VOL. 26, No. 5, May-June, 1953

MATHEMATICS magazine

MATHEMATICS MAGAZINE

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

Address editorial correspondence to Glenn James, special papers to the editors of the departments for which 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. We prefer that, in technical papers, the usual introduction be preceded by a *foreword* which states in simple terms what the paper is about.

The Mathematics Magazine is published at Pacoima, California by the managing editor, bi-monthly except July-August. Ordinary subscriptions are 1 yr. $3.00; 2 yrs. $5.75; 3 yrs. $8.50; 4 yrs. $11.00; 5 yrs. $13.00. Sponsoring subscriptions are $10.00; single copies 65¢. Reprints: 25 free to authors; 100 @ $1.25 per page with a minimum of $6.00. All reprints are bound.

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

Advertisers should contact the managing editor.

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

MATHEMATICS MAGAZINE

Vol. 26, No. 5, May-June, 1953

CONTENTS

When we started publication of the Mathematics Magazine we undertook to have our research articles judged by our readers, as is done in some other fields. However usage seems to have developed the feeling that when an article in mathematics has been published it is no longer a subject for critical analysis. In fact we have received only one negative criticism and it was neither analytical not constructive. Careful reading of the large number of papers we receive quickly became impossible for the editors. Hence we have asked and received the valuable services of the following referees, for which we are very grateful:

E. F. Beckenbach, Clifford Bell, Brockway McMillan, Herbert Busemann, Bernard Friedman, Alford Horn, D. H. Lehmer, Norman Levinson, R. S. Phillips, J. F. Randolph, E. Snapper, I. S. Sokolnikoff, E. G. Straus, J. Dean Swift, A. E. Taylor, W. R. Wasow, and A. L. Whiteman.

Editor

OUR CONTRIBUTORS

John Hamilton Curtiss was born in Evanston, Illinois in 1909. A graduate of Northwestern University (A.B. '30), the University of Iowa (M.S. '31) and Harvard University (Ph.D. '33), he taught at Johns Hopkins and Cornell Universities from 1935 to 1943. He served as a naval officer from 1943-46 and became Assistant to the Director, National Bureau of Standards in 1946. Since 1947 he has held the position of Chief of the National Applied Mathematics Laboratories, National Bureau of Standards.

Stanley E. Rauch, Associate Professor of Mathematics, University of California, Santa Barbara, was born in Portland, Oregon in 1916. After graduating from Reed College (B.A. '37), the University of Washington (M.S. '39) and Stanford University (Ph.D. '41), he taught mathematics and physics at the University of Oregon and then joined the staff of the University of California Radiation Laboratory. In 1945 he was appointed to the faculty of the University of California, Santa Barbara and became Chairman of the Department of Mathematics in 1948.

Thomas Griselle was born in Upper Sandusky, Ohio in 1891. A musician by profession, Mr. Griselle is a graduate of the College of Music of Cincinnati, and studied composition with Nádia Boulanger and Arnold Schoenberg in Europe. He is a pianist and composer, with compositions including songs, piano pieces, chamber music and works for orchestra. His "Two American Sketches" won First Prize in the Victor Talking Machine Contest in 1928. His interest in mathematics is a hobby.

ELEMENTS OF A MATHEMATICAL THEORY OF PROBABILITY

J. H. Curtiss

1. *Introduction.* It seems to be a characteristic of maturity in human thought, whether it be concerned with the sciences or with everyday affairs, ultimately to recognize that everything is subject to variation and change. The realization finally dawns that conclusions are never final and predictions are never sure, however much they may be supported by contemporary evidence; and that often it is best to state them in a form which emphasizes - even quantifies - their uncertainty.

This pattern seems to be deeply bound up with certain principles of human behavior. Primitive or ignorant men, when confronted with poorly understood, uncontrollable situations, turn in their anxiety to ritual, to authority, to revelation. This stage of thinking may eventually be accompanied (or even replaced) by a rationalism of high intellectual order. But there is little room for organized concepts of chance here.

As factual knowledge grows in the community, and with it, control over environment, the authoritarian point of view tends to be replaced by the empirical approach. The trouble is that empiricism, honestly and studiously pursued, keeps running into baffling problems of change and chance. At first, it may be neither wise nor convenient to give open recognition to such problems, but sooner or later they simply must be faced.

All this can be illustrated in many ways; for example, by reciting the history of witchcraft, or of religion, or of various special sciences, such as agriculture or weather forecasting. There is space here only to look at a very broad instance, and then to touch on some typical special cases.

The broad instance is that of the philosophy of the natural sciences in the western world. Omitting the earliest beginnings in superstition and theology, we start with the classical concept of the natural science; which goes back to the Greeks and which attained towering achievements in the seventeenth and eighteenth centuries. It was that of a body of unique, clearly defined, unchanging laws. Human frailty being what it is, it was possible to apprehend the system only little by little; but it was supposed to exist in a complete and perfect state just the same. It would presently become clear by the exercise of the processes of reasoning alone, just as a field of mathematics is clarified. Experimental data could sometimes be an interesting guide, but for drawing ultimate conclusions they were untrustworthy. The important truths were self-evident, or if not, would become so as the thinking evolved.

Euclidean geometry was in fact considered the model branch of science within this classic conception. When the non-Euclidean geometries were discovered in the nineteenth century, and then were actually employed in theoretical physics, a heavy blow was given to the rationalistic tradition. Many other factors influenced the swing away from rationalism; the story is far too complicated even to sketch here. But the fact is that empiricism has now largely replaced the classic tradition, and with it has finally come rather full recognition of the importance of chance and change. The modern scientist freely admits that even his most strongly supported beliefs and theories have a contingent character. His reliance is now primarily on the *adequacy of methods* rather than on the infallibility and uniqueness of conclusions.*

In a scientific world dominated by number and measurement, even such an elusive concept as uncertainty seems to demand a scale of measurement, and a clearly understandable definition, and perhaps even a theory. Actually the crude elements of a quantitative approach to probability have been present for a long time, and there was a substantial development of the theory long before the rationalistic doctrines in the philosophy of science were thrown out. Nagel [13] discerns certain statistical notions in the writings of Aristotle on biology. Cardan in the sixteenth century, Galileo, Pascal, Fermat, and Huygens in the seventeenth century, the Bernoullis, De Moivre, and Bayes in the eighteenth, all solved special problems involving quantitatively measured probabilities.

Their work was brought to a climax in Laplace's *Theorie analytique des probabilities (1812)*. This astonishingly detailed and complete treatise was accompanied by an influential popular exposition, *Essai philosophique sur les probabilities*. Laplace's thinking dominated the field for a hundred years, and it was only in the early part of the twentieth century that substantial advances began to be made again in probability theory.

The motivation for the seventeenth century work in probability theory was largely frivolous, but during the eighteenth century attempts were made to find probabilistic answers to serious questions arising in the theory of observations in public administration, in judicial procedure, and in theoretical physics. These attempts culminated in a string of brilliant successes in the nineteenth century; among them are the Maxwell theory of gases, Boltzman's work; and the Mendelian theory in genetics. The formulation of theories and results in quantitative probabilistic terms has now become a familiar phenomenon in physics, chemistry, biology, and many other branches of science.

2. *The problem of the definition.* Just what is probability? Halmos, a mathematician, says firmly [7], "Probability is a branch of mathematics. It is not a branch of experimental science nor armchair philosophy, it is neither physics nor logic... . The situation is analogous

*Nagel [13], pp. 1-4.

to that in geometry." He then goes on to develop "probability", in his sense, as a branch of mathematical measure theory - a procedure which we follow here when the time comes to turn to mathematics.

But most philosophers, and many scientists, would say that what Halmos was talking about was just a part of probability theory - the part sometimes identified nowadays as "mathematical probability theory". They would consider the mathematician much too ambitious who would preempt the entire field of probability for mathematics - that is, unless he wishes to view mathematics in a greatly extended sense and is willing to take many strange new troubles on his shoulders. The non-mathematical problems involved in probability seem more extensive and more subtle than even those involved in the physical interpretations of geometry and classical analysis.

Consider the following statements in which the term probability or its derivatives appear. It will be noted that they all have quite a familiar sound. They will be divided into two groups.

First group: "It is more probable that it will rain in Los Angeles in January than in June". "The probability that a neutron from a radioactive source placed against a certain lead shield will pass through that shield is 0.346". "The probability that a uniform pair of dice will show a total of seven when tossed is 1/6". "The probability is 0.95 that in random sampling from a normal population with known standard deviation σ and unknown mean μ, the interval $(\bar{x} - 1.96\sigma, \bar{x} + 1.96\sigma)$ will cover μ, where $\bar{x}$ is the arithmetic mean of the sample observations". "The probability that an American male in a certain profession will survive his fiftieth birthday is 0.422".

Second Group: "It is improbable that Bacon wrote Shakespeare". "Probably there is life somewhere in the universe in addition to that on the earth". "It is probable that if Napoleon had not had certain physical ailments, the entire course of history would be changed". "It is more probable that the Aztecs got to Mexico by way of Alaska than by paddling canoes across the Pacific". "Probably this man was murdered". "The theory of evolution is more probable on the evidence than the Biblical theory of creation".

The statements of the first group all have one thing in common: the word probability is being used somehow to indicate the intensity of one's expectation that an event will happen. It is being used in a predictive sense. In the second group, the word seems to have something to do with the adequacy of evidence relating to events or situations which already *have* occurred.

It is the opinion of most students of probability that a more or less satisfactory mathematical interpretation and theory can be adduced for the term probability as used in the statements of the first group. On the other aand, it does not seem to be at all certain that this can be done for the second group*. In any case, with so many implications

*A discussion will be found in Nagel [13], Section III.

to cover, the task of *choosing* and *evaluating* a mathematical model - a project which belongs on the bridge between mathematics and the physical world, rather than in the field of mathematics itself - looks as if it might be quite as difficult intellectually as the formal development of the model chosen.

If there is to be a mathematical theory, it is pretty evident that it would be helpful, *although not necessary*, to have some sort of precise metrical definition of the word "probability" to use inside the mathematical theory as a starting point. Unfortunately neither the users nor the students of probability are agreed among themselves as to how to do this. Certain philosophers, notably Reichenbach, appear to favor the idea of defining probability as the limit (in the classical sense of mathematical analysis) of a sequence of relative frequencies in an infinite reference class. Von Mises [12] in a series of papers which started to appear shortly after World War I, was the first to study this definition mathematically. To make the mathematical theory useful, he found it necessary to restrict the reference class by postulating that certain subsequences of relative frequencies all have the same limit. Other mathematicians subsequently have shown that this restriction is so severe that the definition is now quite generally regarded as unsuccessful in spite of many attempts to fix it up.*

Others like to go back to Laplace's formulation of a probability as the ratio of the cardinal number of two sets of alternatives, with the set corresponding to the denominator including the set corresponding to the numerator. (This is the mathematical content of the definition usually proposed in college algebra books.) Closely related to this is the old definition of "geometric probability" as the ratio of two lengths or two areas, without reference to physical applications. Certain physicists propose that probability shall be a measure of "partial beliefs", given certain evidence. There are numerous other definitions, too, principally proposed by philosophers and requiring the technical language of philosophy to state.

But although the semantical and pragmatic problems involved in probability have not been solved, nevertheless as implied in Section 1, probability concepts are successfully being used throughout modern science and industry. Then too, by the beginning of the present century, quite a large body of mathematical theory had been worked out. The literature of these studies contained a good deal of attractive mathematics, and conveyed a promise of more gold to be found with a little digging. As the new century wore on, more and more mathematicians began taking trips into the gold fields. Under the added stimulus of the applications the mathematical literature began to grow by leaps and bounds.

*See Doob [2], and also the discussion which follows this reference in the same issue of the Annals of Mathematical Statistics.

Now one of the fashions of the day in mathematics is the axiomatic approach. For this to operate effectively in a branch of applied mathematics, it should be based on physical phenomena which are simple, fundamental, and which have been widely experienced. The present-day mathematical probabilists wanted to axiomatize their discipline. What intuitive or physical concept of probability should they choose as the basis? In view of the philosophical difficulties in probability theory touched on above, the decision was not altogether an easy one to make.

The requirements of simplicity and wide understanding, together with the mechanics of many of the applications, dictated the answer. The basic intuitive concept which most of the mathematicians now working in probability theory use is this: Physically speaking, a probability shall be regarded as a stabilized, long-run value of the relative frequency of the occurrence of an event in repeated trials (under "identical" conditions) of an experiment which can result in this and in other events.

The construction of the *mathematical* foundations then proceeds roughly as follows. A fixed number is arbitrarily selected to represent the idealized value of the relative frequency. A calculus of these numbers is then developed on the basis of a Boolean algebra whose elements consist of "events", which are taken as the completely fundamental, undefined elements of the mathematical theory, just as points and straight lines are the undefined elements of Euclidean geometry. It turns out that the calculus actually looks like a rather highly developed special case of the general theory of distributions of mass – a theory which has been studied for a long time in physics and engineering.

This construction does not attempt *at the outset* to provide as faithful a description of the empirical situation as did that of Von Mises. It does not include in its axiomatics, for example, a characterization of "randomness"; but the characterization appears later on in the guise of certain theorems. Incidentally, the simplicity of the axiomatics has the fortunate result that the ensuing calculus can readily be used in connection with a number of other physical interpretations of probability. One of these is the definition used by Laplace involving the ratio of alternatives, and all the classical mathematical work in probability theory can therefore be accepted almost without change*.

*In connection with the axiomatics of mathematical probability, Prof. Mark Kac pointed out to the author that one of David Hilbert's famous twenty-three problems included the task of axiomatizing probability theory. The statement of Hilbert's problem runs as follows [10], p. 306:

"Durch die Untersuchungen über die Grundlagen der Geometrie wird uns die Aufgabe nahe gelegt, *nach diesem Vorbilde die jenigen physikalischen Disziplinen axiomatisch zu behandeln, in denen schon heute die Mathematik eine hervorragende Rolle spielt: dies sind in erster Linie die Wahrscheinlichkeitsrechnung und die Mechanik.*

"Was die Axiome der Wahrscheinlichkeitsrechnung angeht, so scheint es mir wunschenswert, dass mit der logischen Untersuchung derselben zugleich eine strenge und befriedigende Entwicklung der Methode der mittleren Werte in der mathematischen Physik, speziell in der kinetischen

An exposition of the elements of this mathematical theory of probability will now be given. As it unfolds, it is hoped that the mathematical reader, whatever his preconceived notions and prejudices about probability may be, will perceive that the mathematical theory forms a perfectly respectable branch of mathematics. To be sure, it has its own special notation, and its nomenclature has a feature typical of all mathematics intended to be applied, in that purely mathematical objects are sometimes called by names which have distracting physical or psychological connotations (e.g. "moment", "expectation"). But for all that, it is just as rigorous by current mathematical standards as, say, modern topology. The areas of doubt and controversy which have been pointed out earlier do not lie in the realm of mathematics. Their existence only serves warning that however satisfying the theory may be from the mathematical viewpoint, it might end up by not being accepted by the users, at least in its present form. The chances of this look rather small at present, but they are not entirely negligible.

3. *The definition of mathematical probability for the discrete case.** Our starting point is at the principal undefined notions of the mathematical theory. First we introduce the concept of a simple event. Consider the set of absolutely all thinkable or distinguishable outcomes of an experiment or observation. These are the *simple events* of the experiment. They may or may not be naturally characterized by numbers through the conditions of the experiment; this has nothing to do with the concept. For example, in the toss of a coin, the simple events are head and tail. In the experiment consisting of dealing at the card game known as bridge, each of the possible deals is a simple event. In tossing two dice, a red one and a green one, each of the possible throws (such as deuce on the red one and five-spot on the green one) is a simple event. In statistical mechanics, the various possible states of a system correspond to simple events.

It will be convenient to employ geometric terminology. Thus the simple events for an experiment will henceforth be considered to be points in a certain space**, which will be called the *sample space* for the experiment. The sample space is sometimes called the *event space*. In statistical mechanics it is called *phase space*.

The concepts of simple event and of sample space are the basic undefined elements of our theory, in just the same way that points and straight lines are the undefined elements in an axiomatic treatment

Gastheorie Hand in Hand gehe."
Prof. Kac surmises that Hilbert would *not* have been satisfied by the measure-theoretical approach outlined in the sequel in the present paper.

*The treatment in this section and the next roughly parallels that given in Kolmogorov [11], but with certain simplifications. The terminology generally follows Feller [4] and Halmos [7], [8].

**The word *space* is used in this article in the standard sense of modern analysis; it simply means a collection of elements.

of Euclidean geometry.

We have been using the two words "simple event" instead of just the single word "event" because we propose to use the unmodified word "event" to denote something less fundamental; namely a result of an experiment which is obtained if any one of several of the simple events occurs. For example, in the dice-tossing experiment we might be interested in obtaining a total of seven in one toss of the two dice. This occurs if any one of a number of the simple events occurs; they are the various individual throws which add up to a seven. In the card dealing case, it might be convenient to consider all at once the deals in which North is dealt only spades. This is a composite event again - or just an "event" in the terminology we shall adopt - which is brought about if any one of the numerous deals occurs (each of which is a simple event) which put all the spades into North's hand.

Moving over to the geometric picture again, it is natural to represent an event, in the sense just introduced, by the set or aggregate of points in the sample space which consists of all of the simple events whose occurrence makes it occur. Henceforth then, an event will mean the same as a set of points in a sample space.

We shall use capital letters like A to denote individual points (simple events) in the sample space and lower case italicized letters like a, b, ..., to denote sets of points (events) in the sample space. Of course we do not intend by this notation to exclude the case in which a point set a consists of only one point.

If the sample space contains only a finite number of points, or a countable infinity of them (this means that they can be arranged in a sequence A_1, A_2, ...), then the sample space is called *discrete*. We shall confine ourselved to discrete sample spaces in the next few paragraphs.

Suppose that (at least in imagination) an experiment can be repeated over and over in such a way that from intuitive or experimental considerations there is reason to hope that the respective relative frequencies of the occurrence of the simple events will become stabilized near certain fixed numbers as the repetitions proceed. Then with each point A of the sample space, we shall associate a real number $p(A)$, $0 \leqq p(A) \leqq 1$, which is to be regarded physically as the idealized value, or hypothetical value, or proposed value, of the corresponding relative frequency. We call it the *probability* of A.

Following the relative frequency guide, we define the probability of an event a consisting of two or more distinct simple events to be the sum of the probabilities of the component simple events. The probability of a will be written $pr(a)$. The case in which a consists of an infinite sequence A_{j_1}, A_{j_2}, ... is included; the meaning assigned to the infinite series $\sum_k pr(A_{j_k})$ which arises in this connection is the usual one of classical mathematical analysis.

Further following the relative frequency guide, the assignment of the function $p(A)$ is made so that $\sum_j p(A_j) = 1$, where the summation is to be extended over all the points A_1, A_2, ... of the sample space.

The *probability distribution* in the sample space is defined to be $pr(a)$, considered as a function of the set a. Alternatively, in the discrete case here under consideration, it can be defined to be the point function $p(A)$.

We stop here for some observations of a philosophical nature. In the first place, the reader should be careful to distinguish the mathematical elements in the above material from the physical ones, even though we have deliberately put them side by side. All references to relative frequency were only for motivation and guidance as to the appropriate features of the mathematical model. Mathematically, probability as viewed here is merely a non-negative, bounded function defined on the points of the sample space and having certain additive properties. The actual choice of the function $p(A)$ in a given instance, and the psychological implications involved in calling it a probability, are matters which lie outside the realm of mathematics.

But the reader may still wonder if the specification of $p(A)$ is really not the heart of the whole matter; if it cannot be done accurately, then the whole theory may be useless and it is futile to go on and develop it further. There are several answers to this question. In the first place, it is perfectly true that in the present physical formulation, *and also with a formulation based on limits (in the sense of classical mathematical analysis) of relative frequencies in infinite sequences of experiments*, probabilities can never be assigned to events with finality. Therefore with such a theory every probability specification is forced to play the rôle of a *scientific hypothesis*. But researchers in the present-day mathematical theory of probability have been notably successful in developing tools within the theory for the testing of such hypotheses in the light of observed data. With such tests at hand, the hypothesis viewpoint is quite in the spirit of modern empirical science, in which the cycle of hypotheses → theoretical consequences → obtaining data → testing of theory with the data → new hypotheses, is a standard part of the methodology.

Then too, surprisingly useful information about the general behavior of a stochastic* system can sometimes be obtained even with a very crude specification of the initial probabilities. For example, sequences of tosses of a coin exhibit certain interesting regularity and irregularity properties which are pretty much independent of whether or not it is assumed that the probability of a head is ½. And finally there is a big body of asymptotic probability results in which the eventual distributions of certain stochastic systems are shown to be approximately independent of the initial specification.

Another less profound question which may be bothering the reader is this: how does the above definition of probability meet the requirements of the classical elementary combinatorial problems in probability

*Stochastic is a synonym for probabilistic.

theory? The answer is easy to give. In the language of our present discussion, the typical combinatorial problem in the college algebra text-books describes in more or less vague terms an experiment which clearly calls for a finite discrete sample space. Considerations of symmetry, and of experience (they are sometimes dressed up in some rather shaky philosophy which goes back to the eighteenth century) suggest that the appropriate probability specification is one which assigns equal probabilities to all the points in the sample space. The problem then asks for the probability of a certain event a. The solution is merely a matter of counting up the sample points which comprise a. The task can be by no means trivial, and is facilitated by using the various tools of combinatorial analysis.

All the classical "urn problems" are of this character. For example, there is the old urn problem (translated here from urn language to be more entertaining to the reader) which asks for the probability that if a person writes n letters and puts them "at random" in the addressed envelopes, all the letters go astray. The sample space consists of the $n!$ assignments of the letters to the envelopes; the implication is that all assignments are to be probabilized equally. The answer involves some rather careful counting, and comes out to be (curiously enough) approximately $1/e$ for large n, where e is the well-known constant arising in elementary calculus.

The hypothetical nature of the probability specification is possibly somewhat obscured in the case of such impractical problems, but it would rapidly come to the fore if one tried to use the answers to them in real-life gambling operations. Many more serious examples involving the same techniques, but taken from physics and other branches of science, will be found in the newest books on mathematical probability, notably those of Feller [4] and Fortet [5]. In these applications, the hypothetical character of the probability specification is much more apparent, because the specification is clearly a part of the hypotheses of a physical theory which must stand or fall on the evidence of experimental data.

4. *Extension to more general sample spaces.* The discrete type of sample space which has been considered so far is adequate for many purposes, but it is not a suitable basis for a general theory. In the first place, if we are going to set up an acceptable model for the applications in the natural and social sciences, we must provide for events which consist of measurements made, at least conceptually, on a continuous scale. The mathematician may argue all he wishes that, in practice, measurements will always be discrete because they will always be rounded off. His colleagues in the other sciences are not going to be satisfied with purely discrete mathematical theories – not yet, anyhow. They have good arguments on their side from the mathematical viewpoint, too; continuous formulations are often much easier to manipulate than discrete ones.

But the idea of non-discrete sample spaces turns out to be much more than a gadget added to the theory to please the clients; it is necessary to attain completeness in the study of certain familiar experiments which, when considered abstractly, are seen to consist of a non-countable number of simple events. An example is given by the game of craps. Another is the problem of the number of tosses of a coin required to get the first head. To see why the simple events are uncountable in such cases, consider the latter problem, and suppose that no *a priori* restrictions are imposed on the number of tosses. If we denote a head by 1 and a tail by 0, a typical infinite sequence of tosses could be written symbolically as 1 1 0 1 0 0 0 1 0 1 1 ..., and this must be regarded as one of the simple events of the experiment. With a binary point out in front of the symbol, it can be regarded as a binary number on the unit interval. The sample space then consists of all such binary numbers; that is, of all the points in the unit interval. Now it is well known that the points of the unit interval are not countable, so a more general kind of sample space is called for than that which we have heretofore considered.*

True, this example (and that of the crap game) can be given a satisfactory mathematical treatment by restricting the number of throws *a priori* to a finite number - that would have to be done in real life anyway - and then resorting to limiting processes. But to study the structure of "random" sequences like the sequence of 0's and 1's above and to deal satisfactorily with questions of the type which Von Mises attempted to answer in his axiomatics, it is quite necessary to free the theory from any obvious restrictions as to finiteness. This becomes even more urgent in the case of certain important applications to economics and physics, where it is necessary to think of a stochastic system evolving continuously in time.

We shall accordingly now outline an approach which provides a wide generalization of the definition of mathematical probability previously given.

The clue to the way to proceed lies in a closer study of the idea of an event. We have identified an event with a set of points in the sample space. This set consists of the simple events, the occurrence of any one of which makes the event occur. Now to every event a, we sense that there must correspond another event a' which has the property that if a does not occur, a' does, and vice versa. Thus we are led to define for each event a its *complementary event* a', which consists of all of the points of the sample space not in a'. Another intuitive idea is that of at least one of two events a and b occurring. This will happen if the experiment results in a simple event contained in either

*Incidentally, this sort of correspondence between sequences of trials and numbers on the unit interval has providdd a useful connection between number theory and probability theory; see Feller [4] pp. 161-163.

a or b or in the set common to both. We shall therefore define for every pair of events a and b another event denoted by $a \cup b$, which is identified with the set of all points in the sample space which belong to at least one of the sets a and b. In point set theory, $a \cup b$ is called the *union* of a and b. Of course the extension to a *finite* collection of events is immediate; $a \cup b \cup c \cup \ldots$ means the aggregate of points belonging to at least one of the sets $a, b, c, \ldots$.

Still another intuitive idea associated with two events a and b is that of both of them occurring. This will take place if any simple event occurs for which the corresponding point lies in both a and b simultaneously. We shall formalize this by defining for each pair a and b a new event $a \cap b$, represented by the set of all points in the sample space common to both a and b. In point set terminology, $a \cap b$ is the *intersection* of a and b. Similarly, if $a, b, c, \ldots$, is a finite collection of events, $a \cap b \cap c \cap \ldots$, represents the simultaneous occurrence of all of these events.

Of course, $a \cap b$ may be empty; this is the case in which a and b are disjoint or are *mutually exclusive*. It is convenient to enlarge our concept of event slightly to include this case too. We therefore invent the "impossible" event o, and say that $a = o$ means that the set a is empty.

It is also convenient to have a symbol, say e, for the "certain" event consisting of all the points in the sample space.

Now the really important thing about the operations $'$, $\cup$, $\cap$, which we have just introduced, is that they clearly must obey certain algebraic laws. For example, $a \cup b = b \cup a$, with a similar relation for $\cap$; these are commutative laws. There is an associative law, also, for each of these two symbols. There are various obvious relations between the symbols; for example, $(a \cap b)' = a' \cup b'$, $a \cap (b \cup c) = (a \cap b \cup (a \cap c)$. The latter is a distributive law.

It now begins to appear that the naive, seemingly formless idea of an event is not so innocent as it seemed at first, and that a mathematical treatment is possible. There is in fact an algebra of events. This algebra is of a recognizable type which has long been studied. If a non-empty class α of elements is such that (i) if two elements a and b are both in α, then $a \cup b$ is also in α, and (ii) if a is in α then a' is in α, where $\cup$ and $'$ satisfy the algebraic laws sketched above, then this class α is called a *Boolean algebra*. It is clear that with the extensions of the concept of event given above, the totality of events for any experiment may be viewed as the elements of a Boolean algebra.

To take care of certain complicated situations such as that of the infinite series of coin tosses mentioned above, it turns out that a restriction must be imposed on the class of admissible events. This restriction consists of requiring that the union of an infinite sequence of events shall have a well-defined meaning and shall itself be an event in the algebra. In the coin tossing case, it is reasonable to talk about the particular event a of the coin never turning up a head at all. If b_j is the event that head turns up for the first time on the j-th roll, then it is pretty obvious that a' (the event that the coin *does* eventually turn up a head) should be the union of b_1, b_2, ... , provided we can assign a precise meaning to that. (It is true that this example is somewhat artificial, as it can be handled by limiting processes and anyhow we feel that here $a = o$, $a' = e$, which we have already agreed to put in our algebra. Further discussion of the need for the restriction will be found below.)

To put this new restriction in precise terms, we first introduce another fundamental symbol. If a and b are any two events which satisfy the relation $a \cup b = b$, we write $a \subset b$ and say (in the language of events) "a implies b," or (in the language of point sets) "a is contained in b," or "a is smaller than b." We consider $a = b$ to be a special case of $a \subset b$. Given an aggregate of elements, it is now possible to talk about the smallest one. If it exists, it is an element a in the aggregate such that if b is any other element in the aggregate, then $a \subset b$.

Now consider an arbitrary infinite sequence a_1, a_2, ... of elements of a Boolean algebra. Let b be an element of the algebra which contains every one of the elements a_1, a_2, ... Consider the class of all such elements b. If this class contains a smallest one, say a, then we define $a_1 \cup a_2 \cup a_3 \cup \ldots$ to mean a. If such an element a exists for every such sequence a_1, a_2, ... , the algebra is called a σ-algebra.

Our restriction on the class of admissible events is that it shall always be a σ-algebra.

We are now ready to define mathematical probability for general sample spaces. Once again, we shall let the obvious properties of relative frequencies guide us, and furnish motivation. Consider the σ-algebra of the events for a given experiment. Suppose that the experiment can be repeated over and over (at least in the imagination) in such a way that the relative frequency of occurrence of any event a will become stabilized about a certain fixed value as the repetitions proceed. Then with each event a, we associate a real number, $Pr(a)$, $0 \leqq Pr(a) \leqq 1$, which is regarded physically as the idealized value, or hypothetical value, or proposed value, of the corresponding relative frequency. This number is called the probability of a.

It will be remembered that the σ-algebra of events contains the impossible event o and the certain event e. To these we assign respectively the values $Pr(o) = 0$, $Pr(e) = 1$.

Again following the relative frequency guide, if a and b are any two mutually exclusive events, then we *define* $Pr(a \cup b) = Pr(a) + Pr(b)$.

The properties of $Pr(a)$ postulated so far imply that if a_1, a_2, ... is any infinite sequence of mutually exclusive events, then $Pr(a_1 \cup a_2 \cup \ldots a_n) = Pr(a_1) + Pr(a_2) + \ldots + Pr(a_n)$ for n finite. Furthermore the limit as n becomes infinite of the sum on the right-hand side must exist, because the sum is monotonically non-decreasing and yet bounded by unity. In other words, the infinite series $Pr(a_1) + Pr(a_2) + \ldots$ has a sum in the usual sense. Now it will be remembered that $a_1 \cup a_2 \cup \ldots$ is a member of the σ-algebra, and so a probability $Pr(a_1 \cup a_2 \cup \ldots)$ has been assigned to it. *Our final stipulation* regarding the set-function $Pr(a)$ *is that the sum of the convergent series* $Pr(a_1) + Pr(a_2) + \ldots$ *is precisely* equal to $Pr(a_1 \cup a_2 \cup \ldots)$. In technical language, this means that the set-function $Pr(a)$ is countably additive.

The set-function $Pr(a)$ is called the *probability distribution* for the experiment.

At the beginning of Section 2 of this article, we asked the question, "Just what is probability?" The developments of the present section have put us in no better position to answer this question in its philosophical sense. But we can now give a completely precise answer to the question, "Just what is mathematical probability in the present formulation?" *It is simply a countably additive, non-negative set function, bounded from above by unity, defined on the elements of a σ-algebra.*

It is easily seen that the definition of probability for discrete sample spaces is merely a special case of the general definition given immediately above. The comments made in the discrete case in connection with the problem of specifying the probability distribution still hold true. In the applications, the specification is always in the nature of a scientific hypothesis.

In mathematical statistics, where testing of hypotheses is of primary interest, it is a common practice to specify for an experiment not just a single probability distribution $Pr(a)$, but all at once a whole family of probability distributions $Pr(a;\theta)$. The parameter θ, which may be multidimensional, varies on a space Ω of admissible hypotheses".

The extreme values 0 and 1 of $Pr(a)$ require some attention. We do not exclude the possibility of assigning the probability 0 to a set a other than o, if it seems reasonable to do so. This would mean in practice that the occurrence of a is thought to be unlikely, but not necessarily impossible. Similarly, we do not exclude the assignment of a probability of 1 to an event other than e. It should be remembered that a probability is not supposed to reproduce a precisely determined value of a relative frequency, and – we say this again, at the risk of being repetitious – in applying the present theory every mathematical probability is to be regarded as only a scientific hypothesis.

The upper bound of unity on $Pr(a)$ is of course suggested by the situation for relative frequencies. However, if other intuitive or physical concepts of probability are used in connection with the algebra of events which we built up in the preceding section, then there may be no reason why the scale of $Pr(a)$ should be limited to the unit interval. In fact, it might even be convenient to put the values of $Pr(a)$ in the complex plane*.

There is a close relationship between mathematical probability as defined above and Lebesgue measure; an elementary discussion will be found in the article of Halmos [7].

It is worth emphasizing, although it is a rather technical point, that in the present treatment a probability has not been assigned to *all* events a in a given sample space, but just to those which belong to a σ-algebra. Thus in the case of the experiment consisting of an infinite sequence of coin tosses, suppose that the sample space is chosen to be the unit interval. Then not all point sets (that is, events) in the interval would be "probabilized" if the present definition of probability is used, but only the point sets belonging to a σ-algebra on this interval.

If the restriction that the algebra of events to be probabilized must be a σ-algebra is dropped, and if the sample space is a relatively complicated one (for instance, if it can be put into one-to-one correspondence with the unit interval) then there may be events on which a non-negative, additive set function just cannot be defined in a unique and non-contradictory way. The question is much the same as that of the existence of non-measurable sets in the Lebesgue theory, which is discussed in detail in the modern measure-theory texts (see for example [8], pp. 67-72). Of course, such "non-probabilizable" events may be artificial, and even pathological, and certainly without physical interpretation, but that would not make the mathematician any happier if he had to deal with an inconsistent theory.

On the other hand, there are sophisticated cases in which the restriction to a σ-algebra is certainly too severe. The lifting of the restriction has recently been the subject of a good deal of research, and certain technical questions about the existence of "non-probabilizable" events (using definitions of probability suitably modified over the one given here) have not yet been settled.**

5. *Further definitions; elementary theorems.* Using only the definitions given in Section 4 together with the algebra of events, certain more or less obvious theorems can be proved. We shall exhibit only one of them - one of considerable importance in the theory.

Notice the identity $a \cup b = a \cap b' + a \cap b + a' \cap b = a + a' \cap b = b + b' \cap a$. This is derivable from simpler, more fundamental identities, but if the reader will translate it aloud into event language, it will be very obvious. By the additive property of probability, we have,

*In quantum theory in physics, certain complex functions are introduced called *probability amplitudes* which are given a probabilistic significance. The squares of their absolute values are treated as ordinary probability densities (this term is defined below in section 8). See Heisenberg [9], Chap. IV.

**See [5], footnotes on pp. 50, 53, 164.

$$Pr(a \cup b) = Pr(a \cap b) + Pr(b' \cap a) + Pr(a' \cap b),$$

$$Pr(a \cup b) = Pr(a) + Pr(a' \cap b),$$

$$Pr(a \cup b) = Pr(b) + Pr(b' \cap a).$$

Combining these, we obtain:

THEOREM 1. $Pr(a \cup b) = Pr(a) + Pr(b) - Pr(a \cap b)$.

This says that the probability that one or the other of two events occurs is equal to the sum of the probabilities that they occur individually minus the probability that they both occur simultaneously. It clearly reduces to a mere statement of the additive property of probability if $Pr(a \cap b) = 0$.

Consider now the quantity $Pr(a \cap b)$; that is, the probability of the two indicated events occurring simultaneously. It is natural to put a slightly different twist on this concept, and look into the possibility of defining the probability of the event which consists of the occurrence of b when a is known to have occurred. Once again, relative frequencies will be our guide. If in N trials of an experiment, a occurs $N(a)$ times, b occurs $N(b)$ times, and they both occur together $N(b \cap a)$ times, then $N(b \cap a)/N(a)$ is the relative frequency of occurrence of b within the sequence of trials in which a occurred. This fraction can be written as

$$\frac{N(b \cap a)}{N(a)} = \frac{\dfrac{N(b \cap a)}{N}}{\dfrac{N(a)}{N}},$$

and we have already defined the abstractions of the numerator and denominator on the right side as respectively $Pr(b \cap a)$ and $Pr(a)$.

We are thus led to the definition of the conditional probability of b, given a, as $Pr(b \mid a) = Pr(b \cap a)/Pr(a)$. The tacit assumption is always made in defining conditional probability as a quotient that the probability in the denominator is not equal to zero.

With a held fixed, it is easy to see that the class of events of the type $b \cap a$ forms a σ-algebra, with $e = a \cap a$. If $b \cap a = o$, then $Pr(b \mid a) = 0/Pr(a) = 0$. If $b = a$, $Pr(b \mid a) = Pr(a)/Pr(a) = 1$. Also, $Pr(b \mid a)$ satisfies the addition rule of probabilities, because by the distributive law of the event algebra*

$$Pr(b_1 \cup b_2 \cup b_3 \cup \ldots \mid a) = Pr[(b_1 \cup b_2 \cup b_3 \cup \ldots) \cap a]/Pr(a)$$

$$= Pr[(b_1 \cap a) \cup (b_2 \cup a) \cup \ldots]/Pr(a)$$

$$= Pr(b_1 \cap a)/Pr(a) + Pr(b_2 \cap a)/Pr(a) + \ldots .$$

* It is used to pass between the second and third members of the equation below.

Since $a = b \cap a + b' \cap a$, it follows that another way to write $Pr(b \mid a)$ is $Pr(b \cap a)/[Pr(b \cap a) + Pr(b' \cap a)]$; this shows that $Pr(b \mid a) \leqq 1$.

The upshot of all this is that conditional probability has all the mathematical properties of just plain probability, and satisfies all the general theorems.

Conditional probability is a concept of central importance from both the philosophical and the mathematical point of view. It is not an exaggeration to say that practically all of the philosophical and psychological theories of probability somewhere harbor the idea that the intensity of one's belief in the occurrence of an uncertain event, or in the existence of a situation, or in the "truth" of a statement, is always relative to the evidence at hand. In a sense, then, probability is almost always used in a conditional sense by all human beings. From a less philosophical viewpoint, conditional probability is essential in various ways for the development of the mathematical model. The most fundamental application is that it provides the simplest way to approach the very important concept of statistical independence. Then too, it is one of the tools required for the study of time series and sequences of associated or correlated events - phenomena which under the name of "stochastic processes" are receiving a great deal of attention in the literature today.

There are several ways to rewrite the definition of conditional probability. Because of their historical interest we shall elevate two of the formulas to the dignity of theorems. The first is what is sometimes known as the theorem, or law, of compound probabilities:

THEOREM 2. $Pr(a \cap b) = Pr(a)Pr(b \mid a)$.

The second formula has the distinction, unusual in a discipline so long associated with games of chance, of being named after an English clergyman. Let $b_1, b_2, \ldots$ be a sequence of events which are mutually exclusive and such that $e = b_1 \cup b_2 \cup \ldots$. Then $a = a \cap e = (a \cap b_1) \cup (a \cap b_2) \cup \ldots$. Also, $(a \cap b_i) \cap (a \cap b_j) = o$, $i \neq j$, and so by the additive property of probability, $Pr(a) = \sum_j Pr(a \cap b_j)$. We substitute this into the denominator of the fraction which defines $Pr(b_k \mid a)$ and then use Theorem 2 on both numerator and denominator, obtaining

$$\textit{THEOREM 3.} \quad Pr(b_k \mid a) = \frac{Pr(b_k)Pr(a \mid b_k)}{\sum_j Pr(b_j)Pr(a \mid b_j)}$$

This is called Bayes' theorem. It has caused a great deal of excitement and controversy in its day. For a long time it was the principal tool of statistical inference. The classical interpretation was to view the events b_k as "causes" and the event a as the result of an experiment affected by these causes. The formula then was supposed to give the probability that b_k caused a. In many cases it is quite possible to deduce reasonable values for the probabilities $Pr(a \mid b_k)$ which appear in the formula, but the hitch in practice lies in properly estimating the "*a priori* probabilities" $Pr(b_j)$, $j = 1, 2, \ldots$. The limitations of Bayes' theorem are now pretty well understood, and other much better techniques for statistical inference are now available.*

We shall now define statistical independence; that is, independence in the probabilitistic (as opposed to functional) sense. This is the last of our really fundamental definitions.

The intuitive meaning of the independence of the events a and b is that the knowledge that a has occurred in no way affects our expectations as to the occurrence of b. Returning to the relative frequency guide, this should mean that $N(b)/N$ and $N(b \cap a)/N(a)$ ought to be about the same; that is, the occurrence or non-occurrence of a ought to have no bearing on the frequency of occurrence of b. This leads us to make the following definition: *The events a and b are statistically independent if and only if* $Pr(b \mid a) = P(b)$.

Another way to write this, in view of Theorem 2, is $Pr(a \cap b) = Pr(a)Pr(b)$. From this it is quickly seen that if a and b are independent, then not only do we have $Pr(b \mid a) = Pr(b)$, but also $Pr(a \mid b) = Pr(a)$.

The appropriate extension to the case of three events a, b, and c might seem to be via the route of requiring them to be independent in pairs, but it turns out that this is not enough for a satisfactory definition. For example, it does not insure that $Pr(a \cap b \cap c)/Pr(a) = Pr(b \cap c)$.** The easy way out of the difficulty is to say that three events are independent if (1) they are independent in pairs, and (2) $Pr(a \cap b \cap c) = Pr(a)Pr(b)Pr(c)$. The generalization to more than three events proceeds similarly.

6. *Random Variables*. Most of the results of modern mathematical probability theory are stated in a convenient although special language: that of *random variables*. It is a terminology which mathematical purists have always found rather irritating, because random variables are not independent variables in the strict mathematical sense (they are functions), and on closer acquaintance they do not seem to be very random, whatever that may mean. Furthermore, the standard notation used to represent a random variable is somewhat inexplicit.

Nevertheless the concept is intuitively helpful and satisfying. A random variable roughly means a quantity which is associated with the simple events of an experiment to which a probability distribution

*See Cramer [1], in particular pp. 507 ff. It is only fair to state that with other physical concepts of probability than the one used in the present article, Bayes' theorem can be much more important.

**See Feller [4], p. 87.

has been assigned; and this quantity assumes various values with the probabilities derived from the association. In brief, it is a variable quantity whose values are determined by chance.

To arrive at a more precise definition, we must make use of an idea borrowed from the modern theory of functions of a real variable; namely, that of a measurable function.* Consider the sample space for an experiment, and the associated σ-algebra of events. By a measurable function in the present context, we mean a vector function $\mathbf{X} = \mathbf{X}(A) = [X_1(A), X_2(A), \ldots, X_k(A)]$ whose values lie in a Euclidean k-space, and which is defined and finite (that is, none of the components $X_j(A)$ is infinite) for each point A in the sample space, and which has the property that the set of points A in the sample space such that simultaneously $X_1(A) \leqq C_1$, $X_2(A) \leqq C_2$, $\ldots$, $X_k(A) \leqq C_k$, is always a member of the σ-algebra for any and all values of the C's, including infinite ones.

Any such measurable function defined on a sample space will be called a random variable.

A random variable is always taken to be a single-valued function, in the sense that to any point A in the sample space there corresponds just one point in the k-dimensional Euclidean space. This plays an essential rôle later on. If $k = 2$, the random variable may be presented in complex-variable notation as $\mathbf{X} = X_1 + iX_2$; there is no change in the definition of measurability if this is done. If $k = 1$, we drop the bold-face type and use the same notation for the random variable as was used above for the components.

It is not unusual to encounter situations in which several different random variables are defined all at once over a sample space. If they are all one-dimensional random variables, they may be thought of as being the components of a single vector random variable. However, a warning must be given with respect to the interpretation of vector random variables constructed in this way. Consider a concrete example. A coin is tossed twice. The simple events of the experiment, in an obvious code, are HH, HT, TH, TT. Let X and Y be two random variables set up simultaneously on the sample space, with X assuming the values 1, 2, 3, 4 on HH, HT, TH, TT respectively, and Y the values 7, 8, 9, 10 on HH, HT, TH, TT. Now (X,Y) could be considered as a vector random variable whose space was the XY-plane, but *the values that this vector variable takes on as the argument ranges over the sample space are only the values* (1,7), (2,8), (3,9), (4,10). Other points of the plane, such as (1,8), even though they lie in the product space**

*See the article [6] by J. W. Green in this series.

**Given a space of elements (or points) X, and another space of elements (or points) Y, their product space (or Cartesian product space) is simply the set of all ordered pairs (X,Y). Thus the Euclidean plane is the product space of two coordinate axes. The Cartesian product of more than two spaces is defined similarly.

of the values of X and Y, would not be included among the values of the vector variable. (This product space would pertain to a new, more complicated experiment.)

A value of a random variable **X** obtained by actually performing the underlying experiment, getting the event A, and therefrom deriving a numerical value for **X** through the functional correspondence $\mathbf{X} = \mathbf{X}(A)$, is called a *determination* of **X**, or an *observation* on **X**, or a *realization* of **X**. This terminology, of course, pertains to concepts which lie on the bridge between mathematics and the physical world, and are not a part of the mathematical structure which we are now erecting.

Random variables are very often initially introduced into a probability problem just to provide numerical labels for the points of the sample space. This step becomes so natural to workers in probability and statistics that they often take it unconsciously before starting a discussion. The more fundamental concept of sample space is almost never mentioned in the literature of the mathematical theory of probability except in discussions of the foundations.

The transition from sample space to random variables in many cases is made more painless by the circumstance that the most reasonable system of coding for an experiment often drops a random variable in one's lap, so to speak. An example is the experiment consisting of the toss of a die. Here the obvious way to designate the six simple events $A_1, A_2, \ldots, A_6$ is by the number of spots shown; but this sets up a random variable on the sample space at once. Of course it may be desirable to define simultaneously a number of other random variables on this same sample space; one of them, for instance, might be that given through the formula $Y = \exp(iX)$, where X is the number of spots shown.

More complicated examples along this line are those given by "probabilizable" physical measurements made on a continuous scale. If the measurements themselves are numbers, as they usually are, they furnish random variables at once. A mathematical justification for short-circuiting the sample space concept in such a case, and in many other situations as well, will be given in a moment.

The motion of random variable as defined above, like that of sample space, is sterile without the associated concept of the probability distribution of a random variable. The general idea is simply that this distribution is to be derived from that in the sample space by making corresponding sets in the space of the random variable and in the sample space have the same probability. To do this consistently requires some little attention to detail.

To be specific, let $\mathbf{X} = \mathbf{X}(A)$ be a measurable vector function defined over a sample space on which a probability distribution $Pr(a)$ has been set up. For convenience we enlarge (if necessary) the space consisting of all the values taken on by $\mathbf{X}(A)$, as A ranges over the sample space, so that it is an entire k-dimensional Euclidean space, which we henceforth call the **X**-space. By the inverse image of a set **x** in the X-space, we mean the set of points $a_{\mathbf{x}}$ in the sample space such that if a point A lies in $a_{\mathbf{x}}$ (written $A \in a_{\mathbf{x}}$) then $\mathrm{X}(A)$ lies in $\mathbf{x}$. Notice that by the definition of single-valued function, every A has one and only one X, although a single value of X may correspond to several points A; also that if the values of $\mathrm{X}(A)$ do not fill it up the X-space completely as A ranges over the sample space, there will be some sets $\mathbf{x}$ with empty inverse images. The inverse image of an empty set will be, by agreement, the empty set. The inverse image of the entire X-space is the entire sample space, or e, in our previous terminology.

We must first restrict the class of sets $\mathbf{x}$ in the X-space to be "probabilized", because otherwise their inverse images might not belong to the σ-algebra in the sample space. (It will be recalled that the definition of measurable function insures that the inverse images of X-sets only of a certain extremely simple type belong to the σ-algebra in the sample space.) The proper restriction is to the family of Borel sets of the X-space. These are defined as the σ-algebra of sets of Euclidean k-space obtained by applying the operations $\cup, \cap, '$, a finite or countably infinite number of times to X-sets of the simple type appearing in our definition of measurable function. It follows at once that every Borel set $\mathbf{x}$ has an inverse image $a_{\mathbf{x}}$ belonging to the σ-algebra in the sample space.

After all these preliminaries, the probability distribution $p(\mathbf{x})$ of the random variable $\mathbf{x} = \mathrm{X}(A)$ will now be defined. Assuming that $\mathbf{x}$ is a member of the family of Borel sets in X-space, the definition is given simply by the equation $P(\mathbf{x}) = Pr(a_{\mathbf{x}})$.

A few moments' thought on the part of the reader will assure him that this equation defines $P(\mathbf{x})$ uniquely and without contradiction as a countably additive set function, with $0 \leqq P(\mathbf{x}) \leqq 1$. It assigns a probability of zero to the empty set (and also to any set whose inverse image is empty), and a probability of unity to the whole X-space. In other words, $P(\mathbf{x})$ has all the characteristic properties of the probability distribution $Pr(a)$ in the sample space. The probability distribution of X, as so defined, is said to be *induced* by that in the sample space.

No modification to the procedure for inducing distributions is required for random variables which are complex-valued functions. They are treated simply like vector functions with $k = 2$.

If $Y = Y(\mathrm{X})$ is a single-valued Borel-measurable function of X (meaning that the inverse image of the interval $Y \leqq Y_0$ is always a Borel set for every Y_0), then clearly $Y = Y(\mathrm{X}(A))$ is a measurable function defined on the sample space, so Y is itself a random variable. Its distribution may be derived from that of X (or from that in the sample space – the result is the same) in exactly the same way as that in which the distribution of $\mathrm{X}(A)$ was obtained. The extension of these remarks to vector functions $Y = Y(\mathrm{X})$ is immediate.

Although a Borel-measurable function $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$ maps any Borel set of the $\mathbf{Y}$-space back onto a Borel set of the $\mathbf{X}$-space, this is not necessarily true for a Lebesgue-measurable function and Lebesgue-measurable sets. This is one of the reasons for using Borel sets and Borel-measurable functions in probability theory.

The mechanism by which the probability distribution in the sample space induces a probability distribution in the $\mathbf{X}$-space depends heavily on the single-valued character of the function $\mathbf{X} = \mathbf{X}(A)$. This prevents two disjoint sets $\mathbf{x}_1$ and $\mathbf{x}_2$ from having the same inverse image $a_{\mathbf{x}}$. If this had been allowed to happen, then there obviously would have been difficulty in apportioning $Pr(a_{\mathbf{x}})$ to $\mathbf{x}_1$ and $\mathbf{x}_2$ individually.

To put this all in another way, if $\mathbf{X} = \mathbf{X}(A)$ is single-valued but not inversely single-valued, the probability distribution of $\mathbf{X}$ can be induced as above, but the process cannot be made to work the other way: that is, given the distribution of $\mathbf{X}$, this distribution does not induce a unique distribution in the sample space. On the other hand, if $\mathbf{X} = \mathbf{X}(A)$ is inversely single-valued, as would be the case if $\mathbf{X}(A)$ were introduced merely to label the individual points of the sample space with distinct numerical labels, then by reversing the rôles of the sample space and the $\mathbf{X}$-space in the above discussion, the probability distribution in the sample space can be uniquely derived from that in the $\mathbf{X}$-space. Under these circumstances, it makes no difference at all which distribution is specified first. Similar remarks apply to the case of a function $\mathbf{Y} = \mathbf{Y}(\mathbf{X})$, with the rôles of A and $\mathbf{X}$ assumed by $\mathbf{X}$ and $\mathbf{Y}$.

This is the mathematical justification promised earlier in this section for putting primary emphasis on random variables instead of sample spaces when working in probability and statistics, even if the concept of sample space is the more fundamental one. Most sample spaces arising in practice are of a character such that their points can be given distinct numerical labels whose values lie in the space of all real numbers, or in Cartesian products* of spaces of real numbers. As stated before, this process is so natural that it is often performed unconsciously before starting a problem. In fact in many cases it would be merely pedantic to distinguish between the sample space and the space of the labeling random variables. And the moment the labeling has been done, thenceforth the corresponding random variables might as well occupy the center of the stage as far as probability considerations are concerned.

A constant C can be treated as a random variable because it can be viewed as a function. In fact, it often is so treated in the literature of mathematical probability. Its distribution, when induced according

*Actually, uncountably many factors may be needed; that is, sample spaces arise in practice which are abstractly identical with function spaces. Two methods of specifying probability distribution in function spaces are described and compared in [3].

to the above rules by that in any sample space, is always one in which a probability of unity is assigned to the point C, and zero to any set not containing C.

The modern literature of probability and statistics contains countless examples of random variables and induced probability distributions. Here we shall have space only for some very simple concrete illustrations of the process of inducing a probability distribution for a random variable. They will all be based on the coin tossing experiment mentioned earlier, in which the simple events were HH, HT, TH, TT.

A natural probability specification in the sample space for this experiment is one which assigns equal probability to each of the simple events. This implies that each of them gets the probability 1/4. One of the various random variables proposed previously merely labeled these points respectively 1, 2, 3, 4. Its induced probability distribution is therefore given by the equations $P(j) = 1/4$, $j = 1, 2, 3, 4$. Another possible random variable, not mentioned previously, is a variable X which assumes the value 1 on any of the three points in the sample space containing H in its code, and 0 on TT. The total probability attached to the inverse image of $X = 1$ is 3/4, so the distribution of X is $P(1) = 3/4$, $P(0) = 1/4$.

Still another possible random variable is a vector function which plots the simple events as four points on the Euclidean plane as follows: (0,0) for TT, (1,0) for HT, (0,1) for TH, (1,1) for HH. Its induced probability distribution is given by $P(j,k) = 1/4$, $j = 0, 1$, $k = 0, 1$.

REFERENCES

1. Cramer, Harald. Mathematical Methods of Statistics. Princeton, 1946.
2. Doob, J. L. Probability as a Measure. Annals of Math. Stat., vol. 12 (1941) pp. 206-214.
3. Doob, J. L. and Ambrose, W. On Two Formulations of the Theory of Stochastic Processes Depending upon a Continuous Parameter. Annals of Math., vol. 41 (1940) 737-745.
4. Feller, W. An Introduction to Probability Theory and its Applications, New York, (1950).
5. Fortet, Robert. Calcul des Probabilités, Paris, 1950.
6. Green, John W. The Theory of Functions of a Real Variable. Mathematics Magazine, vol. 24 (1951) 209-217.
7. Halmos, Paul R. The Foundations of Probability. American Math. Monthly, vol. 51 (1944) 493-510.
8. Halmos, Paul R. Measure Theory, New York, 1950.
9. Heisenberg, W. The Physical Principles of Quantum Theory. Chicago, 1930.
10. Hilbert, David, Gesammelte Abhandlungen, Berlin, 1935.
11. Kolmogorov, A. N. Foundations of the Theory of Probability. New York, 1950.
12. von Mises, R. Wahrscheinlichkeitsrechnung. Leipzig u. Wien, (1931)
13. Nagel, Ernest. Principles of the Theory of Probability, International Encyclopedia of Unified Science, vol. I, No. 6, Chicago, III, 1939.

Harvard University and

The National Bureau of Standards

CYCLOIDAL MOTION OF ELECTRONS

S. E. Rauch

1. Introduction.

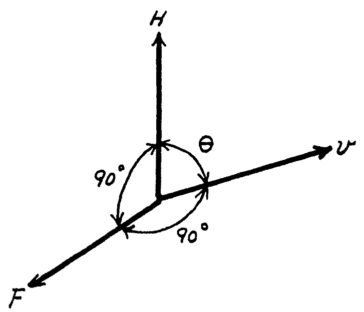
The cycloid has found its way into important applications in the present fields of atomic physics and electronics. Engineering problems which consider the motion of charged particles when moving through a region of crossed electric and magnetic fields deal directly with cycloidal paths and velocities. The following article analyzes the characteristics of the cycloidal path and illustrates briefly two applications.

Let the following concerning forces acting upon an electron be assumed as hypotheses:

a. When a force acts upon an object it produces an acceleration in the direction of the force with a magnitude proportional to the force. Units are so defined that F(dynes) = m(grams) $\cdot$ a(cm./sec.^{2}), where m is the mass being accelerated and a is the acceleration. In vector notation Newton's law of motion is simply stated by $\vec{F} = m\vec{a}$.

b. When an electron having e charge units is in an electric field E, it is accelerated in the direction of the electric field by a force F(dynes) = eE, where the units are chosen for e and E so that the equality is satisfied. In vector notation the above becomes $\vec{F} = e\vec{E}$.

c. When a charged particle having e charge units moves with a velocity v(cm./sec.) relative to the observer across a magnetic field H, a



force is exerted perpendicular to the plane defined by the direction of H and v. The magnitude $F = (Hev/c)\sin\theta$ is in dynes, where c is the velocity of light in cm./sec., θ the angle between H and v, H is gauss units so defined to satisfy the equality. In vector notation the above is summarized by $\vec{F} = (e/c) \cdot (\vec{v} \times \vec{H})$.

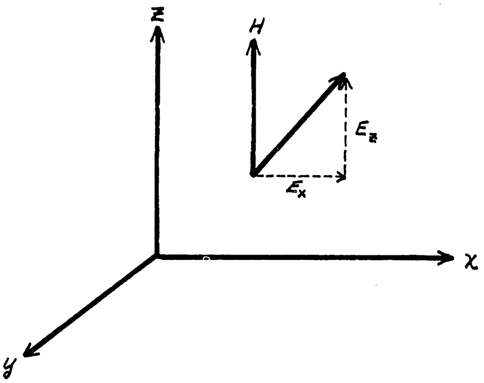
2. Equations of Motion

The general problem to be studied can be discussed in terms of results of a simple case. For this purpose let us first study the motion of an electron when moving in a region with a uniform electric field produced by a parallel plate condenser superimposed on a uniform magnetic field between plane, parallel poles. Define reference axes in this region such that E has no component in the y direction, and H is parallel to the positive z axis. On the basis of the preceding remarks regarding forces acting upon an electron which is moving in an electric and magnetic field, the equations of motion are:

(2.1) $$m\, d^2z/dt^2 = eE_z,$$

(2.2) $$m\, d^2y/dt^2 = -(He/c)\cdot dx/dt,$$

(2.3) $$m\, d^2x/dt^2 = eE_x + (He/c)\cdot dy/dt$$



Integration of (2.1) with respect to t yields

(2.4) $$z = z_0 + z_0' t + (eE_z/2m)t^2,$$

where $z = z_0$, $dz/dt = z_0'$ when $t = 0$. Let $v_x = dx/dt$ and $v_y = dy/dt$. Upon differentiation of (2.3) one obtains

$$m\, dv_x^2/dt^2 = (He/c)\cdot dv_y/dt,$$

and upon substituting into (2.2) it is found that

$$d^2v_x/dt^2 + (eH/mc)^2 v_x = 0.$$

The solution of the latter is

$$v_x = A\cdot\cos\gamma t + B\cdot\sin\gamma t,$$

where $\gamma = eH/mc$, and an integration with respect to t yields

(2.5) $$x = x_0 + (A/\gamma)\sin\gamma t - (B/\gamma)\cos\gamma t.$$

according to (2.3),

$$dy/dt = (1/\gamma)d^2x/dt^2 - cE_x/H;$$

therefore upon integrating with respect to t, assuming $y = y_0$ when $t = 0$, one finds

$$y = y_0 - (cE_x/H)t + (1/\gamma)dx/dt,$$

$$= y_0 - (cE_x/H)t + (A/\gamma)\cos\gamma t + (B/\gamma)\sin\gamma t,$$

(2.6) $$y = y_0 - Ut + (A/\gamma)\cos\gamma t + (B/\gamma)\sin\gamma t,$$

where $U = cE_x/H$.

As further initial conditions let $v_x = x_0'$, $v_y = y_0'$ when $t = 0$. Consequently $A = x_0'$ and $B = y_0' + U$. The equations of motion become

$$z = z_0 + z_0't + (eE_z/2m)t^2,$$

$$x = x_0 + (x_0'/\gamma)\sin\gamma t - [(y_0' + U)/\gamma]\cos\gamma t,$$

$$y = y_0 + (x_0'/\gamma)\cos\gamma t + [(y_0' + U)/\gamma]\sin\gamma t - Ut.$$

If θ_0 is so defined that

$$\sin\theta_0 = x_0'/[x_0'^2 + (y_0' + U)^2]^{1/2} \text{ and } \cos\theta_0 = (y_0' + U)/[x_0'^2 + (y_0' + U)^2]^{1/2},$$

then the equations of motion can be written in the final form:

(2.7) $$z = z_0 + z_0't + (eE_z/2m)t^2,$$

(2.8) $$x = x_0 + (1/\gamma)[x_0'^2 + (y_0' + U)^2]^{1/2}\cos(\gamma t + \theta_0),$$

(2.9) $$y = y_0 + (1/\gamma)[x_0'^2 + (y_0' + U)^2]^{1/2}\sin(\gamma t + \theta_0) - Ut,$$

where $\gamma = eH/mc$, $U = cE_x/H$, $\theta_0 = \arctan x_0'/(y_0' + U)$.

3. Parametric Equations of the Cycloid.

Let us next consider a circle, radius r, rolling along the y axis with an angular velocity γ. The parametric equations for the locus of a point at a distance d from the center of the circle are:

$$y = r\gamma t - d \cdot \sin\gamma t,$$

$$x = r - d \cdot \cos\gamma t,$$

where $y = 0$, $x = r - d$ when $t = 0$. By definition the locus of the point is a prolate cycloid if $d > r$, cycloid if $d = r$, and a curtate cycloid if $d < r$.

A more general form of these equations can be quickly established by considering first a translation of axes. Then

$$y = y_0 + r\gamma t - d \cdot \sin \gamma t,$$

$$x = x_0 + r - d \cdot \cos \gamma t.$$

It $t \neq 0$ when $y = 0$, then a phase angle θ_0 can be defined to satisfy

$$y = y_0 + r(\gamma t + \theta_0) - d \cdot \sin(\gamma t + \theta_0),$$

$$x = x_0 + r - d \cdot \cos(\gamma t + \theta_0).$$

A comparison of the equations of motion (2.7). (2.8). (2.9) with the above equations shows that the electron has two characteristic motions:

(1) a component of motion producing a parabolic path parallel to the magnetic field,

(2) a cycloidal component of motion perpendicular to the magnetic field.

In regards to the cycloidal motion certain quantities to which the physicist refers in his analysis of physical problems are important to mention. The definitions are given in terms of the rolling circle producing the cycloid but they can be easily specialized to discuss the motion of electrons in uniform crossed electric and magnetic fields.

Frequency ν. The angular velocity of the rolling circle associated with the cycloid is γ. Frequency, which is defined as the number of complete revolutions of the rolling circle per second, is

$$\nu = \gamma/2\pi.$$

Period T. The period is defined as the time required for a complete revolution of the rolling circle. Thus

$$T = 2\pi/\gamma.$$

Drift velocity U. The drift velocity is defined as the linear velocity of the center of the rolling circle and in magnitude is

$$U = r\gamma.$$

Amplitude d. The amplitude d is the distance from the center of the rolling circle to the point on the extended radius whose locus produces the cycloid. It is clear from (2.8) that

$$d = (1/\gamma)[x_0'^2 + (y_0' + U)^2]^{1/2}.$$

If x_0' and y_0' are expressed in units of U, then d can be written in the form

$$d = r[x_0'^2 + (y_0' + 1)^2]^{1/2}.$$

It at once follows that the locus is a prolate cycloid, cycloid, or curtate cycloid according to whether $x_0'^2 + (y_0' + 1)^2$ is greater than, equal to, or less than 1 respectively.

Path length per cycle λ. With the use of the familiar calculus definition of length one easily obtains

$$\begin{aligned} ds &= [x_0'^2 + (y_0' + U)^2 + U^2 - 2U[x_0'^2 + (y_0' + U)^2]^{1/2}\cos(\gamma t + \theta_0)]^{1/2}dt, \\ &= [d^2\gamma^2 + U^2 - 2Ud\gamma \cdot \cos(\gamma t + \theta_0)]^{1/2}dt, \\ &= [d^2\gamma^2 + r^2\gamma^2 - 2dr\gamma^2\cos(\gamma t + \theta_0)]^{1/2}dt, \\ &= \gamma[d^2 + r^2 - 2dr \cdot \cos(\gamma t + \theta_0)]^{1/2}dt. \end{aligned}$$

Integration with respect to t between the limits of 0 and T yields

$$\lambda = \gamma \cdot \int_0^T [d^2 + r^2 - 2dr \cdot \cos(\gamma t + \theta_0)]^{1/2}dt.$$

For the special case where the initial velocities vanish, yielding $d = r$, a well known result is obtained.

$$\lambda = r\gamma(2)^{1/2}\int_0^T [1 - \cos\gamma t]^{1/2}dt,$$

$$\lambda = 8r.$$

By substituting the maximum and minimum values of ±1 for $\cos(\gamma t + \theta_0)$ it is possible to establish rough bounds between which λ must lie, that is $|d - r|T < \lambda < |d + r|T$.

Average squared velocity per cycle, v^2/cycle. The use of $(dv)^2 = (dx/dt)^2 + (dy/dt)^2$ and the parametric equations (2.8) and (2.9) leads to the expression

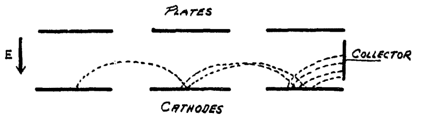
$$\begin{aligned} v^2/\text{cycle} &= (2\gamma^2/T)\int_0^{T/2} [d^2 + r^2 - 2dr \cdot \cos(\gamma t + \theta_0)]dt, \\ &= \gamma^2[d^2 + r^2 + (4dr/\pi)\sin\theta_0]. \end{aligned}$$

For the special case where the initial velocities are zero in magnitude, the result simplifies to

$$v^2/\text{cycle} = 2r^2\gamma^2 = 2U^2.$$

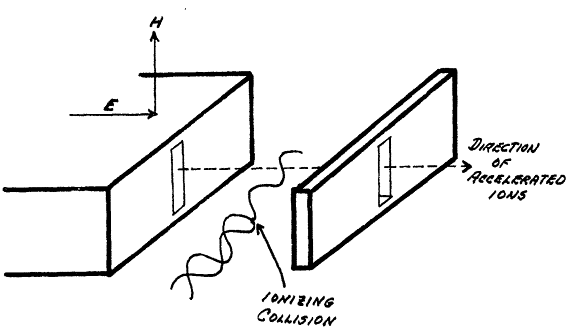
4. Applications.

Uniform crossed electric and magnetic fields have been employed in numerous electronic devices. A typical example is provided by the magnetic electron multiplier [1]. Electrons leaving a photocathode C are accelerated across an electric field towards the plate with P volts. A uniform magnetic field normal to the plane of the adjacent drawing turns the electrons in cycloidal paths to the next secondary



emission target. The electron multiplier seeks to produce a large electron flow upon the collector for only a small initial emission of electrons from a cathode which emits electrons when light falls upon its surface. When the electrons strike the next secondary emission plate, they are able to eject electrons from the surface, thereby multiplying the number of electrons several times. A similar field combination is employed in the television pickup tube, called the Orthicon [2]. The knowledge of the cycloidal paths and the corresponding velocities is necessary in obtaining the maximum performance for the equipment.

In contrast to the previous example electron multiplication can be a most serious engineering problem. Positive ion sources are used in electromagnetic separation of isotopes. In order to accelerate positive ions a uniform electric field is used in the presence of a magnetic field. Stray electrons in these regions are able to multiply exponentially as their path lengths increase. Usually the ability to ionize increases with path lengths, velocity, and gas pressure.



Eventually these electrons strike more positive surfaces, usually the ion source mechanism itself, thereby causing serious heating, destruction of metal, and a lowering of the positive charge on the unit.

The problem of the physicist is to construct the positive ion source such that the cycloidal path lengths and velocities of the electrons are kept at a minimum.

In order to assist the fuller appreciation of the magnitudes involved in the cycloidal paths of the electrons, a typical example of electric and magnetic fields used for the electromagnetic separation of isotopes is offered. For convenience let $x_0 = y_0 = 0$ and $E_x = 3000$ volts/cm. = 10 esu/cm., $E_z = 300$ volts/cm. = 1 esu/cm., $H = 3000$ gauss, $c = 3 \cdot 10^{10}$ cm./sec., $e/m = 5.3^{17} \cdot 10^{17}$ esu/gram. From the results obtained in section 3, the following magnitudes for the characteristics of the cycloidal motion are derived:

$$\nu = eH/2\pi mc = 8.4 \cdot 10^9 \text{ cycles/sec.},$$

$$T = 2\pi mc/eH = 1.2 \cdot 10^{-10} \text{ sec./cycle},$$

$$U = cE_x/H = 10^8 \text{ cm./sec.},$$

$$r = (m/e)(c/H)^2 E_x = 1.9 \cdot 10^{-3} \text{ cms.}$$

For the special cases of negligible initial velocities with respect to the drift velocity U, one finds in addition that

$$\lambda = 1.5 \cdot 10^{-2} \text{ cms.} = .15\text{mm.},$$

$$[v^2/\text{cycle}]^{1/2} = 1.4 \cdot 10^8 \text{ cm./sec.}$$

The magnitude of r suggests that the fluctuation of the electron in the direction of the electric field due to the cycloidal motion is very small. For the example above $E_x = 3000$ volts/cm.; thus a change of $1.9 \cdot 10^{-3}$ cms. is equivalent to only 5.7 volts. These results lead to the approximate statement of the physicist that the electron moves along the equipotentials with a drift velocity U. This latter statement of course does not take into account the effects of the electric field parallel to the magnetic field or that the actual amplitude d is cetermined by the initial velocities x_0' and y_0'. It is quite possible in a physical problem to have d much greater than r.

In regards to this latter point a final remark is worthwhile. When an electron has an ionizing collision with a gas molecule, or when upon striking a metallic surface an electron is ejected, there is an energy transfer to the created electrons which gives them kinetic energies greater than 10^8 cm./sec. This initial velocity can be in any direction relative to the electric or magnetic fields. Thus it is not necessarily correct to assume that x_0' and y_0' are negligible with respect to U. To illustrate the relative importance of the initial velocities a final example is illustrated. It is assumed that $x_0 = y_0 = x' = 0$. As a consequence $\theta_0 = 0$. Equations (2.8) and (2.9) simplify to

$$x = r(y_0' + 1)\cos \gamma t,$$

$$y = r(y_0' + 1)\sin \gamma t - Ut,$$

if y_0' is expressed in units of U. The cycloidal path as a function of the magnitude of y_0' is shown in figure 1.

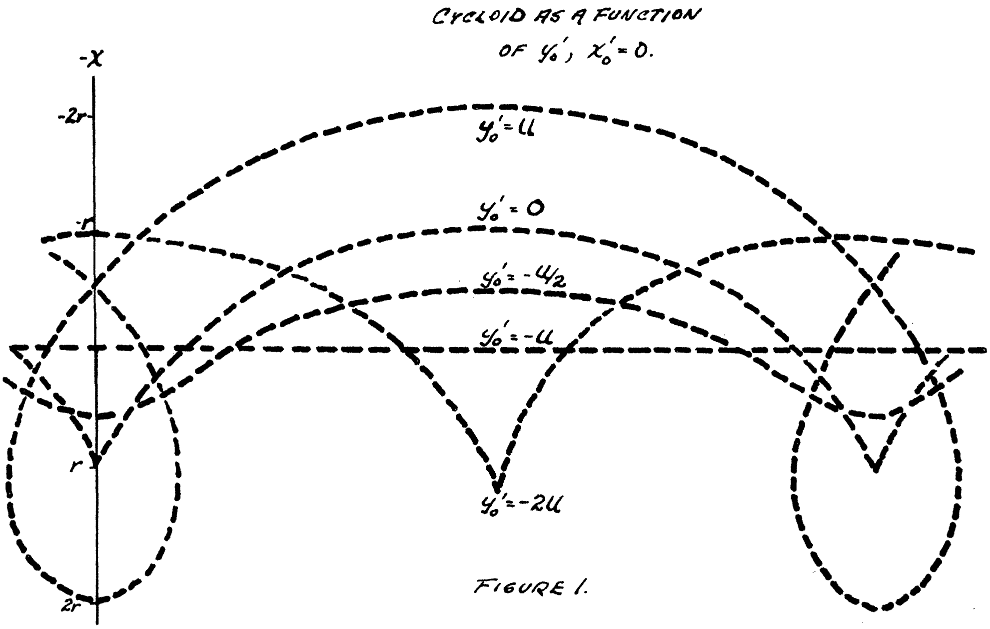


FIGURE 1.

REFERENCES

1. V. K. Zworykin, G. A. Morton, L. Malter, "The secondary emission multiplier - a new electronic device," Proc. Inst. Radio Engrs., Vol. 24, pp. 351 - 375, March 1936.

2. A. Rose and H. Iams, "Television pickup tubes using low velocity electron beam", Proc. Inst. Radio Engrs., Vol. 27, pp. 547 - 555, September 1939.

3. Page and Adams, "Principles of Electricity" , D. Van Nostrand Col, 1931, pp. 278 - 288.

4. Zworykin, Morton, "Electron Optics and the Electron Microscope" , John Wiley and Sons, 1945, pp. 520 - 527.

5. S. E. Rauch, "Mathematical Data for Electron Drain on Positive Ion Sources," Journal of Applied Physics, Vol. 22, No. 9, pp. 1128-1132, September 1951.

University of California, Santa Barbara

PROOF OF FERMAT'S LAST THEOREM FOR n = 2(8a + 1)

Thomas Griselle

THEOREM: No integral solution exists for

(1) $$x^{2m} + y^{2m} = z^{2m}.$$

m being a prime integer $>$ 2 and not a divisor of x, y, or z, unless m is of the form $8a + 1$.

PROOF: It suffices to assume that (1) has a primitive solution (one in which x, y, z are coprime), x being even, y and z being odd. We cannot assume that x and y are odd and that z is even, because the sum of two odd squares is not divisible by 4, as shown by the identity

$$(2a + 1)^2 + (2b + 1)^2 = 2[2(a^2 + a + b^2 + b) + 1].$$

Transposing the terms of (1) and factoring, we have

$$(z^2 - y^2)(z^{2(m-1)} + y^2 z^{2(m-2)} + \cdots + y^{2(m-1)}) = x^{2m}$$

It is obvious that the factor $x^2 - y^2$ is even. Having m odd terms, the second factor is odd. The G.C.D. of these two factors is m or 1*. Assuming that x is not divisible by m, it follows that neither factor is divisible by m. The G.C.D. then is 1. Hence, each is an integral $2m$th power. Let

(2) $$z^{2(m-1)} + y^2 z^{2(m-2)} + \cdots + y^{2(m-1)} = k^{2m}.$$

k being a suitable odd integer.

Being an odd square, each term of (2) is of the form $8a + 1$, as shown by the identity

$$(2b + 1)^2 = 4b(b + 1) + 1.$$

Thus, (2) is of the form

(2a) $$8(a_1 + a_2 + \cdots + a_m) + m = 8A + 1,$$

which, of course, cannot be satisfied unless m is of the form $8a + 1$.

Incidently, this argument proves that no primitive solution exists for

$$x^2 + y^{2m} = z^{2m}.$$

x being even, y and z being odd, m being a prime integer $>$ 2 and not a divisor of x, unless m is of the form $8a + 1$.

*See pages 87 – 88 of Diophantine Analysis, Carmichael.

MISCELLANEOUS NOTES

Edited by

Charles K. Robbins

Articles intended for this department should be sent to Charles K. Robbins, Department of Mathematics, Purdue University, Lafayette, Ind.

A FORMULA FOR THE CALCULATION OF THE INERTIA MOMENT OF SOME GEOMETRICAL SOLIDS

Christos N. Kefalas

The teaching in Colleges of the moment of inertia of the various solids, generally meets with difficulties in view of the necessity to use a specific formula in each particular case. The object of this study is to provide a standard formula covering several solids.

Figure (1) represents a solid bounded by a surface, each plane section of which is parallel to the xy-plane and is at a distance z from it, and consists of regular N-gons of a radius $R = f(z)$.

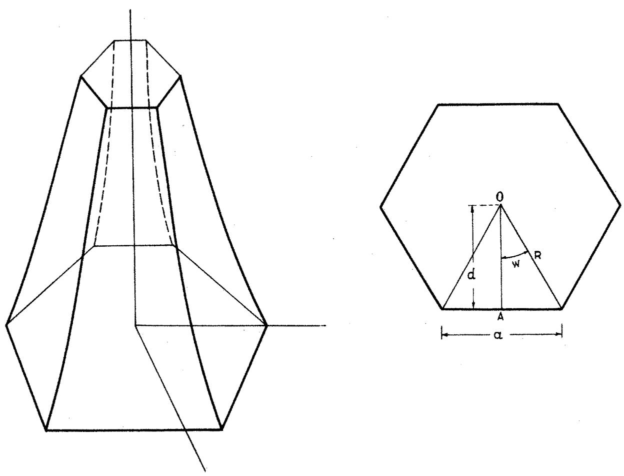


Figure 1

We will have for each section $w = \pi/N$, $a = 2R \sin w$, $d = R \cos w$. Let d = the density.

The moment of inertia of this body to the z-axis is given by

$$(1) \qquad I = D\iiint(x^2 + y^2)\,dx\,dy\,dz = D\int_0^2 dz\iint(x^2 + y^2)\,dx\,dy$$

and for the sake of symmetry, we will have

$$T = \iint_{N\text{-GON}} (x^2 + y^2)\,dx\,dy = 2N \iint_{\text{TRIANGLE } OAB} (x^2 + y^2)\,dx\,dy$$

But $y = x \tan w$, therefore

$$T = 2N \int_0^{R \cos w} dx \int_0^{n \tan w} (x^2 + y^2)\,dy = 2N \int_0^{R \cos w} (x^3 \tan w + \frac{x^3 \tan^3 w}{3})\,dx$$

$$= 2N \frac{R^4}{4} \sin w \cos w \,(\cos^2 w + \frac{1}{3} \sin^2 w)$$

and $$I = \frac{N}{2} D \sin w \cos w \,(\cos^2 w + \frac{1}{3} \sin^2 w) \int_0^h R^4\,dz$$

Given (3) $R^n + bz^m = c$, where b, c, m, and n are real numbers, we will have

$$I = \frac{N}{2} D \sin w \cos w \,(\cos^2 w + ½ \sin^2 w) \int_0^h (c - bz^m)^{\frac{4}{n}}\,dz$$

For $c \neq 0$, $b \neq 0$ and $\left|\frac{b}{c} h\right| < 1$, according to the Binomial theorem, we will have

(5) $$I = \frac{N}{2} D \sin w \cos w \,(\cos^2 w + \frac{1}{3} \sin^2 w)\, b^k \left[h\left(\frac{c}{b}\right)^k - k\left(\frac{c}{b}\right)^{k-1} \frac{h^{m+1}}{m+1} + \frac{k(k-1)}{1 \cdot 2} \left(\frac{c}{b}\right)^{k-2} \frac{h^{2m+1}}{2m+1} + \cdots\right.$$

When $n = 1, 2, 4$, then $k = 4, 2, 1$, and the above formula provides the inertia moment of some geometrical solids.

It is obvious that when $N \to \infty$, $\frac{N}{2} \sin w \cos w \,(\cos^2 w + ½ \sin^2 w)$ has $\frac{\pi}{2}$ for limit; actually, $N \sin w \cos w = \frac{N}{2} \sin 2w = \frac{N}{2} \sin \frac{2\pi}{N}$ and, according to Hospital's formula $\left[-\frac{2\pi}{N^2} \cos \frac{2\pi}{N}\right] : \left[-\frac{2}{N^2}\right]$ has π for limit when $N \to \infty$, and $\cos^2 w + \frac{1}{3} \sin^2 w$ has unit for limit.

Applications

A) Moment of inertia, about a line.

1) Regular prism and right circular cylinder to their axes.

In this case, relation (3) becomes $R = r$, $b = 0$, $c = r$, $n = 1$ and $m = 0$.

Hence, for a regular prism we have

$$I = \frac{DN}{2} \sin w \cdot \cos w(\cos^2 w + \frac{1}{3} \sin^2 w) r^4 \cdot h$$

and for a right circular

$$I = \frac{D \cdot \pi}{2} r^4 h.$$

2) A right pyramid and a right circular cone to their axes.

In this case, relation (3) becomes

$$\frac{z}{R} = \frac{h}{r} \text{ or } R = \frac{r}{h} z \text{, i.e. } b = \frac{-r}{h}, \; c = 0, \; m = n = 1.$$

Hence for a right pyramid we have

$$I = \frac{D \cdot N}{2} \sin w \cdot \cos w(\cos^2 w + \frac{1}{3} \sin^2 w)(\frac{r}{h})^4 \cdot \frac{h^5}{5}$$

$$\text{or } I = \frac{DN}{10} \sin w \cos w(\cos^2 w + \frac{1}{3} \sin^2 w) h \cdot r^4$$

and for a right circular cone

$$I = \frac{D \cdot \pi}{10} \cdot h \cdot r^4$$

3) Frustum of a regular pyramid and frustum of a right circular cone to their axes.

In this case, relation (3) becomes

$$R + \frac{r_2 - r_1}{h} \cdot z = r_2, \quad \text{i.e. } b = \frac{r_2 - r_1}{h}, \; c = r_2, \; \frac{c}{b} = \frac{r_2 h}{r_2 - r_1}$$

and $m = n = 1$

Hence for a frustum regular pyramid

$$I = \frac{D \cdot N}{2} \sin w \cdot \cos w \left[\cos^2 w + \frac{1}{2} \sin^2 w\right] \left[\frac{r_2 - r_1}{h}\right]^4 \left[\left[\frac{r_2 h}{r_2 - r_1}\right]^4 \cdot h - 4 \cdot \left[\frac{r_2 h}{r_2 - r_1}\right]^3 \cdot \frac{h^2}{2} + \frac{4 \cdot 3}{1 \cdot 2} \left[\frac{r_2 h}{r_2 - r_1}\right]^2 \cdot \frac{h^3}{3} - \frac{4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3} \frac{r_2 \cdot h}{r_2 - r_1} \cdot \frac{h^4}{4} + \frac{h^5}{5}\right]$$

$$= \frac{D \cdot N}{10} h \sin w \cdot \cos w(\cos^2 + \frac{1}{2} \sin^2 w)(r_2^4 + r_2^3 r_1 + r_2^2 \cdot r_1^2 + r_2 r_1^3 + r_1^4), \text{ or}$$

$$I = \frac{D\cdot N\cdot h}{10} \sin w\cdot\cos w\ (\cos^2 w + \frac{1}{3}\sin^2 w)\ \frac{r_2^5 - r_1^5}{r_2 - r_1}$$

and for frustum right circular cone

$$I = \frac{D\cdot\pi}{10}\, h\, \frac{r_2^5 - r_1^5}{r_2 - r_1}$$

4) Sphere and Hemisphere to their axes
In this case, the relation (3) becomes

$$R^2 + z^2 = r^2 \text{ i.e. } b = 1,\ c = r \text{ and } m = n = 2$$

Hence $I = \frac{D\cdot\pi}{2}\left[r^4 h - 2r^2\cdot\frac{h^3}{3} + \frac{2\cdot 1}{1\cdot 2}\cdot\frac{h^5}{5}\right]$ and, as for the hemisphere $h = r$, we have $I = \frac{4}{15} D\cdot\pi r^5$, and, for the sake of symmetry, for Sphere

$$I = \frac{8}{15} D\cdot\pi\cdot r^5$$

5) Ellipsoid of revolution to its axis.
In this case, relation (3) becomes

$$\frac{R^2}{A^2} + \frac{Z^2}{B^2} = 1 \quad \text{or} \quad R^2 + \frac{A^2}{B^2}\cdot Z^2 = A^2$$

i.e. $$b = \frac{A^2}{B^2},\ c = A^2,\ \frac{c}{b} = B^2 \text{ and } m = n = 2$$

Hence $$I = \frac{D\cdot\pi}{2}\cdot\frac{A^4}{B^4}\left[B^4 h - 2B^2\frac{h^3}{3} + \frac{2\cdot 1}{1\cdot 2}\,\frac{h^5}{5}\right]$$

But $h = B$ hence

$$I = \frac{8}{15} D\cdot\pi\cdot A^4\cdot B$$

6) Paraboloid of revolution to its axis
In this case relation (3) becomes

$$R^2 - pz = 0, \quad \text{i.e.} \quad b = -p,\ c = 0,\ m = 1 \text{ and } n = 2$$

$$I = \frac{D}{6}\pi p^2 h^3$$

7) Hyperboloid of revolution, to its axis

In this case, relation (3) becomes

$$\frac{R^2}{A^2} - \frac{Z^2}{B^2} = 1 \quad \text{or} \quad R^2 - \frac{A^2}{B^2} Z^2 = A^2$$

i.e. $b = -\frac{A^2}{B^2}$, $c = A^2$, $\frac{c}{b} = -B^2$ and $m = n = 2$.

Hence

$$I = D\frac{28}{15}\cdot \pi \cdot A^4 \cdot B \quad \text{if} \quad h = B.$$

B) A similar calculation should apply respectively to a plane and to a point.

Athens, Greece.

THE RULE OF DOUBLE FALSE

Harold E. Bowie

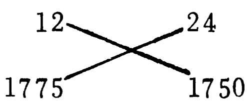
My students became interested as to why the Rule of Double False used by Robert Recorde and other early mathematicians gave them correct answers. They found that arithmetic explanations were difficult because of their tendency to think about problems in the language of algebra.

In our explanations which we developed by methods of the elementary algebra of today, we used a problem described by Vera Sanford in her book *A Short History of Mathematics*.*

"This particular problem and its solution are from Robert Recorde's *Ground of Artes* (*c*. 1542).

> One man said to another, I think you had this year two thousand Lambes: so had I said to the other; but what with paying the tythe of them, and then the several losses they are much abated: for at one time I lost half as many as I have now left, and at another time the third part of so many, and the third time $\frac{1}{4}$ so many. Now guesse you how many are left.

It was clear that after the tithe was deducted, 1800 lambs were left. If the man had had 12 at the end, he would have had 12 + 6 + 4 + 3 or 25 at the beginning. This is 1775 too few. On the other hand, if he had 24 at the end, he would have had 24 + 12 + 8 + 6 or 50 at the beginning which again is 1750 too few. The guesses and errors are then written down and the cross-products are found by multiplying along the guide lines thus:



Then the difference of these products is divided by the difference of the guesses and the quotient is the required number. In this case it is

$$\frac{42600 - 21000}{1775 - 1750} = \frac{21600}{25} = 864."$$

Our explanation follows:
Let x = the number left.
Then

(1) $$x + \frac{x}{2} + \frac{x}{3} + \frac{x}{4} = 1800$$

(2) $$12 + 6 + 4 + 3 = 25$$

(3) $$24 + 12 + 8 + 6 = 50$$

Subtracting (2) from (1) we have

(4) $$(x - 12)(1 + 1/2 + 1/3 + 1/4) = 1775$$

Subtracting (3) from (1) we have

(5) $$(x - 24)(1 + 1/2 + 1/3 + 1/4) = 1750$$

Dividing (4) by (5) we obtain (as $x \neq 24$)

$$\frac{x - 12}{x - 24} = \frac{1775}{1750}$$

Then

$$1750x - 12.1750 = 1775x - 24.1775$$

$$x(1775 - 1750) = 24.1775 - 12.1750$$

$$x = \frac{24.1775 - 12.1750}{1775 - 1750}$$

Which justifies the Rule of Double False as the method of explanation is evidently perfectly general.

In the same volume by Miss Sanford we find a problem illustrating the Rule of False Position.

"To solve the equation $x + 1/7x = 19$, the unknown number x is assumed to be 7. Then the sum of the number and its seventh part will be 8, and the number of the equation is the same multiple of 7 that 19 is of the guessed number 8."

We give an explanation of this simpler rule for the general case.

(1) $$x + 1/7\, x = 19$$

Let $x = K$ be our guess
Then

(2) $$K + 1/7\, K = C$$

From (1)

$$x(1 + 1/7) = 19$$

From (2)

$$K(1 + 1/7) = C$$

Dividing with $K \neq 0$ we obtain

$$\frac{x}{K} = \frac{19}{C}$$

*See *A Short History of Mathematics* by Vera Sanford, Boston, (1930), page 160.

American International College

THEOREM: Every integer greater than 2 is a member of a Pythagorean triplet; every integer of the form $(2n + 1)$ is a member of a primitive Pythagorean triplet.

(Note: For convenience, when the word, "integer", is used in the following, it refers to an integer greater than 2.)

The integers, a, b, c, form a Pythagorean triplet when they satisfy the equation:

$$a^2 + b^2 = c^2$$

When they have no common divisor, the triplet is primitive.

The equation

$$a^2 + 2ab + b^2 = (a + b)^2$$

can be reduced to a Pythagorean triplet whenever $(2ab + b^2)$ is a rational square. Members of such a triplet are a, $(2ab + b^2)^{1/2}$, $(a + b)$.

Every integer is of the form $(2n + 1)$ or $(2n + 2)$ and has a square of similar form, with the square of $(2n + 2)$ being also of the form $(4n + 4)$. As a consequence, the square of any integer is of the form $(2ab + b^2)$ and the integer is, therefore, a member of a Pythagorean triplet, consisting of:

$$\tfrac{1}{2}(b^2 - 1),\ b,\ \tfrac{1}{2}(b^2 - 1) + 1, \text{ where } b \text{ is of form } (2n + 1)$$

$\frac{1}{4}(b^2 - 4)$, b, $\frac{1}{4}(b^2 - 4) + 2$, where b is cf form $(2n + 2)$

If b is of the form $(2n + 1)$, the triplet can be expressed in terms of a as follows: a, $(2a + 1)^{1/2}$, $(a + 1)$. As a, $(2a + 1)$, and $(a + 1)$ are prime to one another, $(2a + 1)^{1/2}$ is prime to a and to $(a + 1)$.

James E. Foster

MATHEMATICS CONTEST

Sponsored by the Euclid Circle of the College of Saint Rose.

Letters explaining the purpose of the contest (that is, to engender a keener interest in mathematics), registration blanks, and rules for the contest were sent to approximately 65 schools in the capitol area near Albany, N.Y. They were sent to the principal with attention line addressed to the chairman of the mathematics department. The registration blank was sent back within a month to give us an idea of the number of prospective entries.

According to the rules, any original geometric design would be accepted. 9″ x 12″ paper was the largest; color, pen, pencil, paint were left to the students discretion.

Prizes were: 1st \$10; 2nd \$5 and the judges named 10 honorable mentions - these were given cards naming their distinction, contest name and date, and their name.

Our response was approximately 180 entries from 16 schools.

The judges were: Chairman of the Board of Regents' Department of Mathematics of N.Y. State; Chairman of the Mathematics Department at the N.Y. State College of Teachers in Albany; Professor of Art at the College of Saint Rose.

The first prize winning entry was in perspective and showed a complete range of mathematics. That is, the many "figures", the circle, point, straight line, curve, triangle, sphere, pyramid, cube, etc., were blended into the design. It was shaded in black, gray, and white.

Second prize appeared to be rather modern. There was a pillar-like design to one side and beside it a sphere with a pyramid cut in. This was in black, red, and white.

These two designs were rather intricate as well as neat. Others ranged from a curtain design pattern to perspective modern drawings.

Reported by Charlene Lysick, Albany, N.Y.

CURRENT PAPERS AND BOOKS

Edited by

H. V. Craig

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

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

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

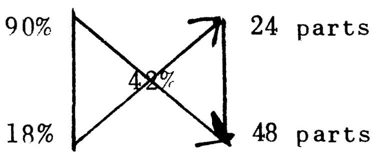
Mr. B. E. Mitchell
Alabama Polytechnic Institute

Dear Sir: I am interested in Mathematics as a recreation, and I came across your article in the Jan.-Feb. number of "Mathematics Magazine."

Pharmacists are frequently compelled to solve problems of the type you present in filling the car radiator.

An old pharmacist showed me a method which, while largely "mechanical" in nature, expedites and simplifies these problems. I have since found the method described in some books on pharmacy under the title "ALLEGATION."

The problem you present would be solved by "ALLEGATION" in this manner.



The higher percent is placed above the lower. The desired percent is placed in the middle. Subtractions are made as indicated by the arrows.

The final product to obtain a mixture 42% alcohol is thus

24 parts 90%

48 parts 18%

or a ratio of 24:48 = 1:2.

Since the total is 3 parts = 21 qts., 1/3 must be 90% = 7 qts.—must be drained. 2/3 must be 18% = 14 qts.

Extending the application of the principal:

Example: In what proportion must 8%, 10%, 16%, 18% solutions be mixed to give a 14% solution?

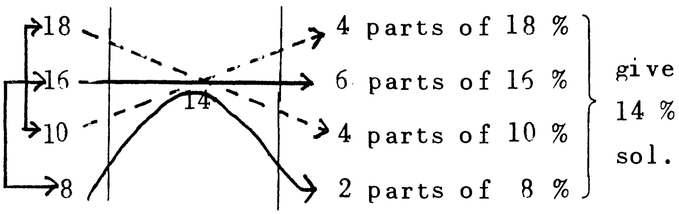
Procedure:

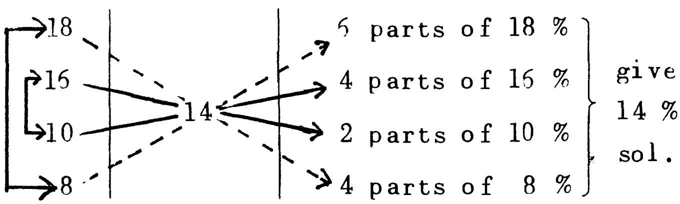
(1) Write the percentages in diminishing percent in the left column.

(2) Write the final (desired) percent in the middle.

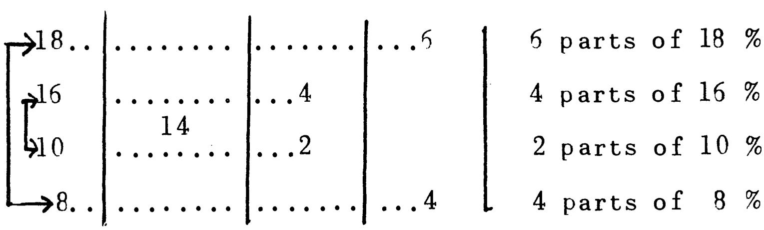
(3) Connect each couptet consisting of a higher and lower percent and perform the subtraction as in the previous problem.

(Note: The multiplicity of original solutions permits several procedures each yielding the same final result. Below are illustrations of *two* possible procedures.)





Another way in which this may be set down is as below:



In regards to your description of the "French Method of Long Division" I am unable to follow you in your description. Will you favor me with a more detailed description. I "bog down" at the following indicated points:

$$\begin{array}{r} 107 \\ 236\,/\,\overline{25456} \\ 1856 \\ 204 \end{array}$$

You say:

"1x6 is 6 and 8 makes 14..."

Questions

Why add 8?
Just where does the 14 come in?

" ... Write 8 and carry one ..." { Write 8 where ?
Carry the 1 where ?

(From here on it just seems you grab figures out of thin air and I am completely lost in a Minoan Mathematical Maze.)

I would deeply appreciate clarification. You will forgive my native ineptitude for figures.

Thanking you for any courtesies extended to me, I am,

Sincerely yours,

Arithmetic vs Algebra

The Mitchell article on page 153 of the Jan.-Feb. number speaks of the subtraction of $1000\,d = 123.\dot{4}\dot{5}$ from $10\,000\,d = 12345.\dot{4}\dot{5}$ and subsequent division of the resulting *equation* by 9000 as an "arithmetic method", in contrast apparently to an "algebraic method".

This raises the question: how does the arithmetic method of attacking a problem differ from that of algebra? Is it not essentially as follows:

1. Arithmetic takes the given numbers and operates with them, obtaining interpretable sums, products, etc., until the desired number is reached.

2. Algebra states the relations between the numbers with which the problem is concerned (supplying letters for numbers not given) and operates on the resulting equations in such a way that some equation is found to give the desired unknown explicitly.

May we illustrate the two methods by this simple problem: the sum of two numbers is 18, and 2 times one of them is the other.

1. By arithmetic. The 2 times and the one make 3. So 18 is 3 times the one. Then $18 \div 3 = 6$, and $2 \times 6 = 12$. So the numbers are 6 and 12.

2. By algebra. The problem says that $18 = a + b$ (a and b are the two numbers), and that $a = 2b$. By substitution, $18 = 2b + b$, so $3b = 18$, $b = 6$, and $a = 12$.

We must remark that those who teach algebra to beginners usually obscure the strictly algebraic method by doing such a problem partly by arithmetic, saying at once that $2b + b = 18$, which is a relation not explicitly given in the wording. Such obscuration commonly results from the (regrettable!) attempt to "solve by *one* unknown" a problem in which there is more than one unknown number mentioned or implied in the problem.

William R. Ransom, Tufts College

Analytic Geometry and Calculus. Frederick H. Miller. xii plus 658 pp. $5.00. John Wiley & Sons, Inc. New York, 1949.

This text is suitable for a two or three semester course. It correlates the subjects of plane and solid analytic geometry, differential calculus, and integral calculus, and presents the fundamentals of calculus early enough to be of use in other subjects; this early presentation is of particular advantage in engineering and other science courses.

Among the outstanding features of the book the reviewer notes the more than usually comprehensive study of maxima and minima; the careful motivation of the study of conic sections with emphasis on their definition in terms of eccentricity; detailed consideration of graphs of equations in polar coordinates; special treatment of the limit of $(\sin x)/x$ as x approaches zero and the limit which serves as Napierian base; a treatment of integration in accord with the analytical method of more advanced analysis; an appendix including useful symbols and tables, essential definitions and facts of more elementary mathematics; 3025 exercises together with answers to the odd-numbered exercises; chapter summaries useful both to the teacher as an outline for his course and to the student as a basis for review.

The author presents specific topics much as he has presented them in his other textbooks. The treatment is sufficiently rigorous to serve the student whose primary interest is in pure mathematics and at the same time it meets the need of the student who wishes to use mathematics in science or engineering. The teacher will welcome the clear exposition and will find among the numerous problems many suitable for the average student and others which should stimulate the interest and effort of the abler members of his class.

Helen G. Russell

Vorlesungen uber Differential-und Integralrechnung, Zweiter Band. By A. Ostrowski, Birkhauser, Basel, 1951, 482 pp. 67 Swiss fr.

This volume is a rigorous, detailed, and lucid treatment of important topics in the field of the differential calculus of several variables. Among the topics treated we find: infinite sets, functions on sets, infinite sequences and series, differentiation of functions of several variables, implicit function theory, numerical approximation methods, vector algebra and differential calculus (in Gibb's notation) and differential geometry of curves and surfaces. This easy reading and elegant work should be a valuable reference volume for calculus instructors.

Homer V. Craig

PROBLEMS AND QUESTIONS

Edited by

C. W. Trigg, Los Angeles City College

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

Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and twice the size desired for reproduction. Readers are invited to offer heuristic discussions in addition to formal solutions.

Send all communications for this department to *R.E. Horton, Los Angeles City College, 855 N. Vermont Ave., Los Angeles 29, California.*

PROPOSALS

168. *Proposed by Frank C. Gentry, University of New Mexico.*

If P is any point in the plane of a triangle ABC and if P', P'' and P''' are the harmonic conjugates of P with respect to ABC, then the perpendiculars let fall from P', P'' and P''' on BC, CA and AB respectively are concurrent when P is at the circumcenter O, the centroid G, or the symmedian point K.

169. *Proposed by Norman Anning, University of Michigan.*

One student tries to adjust k to make $4x + y = 24$ a tangent to $y = kx^2$ and gets one answer. Another tries to make the same line a tangent to $x^2 = ky$ and gets two answers. Who is right and what is wrong?

170. *Proposed by R. E. Horton, Lackland Air Force Base, Texas.*

A fighter plane flies under the following conditions:

Fuel Consumption: 100 gallons per hour.

Fuel Capacity: Main tank holds 250 gallons.
Two auxiliary wing tanks hold 100 gallons each.

Cruising air speed: Initial air speed is 400 miles per hour.
Air speed increases .125 mph per gallon of fuel consumed.
Air speed increases 5% when wing tanks are dropped simultaneously.

1) What is the effective range of the plane if a 20% reserve of fuel must be kept and the wing tanks are dropped simultaneously when both are empty?
2) What is the effective radius of action North with a wind of 30 mph from the South, assuming a fuel reserve of 20% and wing tanks dropped simultaneously when both are empty? (Radius of action is the distance a plane can fly and still return to its base).
3) Under what wind conditions will the time out exactly equal the time back on a radius of action problem for this plane?

171. *Proposed by N. A. Court, University of Oklahoma.*

The four spheres having for great circles the polar circles of an orthocentric tetrahedron (T) have for orthogonal center the orthocenter of the medial tetrahedron of (T).

172. *Proposed by H. E. Fettis, Wright-Patterson Air Force Base, Dayton, Ohio.*

Evaluate: $$\lim_{a\to 0}\,[\pi^2\csc^2\pi a - 1/a^2].$$

173. *Proposed by F. J. Duarte, Caracas, Venezuela.*

Prove that (1) the equation $x^3 - 6abx - 3ab(a + b) = 0$ has no solution in integers; (2) the equation $2x^3 - 6abx - 3ab(a + b) = 0$ has an infinite number of integer solutions.

174. *Proposed by J. E. Foster, Evanston, Illinois.*

The equation, $2^p + 1 = 3p'$, where p and p' are odd primes, has been empirically confirmed for values of p through 17, as follows:

p	3	5	7	11	13	17
p'	3	11	43	683	2731	43691

Are there solutions of the equation for $p > 17$?

SOLUTIONS

Late Solutions

139. *M. S. Klamkin, Polytechnic Institute of Brooklyn, N. Y.*

145. *Robert Bonic, University of Chicago, Illinois.*

120° Triangles with Integer Sides

147. [November 1952] *Proposed by Leon Bankoff, Los Angeles, California.*

In a triangle with integer sides, the side opposite the 120° angle is 1729. Find all possible values of the pair of other sides.

Solution by E. P. Starke, Rutgers University. By the law of cosines the side opposite the 120° angle is $(a^2 + ab + b^2)^{1/2} = 1729$, where a, b are the other sides. We have

$$a^2 + ab + b^2 = (1729)^2 = 7^2 \cdot 13^2 \cdot 19^2.$$

Now numbers of the form $x^2 + xy + y^2$ are multiplicative, i.e. the product of two is a third of the same form. The relation is

$$(a^2 + ab + b^2)(c^2 + cd + d^2) = r^2 + rs + s^2,$$

$$r = ac + bd + ad, \qquad s = bc - ad, \qquad (A)$$

and a second choice of r, s which results from interchanging a, b and/or interchanging c, d, but so that $bc - ad \geqq 0$.

For convenience write $a^2 + ab + b^2 = (a, b)$. Then $7 = (2, 1)$, $13 = (3,1)$, $19 = (3,2)$. Using (A) we get in succession:

$$7^2 = (7,0) = (5,3), \quad 13^2 = (13,0) = (8,7), \quad 19^2 = (19,0) = (16,5),$$

$$7^2 \cdot 13^2 = (91,0) = (56,49), \quad (65,39), \quad (85,11) = (80,19), \text{ whence}$$

$$7^2 \cdot 13^2 \cdot 19^2 = (1729,0), \quad (1064,931) = (1235,741) = (1615,209)$$

$$= (1520,361) = (1456,455) = (1309,651) = (1421,504) = (1144,845)$$

$$= (1560,299) = (1305,656) = (1591,249) = (1679,96) = (1185,799).$$

Conversely it can be shown that if a product $M \cdot N = x^2 + xy + y^2$, then there exist non-negative integers a, b, c, d such that $M = a^2 + ab + b^2$, $N = c^2 + cd + d^2$, $x = ac + bd + ad$, $y = bc - ad \geqq 0$. Thus the thirteen triangles found above, omitting $(1729,0)$, are the only ones possible.

Also solved by *Ward Bouwsma, Calvin College, Grand Rapids, Michigan; Sam Kravitz, East Cleveland, Ohio; and the proposer*. Partially solved by *G. G. Becknell, University of Tampa, Florida; M. S. Klamkin, Polytechnic Institute of Brooklyn; and Prasert Na Nagara, College of Agriculture, Thailand.*

For other discussions dealing with triangles having integer sides and one angle of 120°, see *L'Arith. de S. Stevin*, par A. Girard, Leide, (1625), 676; *Les Oeuvres Math. de S. Stevin,* par A. Girard, (1634), 169; H. Böttcher, *Unterrichtsblätter für Math. u. Naturwiss.*, **19**, 132-3, (1913); *American Mathematical Monthly*, **44**, 113, (1937); **48**, 707, (1941).

A "Lewis Carroll" Pillow Problem

148. [Nov. 1952] *Proposed by D. L. MacKay, Manchester Depot, Vt.*

Upon the sides of triangle ABC the squares $ABDE$, $BCFG$, $ACHL$ are constructed exterior to the triangle. Construct triangle ABC given the A', B', C' which are the intersections of DE and HL, ED and FG, GF and LH, respectively.

I. *Solution by Prasert Na Nagara, College of Agriculture, Thailand.* Produce $A'B'$ to S making $B'S = A'B'$ and $A'C'$ to U making $C'U = A'C'$. Construct the circle (T) through S, B', C' and the circle (V) through B', C', U. To (T) and (V) at B' and C' respectively, construct tangents which will intersect at O. Let P, Q, R be the feet of the perpendiculars from O to $B'C'$, $C'A'$, $A'B'$. Construct d, the fourth proportional to $(OP + B'C')$, $B'C'$, and OP. At distance d from $B'C'$ draw a parallel meeting $B'O$ and $C'O$ in B and C. Through C draw a parallel to $A'C'$ meeting OA' in A.

Proof. Let the feet of the perpendiculars from A, B to $A'B'$ be E, D; of those from B, C to $B'C'$ be G, F; and of those from C, A to $C'A'$ be H, L.

$$\frac{OP}{OR} = \frac{\sin OB'P}{\sin OB'R} = \frac{\sin \frac{1}{2} C'TB'}{\sin \frac{1}{2} STB'} = \frac{\frac{1}{2} a'}{\frac{1}{2} c'} = \frac{a'}{c'}.$$

Similarly, $OP/OQ = a'/b'$.

Now $d = a'(OP)/(OP + a')$, so

$$\frac{BC}{a'} = \frac{OB}{OB'} = \frac{OB' - B'B}{OB'} = 1 - \frac{B'B}{OB'} = 1 - \frac{d}{OP} = 1 - \frac{a'}{OP + a'}, = \frac{OP}{OP + a'}\ \frac{d}{a'}.$$

Hence $BC = d$.

$$a'/(OP + a') = 1/(1 + OP/a') = 1/(1 + OR/c') = c'/(OR + c').$$

$$BD/OR = B'B/B'O = d/OP = c'/(OR + c').$$

$$AB/c' = OB/OB' = 1 - d/OP = 1 - c'/(OR + c') = OR/(OR + c').$$

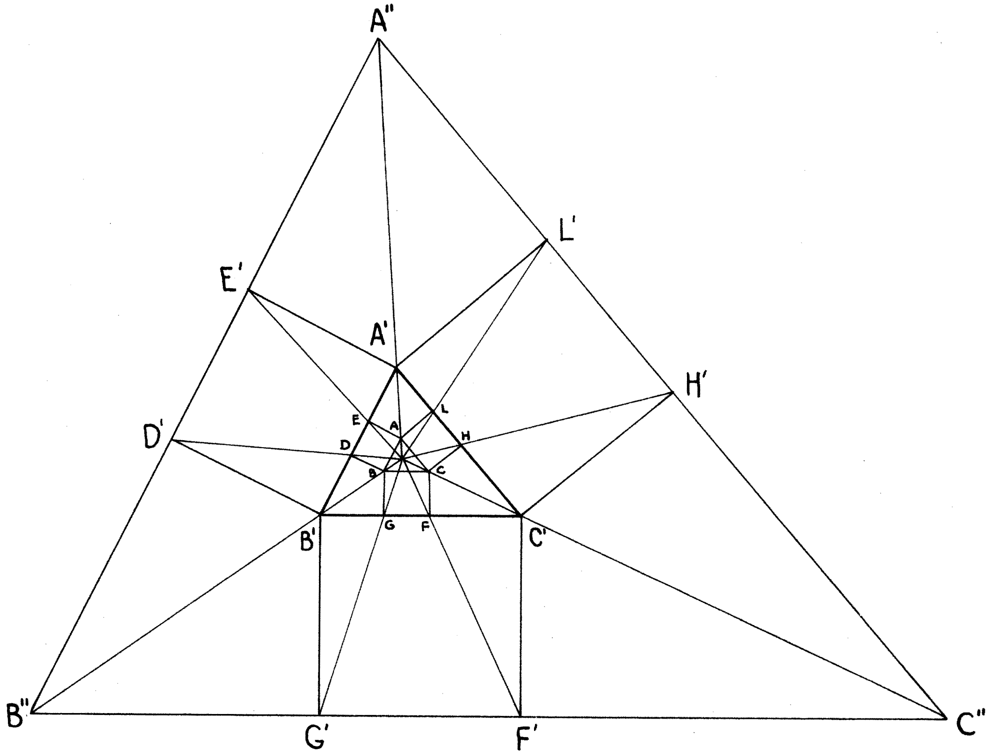
Hence $AB = c'(OR)/(OR + c') = BD$.

$OC/OC' = Bc/a' = OB/OB' = AB/c' = OA/OA'$, so CA is parallel to b'.

It follows that $AC = AL$.

II. *Solution by John Jones, Jr., University of North Carolina.* The triangles ABC and $A'B'C'$ are homothetic, so AA', BB', CC' intersect in the homothetic center Q. Now the distance from A to $A'C'$ is equal to b and the distance from A to $A'B'$ is equal to c. Hence a point on AA' may be constructed by drawing a line parallel to $A'B'$ at a distance $A'B'$ from it, and a line parallel to $A'C'$ at a distance $A'C'$ from it. Thus AA', and in like manner, BB' may be determined. So Q, the intersection of the symmedians AA' and BB', is the symmedian point of the triangles. Inscribe a square in triangle $A'QB'$ with one side lying on $A'B'$. The other two corners of the square will determine vertices A and B, on $A'Q$ and QB', respectively. Through A and B draw parallels to $A'C'$ and $B'C'$, respectively. These parallels will intersect in C.

III. *Solution by Ward Bouwsma, Calvin College, Grand Rapids, Michigan.* Upon the sides of triangle $A'B'C'$ construct squares $A'B'D'E'$, $B'C'F'G'$, $A'C'H'L'$. Then the intersections of $D'E'$ and $H'L'$, $E'D'$ and $F'G'$, $G'F'$ and $L'H'$ determine vertices A', B', C' of a triangle similar to $A''B''C''$ and hence to ABC. From A' lay off on $A'B'$ the fourth proportional, $A'E$, to $A''B''$, $A''E'$, and $A'B'$. At E, draw a perpendicular into triangle $A'B'C'$. On this perpendicular, from E lay off EA, the fourth proportional to $A''B''$, $E'A'$ and $A'B'$. In like manner, B and C can be located.



IV. *Solution by Leon Bankoff, Los Angeles, California.* Draw triangle $A''B''C''$ as in **III**. Draw the symmedians $A''A'$, $B''B'$, $C''C'$ which intersect in Q. Connect Q to D', E', L', H', F', G'. These connectors cut the sides of triangle $A'B'C'$ in D, E, L, H, F, G respectively. perpendiculars erected at these points to the sides of $A'B'C'$ intersect, by twos on the symmedians, in A, B, C. The validity of this construction follows from well-known properties of homothetic figures.

The problem itself has always aroused interest. Five solvers successfully attacked it for the right triangle in Question 1615 of the *Ladies Diary* of 1838. Charles L. Dodgson gives a geometric and a trigonometric solution as Problem 57 in *Pillow Problems*, 3rd Edition, Macmillan, London, 1894. Most modern geometries deal with it or its essentials, for example, Altshiller-Court, College Geometry, Johnson (1925), page 232.

Also solved by *Prasert Na Nagara, College of Agriculture, Thailand* (a second solution); *Charles Salkind, Polytechnic Institute of Brooklyn; A. Sisk, Maryville College; and the proposer* who remarked that the problem is based on the theorem regarding the concurrency of AA', BB', CC' given by E. W. Grebe in 1847. It is the basis of the German claim that the symmedian point should be named Grebe's point.

A Gambling Game

149. [November 1952] *Proposed by L. C. May, John Muir College, Pasadena, California.*

A has a gambling device so arranged that he always wins 3 times out of every sequence of 5 plays. B suspects that the game is "fixed" and demands that A always wager half of his resources, which B will match with an equal amount. B's funds are assumed to be unlimited.

(1) Show that A always loses over a sequence, regardless of the order of his 3 winning and 2 losing plays.

(2) Determine the best fixed percentage of his resources for A to wager if he wins 3 out of every 5 plays.

(3) If A must always wager 1/2 of his resources, determine to the nearest integer the number of times he must win out of a sequence of one hundred plays in order to break even.

Solution by H. M. Gehman, University of Buffalo, New York.

(1) If A always wagers half his resources, when he wins his resources are multiplied by 3/2 and when he loses they are multiplied by 1/2. Thus a sequence of 3 winning and 2 losing plays *in any order* multiplies A's resources by $(3/2)^3(1/2)^2 = 27/32$, which is to his disadvantage.

(2) If A always wagers x of his resources, when he wins his resources are multiplied by $(1 + x)$ and when he loses they are multiplied by $(1 - x)$. Thus a sequence of 3 winning and 2 losing plays multiplies A's resources by $(1 + x)^3(1 - x)^2$. This function has a maximum when $x = 1/5$. Hence if A always wagers 20% of his resources, each sequence of 5 plays multiplies his resources by $3456/3125 = 1.10592$. This is A's most advantageous choice of the percent he should wager.

(3) If A wagers half of his resources and wins t times in 100 plays, in order to break even t must satisfy the equation: $(3/2)^t(1/2)^{100-t} = 1$ or $3^t = 2^{100}$. The nearest integral root of this equation is 63. But if A wins only 63 times his resources will be decreased; if he wins 64 times they will be increased.

Also solved by *Robert Bonic, University of Chicago; J. M. Howell, Los Angeles City College; Sam Kravitz, East Cleveland, Ohio; Prasert Na Nagara, College of Agriculture, Thailand; and the proposer.*

A Parallelogram Associated with the Triangle

150. [November 1952] *Proposed by P. D. Thomas, U. S. Coast and Geodetic Survey.*

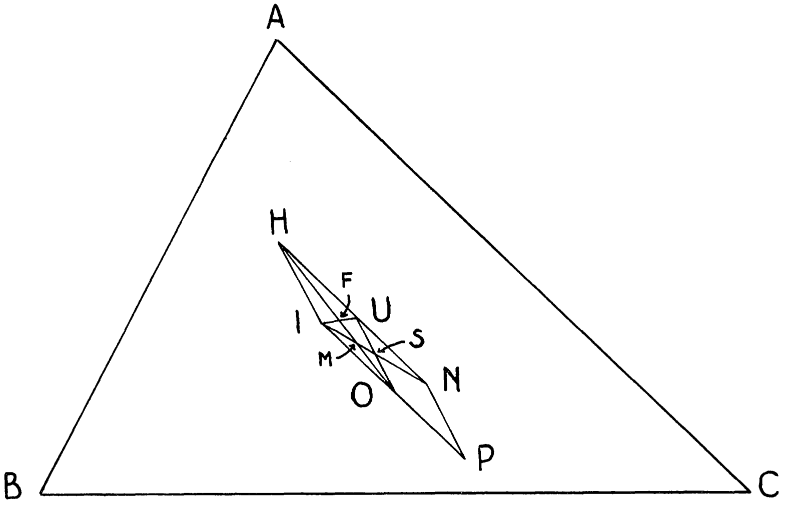
The lines joining the vertices of a triangle to the internal points of contact of the escribed circles meet in a point. The perpendiculars upon the sides of the triangle from the excenters meet in a point. Show that these two points together with the orthocenter and the incenter of the triangle are the vertices of a parallelogram.

Solution by Leon Bankoff, Los Angeles, California. Using the conventional notation, H for orthocenter, I for incenter and O for circumcenter, let N be the Nagel point (the intersection of the lines joining the vertices to the internal points of contact of the escribed circles), and call the intersection of the perpendiculars upon the sides of the triangle from the excenters, P. Denote the midpoint of HN by U, the centroid by M, the center of the nine-point circle by F, and the center of the Spieker circle by S.

IO is parallel to HN, and $IO = HN/2$ (Johnson, *Modern Geometry*, (1929), page 226). Now O is the midpoint of IP (Altshiller-Court, *College Geometry*, (1950), page 105). Hence $HN = IP$, and $HIPN$ is a parallelogram.

Furthermore, S is the midpoint of IN (Johnson, page 226), so IN, HP and OU bisect each other in S. Also, HO, the Euler line, is bisected by F, which is also the midpoint of IU. Now M trisects HO and hence OF and IN, so M is the centroid of IOU. Again, HN is a diameter of the Fuhrmann circle (Johnson, page 228), so U is the circumcenter of the Fuhrmann triangle.

Also solved by *the proposer.*



The General Term of a Sequence

151. [November 1952] *Proposed by Dewey Duncan, East Los Angeles Junior College.*

In a recent text on *Backgrounds for Secondary Mathematics Teachers* the following statement appears: "Consider the sequence 1, 2, 3, 2 1/2, 2, 1 1/2, 1 3/4, 2, 2 1/4, 2 1/8, $\cdots$. It has the limit 2, as can be shown by locating the values on a number scale. There is no single formula for this sequence." Refute this last assertion by example.

I. *Discussion by M. S. Klamkin, Polytechnic Institute of Brooklyn.* This problem illustrates an error found in many elementary texts dealing with sequences and in many Civil Service examinations and intelligence tests. One cannot say anything about the next term of a sequence nor about its limit, if only a finite number of terms is given. The sequence is not uniquely determined unless the law of formation is stated or a general term is specified. Thus there is an infinity of formulas applicable to the given sequence. Furthermore, the limit is not necessarily 2, it can be any arbitrary number. Also, the series may diverge or even oscillate, depending upon the law of formation.

The fundamental principle involved is clearly shown by the following more general problem. Let us fit an n-th term formula to a sequence of which any r terms are given. Let $a_{n_1}, a_{n_2}, \cdots, a_{n_r}$ be the $n_1, n_2, \cdots, n_r$-th given terms of the sequence. A general formula which fits this sequence is

$$a_n = \sum_{\substack{\text{cyclic} \\ 1,\ 2,\ \cdots\ .\ r}} \frac{a_{n_1}\phi(n_1)}{\phi(n)} \cdot \frac{(n-n_2)(n-n_3)\cdots(n-n_r)}{(n_1-n_2)(n_1-n_3)\cdots(n_1-n_r)},$$

where $\phi(n_k) \neq 0$, $k = 1, 2, \cdots, r$, and otherwise $\phi(n)$ is arbitrary.

If we take $\phi(n)$ of order n^{r-1} then a_n approaches a limit. If we take $\phi(n)$ of order less than n^{r-1} then a_n approaches ∞. Finally, if we take $\phi(n)$ to be oscillatory, then a_n oscillates, e.g., if $\phi(n) = (-1)^n$.

Editorial Note: The author of the text in question has written that he used "single" as synonymous to "simple" in this situation. Those who submitted formulas accepted the limit 2 as well as the most obvious law of formation. Almost every form received was different, however upon simplification they fell into typical patterns which follow.

II. *Solution by Harry M. Gehman, University of Buffalo, New York.* We shall use a heuristic method to find a formula for a_n, the n-th term of the given sequence (A).

By subtracting 2, the limit of (A), from each term, we can see how the terms of (A) oscillate about the limit. This gives us the sequence:

$$-1,\ 0,\ 1,\ 1/2,\ 0,\ -1/2,\ -1/4,\ 0,\ 1/4,\ 1/8,\ \cdots. \qquad (B)$$

The terms of (B) may be grouped into sets of three with a common denominator 2^t, where $t = [(n - 1)/3]$ and $[x]$ denotes the largest integer $\leqq x$. Multiplying each term of (B) by 2^t results in the sequence:

$$-1,\ 0,\ 1,\ 1,\ 0,\ -1,\ -1,\ 0,\ 1,\ 1,\ \cdots. \qquad (C)$$

The periodicity of (C) suggests some type of sine curve. A little experimentation shows that the terms of (C) are given by the formula: $c_n = (2/\sqrt{3})\ \sin(n - 2)\pi/3$.

Taking into account the operations which led to (C), we find that each term of sequence (A) is given by the formula:

$a_n = 2 + \dfrac{\sin(n - 2)\pi/3}{2^u\sqrt{3}}$, where $u = [(n - 4)/3]$ and $[x]$ denotes the largest integer $\leqq x$. This may also be written in the form

$$2 - (1/2)^{[(n-1)/3]}(2/\sqrt{3})\sin(n + 1)\pi/3.$$

Others who submitted results essentially in this form are *H. W. Gould, Portsmouth, Va.; J. D. E. Konhauser, State College, Pa.; William Leong, Student, University of California at Berkeley; H. C. Parrish, North Texas State College; and the proposer.*

III. Form essentially submitted by *Arthur Gregory, Albuquerque, New Mexico; Prasert Na Nagara, College of Agriculture, Thailand; and Mason Phelps, Student, Harvard University.*

$$a_n = \sum_{k=1}^{n} \left(-\frac{1}{2}\right)^{[(k-1)/3]}$$

IV. Second form submitted by *H. W. Gould, Portsmouth, Va.*

$$a_n = 2 - \{(-1)^{[(n+1)/3]} + (-1)^{[n/3]}\}/2^{[(n+2)/3]}.$$

V. Form essentially submitted by *Fred Marer, Los Angeles City College; Charles Salkind, Polytechnic Institute of Brooklyn;* and *L. A. Ringenberg, Eastern Illinois State College.*

$$a_n = 2 + (n - 2 - 3[n/3])(-1/2)^{[n/3]}.$$

VI. Second and third forms submitted by *William Leong, Student, University of California at Berkeley.*

$$a_n = 2 + \{3[(n - 1)/3] - n + 2\}(-1)^r \left(\frac{1}{2}\right)^{[(n-1)/3]}.$$

$$a_n = 2 - \frac{2}{\sqrt{3}}\{\sin \pi (n + 1)/3\} \left(\frac{1}{2}\right)^m, \text{ where}$$

$$m = \frac{1}{3} \{n - 2 - \frac{2}{\sqrt{3}} (-1)^n \sin \pi(n + 1)/3\}.$$

The last formula has the advantage that it contains no symbols that would be mysterious to one who has been exposed to just high school mathematics.

VII. Form submitted by *H. F. Fehr, Teachers College, Columbia University, N.Y.*

$$a_n = 2 - \{\frac{2}{\sqrt{3}} \sin (n + 1) \pi/3\} \left(\frac{1}{2}\right)^p, \text{ where}$$

$$p = \{2(n - 1)\sin 2\pi n/3 + (n - 3)\tan 2(n - 1)\pi/3\}/3\sqrt{3}.$$

VIII. *Solution by Vern Hoggatt, Oregon State College* and *E. G. Goman, College of Puget Sound.*

$$a_n = \left\{2 - \left(-\frac{1}{2}\right)^{\frac{n-1}{3}}\right\} \sum_{k=0}^{n} \delta_n^{3k+1} + 2 \sum_{k=0}^{n} \delta_n^{3k+2} + \left\{2 + \left(-\frac{1}{2}\right)^{\frac{n-3}{3}}\right\} \sum_{k=0}^{n} \delta_n^{3k},$$

where the δ_j^i are Kronecker Deltas and are equal to 1 or 0 according as $i = j$ or $i \neq j$.

A Faulty Evaluation

152. [November 1952] *Proposed by Malcolm Robertson, Rutgers University.*

In finding the area, $A = \pi ab$, of the ellipse $\rho^2 = a^2b^2/(a^2\sin^2\theta + b^2\cos^2\theta)$ a student gets an incorrect answer as follows:

$$A = ½ \int_0^{2\pi} \rho^2 d\theta = \frac{ab^2}{2} \int_0^{2\pi} \frac{a \sec^2\theta \, d\theta}{b^2 + (a \tan \theta)^2}$$

$$= \frac{ab^2}{2} \left\{\frac{1}{b} \arctan \left(\frac{a}{b} \tan \theta\right)\right\}_0^{2\pi} = \frac{ab}{2} (\arctan 0 - \arctan 0) = 0.$$

Detect and explain the source of error.

Solution by Charles Salkind, Polytechnic Institute of Brooklyn, N.Y. The student made the not uncommon error of integrating over essential discontinuities at $\pi/2$ and $3\pi/2$. All would have been well if he had written

$$A = \tfrac{1}{2}\int_0^{2\pi} \rho^2\, d\theta = \frac{4\,ab^2}{2}\int_0^{\pi/2} \frac{a\sec^2\theta\, d\theta}{b^2 + (a\tan\theta)^2}$$

$$2ab^2\left\{\frac{1}{b}\arctan\left[\frac{a}{b}\tan\theta\right]\right\}_0^{\pi/2} = \frac{2ab^2}{b}(\pi/2 - 0) = \pi ab.$$

Also solved by *Louis Berkofsky, Lexington, Mass.; Arthur Gregory, Albuquerque, N. Mex.; M. S. Klamkin, Polytechnic Institute of Brooklyn, N.Y.; Prasert Na Nagara, College of Agriculture, Thailand.*

QUICKIES

From time to time this department will publish problems which may be solved by laborious methods, but which with the proper insight may be disposed of with dispatch. Readers are urged to submit their favorite problems of this type, together with the elegant solution and the source, if known.

Q 28. [March 1951] *J. M. Howell* offers an alternate method of simplification: $(27 + 8i)/(3 + 2i^3) = (27 + 8i^9)/(3 + 2i^3) = 9 - 6i^3 + 4i^6 = 5 + 6i$.

"**Q 57.** [March 1952] Prove that the derivative of an even function is odd and vice versa." *M. S. Klamkin* offers this alternate solution: $E(x) = [E(x) + E(-x)]/2$ so $E'(x) = [E'(x) - E'(-x)/2$ = odd. Also, $O(x) = [O(x) - O(-x)]/2$ so $O'(x) = [O'(x) + O'(-x)]/2$ = even. [In this proof, the notation $N'(y)$ means $d[N(y)]/dy$.]

Q 88. Given $f(x) = x^{10} + x^8 + x^6 + \cdots + 1$, show that $f(2i) \equiv 0 \pmod 9$. [*Submitted by T. C. Wilderman.*]

Q 89. A circle of radius 15 intersects another circle, radius 20, at right angles. What is the difference of areas of the non-overlapping portions? [*Joseph Kennedy in School Science and Mathematics*, **52**, 162, February 1952.]

Q 90. Prove that the sum of the vectors from the center, O, of a regular n-gon to its vertices is zero. [*Submitted by Richard Couchman.*]

Q 91. Prove that $\log_{10} 2$ is irrational, without assuming a knowledge that any k-th root of 10 is irrational. [*Submitted by M. P. Fobes.*]

Q 92. If p is an odd prime, all quadratic residues of p are congruent to $1^2, 2^2, \cdots, [(p-1)/2]^2$ modulo p. For what values of p are the quadratic residues equal to 1^2, 2^2, $\cdots$, $[(p - 1)/2]^2$? [*Submitted by V. C. Harris.*]

Q 93. If $(1 + x)^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ and n is odd, then $a_0^{\,k} - a_1^{\,k} + a_2^{\,k} - \cdots + (-1)^n a_n^{\,k} = 0$. [*Submitted by B. K. Gold.*]

Q 94. Find three integers whose sum is 117 and whose squares are in arithmetic progression. [*Edwin Tabor in THE BAT, No. 47, page 326, November 1947.*]

ANSWERS

A 88. Since $f(x) = (x^{12} - 1)/(x^2 - 1)$, then $f(2i) = [(2i)^{12} - 1]/[(2i)^2 - 1] = (4096 - 1)/(-4 - 1) = -819 = -9(91)$.

A 89. If two circles of areas a and b have a common area x, then the non-overlapping portions are $a - x$ and $b - x$. Then the difference of the non-overlapping portions is $a - b$, the difference of the areas of the circles. In the specific case, the difference is $\pi[(20)^2 - (15)^2]$ or 175π.

A 90. Let R be the resultant of the vectors. About O, rotate the configuration through $2\pi/n$ radians, bringing it into coincidence with its former position. Note that R also rotates through $2\pi/n$ to become R'. Clearly, $R = R'$, but since their directions differ, then $R = R' = 0$.

A 91. Suppose $\log_{10} 2 = a/b$, with a and b integers and the fraction in lowest terms. Then $10^{a/b} = 2$, whence $10^a = 2^b$, or $2^a 5^a = 2^b$. That is, $5^a = 2^{b-a}$, where, since a/b must be a proper fraction, both exponents are positive. However, the left hand side of the equation is odd and the right hand side is even, so the assumption has led to a contradiction.

A 92. We must have $[(p - 1)/2]^2 < p$. This implies $p^2 - 6p + 1 < 0$. Hence, $3 - 2\sqrt{2} < p < 3 + 2\sqrt{2}$ and the values of p are 3 and 5.

A 93. In the binomial expansion, $a_i = a_{n-i}$. If n is odd, the number of terms is even and all the coefficients may be so paired that $a_i^{\,k} - a_{n-i}^{k} = 0$. Hence the proposition.

A 94. The first seven integer squares are 1, 4, 9, 16, 25, 36, and 49. From this it is evident that the three smallest integers whose squares are in A.P. are 1, 5 and 7. Now $1 + 5 + 7 = 13$ and $117 \div 13 = 9$. Hence the integers sought are (1)(9), (5)(9), and (7)(9) or 9, 45, and 63.

TRICKIES

A trickie is a problem whose solution depends upon the perception of the key word, phrase or idea rather than upon a mathematical routine. Send us your favorite trickies.

T 6. In a bureau drawer in a dark room there are 26 grey and 26 blue socks. How many socks must a man take in order to be sure of having a pair of the same color? [*Submitted by R. E. Winger.*]

T 7. I invite you to play the following card game: Shuffle an ordinary deck of cards, and turn them face up in pairs. If both cards of a pair are black, you get them. If both are red, I get them, and if one is red and one black the pair belongs to neither of us. You pay one dollar for the privilege of playing the game. When the game is over you pay nothing if you have at most the same number of cards as I, and for every card that you have more than I, I will pay you three dollars. Would you care to play with me? [*Submitted by Leo Moser.*]

T 8. If you traveled from one town to another at 30 m.p.h., at what speed must you travel on the return trip in order to average 60 m.p.h. for the entire journey? [*Submitted by J. M. Howell*]

T 9. Weary Willie went to the zoo to feed the elephants. Buying a bag of peanuts and wanting to treat each and every elephant alike, he took out his notebook and did a little figuring. He found that if he gave every elephant 7 peanuts, he'd have 6 peanuts left over. On the other hand, if he fed every elephant 9 peanuts, there'd be 2 peanuts over. How many peanuts should weary Willie have given each of these elephants to come out even? [*Monte Dernham in THE BAT, No. 66, page 502, June 1949.*]

SOLUTIONS

S 6. Three. If the first two chosen are not of the same color, the next choice must match one of them.

S 7. You shouldn't. Since the neutral cards are paired, the number of cards in your pile must always equal the number in mine, so you will lose a dollar each time you play.

S 8. If the distance between the towns is x miles, then the time on the outward trip is $x/30$ hours. The time for the entire journey is $2x/60$ hours. Since these times are equal, no time is available for the return journey. The situation is impossible.

S 9. Ten. If feeding each elephant 2 additional peanuts reduces the left-overs by exactly 4 peanuts, there must have been 2 elephants and 20 peanuts.

FALSIES

A falsie is a problem for which a correct solution is obtained by illegal operations, or an incorrect result is secured by apparently legal processes. For each of the following falsies, can your offer an explanation? Send in your favorite falsies.

F 7. A student solved the equation $7 \sin A = 3$ as follows:

$$A = \frac{3}{7 \sin} = \sin^{-1} \frac{3}{7}.$$

[*Submitted by Norman Anning.*]

F 8. The value of the expression

$$(x^4 + x^3 + x^2 + x + 1 + \frac{2}{x - 1})(x^4 - x^3 + x^2 - x + 1 - \frac{2}{x + 1})$$

will not be changed if we suppress the two fractions. [*M. Kraitchik, Mathematical Recreations, Norton (1942), page 42.*]

F 9. Here is a student's method of proving one identity:

$$\frac{\sin 6\theta + \sin 2\theta}{\cos 6\theta + \cos 2\theta} = \frac{\sin \frac{1}{2}(6\theta + 2\theta)}{\cos \frac{1}{2}(6\theta + 2\theta)} = \frac{\sin 4\theta}{\cos 4\theta} = \tan 4\theta.$$

[*Submitted by J. M. Howell.*]

F 10. Three coins are tossed at once. We can say with assurance that of the three coins tossed, two of them must come down alike – both heads or both tails. What of the third coin? The probability that it is heads is ½; that it is tails, also ½. In either case the probability that it is the *same* as the other two is ½. Consequently the probability that all three are alike is ½. [*E. P. Northrop, Riddles in Mathematics, Van Nostrand (1944), page 172.*]

EXPLANATIONS

E 7. Though apparently correct algebra, and atrocious trigonometry, the possibility of this occurence is an argument for the use of "arcsin" rather than "$\sin^{-1}$".

E 8. If each factor be written as a fraction, we have

$$\frac{x^5 + 1}{x - 1} \cdot \frac{x^5 - 1}{x + 1} = \frac{x^5 - 1}{x - 1} \cdot \frac{x^5 + 1}{x + 1} = (x^4 + x^3 + x^2 + x + 1)(x^4 - x^3 + x^2 - x + 1).$$

Thus in the expansion of the original expression, the sum of all the fractional terms is identically zero.

E 9. Since $\sin A + \sin B = 2 \sin \frac{1}{2}(A + B) \cos \frac{1}{2}(A - B)$ and $\cos A + \cos B = 2 \cos \frac{1}{2}(A + B) \cos \frac{1}{2}(A - B)$, the student merely failed to give one essential step. Unfortunately, his process will always lead to a correct result.

E 10. The assumption that it is just as likely for the third coin considered to be like the first two as to be unlike them is not valid. This may be seen by examining the 8 equally like ways that the coins can fall. Since HHH and TTT can each occur in only one way, the correct probability is 2/8 or 1/4.

HERMAN LYLE SMITH

Herman Lyle Smith, Professor of Mathematics at Louisiana State University, died on June 13, 1950.

He was born at Pittwood, Illinois, on July 7, 1892. From the University of Chicago he received the degrees of B. S. (1914), M. S. (1916), Ph. D. (1926). Except for World War I service in the ordnance department at Washington in 1918-19 with Major F. R. Moulton, his life was devoted to teaching and research. He was an instructor at Northwestern (1915-16), Princeton (1916-18), Cornell (1919), Wisconsin (1919-21); professor at the University of the Phillipines (1931-24), assistant professor at the University of Minnesota (1924-26) and assistant professor and later professor at Louisiana State University (1926-31, 1931-50).

The name of H. L. Smith is constantly before us in connection with the "Moore-Smith limit". [1] It is quite possible that more mention of that general concept of limit will appear in modern expository texts of integration theory. His paper on the Stieltjes Integral [2] makes good use of it. In a later paper on the General Theory of Limits [3] he gave a generalization of the earlier concept.

He was active for many years as an editor of the National Mathematics Magazine. Those who were associated with him appreciated greatly his keen interest and high standards.

[1]. Am. Journal of Mathematics, Vol. XLIX (1922)
[2]. Trans. Am. Math. Soc,, Vol. 27 (1925)
[3]. National Math. Magazine, Vol. XII (1937-38)

W. E. Byrne