_i, and three of these points are evidently the midpoints of the sides of the triangle. Thus the deltoid which touches the sides of a triangle and any pedal line is circumscribed to the ninepoint circle and has a fixed position determined by the direction of the trisectors of the angles joining the nine-point center to the midpoints of the sides, and must therefore be tangent to all pedal lines of the triangle.

(2)   Theorems 1 and 2 may be combined to furnish a proof of the following property of the quadrilateral due to Howard Eves.[2]

"If one side of a given complete quadrilateral is parallel to the Euler line of the triangle formed by the remaining three sides, the same is true for every side of the quadrilateral."

For, let it be given that O_iH_i is parallel to the side (i). Then, since the perpendiculars to the Euler lines at the four nine-point centers meet at M, and the perpendiculars from the nine-point centers to the sides meet at R, and since O_iH_i is parllel to side (i), it follows that R lies on N_iM. But RM is also perpendicular to the line of orthocenters at M, so that R and M must coincide. This, of course, requires that each of the other Euler lines be parallel to its respective side, as stated in the theorem.

The above argument is evidently not valid in the exceptional case when the given Euler line coincides with the line of Orthocenters. However, in this case it may be argued that since O_i lies on the line of orthocenters, P must lie on the side (i), and since PO_iRQ_i is a parallelogram, R must lie on O_iH_i, or, R must coincide with M.

REFERENCES

(1)   Morley, Frank and Morley F.V. *Inversive Geometry*. Ginn and Co., 1933.
       (See Art. 141 for proofs of theorems 1, 2, and 3 by analytic methods)
(2)   Johnson, Roger A., *Modern Geometry*. Houghton Mifflin Co., 1939.

[1] The geometric properties of this curve are discussed in the article *Geometric Properties of the Deltoid*, National Mathematics Magazine, Vol. XIX, No. 7.

[2] National Mathematics Magazine, Vol. XIX, No. 8, Problem No. 617.

ROOKS AND RHYMES

H. W. Becker

Kaplansky and Riordan have "unified and generalized" various results in statistics, in terms of the Problem of the Rooks, on a trapezoidal chessboard. A key theorem is: "the number of ways of putting c non-attacking rooks on a right-angled isosceles triangle of side $r-1$ is the Stirling number $\Delta^{r-c}\,0^r/(r-c)!$" = $_{r-c}@_r$. In other words, this is the number of selections of c points on such a board, such that none have any row or column index in common. [1]

We will exhibit well-ordered tables of these point sets thru $r = 5$, in 1 to 1 correspondence with the sequations (rhyme schemes) also enumerated by Stirling numbers [2,3]. We will then formulate other classifications, with respect to: 2) row location of the topmost rook; 3) number of rooks on the principal diagonal; 4) column vacancies; 5) column location of the bottom rook. Each of these except 3), of course has a dual interpretation, under interchange of row and column.

In Table I, the row index precedes the column index, and a vacant board is indicated by 0. The index 11 is at the top of the board, away from the observer, with rr at the bottom diagonal corner nearby on the right.

The sequations have further isomorphs in the substitution cycles [4]. All the rhyme repetitions of the kth letter correspond to all the letters in the kth substitution cycle. Thus the sequation $aaaaa$ corresponds to the distribution or substitution cycle $(abcde)$, while the sequation $abcde$ corresponds to the distribution $a/b/c/d/e$, etc. These correspondences are too easily read off by inspection, to need printing here.

```
a        a  a        a  a  a  a  a        a  a  a  a  a  a  a  a  a  a  a  a  a  a  a
         a  b        a  a  b  b  b        a  a  a  a  a  b  b  b  b  b  b  b  b  b  b
0                    a  b  a  b  c        a  a  b  b  b  a  a  a  b  b  b  c  c  c  c
    11   0           11 11 21 22  0       a  b  a  b  c  a  b  c  a  b  c  a  b  c  d
         22                               11 11 11 11 11 21 21 21 22 22 22 31 32 33  0
                                          22 22 32 33       32 33       31 33
                                          33
```

```
a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a
a  a  a  a  a  a  a  a  a  a  a  a  a  b  b  b  b  b  b  b  b  b  b  b  b  b
a  a  a  a  a  b  b  b  b  b  b  b  b  b  b  a  a  a  b  b  b  a  a  a  a  b
a  a  b  b  b  a  a  a  b  b  b  c  c  c  c  a  a  a  b  b  b  c  c  c  c  a
a  b  a  b  c  a  b  c  a  b  c  a  b  c  d  a  b  c  a  b  c  a  b  c  d  a
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 21 21 21 21 21 21 21 21 21 21 22
22 22 22 22 22 32 32 32 33 33 33 42 43 44    32 32 32 33 33 33 42 43 44    31
33 33 43 44    43 44    42 44                43 44    42 44                43
44
```

```
a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a  a
b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b  b
b  b  b  b  b  b  b  b  b  c  c  c  c  c  c  c  c  c  c  c  c  c  c  c  c  c
a  a  b  b  b  c  c  c  c  a  a  a  a  b  b  b  b  c  c  c  c  d  d  d  d  d
b  c  a  b  c  a  b  c  d  a  b  c  d  a  b  c  d  a  b  c  d  a  b  c  d  e
22 22 22 22 22 22 22 22 22 31 31 31 31 32 32 32 32 33 33 33 33 41 42 43 44  0
31 31 33 33 33 41 43 44    42 43 44    41 43 44    41 42 44
44    41 44
```

Table I

The sequations are in lexical order, as they would appear in a dictionary, or directory. Each set of c rooks on the r-1 board is made to correspond with a sequation in $_{r-c}@_r$, that is, one having r letters, $r-c$ different. Whenever a sequation increases its range, goes from $_c@_r$ to $_{1+c}@_{r+1}$, the corresponding set of r-c rooks is unaltered on the larger board. Otherwise, when it goes only to $_c@_{r+1}$ there will be an additional rook which may be located in any of c squares on the added bottom row. That is the significance of the recurrence

$$(1)\quad _{r-c}@_r = {}_{r-c-1}@_{r-1} + (r-c)\cdot{}_{r-c}@_{r-1} = {}_cR_{r-1} = {}_{c-1}R_{r-2}\cdot(r-c) + {}_cR_{r-2}.$$

We observe that rook sets totaling $^{(q)}R_r$, the upper one of which is in the qth row of the r-board, occur together in the lexicon, and correspond to sequations whose first q letters are distinct, followed by a repetition. These last are known to be enumerated by

$$(2)\quad ^{(q)}R_r = q\cdot@_{r-q}(q) = q\cdot(@ + q)^{r-q}$$

We adopt the convention in Table II that $^{(r+1)}R_r = 1$, the vacant board.

q	1	2	3	4	5	6	7		0	1	2	3	4	5	6
r															
0	1								1						
1	1	1							1	1					
2	2	2	1						2	2	1				
3	5	6	3	1					5	6	3	1			
4	15	20	12	4	1				15	20	12	4	1		
5	52	74	51	20	5	1			52	75	50	20	5	1	
6	203	302	231	104	30	6	1		203	312	225	100	30	6	1

Table II, $^{(q)}R_r$ Table III, ^qR_r

Tables II and III are identical thru $r = 4$, then deviate more and more. It is inherent in the mode of ordering the rook and rhyme schemes, that whenever we append a letter just equal to the range or previous high letter to a sequation, we add a rook in the outer or principal diagonal of the corresponding rook pattern. Then if ^qR_r = the number of rook sets on an r-board, with q rooks on the principal diagonal, we have, (r,q) being a binomial coefficient $(r)_q/q!$

$$(3)\quad ^qR_r = (r,q)\cdot R_{r-q} = {}^{q+1}@_{r+1} = (r,q)\cdot@_{r-q}$$

The latter expression denotes the number of sequations of $r+1$ letters having $q+1$ letters just equal to the previous high letter of the sequation. This table has two other rhyme interpretations: with $q+1$ as the number of a's (also the number of substitution cycles of $r+1$ letters with $q+1$ letters in the a-cycle); and—the rows written in reverse order—$q+1$ is the number of changes, or non-repetitive pairs of consecutive letters.

The evolution of (3) is: add q outer diagonals to the $(r-q)$-board, and permute the q rooks on the principal diagonal and intersecting column-rows, in all $(r-q)$ ways amongst the patterns of R_{r-q}.

Let $_c{}'R_r$ = the number of patterns on the r-board in which the cth column is empty. If we adopt the convention that $_{r+1}{}'R_r = R_r$, Table IV is the same as the difference table of $R_r = @_{r+1}$ with the rows and columns interchanged. Or, in terms of the operators ∇ and E such that $\nabla U_n = U_n\ U_{n-1}$, and $E^{-1} U_n = U_{n-1}$, so $\nabla = 1 - E^{-1}$,

$$(4.1)\qquad _c{}'R_r = \nabla^{r-c-1} R_r = (1 - E^{-1})^{r-c-1} R_r = (1 + E^{-1})^{c-1} R_{r-1}$$

These follow by induction, working either way from the obvious

$$(4.2)\qquad _r{}'R_r = \nabla R_r = R_r - R_{r-1},$$

$$(4.3)\qquad _1{}'R_r = R_{r-1}$$

The process is the same as in the Problem of the Incompatible Mechanics [5], whence the correspondence: $_c{}'R_r$ = the number of non-attacking rook patterns on a right triangular chessboard of side r and cth column vacant = the number of organizations of $r+1$ men into crews under the restriction that one man is incompatible with and must be segregated from $r-c$ other men.

c'	1	2	3	4	5	6	7		(c)	0	1	2	3	4	5	6
r																
0	1									1						
1	1	2								1	1					
2	2	3	5							1	2	2				
3	5	7	10	15						1	4	5	5			
4	15	20	27	37	52					1	9	12	15	15		
5	52	67	87	114	151	203				1	24	32	42	52	52	
6	203	255	322	409	523	674	877			1	76	99	129	166	203	203

Table IV, $_c{}'R_r$ Table V, $_{(c)}R_r$

Let $_{(c)}R_r$ = the number of patterns on the r-board such that the bottom rook is in the cth column. Then it is plain that

$$(5.1)\qquad _{(0)}R_r = 1,\quad _{(c)}R_r = {}_{(c)}R_{r-1} + {}_c{}'R_{r-1} = {}_{r-1}\Sigma_{s=c-1}\ {}_c{}'R_s,$$

the summation being between $c-1$ and $r-1$, a notation which facilitates printing. The $_{(c)}R_r$ may be given absolute evaluations as polynomials in R or $@$. Abbreviating $_r\Sigma_{c=1}\ @_c = @_r)$, we have

$$(5.2)\qquad _{(2)}R_{r+2} = @_r) + 1,$$

$$(5.3) \qquad _{(3)}R_{r+2} = @_{r)} + @_{r-1)} + 1$$

$$(5.4) \qquad _{(4)}R_{r+2} = @_{r)} + 2@_{r-1)} + @_{r-2)} - 1$$

$$(5.5) \qquad _{(r)}R_r = _{(r-1)}R_r = @_r$$

$$(5.6) \qquad _{(r-2)}R_r = @_r - @_{r-1} + @_{r-2}$$

$$(5.7) \qquad _{(r-3)}R_r = @_r - 2@_{r-1} + 2@_{r-2}$$

BIBLIOGRAPHY

1.   I. Kaplansky and J. Riordan, *The Problem of the Rooks and its Applications,*
        Duke Math Jr., 13 (1946) 259-268.
2.   H. W. Becker, Am. Math. Monthly, 48 (1941) 701.
3.   H. W. Becker, Bull. Am. Math. Society, 52 (1946) 415.
4.   J. Touchard, *Sur les Cycles des Substitutions,* Acta Math. 70 (1939) 249.
5.   D. H. Browne, Am. Math. Monthly, 51 (1944) 47.
6.   M. Cotlar, *A Generalization of the Factorials,* Math. Reviews, 7 (1946) 242.
7.   W. W. R. Ball, *Mathematical Recreations and Essays.*
8.   Whitworth, *Choice and Chance.*

Omaha, Nebraska

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.

The purpose and policies of the first division of this department (Comments on Papers) derive directly from the major objective of the MATHEMATICS MAGAZINE which is to encourage research and the production of superior expository articles by providing the means for prompt publication.

In order that errors may be corrected, results extended, and interesting aspects further illuminated, comments on published papers in all departments are invited. Comments which express conclusions at variance with those of the paper under review should be submitted in duplicate. One copy will be sent to the author of the original article for rebuttal.

Communications intended for this department should be addressed to

H. V. Craig, Department of Applied Mathematics,
University of Texas, Austin 12, Texas.

Elementary Differential Equations (third edition). By L. M. Kells
xiv + 312 pages, $3.00, McGraw-Hill Book Co., New York, N. Y., 1947.

In this third edition, a certain amount of theoretical material has been added, the supply of problems has been improved, and more applications have been included. But the author still fails to distinguish between a function and an equation (p. 4), he still does not explain why one must "multiply by dx" before integrating (p. 10), and he still confuses ordinary and line integrals (p. 25).

The book has been considerably embellished.

C. C. Torrance

Theory of Functions. By J. F. Ritt.
New York, King's Crown Press, 1947. x + 181 pp. $3.00.

In this text is presented an outline of a year's course in function-theory which the author has given for many years at Columbia University. While the emphasis is on the complex variable, about one-third of the book is devoted to the real variable.

The text is divided into forty-five short chapters. The first part of the book is concerned with the fundamental concepts of number, limit, function, continuity, derivative, Riemann integration, infinite series and infinite sequences. The author bases the real number system upon the concept of infinite decimals rather than upon the more conventional ideas of Dedekind and Cantor. This approach seems to be an excellent pedagogical device; many theorems thus become obvious to the student. Enough topology is introduced for the author to be able to consider regions bounded by Jordan curves. Thus the Cauchy integral theorem is proved first for triangles, then the proof is extended to polygons, rectifiable curves, and finally to Jordan curves. A systematic study of analytic functions makes up the latter part of the book. Expansion in Taylor's series and Laurent series, singularities, the Weierstrass factorization

theorem, residues, and analytic continuation are all given brief but adequate treatment. The beginning graduate student will be interested in finding in this book four different proofs of the fundamental theorem of algebra.

There are many good examples but no exercises for solution. There is no index. Numerous misprints would tend to hamper anyone who might attempt to use this book as a text for self-study. But the terseness of presentation indicates that this book is *not* intended to be studied without a teacher. In the hands of a competent teacher who will amplify the text and can supply whatever exercises seem necessary, Ritt's book should prove entirely satisfactory as a textbook for a course in the Theory of Functions.

<div style="text-align: right;">

H. M. Gehman
University of Buffalo

</div>

THE INTERNATIONAL CONGRESS OF MATHEMATICIANS

CAMBRIDGE, MASSACHUSETTS, U. S. A., August 30-September 6, 1950

An international Congress of Mathematicians will be held in Cambridge, Massachusetts, in 1950 under the auspices of the American Mathematical Society.

Time and Place. The dates for the Congress have been fixed as August 30-September 6, 1950. Harvard University will be the principal host institution. A number of other institutions in metropolitan Boston will join in the entertainment of Congress visitors by arranging special features on their campuses.

Type of Congress. Following the established custom, the Organizing Committee plans to have a number of invited hour addresses by outstanding mathematicians. In addition, sectional meetings for the presentation of contributed papers not included in Conference programs will be held in the following fields: I, Algebra and Theory of Numbers; II, Analysis; III, Geometry and Topology; IV, Probability and Statistics, Actuarial Science, Economics; V, Mathematical Physics and Applied Mathematics; VI, Logic and Philosophy, VII, History and Education.

The official languages of the 1950 Congress will be English, French, German, Italian, and Russian.

Entertainment. Harvard University has offered the use of its dormitories and dining rooms for mathematicians and their guests for the period of the Congress. The Organizing Committee hopes that it will be possible to furnish room and board without charge to all mathematicians from outside continental North America who are members of the Congress. Congress membership fees and rates for room and board will be announced well in advance of the opening of the Congress.

The Entertainment Committee, of which Professor L. H. Loomis of Harvard University is Chairman, is planning many interesting features, including a reception, garden party, symphony concert, and banquet.

Information. Detailed information will be sent in due course to individual members of the American Mathematical Society and to foreign mathematical societies and academies. Others interested in receiving information may file their names in the office of the Society, and such persons will receive from time to time information regarding the program and arrangements.

Communications whould be addressed to the American Mathematics Society, 531 West 116th Street, New York City 27, U.S.A.

<div style="text-align: right;">

The Organizing Committee

</div>

Edited by

Joseph Seidlin, L. J. Adams 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.

THE GIST OF THE CALCULUS

by Glenn James

Calculus is the father of our modern mechanical conveniences and the cornerstone of the great structure of mathematics that has been built during the last few decades. Moreover its type of thinking is indispensable to any scientist and indeed to anyone who tries to understand something about his automobile and the many labor saving gadgets that he uses almost daily.

This discussion approaches the concepts and processes of the calculus through simple, familiar experiences, along a road that can be followed in the main by any thoughtful person who can conceive of a letter as denoting any one of a set of numbers*.

When one is climbing a mountain he is much concerned about the slope of the path he is following, that is, about how high he has to raise a tired foot to proceed forward. Everyone has some notion of this term, *slope*; that the greater the slope the more difficult is the climbing, and that if the slope is zero the path is horizontal. However, our casual estimates of slopes are not dependable. Most of us have driven along highways which seemed to be ascending although actually our cars would coast. We can not, for instance, build satisfactory railroads on the basis of such hazy notions of *slope*. This concept must be precisely defined so that when the maximum load to be hauled and the maximum power available are known we can build the track with such slope that the load can be hauled over the mountain, spiraling up if necessary.

This rather simple problem of defining precisely the slope of a path and finding how great it is at each point of the path is the basis of the calculus. Its solution could be put on a half page. Yet growth from this nucleus has made possible the phenominal advances in mechanical achievement of the last century and has assisted greatly in the development of all sciences, especially physics and chemistry. The justification of this statement will appear as we proceed through the solution of the above problem to the unfolding of the calculus.

When passing along a straight path from one point to another, say P to P', it is customary to denote by dy the vertical rise occuring, as in Fig. I, and by dx the corresponding horizontal change. The ratio $\dfrac{dy}{dx}$

*See "Fundamentals of Beginning Algebra" by E. Justin Hills, Mathematics Magazine, Vol. XXI, No. 4.

is then said to be *the slope of the straight line* LL', or the rate of change of the vertical rise with respect to the corresponding horizontal change.
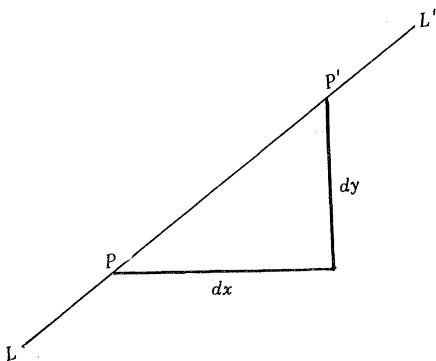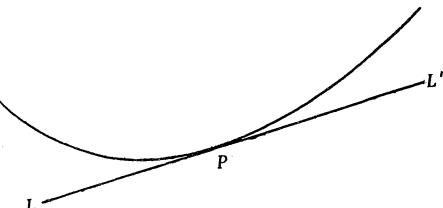

Fig. I   Fig. II

If, as in Fig. II, a straight line LL' just grazes a curved path, say at P, like an imagined straight-edge laid on the path, then the slope of this line is said to be the slope of the curved path at the point where the line touches it. This slope is thought of as the *instantaneous rate of change of the vertical* with respect to the horizontal at the point P. The line is called *the tangent line to the curve at the point P*.

This concept of instantaneous rate of change is suggested by the physical law that if a particle is moving in a curved path and if all constraining forces are suddenly removed the particle will leave the path in the direction of the tangent to the path at the point which the particle occupied when the constraining forces were removed. This law is illustrated by the sling with which David killed Goliath, by sparks flying off an emery wheel and by numerous other mechanical devices.

We will now exhibit the method by which the slope of the tangent line to a known curve at a given point can be found. (A curve is *known* when the vertical rise from some fixed horizontal line can be found at any point where the horizontal distance from some vertical line is given. The vertical rise is usually called the *ordinate* of the point and denoted by y and the horizontal distance is called the *abscissa* and denoted by x. See Fig. III. Here $y = x^2$. The curve is called the *graph of the equation*, and the abscissa and ordinate of a point are, collectively, called the *coordinates* of that point.)

Suppose we seek the slope of the graph of $y = x^2$ at the point P where $x = 3$. We choose an arbitrary point P' different from P (Fig. IV) and draw a line, called the secant, through P and P'. As indicated in the figure, we denote the difference between the abscissas of P' and P by some letter, say h. Now if we move P' toward P, i.e. make h approach zero, this secant will move toward the line LL' which is tangent to the curve at P.
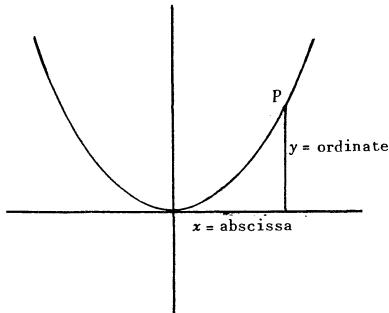
Fig. III
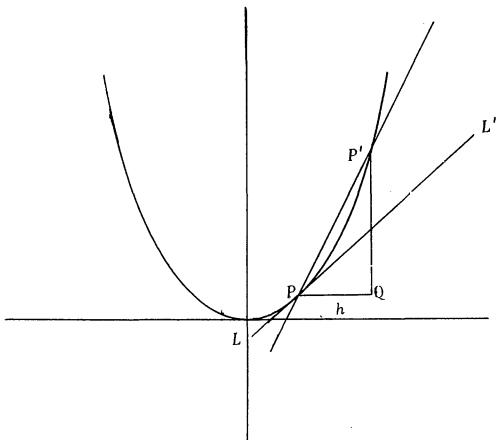
Fig. IV

Hence in order to find the slope of the tangent line at P we need merely formulate the slope of this secant and find what this slope approaches when P' moves toward coincidence with P. This can always be done when the path is *smooth* and *continuous*. However if P were just over the edge of a precipice our scheme isn't good. The road sign that a mathematician would put by such a path would be "This Path is Discontinuous."

At the point where $x = 3$ we find from $y = x^2$ that $y = 3^2$. At the point which is h units from P we have $x = 3 + h$, whence $y = (3 + h)^2$. Hence the *vertical rise* QP' is $(3 + h)^2 - 9$. The corresponding *horizontal change* being h, the slope of the secant is

$$\frac{(3 + h)^2 - 9}{h}.$$

Since $(3 + h)^2 = 9 + 6h + h^2$ this slope can be written

$$\frac{9 + 6h + h^2 - 9}{h}$$

and since $9 - 9 = 0$, we can divide h out of numerator and denominator and thus show that

$$\frac{9 + 6h + h^2 - 9}{h} = 6 + h.$$

For the reader who is not real familiar with algebra it is helpful to test this statement for a few values of h, say 1, .1, .01. When

$$h = 1, \quad \frac{9 + 6h + h^2 - 9}{h} = \frac{9 + 6 + 1 - 9}{1} = 7 = 6 + 1;$$

when $h = .1,$ $\quad \frac{9 + .6 + .01 - 9}{.1} \quad \frac{.61}{.1} = 6.1 = 6 + .1,$

and when $h = .01$, $\dfrac{9 + .06 + .0001 - 9}{.01} = \dfrac{.0601}{.01} = 6.01 = 6 + .01$.

When P' approaches P, that is, when h approaches zero, $6 + h$ approaches 6. Hence *the slope of the curve $y = x^2$ at the point where $x = 3$ is 6* or in simpler terminology $\dfrac{dy}{dx} = 6$ when $x = 3$. If instead of $x = 3$ we had taken $x = c$ where c is any number whatever, the vertical rise would have been $(c + h)^2 - c$ and the slope of the secant would have been

$$\dfrac{(c + h)^2 - c^2}{h}$$

Simplifying as before we obtain $2c + h$, whence $\dfrac{dy}{dx} = 2c$ when $x = c$. Since c is any value of x whatever, we could just as well use x itself, provided we treat x as fixed at the point P. We would then say that the slope of the curve $y = x^2$ at the point x is $2x$ or briefly

$$\dfrac{dy}{dx} = 2x.$$

The expression $\dfrac{dy}{dx}$ is spoken of as *the derivative of y with respect to x*. (Since in this problem $y = x^2$, we may say the *derivative of x^2 with respect to x* is $2x$ and write it in the condensed from $\dfrac{dx^2}{dx} = 2x$. Instead of the expression "finding the derivative of" we often say "differentiating." However we need not bother much about synonyms here. These can be found in any standard text book on calculus.)

From the example just worked out it is easy to infer the procedure for finding the derivative of most any expression in x. Those interested in doing a little algebraic work would find it entertaining to show that

$$\dfrac{dx^3}{dx} = 3x^2, \quad \dfrac{dx^4}{dx} = 4x^3, \quad \dfrac{dx^n}{dx} = nx^{n-1}, \text{ and } \dfrac{dkx^n}{dx} = knx^{n-1},$$

k being a constant, e.g. $\dfrac{d3x^2}{dx} = 3 \times 2x = 6x$ (A better terminology than 'expression in x' is *function of x*, written $f(x)$, $F(x)$, $g(x)$, or what not, and read f-function of x, capital F-function of x, g-function of x, etc., or in mathematical vernacular just f of x, F of x, g of x, etc. "Function of" means dependent upon or determined by.)

It is interesting to note that if a line is horizontal, that is, if all of its ordinates are equal, in other words if the equation describing it is $y = k$ where k is any constant, then the slope of the line is zero. Thus *the derivative of any constant is zero*. In the usual notation

$$\dfrac{dk}{dx} = 0,$$

where k is any constant.

If a line is vertical, no matter what vertical rise we take the

corresponding horizontal change is zero. Hence in attempting to formulate the slope of a vertical line we meet the impossible case, met first in arithmetic, of division by zero. Thus we see that the slope of a vertical line is not defined; but this gives us little trouble because it is the *only* line whose slope is not defined. Certainly if there were just one person in the world without a name he would be well designated by his lack of a name.

Again, if the derivative of the sum of two functions of x is sought, we see after a little thoughtful consideration that the total vertical rise due to a change in the horizontal distance is equal to the sum of the separate vertical rises of the functions. Hence *the derivative of the sum of two functions of x is the sum of their derivatives*: e.g. if

$$y = x^2 + x^3, \quad \frac{dy}{dx} = 2x + 3x^2.$$

When y is equal to more complicated expressions in x, greater difficulty is met in simplifying the expressions for the slopes of the secant so as to be able to find the limits of these slopes as the secants approach the tangents. However the *logical steps* used in finding the derivatives are the same in all cases and a table of derivatives of various functions can be found in any text on calculus. It is not necessary that the reader verify these standard formulae. They are as well established as is the fact that our lights begin to shine when we turn on the switch provided the connections are good, and the connections are always good in the standard processes of differentiation.

In defining the derivatives we have talked about mountain paths but we actually used only the *numerical values* of the vertical and horizontal distances that entered into the discussion. The mathematical work would have been exactly the same if we had attached to these numbers other concrete references. We list below a few such references, which have great practical significance.

If y represents the space passed over, by some particle, in time x, then the rate of change of y with respect to x is the velocity, and this is, as we have seen, $\frac{dy}{dx}$. Using a more suggestive notation, if $s = t^2$, the velocity at any time t is $2t$. In finding that $\frac{ds}{dt} = 2t$, there is no possible objection to going back to our mountain path and thinking of s as the vertical rise and of t as the corresponding horizontal distance.

Now consider the relation $v = 2t$. Looking upon v as the vertical rise and t as the corresponding horizontal change, the rate of change of v with respect to t is $\frac{dv}{dt}$, which is 2. But this is the rate of change of the velocity with respect to time which is, of course, the acceleration of the particle. Thus when a particle is moving according to the law $s = t^2$ its velocity is $2t$ and its acceleration is 2. Since v is equal to $\frac{ds}{dt}$, we may write $\frac{dv}{dt}$ in the form $\frac{d}{dt}\frac{ds}{dt}$. The latter is usually written $\frac{d^2s}{dt^2}$ and read *the second derivative of s with respect to t*. We hesitate

to mention what you already see, namely that we can have *third, fourth, fifth, sixth, etc., derivatives,* or as we say in calculus-language, we can have *derivatives of any order.*

Suppose now that x represents one side of a rectangle whose perimeter is fixed, say it is 4. Then the other side is $(4-2x)/2$ or $2-2x$ and the area, which we shall denote by y is $x(2-x)$ or $2x-x^2$. Thus we have the equation

$$y = 2x - x^2$$

in which x is a side of a rectangle whose perimeter is 4 and y is the area of this rectangle. Much valuable information can be gotten from this little equation. In studying it we can, as has been mentioned, think of y as the height and x as the horizontal measurement of a point on a mountain path.

The two points where x is, respectively, 0 and 2 look interesting because y is 0 at both of these points.

The slope of this path at any point x is $\frac{dy}{dx}$ which is $2-2x$. When x is 0 the slope is 2. When x is 2 the slope is -2. The latter means that when x is 2 the ordinate is at that instant *decreasing* at the rate of 2 units to every unit that x increases. See Fig. V.
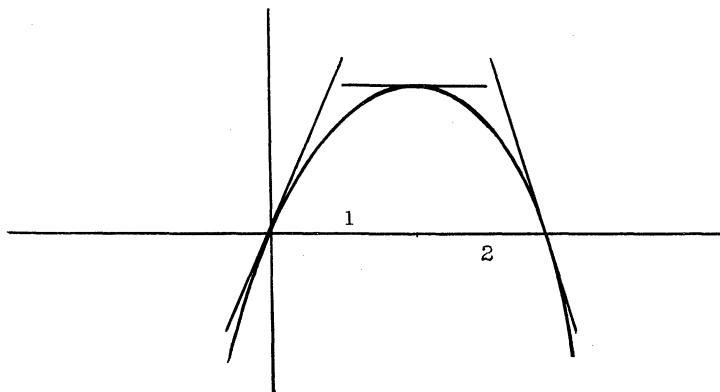
Fig. V

From these observations, we suspect that the curve goes over a hilltop, in other words, that there is a largest rectangle whose perimeter is 4. To hunt for this and similar 'hilltops' is an exciting adventure that results in very useful findings. On such a hilltop the tangent is horizontal, that is, the slope of the curve is zero (we call such a hilltop a *maximum*). Moreover just before and just after a path reaches a maximum its ordinates are less than at the maximum. So we set the slope, that is the derivative of $2x-x^2$, equal to zero and find what value of x makes this equality true; next we substitute this value of x in $2x-x^2$ to find the ordinate at the point where the tangent line is horizontal, and then we examine the ordinates just before and just after we have reached this point. In this problem $\frac{dy}{dx}$ is $2-2x$ hence we have $2-2x = 0$. From this

equation, we find that $x = 1$ is the abscissa of the point where the
tangent line is horizontal. The ordinate at that point is $2 \times 1 - 1 \times 1$
or 1. Then we find, for instance, the ordinates corresponding to $x = .9$
and $x = 1.1$; these are .99 and .98 respectively. Hence we infer that 1 is
the *maximum area* of a rectangle with perimeter 4, that is, *the rectangle
has a maximum area when it is a square*. (Another and better way of
showing that we are at the hilltop is to note that the slope $(2 - 2x)$ is
positive for x a little less than 1 and negative for x a little greater
than 1.) Since we did not stipulate the units of measure used when we
said the perimeter was 4, it follows that no matter whether its
perimeter is 4 feet, 4 inches or what not *any rectangle has its maximum
area when it is a square*. The possible uses of this result are limited
only by the imagination.

It is said that when John D. Rockefeller first got his hand in the oil
business he immediately had the oil barrels and cans made so that for a
given content their surfaces contained a minimum (i.e. least possible)
amount of material. The calculus very easily determines the proper ratio
between the altitude and diameter of any such cylinder. Let r be the
radius of the ends of a can and h its height. Then we seek the ratio
between r and h when the volume is fixed and the surface is a minimum.
From geometry the area of each end is πr^2, the lateral surface is $2\pi rh$
and the volume is $\pi r^2 h$. The total surface is then the sum of the two ends
and the lateral surface. That is, if the total surface is denoted by s,
we have

$$s = 2\pi r^2 + 2\pi rh.$$

Suppose the volume is 1, 1 quart, 1 gallon, 1 barrel or whatever is at
hand. Then $\pi r^2 h = 1$. Dividing this equality by πr^2 gives $h = 1/\pi\, r^2$. We
now substitute this form for h in the formula for s in order to express
s in terms of r, that is, to make changes in s depend only on changes in
r. We thus obtain

$$s = 2\pi r^2 + 2/r.$$

This equation can be thought of as a path for which s is the height
corresponding to the horizontal distance r measured from some fixed point.
We examine this path to find whether or not it has a minimum height, that
is, whether it goes down into a valley and up on the other side. If it
does this, there is a point in the valley where the slope is zero, and
the curve rises on both sides of that point. By steps analogous to those
used in the rectangle problem, we find that *for a minimum surface $h = 2r$*.
When this result is once known the metal saved by using the correct
shape of cylindrical container is net profit. All competitors who fail
to use this and similar results are sure to fail in business, because
they have failed in mathematics.

The preceding examples give us a hint that knowledge of the slopes of
a curve at various points would aid in drawing the curve. Even more help
can be gotten from learning the rate of turning of the tangent as we

move along the curve. This can be stated more explicitly as follows.
Denote the angle between the tangent line and the horizontal by θ, as in
figure VI, and the length of the arc from a fixed point P_0 to P by s, then
the rate of turning of the tangent as we move along the curve is $\frac{d\theta}{ds}$.
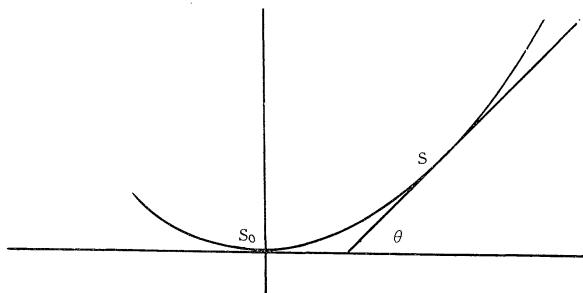This is called the *curvature*

Fig. VI

of the curve at the point under consideration. The *curvature of a circle
is evidently constant*. The curvature of a straight line is zero. For
other curves it changes as one passes from point to point. In picturing
a curve with changing curvature, we think of a circle with the same
curvature as the curve at the point immediately under consideration and
drawn touching the curve at that point. See Fig. VII. This is the circle
which fits closest to the curve at the point under consideration, that
is, it is the circle whose tangent turns at the same rate as the tangent
to the curve is turning when we pass through the point as we move along
the curve.
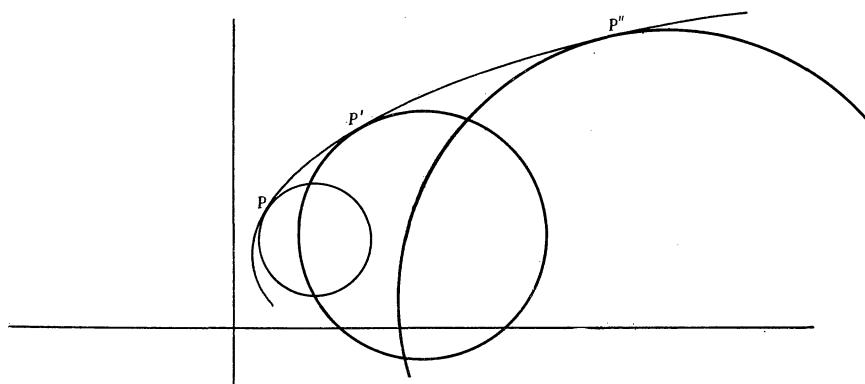
Fig. VII

Such a circle at a point P, is called *the circle of curvature* at the point P and its radius is called the *radius of curvature* of the curve at point P. Given the equation of a curve and the coordinates of a point on this curve one can find the center of the circle of curvature at this point as well as its radius. Formulas for doing these things are in any text on the calculus. Having drawn the circles of curvature for several successive sets of coordinates that satisfy a given equation, one has a pretty fair idea of how a curve appears in the vicinity of these points.

If we think of a circle rolling along a curve continually adjusting its size so that it is always the *circle of curvature* of the curve at the point of contact, (See Fig. VIII), the center of this rolling circle describes another curve which is called the *evolute* of the given curve: relative to its evolute a given curve is called the *involute* of the former.
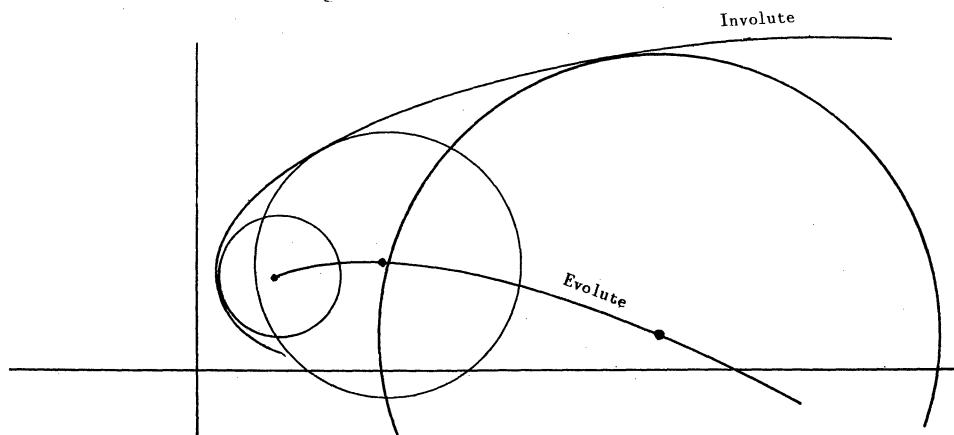
Involute

Evolute

Fig. VIII

It is easily proven that every tangent line to a curve is perpendicular to the tangent to the curve's involute at the point where it cuts it, as indicated in Fig. IX.

Involute

90°

Curve (Evolute)
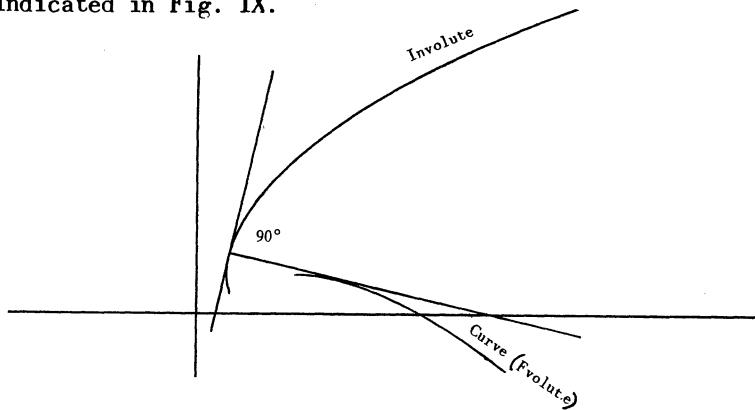
Fig. IX

Because of this property the involute of a circle may be described by
a point on a taut thread as the thread is unwound from the circle (see
Fig. X).
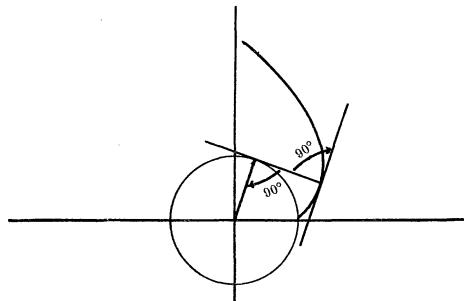
Fig. X

Gear teeth are curved like an involute of a circle. Pressure perpendi-
cular to their surfaces is then directed along the tangent to the circle,
which is of course perpendicular to the radius of the circle. This
gives maximum efficiency and avoids shearing at high speeds.

   We turn now to quite a different use of the calculus. Often times we
have given a point, say P, on a curve and wish to find the *change in the
ordinate* resulting from a small change, say dx, in the abscissa. The
change in the ordinate is denoted by Δy and is called delta y. Now the
corresponding change in the ordinate of the tangent line at P has been
denoted by dy. See Fig. XI.
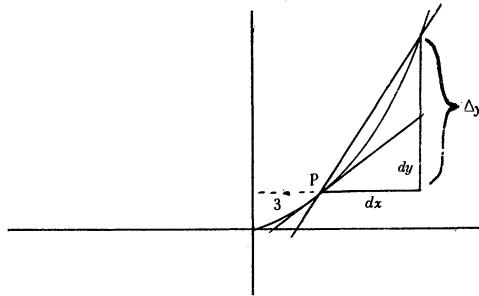
Fig. XI

Thus dy is an approximation to y, a *better* and *better approximation* as
dx decreases. Moreover since the derivative at the point P on the curve
is $\dfrac{dy}{dx}$, it follows that the derivative at this point multiplied by dx is
the approximation to Δy which we are seeking. If $y = f(x)$ is the curve
with which we are dealing and we denote the derivative of $f(x)$
by $f'(x)$, then the above approximation to Δy is given by the formula

   (A) $dy = f'(x)dx$

Fig. XI exhibits the meaning of this formula when used to find the approximate change in y for the curve $y = x^2$ at the point $x = 3$ when dx is 1/32". (If, for instance, x has been measured as 3" with a possible maximum error of 1/32"). Such a problem arises whenever one attempts to measure a side of a square plate and compute the area of the plate from his measurement. In this case

$$dy = f'(x)dx = 2 \, x \, dx = 2 \times 3 \times \frac{1}{32} = \frac{3''}{16}$$

In fact such a problem arises whenever we base any computation on measurement. Measurement is always approximate. One can never wager his life on having an exact measurement. The odds are that a magnifying glass would embarrass him. All one can be sure of is that the distance to be measured is *between two estimates;* and this is where our dx comes in.

Usually the matter of measuring and making computations with the results is not as simple as this case of finding the area of a square from a side. The area of a rectangle, the volume of a cone, etc., depend upon three measurements; and the mass of a block depends upon four measurements, the three that determine its volume and one that determines its density. Thus we need to study functions of several variables.

To find the change of a function of several variables due to changes in all of the variables we combine the changes due to each of them while the others are being held constant. For example, consider that z depends on x and y, written $z = f(x,y)$. Now by fixing y we find the instantaneous rate of change of z with respect to x by merely finding $\frac{dz}{dx}$. We denote this sort of procedure by $\frac{\partial z}{\partial x}$. This new symbol $\frac{\partial z}{\partial x}$ is read *the partial derivative of z with respect to x.* Partial derivatives differ in no way from ordinary derivatives except that they indicate that the other variable or variables, are being held constant at the instant when we are obtaining the rate of change of the function due to a change in the variable which is before our attention. We can make this matter more thinkable by use of a graph. We use the surface of a mountain instead of a single path over it. We think of x and y as being the coordinates of a point in the horizontal plane on which the mountain sits and of z as being measured vertically up. Then as the point in the horizontal plane wanders about all over that plane (often called the xy plane) the top line which represents z traces out the surface of the mountain. The approximate change in z, due to a change, dx, in x while y is held constant is $\frac{\partial z}{\partial x} dx$ by formula (A), and the approximate change in z due to a change, dy, in y while x is being held constant is $\frac{\partial z}{\partial y} dy$. Then the total approximate change in z (denoted by dz) is given by

(B) $$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy .$$

If the surface of the mountain is smooth and continuous, it does not make any difference whether the changes in x and y are made separately or simultaneously, i.e. whether we pass from the point on the mountain

whose height is z to a neighboring point along one path or another.

As an illustration of the use of this formula, suppose we have measured the sides x and y of a rectangle and found them to lie between 2.9 and 3.1, and 5.9 and 6.1, respectively, and it is required to find the approximate maximum possible error in the area if we take $x = 3$ and $y = 6$. We denote the area by z then $z = xy$, and $\frac{\partial z}{\partial x} = y$ and $\frac{\partial z}{\partial y} = x$. Hence $dz = ydx + xdy$. At the point where $x = 3$ and $y = 6$, $\frac{\partial z}{\partial x} = 6$ and $\frac{\partial z}{\partial y} = 3$. Also dx and dy, in this case, are both .1. Hence

$$dz = 6(.1) + 3(.1) = .9.$$

Wonders can be worked with the above formula (B). Suppose, for instance, we know the rates of change of x and y with respect to time, i.e., $\frac{dx}{dt}$ and $\frac{dy}{dt}$, at some point whose coordinates are x and y, then this formula enables us to find the rate of change of z at this point. We merely divide by dt obtaining

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt},$$

then substitute the given values of $\frac{dx}{dt}$ and $\frac{dy}{dt}$ on the right hand side of this result and also substitute the given values of x and y in $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

It is very evident by this time that calculus is a study of varying rates of change. If $s = kt$ then $v = k$, which we already knew before calculus was dreamed of. But if $s = t^2$, then $v = \frac{ds}{dt} = 2t$. Thus the velocity increases or decreases as t changes, so we have spoken of $2t$ as the instantaneous velocity. Our habit of referring the more complex cases back to the simpler ones leads us to try to find a constant velocity which is equivalent to a variable velocity for a given interval of time. Using t and s in place of x and y, suppose a particle travels along the curve $s = t^2$ from the time $t = 2$ sec's to the time $t = 5$ sec's. The space passed over is $5^2 - 2^2$ or 21 and the time consumed is $5 - 2$ or 3 sec's. The constant velocity which would cover this same space in 3 sec's is, plainly, 21/3 or 7 units per sec. The velocity $2t$ is 4 when $t = 2$; at the time $t = 5$ it is 10. At the time when $2t = 7$, i.e. $t = 3.5$, the varying velocity is the same as the constant velocity. This constant velocity is called the *mean velocity* over the interval $t = 2$ to $t = 5$. To generalize this discussion, we use the time interval t_1 to t_2 instead of 2 to 5, and $s = F(t)$ instead of $s = t^2$, then instead of $\frac{ds}{dt} = 2t$ we use the general notation $\frac{ds}{dt} = f(t)$. Whence we can write, in analogy with the equation $2t = 7$,

$$f(T) = \frac{F(t_2) - F(t_1)}{t_2 - t_1},$$

where T is the time at which the varying velocity is the same as the constant velocity required to take the particle over the same space in

time $t_2 - t_1$. A more usual form for this is gotten by replacing the t's by x's and then writing dx instead of $x_2 - x_1$. Thus we have

$$f(X) = \frac{F(x_2) - F(x_1)}{dx} \quad \text{or} \quad f(X)dx = F(x_2) - F(x_1).$$

We illustrate this in Fig. XII where the slope at P is $f(X)$ and is the same as the slope of the secant MN.
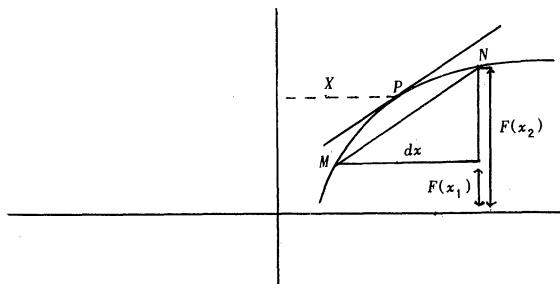
Fig. XII

This relation is called the *mean value theorem*. Out of this simple relation there grow results of the utmost importance in our daily lives. Your time consumed in going into some detail concerning these results will be liberally rewarded.

In geometry we are able to find areas of only certain simple figures such as rectangles, triangles and circles; but with the calculus we can find the areas of almost any closed figure. We shall show how to find the area bounded by the x-axis, the curve $y = f(x)$ and the ordinates of this curve corresponding to $x = a$ and $x = b$. See Fig. XIII.
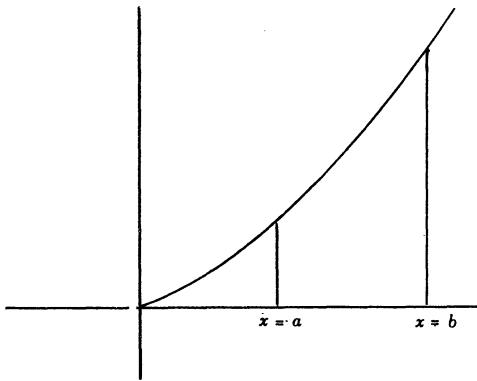
Fig. XIII
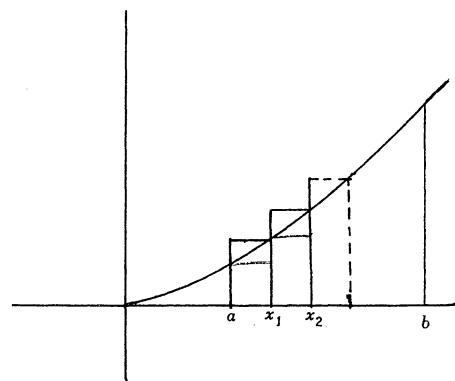
Fig. XIV

We draw a set of rectangles beneath the curve and a set extending above the curve, as shown in Fig. XIV. The sum of the former set is less

than the desired area beneath the curve and the sum of the latter set is greater than this area. If we decrease the widths of these rectangles, the difference of the two sums approaches zero, and of course both sums approach the area beneath the curve *since it is between the two sums.* This fact is needed to *squeeze another set of rectangles whose sum we always know into the desired area.* Suppose, as in Fig. XV, we construct beneath the curve $y = f(x)$ a set of rectangles of equal width $x_1 - a$, $x_2 - x_1$, etc., which we denote by dx, and let X_1, X_2, etc. be the points determined by the mean value theorem when applied to each interval a to x_1, x_1 to x_2, etc. We then have the relations:

$$F(x_1) - F(a) = f(X_1)dx$$
$$F(x_2) - F(x_1) = f(X_2)dx$$
$$F(x_3) - F(x_2) = f(X_3)dx$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$F(b) - F(x_{n-1}) = f(X_n)dx$$

where $\dfrac{dF(x)}{dx} = f(x)$.
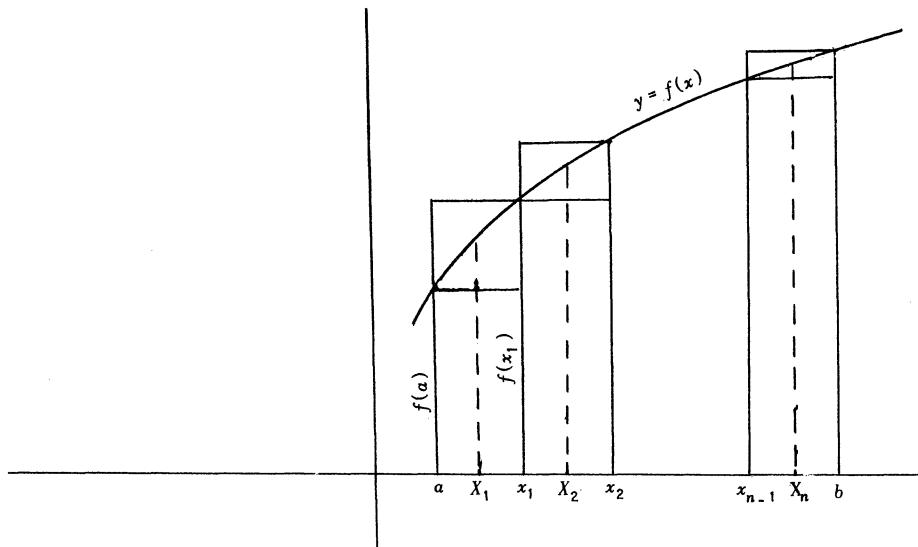
Fig. XV

Now adding these n identities, we have for the sum of the left members $F(b) - F(a)$, since alternate terms cancel out. The sum on the right lies between the sum of the areas of the lesser rectangles and the sum of the areas of the larger ones, both of which approach the area beneath the curve. Hence $F(b) - F(a)$ is the area desired.

Thus we see that *to find the area bounded by the curve* $y = f(x)$ *the ordinates corresponding to* $x = a$ *and* $x = b$ *and the x-axis, we find the function* $F(x)$ *whose derivative is* $f(x)$, *then substitute a and b successively in* $F(x)$ *and take the difference* $F(b) - F(a)$. We denote this rule by $\int_a^b f(x)\, dx$ and call it the *definite integral of* $f(x)$ *between the limits a and b*. The symbol $\int$ is an elongated S denoting "sum". Thus $_a\int^b$ is a certain sum from a to b.

The gist of this process consists of finding $F(x)$ from a given $f(x)$. This step is denoted symbolically by $\int f(x)\, dx$ and called the *indefinite integral of* $f(x)$, or the *primitive of* $f(x)$, it being, as we have said, the function, $F(x)$, whose derivative is $f(x)$, i.e. $\dfrac{dF(x)}{dx} = f(x)$.

For example to find the area between the x axis, the line $y = 2x$ and the ordinates corresponding to $x = 1$ and $x = 5$ we evaluate $\int_1^5 2x\, dx$. See Fig. XVI.
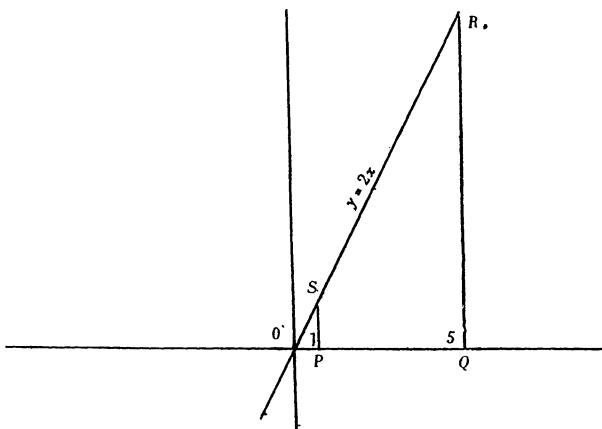
Fig. XVI

The steps used in doing this are usually written

$$\int_1^5 2x\, dx = \left. x^2 \right]_1^5 = 25 - 1 = 24.$$

It is satisfying to check this result by plane geometry. The area of a triangle being one half of the base times the altitude, we have area of $PQRS$ = area of OQR – area of OPS = ½ × 5 × 10 – ½ × 1 × 2 = 24.

When searching for the function whose derivative is $2x$, one may first think of x^2 but $x^2 + 1$, $x^2 + 2$, and in general $x^2 + c$, where c is any constant, all have $2x$ for their derivatives since the derivative of any constant is zero. One can see geometrically that adding a constant to a function merely lifts the curve up vertically and does not change the slope. (See Fig. XVII) However, when evaluating a definite integral, we do not usually write down the c (called the constant of integration) because it always drops out in the process of subtraction.

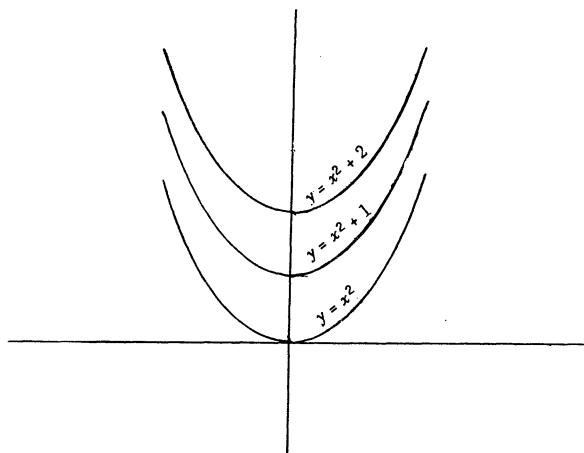Finding the indefinite integral is not merely a step in evaluating

Fig. XVII

a definite integral. We often desire it for other purposes, in which cases the constant of integration may be of great importance. E.g., suppose a particle starting from rest moves with its velocity equal to $2t$ and we wish to find how the space it passes over is related to the time consumed. In other words, given $v = 2t$, it is required to find $f(t)$ such that $s = f(t)$. Now from

$$v = \frac{ds}{dt} \qquad \text{we have} \qquad \frac{ds}{dt} = 2t, \qquad \text{whence} \qquad s = t^2 + c.$$

Since the particle started from rest, when $t = 0$, $s = 0$; hence $c = 0$, and the desired relations between s and t is $s = t^2$.

Although the evaluation of $\int_a^b f(x)\, dx$ made use of the idea of finding an area, it actually depended only on the numerical quantities a, b, and $f(x)$. Consequently interpretations other than finding areas could be given to these quantities and the same evaluation process used, or perhaps it is more meaningful to say that whatever interpretation is given to these quantities, one can think of the process as a matter of finding an area. A few interesting examples will clarify this statement:

Suppose that it is desired to find the mass of a rectangular plate whose sides are 2 and 4 and whose density at each point is equal to the distance of the point from the shorter side. See Fig. XVIII.
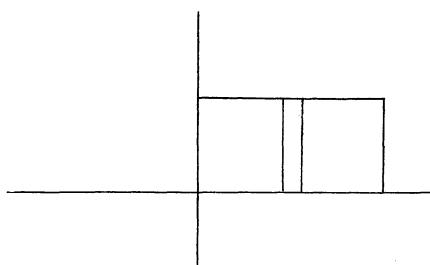

Fig. XVIII

The area of any one of the rectangles of height 2 and width dx is $2dx$ and its mass (density times area, since the plate can be considered to be of unit thickness) is greater than $x\,2\,dx$ and less than $(x+dx)\,2\,dx$. Thus we have two sets of rectangles and the desired mass is equal to the area beneath the line $y = 2x$ above the x-axis and between the ordinates corresponding to $x = 0$ and $x = 4$. (See Fig XIV and the accompanying discussion.) This area is equal to

$$\int_0^4 2x\,dx = x^2\Big]_0^4 = 16$$

which is the mass of the rectangle.

Again, suppose a particle is being dragged along the x-axis by a force which is equal to twice the distance of the particle from the y-axis, and it is desired to find the work done in moving the particle from the point where $x = 0$ to the point where $x = 4$. Since the work done is defined as the *force times the distance passed over*, the work done in moving this particle over a distance dx measured from the point x is greater then $2xdx$ and less then $2(x+dx)\,dx$. Hence the work done in moving the particle from the first point to the second is equal to

$$\int_0^4 2x\,dx = 16.$$

As a third example, suppose it is desired to find the force exerted on the face of a rectangular dam of height 100 feet and width 300 when the water stands to the top of the dam.

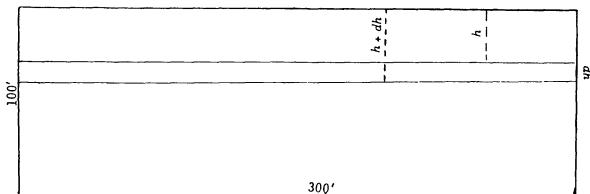The force exerted on any strip of width 300 feet and height dh

Fig. XIX

(see Fig. XIX) is greater than the area of this strip multiplied by the depth of the water to the *top of the strip* and by the weight of a cubic foot of water (63+ lbs) and is less than this product with the depth counted to the *bottom of the strip,* that is, $h+dh$. Hence the force lies between $300\,h\,(63+)\,dh$ and, $300(h+dh)\,(63+)\,dh$. Hence the total force is numerically the same as the area bounded by the line $y = 300(63+)x$, the x-axis and the lines $x = 0$ and $x = 300$. See Fig. XIV and the accompanying discussion. Thus this force is

$$\int_0^{300} 300(63+)x\,dx = 300(63+)\frac{x^2}{2}\Big]_0^{300} = (300)^3\,\frac{63+}{2}.$$

Probably the most used mathematical concept is that of distance

along a line or curve. The latter is frequently estimated by laying a
tape measure along the curve, as a tailor measures a man for a suit.
However, this procedure is not always feasible and rarely sufficiently
accurate for scientific purposes. Suppose for instance we were to try to
solve the famous problem of finding the path down which a particle will
slide from one point to a lower one not vertically beneath it in the
quickest possible time. Here we need a precise formula for the length of
the curve. Since we are familiar with straight line measurements we
naturally build our formula for determining the length of a curve upon
them. As in Fig. XX, divide the interval from $x = a$ to $x = b$ into subin-
tervals each of width dx. Then construct segments of tangents at the
points a, x_1, x_2, etc. each one being terminated by the next ordinate.
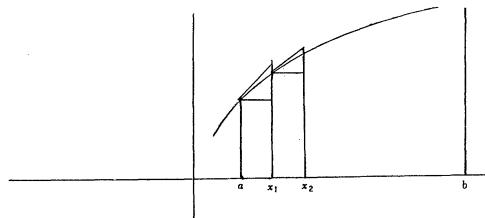
Fig. XX

The sum of these segments comes nearer and nearer to the length of
the curve as we decrease dx. Each of these segments is equal to
$\sqrt{dx^2 + dy^2}$ where x is the abscissa of the point of tangency under
consideration. Now we can take dx out of the radical and write this
in the form $\sqrt{1 + \left(\dfrac{dy}{dx}\right)^2}\, dx$ or in the equivalent form $\sqrt{1 + [f'(x)]^2}\, dx$
where, of course, $f'(x)$ is found from the equation $y = f(x)$ of our curve.
We then sensibly and safely define the length of the curve to be

$$\int_a^b \sqrt{1 + [f'(x)]^2}\, dx$$

Let's consider the special problem in which $y = f(x)$ is $y = 2x$, and the
length of the path from the point where $x = 0$ to the point where $x = 4$ is
desired. Here $\dfrac{dy}{dx} = 2$ and $\sqrt{1 + [f'(x)]^2}$ becomes $\sqrt{1 + 2^2}$. Hence the length
of the curve, in this case a straight line, from $x = 0$ to $x = 4$ is

$$\int_0^4 \sqrt{5}\, dx = \sqrt{5}\ x \Big]_0^4 = 4\sqrt{5}.$$

For more complex forms of $f(x)$ the form $\sqrt{1 + [f'(x)]^2}$ becomes more com-
plicated and the evaluation of the definite integral more difficult,
but you can hire a secretary to use tables and do this part of the
work.

   If we have, in this paper, conveyed a fair understanding of the
meaning of the derivative, and the definite, and indefinite integral,
the reader can forget the details and yet feel himself initiated
into the calculus way of thinking and prepared to pursue a rigorous
course in calculus without much difficulty in understanding the
subject.

FIVE REQUIREMENTS FOR GOOD TEACHING

Talk by W. C. Krathwohl

Speaking before the Division on Educational Methods at the annual meeting of the American Society for Engineering Education at Austin, Texas, Dr. Krathwohl stated that a good teacher satisfies at least five requirements:

"(1) He is enthusiastic about his subject; (2) he knows his subject thoroughly; (3) he is more interested in his students than in the subject; (4) he has a sense of humor without being ridiculous; (5) he has chosen teaching as his occupation because he would rather teach all day long than do anything else in the world."

Beyond these five conditions, Dr. Krathwohl explained, there are the techniques of the occupation which the truly successful teacher must master.

In the matter of relations between the teacher and the student, a good teacher knows that "there is a time to be personal and also a time to be impersonal," Dr. Krathwohl stated. Impersonality is most urgently required when discipline is necessary. "The rule is," he said, "dislike what a student does, but never dislike a student."

Some instructors seek popularity by omitting homework, "but such a course of action results in poor teaching," warned Dr. Krathwohl. "There is no substitute for hard work. One of the most efficient ways of learning a subject is to learn by doing."

He suggests, however, that grading for the course be done solely on the basis of frequent examinations rather than on examinations and homework. "Examinations more accurately reflect what the student has accomplished, and this method eliminates the possibility of dishonesty, allows the bright students to help those not so bright, and encourages an interchange of ideas between students."

The use of examinations as a measure of achievement should be secondary to their use as a review and consolidation of material, Dr. Krathwohl declared. The examination should be moderately difficult but easy enough to give the student the satisfaction of accomplishment. It should contain no trick problems and should be short enough to permit the average student to check his work in the allotted time.

"In order to keep the length of the examination within bounds," he said, "a good rule to follow is that the instructor must be able to work the examination in one-fourth to one-fifth of the time allotted to the student."

Occasionally certain students are unjustly penalized by variations in the grading of examinations due to fatigue of the instructor, Dr. Krathwohl continued. "To avoid this," he said, "examinations should be graded problem by problem rather than paper by paper. The variations will then be more evenly distributed among the students."

All the methods outlined in his talk are the results of trial and error over a period of years in which the better schemes were accepted and poorer ones rejected. Dr. Krathwohl stated.

PROBLEMS AND QUESTIONS

Edited by

C. G. Jaeger and H. J. Hamilton

This department will submit to its readers, for solution, problems which seem to be new, and subject-matter questions of all sorts for readers to answer or discuss, questions that may arise in study, research or in extra-academic applications.

Contributions will be published with or without the proposer's signature, according to the author's instructions.

Although no solutions or answers will normally be published with the offerings, they should be sent to the editors when known.

Send all proposals for this department to the Department of Mathematics, Pomona College, Claremont, California.

SOLUTIONS

No. 16. Proposed by H. E. Bowie, American International College

A circle of radius 3-in. is tangent externally to a rectangle at the midpoint of one end. Two other circles, both of radius 2-in., are tangent externally to the two sides of the rectangle and to the first circle. The rectangle is 2-in. wide. Find the radius of a circle to which the three given circles are tangent internally.

Solution by Howard Eves, Corvallis, Oregon.

Let r be the sought radius and let x be the distance from the center of the sought circle to the center of one of the smaller given circles. Then

$$x + 2 = (4 - \sqrt{x^2 - 9}) + 3.$$

Solving we find $x = 3.4$ in. Therefore $r = 5.4$ in. (exactly).
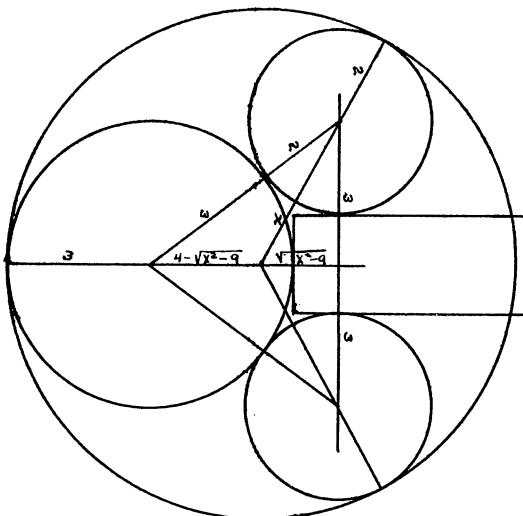
Solved also by K. L. Cappel, San Francisco.

No. 5. Proposed by Victor Thebault, Tennie Sarthe France.

Using once each of the digits 0,1,2,3,4,5,6,7,8,9, form a number, which when increased by one million becomes a perfect square.

Solution by Francis L. Miksa, Aurora, Ill.

If X is the square number then:

$$X^2 \equiv 10^6 \pmod 9.$$

Which gives $X = 9K \pm 1$ as the solution. Also the limits for X will be

$$31995 < X < 99550.$$

Not having a table of squares in this range the writer actually constructed all the squares of from $(9K \pm 1)^2$ by addition method. In the process he found 44 numbers satisfying the problem.

<div align="center">

Form $9K + 1$

$1\ 287\ 953\ 604 + 10^6 = (35\ 902)^2$

$1\ 507\ 234\ 896 + 10^6 = (38\ 836)^2$

$2\ 063\ 157\ 489 + 10^6 = (45\ 433)^2$

$\cdot\ \cdots\ \cdots\ \cdots\ \cdot\ \cdots\ \cdot\ \cdots\ \cdots$

</div>

No. 7. Proposed by Pedro A. Piza, San Juan, Puerto Rico.

Find squares of nine digits $a_1a_2a_3b_1b_2b_3c_1c_2c_3 = D^2$ where $a_1 \neq 0$, so that

$$a_1a_2a_3 = A^2, \quad b_1b_2b_3 = B^2, \quad c_1c_2c_3 = C^2$$
$$a_1a_2a_3 + b_1b_2b_3 + c_1c_2c_3 = F^2,$$

and $c_1c_2c_3b_1b_2b_3a_1a_2a_3 = E^2$.

Solution by the proposer.

Nine-digit squares meeting the requirements of the problem take the form

$$(1000x+y)^2 = 1000\ 000\ x^2 + 1000 \cdot 2xy + y^2.$$

Therefore if x^2 is a three-digit integer of which the first digit is not zero, and if $2xy = t^2$ and y^2 are three-digit integers of which the first two may be zeros, the sum of

<div align="center">

$\underline{x^2}\ \underline{000}\ \underline{000}$

$\underline{2xy}\ \underline{000}$

$\underline{y^2}$

$\overline{}$

$\underline{x^2}\ \underline{2xy}\ \underline{y^2}$

</div>

(where each underlined group represents a three-digit square), will satisfy the problem inasmuch as $x^2 + 2xy + y^2 = (x+y)^2$ and $(1000y+x)^2 = \underline{y^2}\ \underline{2xy}\ \underline{x^2}$, which we call the 'reversal' of $(100x+y)^2$.

In order for x^2 to be a three-digit integer of which the first one is not zero, the value of x^2 must be less than 1000 but not less than 100. Therefore x

must be not less than 10 and not greater then 31. Then y^2 must be less than 1000 and $t^2 = 2xy$ must be an even square not greater than 900. There are the following possible values of x and y:

x	y	$t^2 = 2xy$	x	y	$t^2 = 2xy$
10	5	•100	18	16	576
10	20˙	400	18	25	900
11	22	484	20	10	400
12	6	144	22	11	484
12	24	576	24	3	144
13	26	676	24	12	576
14	7	196	25	2	100
14	28	784	25	8	400
15	30	900	25	18	900
16	2	64	26	13	676
16	8	256	27	6	324
18	1	36	28	14	784
18	4	144	30	15	900
•18	9	324			

There are no solutions for x = 17, 19, 21, 23 and 31.

The distinct nine-digit squares meeting all the conditions of the problem, all of which can be reversed by exchanging x and y in $(1000x+y)^2$, are easily written out.

PROPOSALS

No. 17.   Proposed by Leo Moser, University of Manitoba.

Given an integer of n non-zero digits, show that it is always possible to replace a certain r $(0 \leq r < n)$ of these digits by others (zero not excluded) in such a way that the resulting number is divisible by n.

No. 18.   Proposed by Julius Sumner, Dillard university.

A smooth circular hoop rests on a smooth horizontal table. A small marble is to be projected from a point A on the inner side of this hoop so as to return to point A after two reflections, or rebounds. If e is the coefficient of restitution, and both friction and rolling are neglected, find the angle between the first path and the radius drawn to A.

No. 19.   Proposed by V. Thébault, Tennie, Sarthe, France.

Find a perfect square such that the numbers formed by its digits,

taken in sets of three marked off from the right, and by its square root, form three consecutive terms of an arithmetic progression whose common difference is r^2. i.e. if $n^2 = \overline{abcdef}$, then $\overline{abc}$, $\overline{def}$, n form an arithmetic progression.

No. 20. Proposed by Victor Thébault, Tennie, Sarthe, France.

In any tetrahedron $ABCD$, of centroid G, for which the tetrhedron $GABC$ is trirectangular at G, show that the relations

$$m_a^2 + m_b^2 + m_c^2 = 11\, m_g^2$$
$$\overline{ABD}^2 + \overline{BCD}^2 + \overline{CAD}^2 = 11\, \overline{ABC}^2$$

hold between the lengths m_a, m_b, m_c, m_g of the medians of the tetrahedron $GABC$ drawn from A, B, C, G and the areas $\overline{ABD}$, $\cdots$, of the faces ABD, $\cdots$, of the tetrahedron $ABCD$.

No. 21. Proposed by Julius Sumner, Dillard University.

A plane is inclined to the horizontal at an angle B. At the foot of this plane a particle is projected with velocity V at an angle A with the plane. Find the condition for maximum range.

No. 22. Proposed by Pedro A. Piza, San Juan, Puerto Rico.

Let x and n be any two positive integers and let $\Sigma^n x^2$ stand for the nth iterated summation of all the squares from 1 to x^2 inclusive. For instance $\Sigma\, 4^2 = 30$, $\Sigma^2 4^2 = 50$, (that is, the sum of the sums of all the squares from 1 to 16 inclusive), and $\Sigma^5 4^2 = 156$ (that is, the sum of the sums of the sums of the sums of all the squares from 1 to 16 inclusive). Prove that in general

$$\Sigma^n x^2 = \frac{x(x+1)(x+2)(x+3)\cdots(x+n)(2x+n)}{(n+2)!}.$$

No. 23. Proposed by V. Thébault, Tennie, Sarthe, France.

Given an orthocentric tetrahedron $ABCD$, of orthocenter H, show that the spheres (A), (B), (C), (D) of centers A, B, C, D, orthogonal to a sphere of center H, cut the planes of the faces BCD, CDA, DAB, ABC in four circles which lie on the same sphere.

MATHEMATICAL MISCELLANY

The keen eye of W. R. Ransom (Tufts College) discovered in the public press a tale of a calculating machine company that "offered $1000 to anyone who could square a circle, double a cube, or trisect one angle of a triangle by using only a straight-edge and compass." A fellow in Mathematics sues the company, claiming that he has squared the circle. The judge rules that he hasn't done it. Well, well, well! Here we go again! This seems just like old times. We could give one word of advice to the company, and that is to specify that the straight-edge shall be unmarked. Also let the angle to be trisected be an arbitrary angle. Then the company may rest comfortably in the knowledge that it has been proved that no one of the three constructions can be done. And how we wish that would become generally known.

Colonel Byrne sends us more news of mathematical colleagues in France. Professor Gaston Julia has given five lectures in Switzerland as follows: one at Bale, two at Zurich, one at Lausanne, one at Geneve. Professor P. Montel (retired) has the titles Professeur Honoraire and Doyen Honoraire.

A Remark on Mathematical Induction

Suppose we wish to prove a certain theorem, T. It may happen that the simplest way of proving T is to establish a stronger theorem, T^*, of which T is a simple corollary. That is, it may be easier to prove T^* than it is to prove T without using T^*. This fact is particularly useful in framing a proof which employs mathematical induction, and since, moreover, it is an important notion in many mathematical proofs, it deserves early mention in the classroom treatment of mathematical induction.

Interesting in this connection are the remarks of Felix Bernstein [Bull. Amer. Math. Soc., 52 (1946) Abstract 259, p. 622] who suggests that "the four-color theorem may be a simple consequence of a more inclusive theorem which can be proved by complete induction."

A simple example, suitable for elementary instruction, is the following:

Let $S(n) = 1^2 + 2^2 + \cdots + n^2$ for each positive integer n, and let T be the statement that $n + 1$ divides $6S(n)$. A simple, direct proof of T is perhaps not immediately apparent. Moreover, the fact that $k + 1$ divides $6S(k)$ is not in itself enough to imply that $(k+1)+1$ divides $6S(k+1)$. Hence a proof by induction is evidently not feasible. But now let T^* be the stronger statement that $6S(n) = n(n+1)(2n+1)$. Then T^* can be easily proved by induction and T is obtained as a corollary.

University of Virginia.   V. L. Klee, Jr.

Around 200 former students and colleagues of Professor William D. Reeve, retiring head of the mathematics department at Teachers College, Columbia University, gathered for a testimonial dinner in his honor on July 15th.

It was announced at the dinner that the David Eugene Smith Mathematics Club, which sponsored the testimonial, is accumulating money to set up a William David Reeve Scholarship in the Teaching of Mathematics. The Scholarship will be awarded annually to doctoral students, and approximately $1,500 of a $5,000 goal has been contributed to the fund to date.

William Higgins, president of the Club, was toastmaster, and Dr. Carl N. Shuster of the teaching staff presented Dr. Reeve with a watch. Other College officers who made brief remarks were Associate Dean Hollis L. Caswell; Professors Reeve and John R. Clark, and Instructors Howard Fehr and Nathan Lazar. Also on the program were Dr. Rolland Smith, supervisor of mathematics in the Springfield, Mass., schools and formerly on the Teachers College staff; Dr. Alfonso Elder, president of the North Carolina College, Durham, and Dr. Aaron Bakst, former students, and Miss Anita Feinstein, a current student.

A member of the department since 1923, Dr. Reeve is also the long-time editor of THE MATHEMATICS TEACHER. Previous to his affiliation with Teachers College, he was connected with the University of Chicago and the University of Minnesota. His retirement is effective in the fall.

Classroom Discussion of a Question on Infinite Series

This note relates to the following well known theorem.

Theorem. If an infinite series has the following three properties, it converges: (P_1) *the series alternates;* (P_2) *the limit of the nth term is zero;* (P_3) *the terms never increase numerically when read in order* (the first term, the second term, etc.).

After demonstrating this theorem for a class in *Integral Calculus* recently, we were asked this question: "If exactly one of the properties P_i in the theorem is waived, does there exist a divergent series which has the other two properties?"

Our answer after some reflection, was as follows. The harmonic series, with nth term $u_n = 1/n$ has properties P_2 and P_3 but not P_1, and we know that it diverges. The oscillating series with nth term $u_n = (-1)^{n-1}$ has properties P_1 and P_3 but not P_2, and therefore diverges. The series with nth term

$$u_n = \frac{1}{1 + (-1)^n \sqrt{\left[\frac{n-1}{2} + 2\right]}} ,$$

in which the symbol $\left[\frac{n-1}{2} + 2\right]$ denotes the greatest integer not larger than $\frac{n-1}{2} + 2$, has properties P_1 and P_2 but not P_3; and this series,

$$\left\{\frac{1}{1-\sqrt{2}} + \frac{1}{1+\sqrt{2}}\right\} + \left\{\frac{1}{1-\sqrt{3}} + \frac{1}{1+\sqrt{3}}\right\} + \left\{\frac{1}{1-\sqrt{4}} + \frac{1}{1+\sqrt{4}}\right\} + \cdots$$

$$= -2\left(\quad 1 \quad + \quad \frac{1}{2} \quad + \quad \frac{1}{3} \quad + \cdots\right)$$

diverges because the harmonic series does.

The class expressed complete satisfaction.

Northwestern University   H. A. Simons

Occasionally we shall quote from letters sent to any one of the editors of the *Magazine,* so don't be surprised to meet yourself in print. We shall give signatures only when we have permission to quote. Here is this month's letter:

"I enclose $10 for my current sponsoring subscription to the *Mathematics Magazine.* The new typography is beautiful and distinctive and I'm uncertain as to its comparative readability. The articles are well chosen, for readability and serviceability to all brackets of the profession."

We have received from Professor Cleon C. Richtmeyer of the Department of Mathematics of Central Michigan College of Education an excellent pamphlet, published by the Michigan Section of the Mathematical Association of America, and sent to all high schools in Michigan. Its title is "A Mathematics Student—To Be or Not to Be?" It states to high school students how they may find themselves hampered, when they get to college, if they do not take in high school at least a year of algebra and a year of geometry. An explanation is given of the different college subjects which demand that much mathematics as a prerequisite and of the different professions after college for which a knowledge of mathematics is important.

To Young Instructors of Mathematics

Of course you are teaching mathematics because you enjoy and respect the subject and are eager to have others share your interest. In the day-by-day study of mathematics there are necessarily dull details to be mastered. Keep them from seeming dull by showing implications of the subject beyond the matter in hand. Historical parentheses have their place. So, too, do occasional discussions which seem to you largely made up of half-understood philosophical and mathematical ideas. Such ideas presented by one student may now and then point the way to more profound ideas on the part of another student.

Prepare your work as well as you can, and expect your students to prepare their work as well as they can. Don't expect perfection—keep your perspective on the learning process by continuing to learn in your own particular field of interest, reading the publications of others and trying to write something of your own. Don't expect perfection, but take time to rejoice when a student comes near it. The whole class should feel a sense of satisfaction in the resourcefulness and independent thought of an excellent student and in the beauty of the painstaking step-by-step process by which mathematical results are achieved.

In the main, undergraduates learn by "doing". Give plenty of opportunity for written work outside of class. Give time in class for *thoughtful* work, don't feel that ideas must be offered every minute.

The class may appear more interested if suggestions are popping out, but the good student wants quiet in which to think, and the poor student only gets a sense of rush and discouragement if other students are presenting frequent and important-sounding comments.

Don't expect to know everything and don't pretend you do. You lose your own self-respect for a silly pretense and you lose the respect of the brighter members of the class.

Don't get too discouraged when things go wrong, when you fail to interest the girl who is taking mathematics because her father said she should, or even her friend who "just loved" high school mathematics. Do the best you can to make the subject clear, interesting, exciting. You may fail with some individuals, you may succeed with others, but in neither situation be over-egotistical. Keep your sense of fairness toward yourself and others, and, above all, keep your sense of humor.

Wellesley College   Helen G. Russell

How can you tell a mathematician from a demagogue?
The mathematician postulates.
The demagogue expostulates.

H. W. B.

Absent Minded Professors

When the Mathematics Magazine replaced the National Mathematics Magazine, after the latter had been out of publication for two years, over two hundred subscribers signed for the Mathematics Magazine agreeing to remit upon receipt of the first issue. After receiving four issues, a few had still forgotten. So we sent reminders. Replies were pleasing, especially the following:

"Your letter of the 20th instantly clears up for me a mystery of many months' standing. Each time I received a copy of Mathematics Magazine I was pleasantly surprised. I could not recall having subscribed to the magazine. Perhaps, thought I, some good friend was making me a gift.

Now, it is all clear.

I like the magazine, but 75¢ a copy seems steep, even with todays' prices.* Nevertheless, I am sending you enclosed a check for $6.00 as requested by you.

With best whishes for success, I am
Sincerely yours, "

*This is an error of 25%. The actual cost to subscribers has been 60¢, which is less than the cost to the publisher of each of the first four issues. However we have reduced production costs by buying our own vari-typer, thus avoiding any increase in the subscription price.

First   Year   College
Mathematics   with   Applications

By Daus and Whyburn

This new text presents a coordinated study of college algebra, analytical
trigonometry, and analytical geometry complete in one volume. Emphasis
throughout the book is placed on creating understanding as well as on
learning manipulative techniques. Each topic has been included because
of its immediate applications as well as future needs. These applications
include problems of a geometric character with an applied background,
problems in curve fitting, and elementary electric circuit theory when
related to mathematical problems involving algebra or analytic geometry.
To be published in the fall. $5.00 (probable)

> PAUL H. DAUS is Professor of Mathematics, University of Cali-
> fornia, Los Angeles. WILLIAM M. WHYBURN is Professor of
> Mathematics and President of Texas Technological College.

Analytic   Geometry
Fourth   Edition

By Clyde E. Love

The fourth edition of this text differs from previous editions in both
style and content. Explanations are fuller, and applications and
exercises are more numerous and more varied. Algebraic curves are
introduced early, and less space is devoted to conic sections. A new
chapter has been added on the analytic geometry of trigonometric
functions, and one on exponentials and logarithms.
Published March 23, 1948. $3.50

> CLYDE E. LOVE is Professor of Mathe-
> matics at the University of Michigan.

The   Macmillan   Company
60 Fifth Avenue   New York 11

J. B. LINKER and M. A. HILL, JR., *University of North Carolina*

MATHEMATICS OF FINANCE

In the most concise and teachable manner this new text covers the entire field of a first college course in the mathematics of finance. It is suitable for use in either a semester or a quarter course. Distinguished for its insight into student needs, the book provides many illustrative examples and well-graded practical exercises. It is unique in its extensive treatment of life annuities and life insurance and in its discussion of the symbolic method of studying annuities.

"I like the method of presentation, the content and the exercises in this book....The printing is clear, the pages make a nice appearance, and the use of symbols is consistent."
—*H. A. Dangel, University of Cincinnati*

"The chief methods are treated concisely and clearly, special emphasis is given on applications to the problems of insurance, and the text includes adequate tables for computation."
—*O. W. Albert, University of Redlands*

"This is a most excellent book. It gives what the student of commerce needs for his work for the first semester of Business Mathematics."
—*Thomas C. Amick, Louisburg College*

1948 - 258 *pages, including* 83 *pages tables* - $2.90

CHARLES H. SISAM, *Colorado College*

CONCISE ANALYTIC GEOMETRY

Thoroughly adapted to the short course in analytics, this text presents an excellent, well-rounded, and brief treatment of the subject. Clear and complete in spite of its brevity, it is given great flexibility by the addition of more than 1,100 exercises and problems, graded to meet the needs of students of varying ability.

1946 - 155 *pages* - $2.00

ANALYTIC GEOMETRY

This sound, teachable, and widely used book drills the student in the essentials of plane and solid analytical geometry and stresses types of reasoning vital to later work.

1936 - 310 *pages* - $2.20

CLIFFORD BELL, *University of California, Los Angeles*
TRACY Y. THOMAS, *Indiana University*

ESSENTIALS OF PLANE AND SPHERICAL TRIGONOMETRY

In this enlarged edition of a successful text, the authors have completely revised and expanded their treatment to adapt it to the standard requirements of present-day courses. A special feature is the large number of new problems and exercises presented at the close of each chapter. New tables include an excellent five-place table of the natural functions, a conversion table, a table of roots, and a table of useful constants.

Revised 1946.- 409 *pages, including* 144 *pages tables* - $2.30
without tables - $2.00

HENRY HOLT AND COMPANY
257 FOURTH AVENUE
NEW YORK 10

OUR CONTRIBUTORS

Nilos Sakellariou, Prof., U. of Athens, was born in Hagios Petros, Arcadia, Greece in 1882. He studied at the universities of Athens (Ph. D), Jena and Gottingen and was appointed to his present position in 1918. Dr. Sakellariou has also been a professor of theoretical mechanics at the Greek Naval Academy. A leader in mathematical education in Greece as well as a prominent research mathematician, Prof. Sakellariou is the president of the Greek Mathematical Society and an editor of the Bulletin of that Society. He has written many textbooks ranging from Practical Arithmetic to Lectures on Theoretical Mechanics, as well as a number of papers on the calculus of variations and differential geometry.

H. W. Becker was born in Sioux Co., Nebraska in 1908. After attending the University of Chicago and the University of California, he became engaged in the automatic heating business in Omaha. During the war Mr. Becker was an electronics instructor for the Navy at Mare Island. His memberships in the A.I.E.E., I.R.E., A.M.S., Math. Assn. and the Modern Languages Assoc. reflect his diversified interest in electrical engineering mathematics and poetry. Another hobby is chess.

Henry E. Fettis attended Wittenberg College in Springfield, Ohio (A.B. '39) and taught for three years in schools in the vicinity of Dayton. Since 1942 he has been employed as an Engineering Mathematician in the Dynamics Branch of the Wright Field Aircraft Laboratories. A member of the Math. Assn., he is interested particularly in geometry and aerodynamics.

Glenn James was born in Wabash, Ind. in 1882. He attended the Universities of Indiana (A.B. '05, A.M. '11), Chicago and Columbia (Ph.D. '17) and taught at Michigan State, Purdue and the Carnegie Institute of Technology before joining the faculty of U.C.L.A. in 1922. Dr. James is the author of a number of research papers and articles on teaching. For several years he was editor and publisher of the Peace Digest, and was the editor and compiler (with his son R. C. James) of the Mathematics Dictionary. Since 1947 he has been Managing Editor of the Mathematics Magazine.

A biographical sketch of Prof. *Maurice Fréchet* appeared in the March-April issue of the Mathematics Magazine (Vol. XXI, No. 4)