MATHEMATICS MAGAZINE

CONTENTS

VOLUME 36 • JANUARY 1963 • NUMBER 1

MM

MATHEMATICS MAGAZINE

ROBERT E. HORTON, *Editor*

EDITORIAL CORRESPONDENCE should be sent to the Editor, ROBERT E. HORTON, Department of Mathematics, Los Angeles City College, 855 North Vermont Avenue, Los Angeles 29, California. Articles should be typewritten and double-spaced on $8\frac{1}{2}$ by 11 paper. The greatest possible care should be taken in preparing the manuscript, and authors should keep a complete copy. Figures should be drawn on separate sheets in India ink and of a suitable size for photographing.

NOTICE OF CHANGE OF ADDRESS and other subscription correspondence should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo 14, New York.

ADVERTISING CORRESPONDENCE should be addressed to F. R. OLSON, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo 14, New York.

The MATHEMATICS MAGAZINE is published by the Mathematical Association of America at Buffalo, New York, bi-monthly except July-August. Ordinary subscriptions are: 1 year \$3.00; 2 years \$5.75; 3 years \$8.50; 4 years \$11.00; 5 years \$13.00. Members of the Mathematical Association of America may subscribe at the special rate of 2 years for \$5.00. Single copies are 65¢.

PUBLISHED WITH THE ASSISTANCE OF THE JACOB HOUCK MEMORIAL FUND OF THE MATHEMATICAL ASSOCIATION OF AMERICA

Second class postage paid at Buffalo, New York and additional mailing offices.

MM

MATHEMATICS MAGAZINE

Published by

THE MATHEMATICAL ASSOCIATION OF AMERICA

(INCORPORATED)

A JOURNAL OF COLLEGIATE MATHEMATICS

VOLUME 36

1963

Published at

MENASHA, WIS., AND BUFFALO, N. Y.

STATISTICAL DECISION PROCEDURES IN INDUSTRY

IV. ACCEPTANCE SAMPLING BY VARIABLES

O. B. MOAN, Lockheed Missiles and Space Co., Sunnyvale, Calif.

4.1. Introduction. The third article in this series which appeared in the November, 1962, issue of this magazine discussed the lot-by-lot acceptance decision process when the inspection items were classified as defective or non-defective and the quality of the lot was measured in percent defective. This article discusses an alternate acceptance decision process based on quantitative measurements of a quality characteristic. On the basis of these measurements, a statistical inference is drawn regarding the acceptability of the lot. Such a procedure is known as acceptance sampling by variables.

It is obvious that more information regarding an item is obtained when a quantitative measurement is made than when the item is classified as defective or non-defective. For example, the measured resistances of 150 and 500 ohms provide much more information than saying that the resistances exceeded 100 ohms in both cases. For this same reason, variables inspection of all items in a sample will give more information about the lot than attributes inspection. This greater efficiency of measurements over attributes is also demonstrated in the second article. Therefore, a variables sampling plan usually requires a smaller sample size than an attribute plan with the same degree of protection. However, one should not conclude from this that inspection by variables is always preferred over inspection by attributes. Some disadvantages of variables plans are:

1) The required measuring or test equipment is usually more complex.
2) They often require more inspection skill and time.
3) They are based on more stringent assumptions and are more complicated mathematically. This may require more care and ability in the administration of the plans.
4) They are applicable to only one quality characteristic at a time.

Thus, whenever either a variables or attributes plan can be used, the selection becomes largely one of economics. The one which has the lowest total cost including inspection costs, administration costs, etc. should be used. When both test methods are destructive, the variables method is likely to be cheaper.

In making decisions regarding the acceptance of the lot, using variables plans, the lot quality may be measured in terms of the mean, the standard deviation or the percent defective. This article discusses only sampling procedures based on the mean and percent defective. The percent defective plans have become most popular in industry because: (1) lot quality is customarily defined in this manner and (2) it is also the measure used in attributes plans. When lot quality is measured in terms of the standard deviation, the index of acceptability is the variability of the lot. The reader is referred to page 371 of [1.5] for the details of this procedure. The variables plans discussed in this article require the measurements to be independently distributed normal variables. As in the attributes method, every effort must be made to ensure that the sample is representative of the lot.

4.2. Sampling Plans Based on the Mean. Before discussing the method for establishing a single sampling plan based on the mean, it might be well to expand on the statistical theory discussed in Section 1.7 of the first article. Equation (1.11) represents the normal probability distribution function of the random variable x as:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-(1/2\sigma_x^2)(x-\mu_x)^2}. \tag{1.11}$$

To determine the probability that a randomly selected value of x lies between any two constants a and b, it is necessary to determine the area under the curve between these two constants a and b, i.e.,

$$P(a \leqq x \leqq b) = \int_a^b f(x)dx. \tag{4.1}$$

This integral cannot be written in a simple form. If it is to be tabulated, a separate table would be required for each combination of μ_x and σ_x. The simple transformation

$$t = \frac{x - \mu_x}{\sigma_x} \tag{4.2}$$

reduces equation (1.11) to

$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}. \tag{4.3}$$

It is noted that $\phi(t)$ has the form of a normal probability distribution with mean 0 and standard deviation 1. t is known as a standardized normal deviate. The probabilities associated with the random variable x can be obtained from the standardized normal deviate as follows:

$$P(a \leqq x \leqq b) = \int_a^b f(x)dx = \int_{t_1}^{t_2} \phi(t)dt = P(t_1 \leqq t \leqq t_2) \tag{4.4}$$

where

$$t_1 = \frac{a - \mu_x}{\sigma_x} \quad \text{and} \quad t_2 = \frac{b - \mu_x}{\sigma_x}\cdot$$

Table A, pp. 404–405, [1.5], gives the values of

$$\int_{-\infty}^{t} \phi(t)dt \tag{4.5}$$

for values of t from -3.5 to $+3.5$ in increments of 0.01.

The fact that the variable sampling plans are based on the assumption of normality should not alarm anyone. The quality characteristics of most products are usually distributed close enough to the normal distribution for practical use. The main concern is that the exceptional cases be recognized and that appropriate allowances be made.

When the standard deviation remains unchanged from lot to lot, and the mean varies, the quality of each lot may be given by its mean. In the case of a single upper specification limit, it is desired to determine a sampling plan which protects against too high a mean. A low mean, $\mu_x = \mu_1$, indicates good quality, while a high mean, $\mu_x = \mu_2$, indicates poor quality. The producer (seller) requires a plan that will nearly always accept lots with a low mean. On the other hand, the consumer (buyer) requires a plan which will nearly always reject lots with high mean. Thus, the Producer's and Consumer's Risks are:

α = probability of rejecting good quality when offered.
β = probability of accepting poor quality when offered.

Given these four values, μ_1, μ_2, α and β, it is possible to determine for a single sampling plan the sample size, n, and the necessary criterion, K, for acceptance.

Example: Suppose the upper specification limit on the firing time of a fusing device is 75 milliseconds. The fuses are submitted in lots for acceptance. Based

TABLE 4.1

FUSE FIRING TIME IN MILLISECONDS OF 7 RANDOM SAMPLES FROM EACH OF 20 LOTS

	Lot No.	X, *ms*							$\sum X$	$\overline{X}$	Decision*
Acceptable Quality ($\mu_x=65$)	1	65	64	70	60	66	61	73	459	65.6	A
	2	66	64	60	55	68	64	66	443	63.3	A
	3	69	64	69	70	63	65	66	466	66.7	A
	4	68	66	66	67	67	63	67	464	66.3	A
	5	61	63	65	65	65	59	70	448	64.0	A
	6	67	62	61	68	69	68	67	462	66.0	A
	7	60	64	59	70	61	61	65	440	62.9	A
	8	69	61	73	60	70	65	69	467	66.7	A
	9	62	61	61	65	63	65	60	437	62.4	A
	10	65	60	68	66	67	62	68	456	65.1	A
Rejectable Quality ($\mu_x=69$)	11	64	75	72	71	69	75	75	501	71.6	R
	12	67	67	66	69	71	64	71	475	67.9	R
	13	69	71	68	71	70	69	71	489	69.9	R
	14	67	63	67	72	63	76	68	476	68.0	R
	15	63	69	68	70	67	69	60	466	66.6	A
	16	67	70	70	68	73	68	70	486	69.4	R
	17	74	68	72	69	69	70	70	492	70.3	R
	18	64	74	67	69	74	71	66	485	69.3	R
	19	68	68	65	68	73	67	67	476	68.0	R
	20	76	68	65	68	72	65	72	486	69.4	R

* A—Accept the lot.
R—Reject the lot.

on experience it is known that the standard deviation of each lot is 3.5 milliseconds, i.e., $\sigma_x = 3.5$ ms. When the lot average is 65 milliseconds, which is 2.86 standard deviations below the upper specification limit, the lot is expected to contain only 0.2% defective and will be considered acceptable. Hence, take $\mu_1 = 65$ ms. as acceptable quality. When the lot average is 69 milliseconds, it is 1.71 standard deviations below the upper specification limit. Such lots will contain about 4.4% defective, and are considered rejectable. Therefore, $\mu_2 = 69$ ms. is known as rejectable quality. Suppose the α and β risks are specified as 0.05 and 0.10 respectively, the problem now is to find the proper sample size, n, and acceptance criterion, K, which will give the protection corresponding to the above conditions.

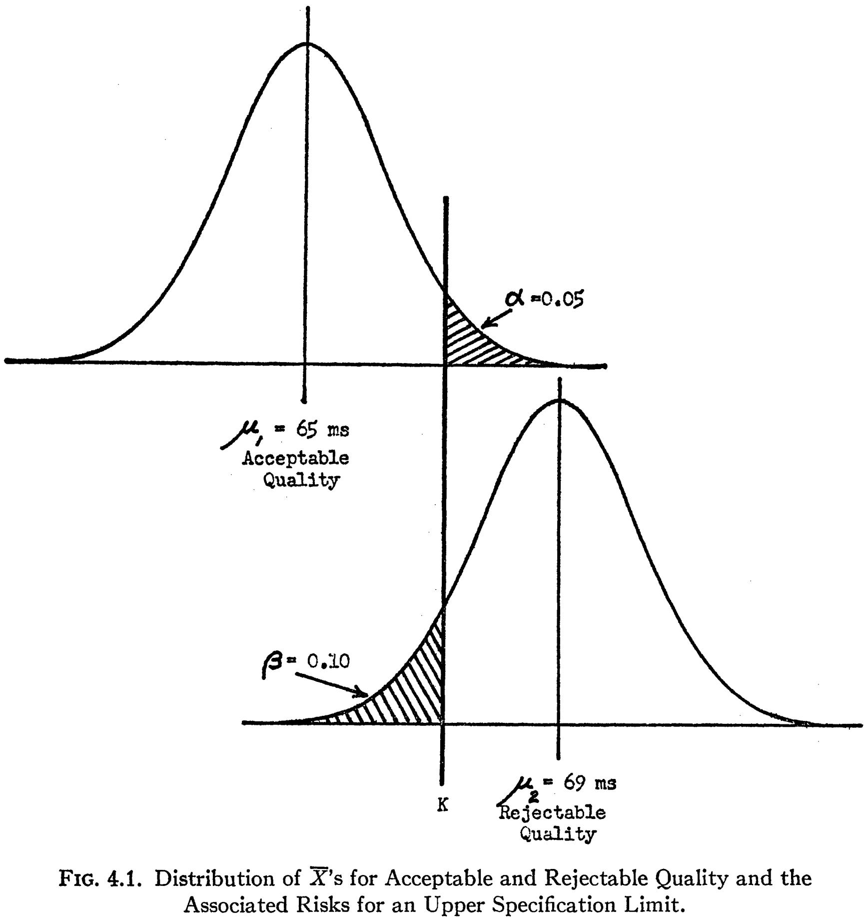


FIG. 4.1. Distribution of $\overline{X}$'s for Acceptable and Rejectable Quality and the Associated Risks for an Upper Specification Limit.

Since the plan is based on averages, the distribution of $\overline{X}$'s must be used in the derivation of the plan. When samples of size n are repeatedly drawn from a normal distribution with mean, μ_x, and standard deviation, σ_x, the distribution function of the sample means, $\overline{X}$, is also normal with mean, $\mu_{\bar{x}} = \mu_x$, and standard deviation, $\sigma_{\bar{x}} = \sigma_x/\sqrt{n}$. (Even when the distribution of the population is not normal, the distribution of $\overline{X}$'s, tends to be near normal [1.6]). Thus, for the acceptable quality, $\mu_1 = 65$ milliseconds, and the rejectable quality, $\mu_2 = 69$ milliseconds, the corresponding $\overline{X}$ distributions have means equal to μ_1 and μ_2 respectively and the same standard deviation $\sigma_x/\sqrt{n}$. These two distributions must be located relative to each other in such a way that the conditions for α and β are also satisfied. See Figure 4.1. Let K be the criterion which discriminates between acceptable and rejectable qualities and it is located such that $\alpha = 0.05$ and $\beta = 0.10$. When $\alpha = 0.05$ the normal deviate $t_\alpha = 1.645$ and when $\beta = 0.10$, $t_\beta = -1.282$. Hence the following two equations may be written:

$$t_\alpha = \quad 1.645 = \frac{K - \mu_{\bar{x}}}{\sigma_{\bar{x}}/\sqrt{n}} = \frac{K - 65}{3.5/\sqrt{n}} \tag{4.5}$$

$$t_\beta = -1.282 = \frac{K - 69}{3.5/\sqrt{n}} \tag{4.6}$$

Solving these two equations for n, one obtains $n = 6.83$. Since sample sizes are always whole numbers it is advisable to round all decimals up to the next whole number. Therefore, let $n = 7$. This method assures that the Producer's and/or Consumer's Risk will not be increased. K is determined by substituting $n = 7$ in either equation (4.5) or (4.6). Substituting in (4.5) preserves α while β is preserved if (4.6) is used. Using (4.5), K becomes 67.2 ms.

Thus the plan is:

Select a random sample of 7 fuses and measure the "firing time," X. Calculate $\overline{X}$. If $\overline{X} \leqq 67.2$ ms, accept the lot, otherwise reject the lot.

The lot quality and its associated probability of acceptance is given by the Operating Characteristic Curve as shown in Fig. 4.2. This Operating Characteristic Curve is computed by using equation (4.4).

To illustrate how this sampling plan functions, 4 dice may be rolled. The mean of the distribution of the sum of 4 dice is 14 and the standard deviation is 3.42. Let each point on any die be 1 millisecond, then

$$\mu_x = 14 \text{ milliseconds}, \quad \sigma_x = 3.42 \text{ milliseconds}.$$

The latter is very close to $\sigma_x = 3.5$ ms, the standard deviation of the lot. For $\overline{X}$, assume that the total points for the 4 dice are the number of milliseconds above some arbitrary value. For the acceptable quality, this value would be 51 ms and for rejectable quality 55 ms.

Table 4.1 gives the results of this simulated fuse inspection. Four dice were thrown seven times to represent the measurements of seven random samples inspected from each lot. This was repeated 20 times, representing 20 lots. For the first 10 lots, the sum was added to 51 and then entered in the table. For lots 11–20, the sum was added to 55.

$\overline{X}$ was computed for each lot and compared with 67.2 ms. Lots with $\overline{X} \leqq 67.2$ ms were accepted while those with $\overline{X} > 67.2$ ms were rejected. The results are shown in the right hand column of the table. None of the lots with $\mu_1 = 65$ ms, acceptable quality, were rejected. This is to be expected since the probability of rejecting a lot which has a mean of 67.2 ms is approximately 0.05. For the rejectable quality, $\mu_2 = 69$ ms, one lot (No. 15) was accepted. This is also to be expected, since the probability of accepting rejectable quality is approximately

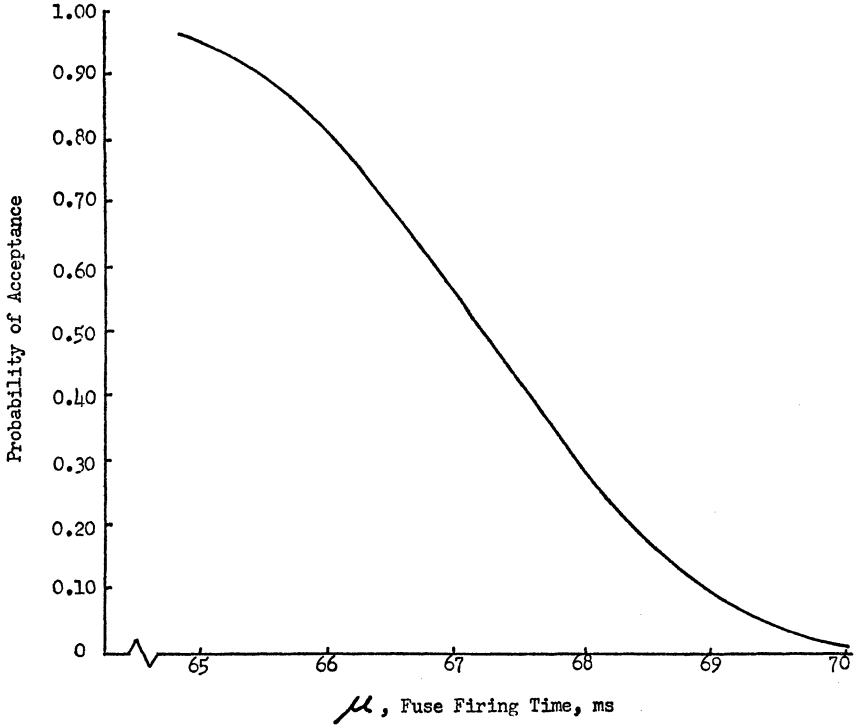


FIG. 4.2. Operating Characteristic Curve for Single Sampling Plan: $n=7$, Accept if $\overline{X} \leq 67.2$ ms, $\sigma_x = 3.5$ ms.

0.10. By adjusting the arbitrary value, other values of μ_x could be tried and results compared with the O.C. Curve (Fig. 4.2).

4.3. Sampling Plans Based on Percent Defective. The aim of these plans is to accept the lot if the percent defective of the lot is less than a predetermined constant. For a single specification limit, an item is defective if

1) its measurement is less than the lower specification limit, L, or
2) its measurement exceeds the upper specification limit, U.

When both the upper and lower specification limits are specified, an item is defective if its measurement is either above U or below L. Therefore, in order to determine the percent defective, it is necessary to examine the relationship between the specification limit(s) and the normal distribution of measurements in the lot. A normal distribution, equation (1.11), is characterized by two quantities, its mean, μ_x, and its standard deviation, σ_x. When σ_x is small the values are tightly clustered around μ_x. For widely scattered values about the mean, σ_x is large. Thus the proportion of items above an upper specification, U, will depend upon how far U is above the mean in terms of the standard deviation. From Fig. 4.3, it is readily apparent that for a fixed mean, less than U, the fraction defective, p'_U, decreases as the standard deviation decreases. For a given standard deviation, the fraction defective, p'_U, decreases as the mean decreases.

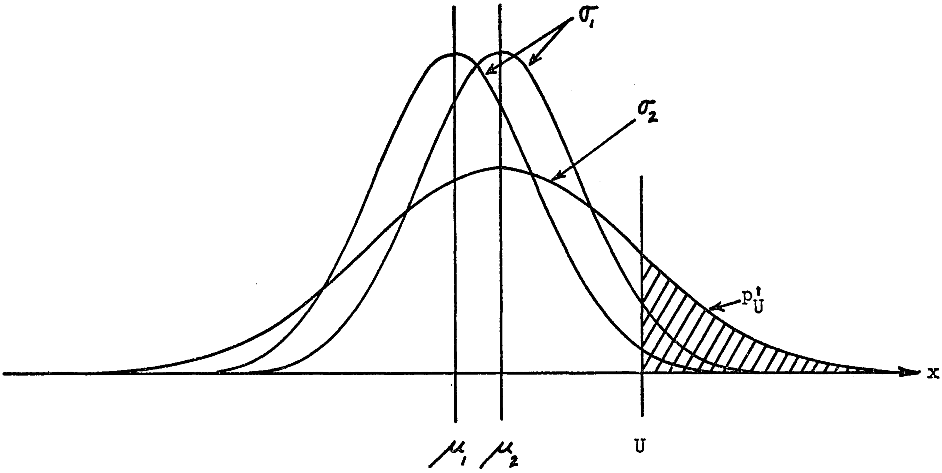


FIG. 4.3. Relationship between the Fraction Defective, p'_U, and the Mean and Standard Deviation for an Upper Specification Limit, U.

Thus, the acceptability of the lot in terms of an acceptable fraction defective will depend upon μ_x and σ_x. To determine the proportion of items above the upper specification limit, U, it is necessary to compute $t=(U-\mu_x)/\sigma_x$. As noted in Section 4.2, Table A in [1.5] gives what portion of the normal distribution is below t. There are many combinations of μ_x and σ_x which result in the same value of t. All of these lots will have the same quality in terms of percent defective. A lot is accepted if its fraction defective is less than a specified value, p_1'. In other words, the lot is acceptable if

$$\mu + K_{p_1'}\sigma \leqq U \tag{4.7}$$

where

$$\int_{kp_1'}^{\infty} \phi(t)dt = p_1'. \tag{4.8}$$

The above discussion, has been in terms of the lot mean, μ_x, and lot standard deviation, σ_x. In actual practice, the lot mean, μ_x, is unknown and the standard deviation, σ_x, is usually unknown. Thus, it is necessary to make estimates of these population parameters using the sample mean and sample standard deviation. To determine whether (4.7) is satisfied, a multiple of the sample standard deviation, s, is subtracted from the sample mean, $\overline{X}$, and this result compared with the limit, U. In selecting the multiple of the sample standard deviation, allowances must be made for the fact that both the sample mean and sample standard deviation are subject to variations from sample to sample. That is, samples will not yield the same value of $(U-\overline{X})/s$. Thus, if $k=(U-\overline{X})/s$, k must be small enough so that a lot of acceptable quality will meet the acceptance criteria,

$$\overline{X} + ks \leqq U \tag{4.9}$$

a high percentage of the time. The sample standard deviation is computed from the n sample measurements as follows

$$(4.10) \qquad s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} = \sqrt{\frac{n\sum_{i=1}^{n} X_i - \left(\sum_{i=1}^{n} X_i\right)^2}{n(n-1)}}$$

A detailed discussion on the determination of k may be found in [4.1] and [4.4].

The above sampling plans are classified into three basic categories:

1) Known standard deviation plans,
2) Unknown standard deviation plans based upon the sample standard deviation,
3) Unknown standard deviation plans based upon the average sample range. As noted in the first article, the range may be used as a measure of the variability. This method is not discussed in this section. The reader is referred to [4.6] for further details.

A summary of the above criteria for acceptance is:

1) Lower Specification Limit, L
   a) Known standard deviation—accept if $\overline{X}-k\sigma \geqq L$, otherwise reject.
   b) Unknown standard deviation—accept if $\overline{X}-ks \geqq L$, otherwise reject.
2) Upper Specification Limit, U
   a) Known standard deviation—accept if $\overline{X}+k\sigma \leqq U$, otherwise reject.
   b) Unknown standard deviation—accept if $\overline{X}+ks \leqq U$, otherwise reject.
3) Both Upper and Lower Specification Limits
   a) Known standard deviation—accept if both $\overline{X}+k\sigma \leqq U$ and $\overline{X}-k\sigma \geqq L$, otherwise reject.
   b) Unknown standard deviation—see [4.7] for a graphical procedure.

Instead of comparing the average plus or minus a multiple of the standard deviation with specification limit(s), the percent defective of the lot may be estimated and compared with a maximum percent defective [4.8].

Let p_U = estimated proportion of items above U

p_L = estimated proportion of items below L

M = maximum allowable fraction defective.

Then, the acceptance procedure becomes:

1) Lower Specification Limit—accept if $p_L \leqq M$, otherwise reject.
2) Upper Specification Limit—accept if $p_U \leqq M$, otherwise reject.
3) Both an Upper and Lower Specification Limit—accept if $p_L + p_U \leqq M$, otherwise reject.

It is evident that these procedures are similar to the attribute plans discussed in the third article.

4.4. MIL-STD-414. Because of the increased use of variable sampling plans, a military standard, MIL-STD-414, [4.2], an alternative to MIL-STD-105C, was issued in late 1957. This standard is divided into the following four sections:

Section A. General Description of Sampling Plans.
Section B. Variability Unknown—Standard Deviation Method.
Section C. Variability Unknown—Range Method.
Section D. Variability Known.

In addition, each of the Sections B, C and D are divided into

Part I. Single Specification Limit.
Part II. Double Specification Limit.
Part III. Estimation of Process Average and the Criteria for Reduced and Tightened Inspection.

Section A. This section is always used in conjunction with the other sections. It provides definitions and general procedures for the sampling plans. Many of the features of this standard are similar to the attribute standard. For example, the sampling plans are indexed by the AQL, inspection level and lot size. After selecting the AQL, the inspection level and the sample size code letter, the sampling plan is selected from the master table in one of the other three sections.

Single Specification Limit—In this case, the acceptability criterion is given in two forms. Form 1 provides the acceptability criterion of the lot without estimating the percent defective. For example, the lot is accepted if $(U-\overline{X})/s \leqq k$. Form 2 requires estimates of the lot percent defective. For example, the lot is accepted if $p_U \leqq M$. Either one may be used since they require the same sample size and possess the same Operating Characteristics Curve. However, MIL-STD-414 states "unless otherwise specified, unknown variability, standard deviation method sampling plans and the acceptability criterion of Form 2 shall be used."

Double Specification Limit—Form 2 must be used in this situation. A single AQL value for the total percent defective or an AQL value for each limit may be specified. For example, the lot is accepted if $p_U + p_L \leqq M$.

Estimation of Process Average—This section gives procedures for estimating the process average and the criteria for tightened and reduced inspection based on the inspection results of preceding lots.

4.5. Other Sampling Plans. A comprehensive set of variables sampling inspection plans are given in [4.5]. Like MIL-STD-414, the lot quality is judged in terms of percent defective. Both single and double sampling plans are presented. Detailed plans are given for use with quality characteristics having a single specification limit. In addition, methods are given for using these plans for quality characteristics which have both a lower and upper specification limit.

It is also possible to devise sequential sampling plans by variables. In this case, measurements are taken one by one, and after each measurement the decision is either to accept, reject or take another measurement. For further information, the reader is referred to [4.3]. The potential advantage of either double or sequential sampling over single sampling is that they result in less average inspection for the same Operating Characteristic Curve.

4.6. Summary. This article has endeavored to outline the general procedure of acceptance sampling by variables. These plans usually require smaller sample sizes than comparable attribute plans. However, when inspection costs are considered, attribute plans may be more economical. One must not forget that this cost of inspection must be balanced against the cost of errors in making a wrong decision. By use of simple statistical methods, sampling plans can be designed to control these risks. Thus, satisfactory estimates of the various costs can be made.

It is imperative that careful analysis of sampling plans be made prior to installation. Associated with this analysis, is the necessity of having well-formulated specifications, accurate measurements, and random sampling.

References

4.1. American Statistical Association, *Acceptance Sampling, A Symposium*, The American Statistical Association, Washington, D.C., 1950.

4.2. MIL-STD-414, *Sampling Procedures and Tables for Inspection by Variables for Percent Defective*, Superintendent of Documents, Washington, D.C., 1957.

4.3. Statistical Research Group, *Sequential Analysis of Statistical Data: Applications*, Columbia University Press, New York, 1945.

4.4. Statistical Research Group, *Techniques of Statistical Analysis*, Columbia University Press, New York, 1947.

4.5. Bowker, A. H. and Goode, H. P., *Sampling Inspection by Variables*, McGraw-Hill Book Company, Inc., New York, 1952.

4.6. Bowker, A. H. and Lieberman, G. J., *Engineering Statistics*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1959.

4.7. Ireson, W. G., *Sampling Tables for Inspection by Variables*, Tech. Report No. 7, Applied Mathematics and Statistics Laboratory, Stanford University, Calif., 1952.

4.8. Lieberman, G. J. and Resnikoff, G. J., *Sampling Plans for Inspection by Variables*, Tech. Report No. 13, Applied Mathematics and Statistics Laboratory, Stanford University, Calif., 1954.

THREE CHEERS FOR D

Hail, limit of the quotient
Of changes evanescent
In things which have dependent fluctuations!

Hail, rating instantaneous
Of changes miscellaneous
Like speed and simultaneous celerations!

Hail, slope of all tangential
Supports, and the essential
Approach to differential operations!

MARLOW SHOLANDER

THE "REFLECTION PROPERTY" OF THE CONICS

R. T. COFFMAN, Richland, Washington and C. S. OGILVY, Hamilton College

The ellipse has the property that a tangent at any point P makes equal angles with the two focal radii to P. This can easily be proven by an appeal to the optical principle of reflection of light from a plane mirror.

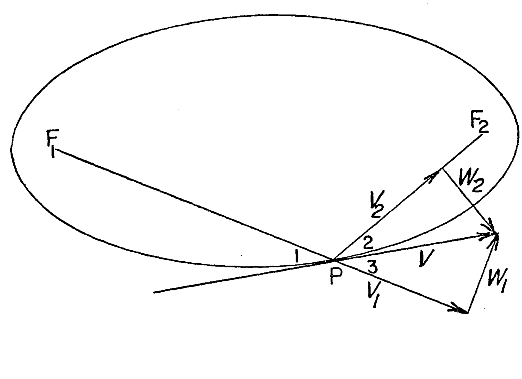
The parabola has the even more familiar property, exploited in every automobile headlight and radar-scope, that the tangent at P makes equal angles with the focal radius to P and the line through P parallel to the parabola's axis. There are various proofs, some by analytic geometry and some depending on calculus, all slightly involved.

The hyperbola also has a corresponding property, not mentioned by some texts and not even known to many students, although one ought to guess it if he has any feeling for the "unity" of the conics. It is elegantly and immediately evident as a consequence of the orthogonality of confocal central conics; but this, again, is a bit beyond the high school level.

We give here an easy method of demonstrating these properties, requiring only the notion of simple velocity vectors. This approach has the advantage of emphasizing that (1) the three different reflection properties of the ellipse, parabola, and hyperbola are essentially the same property, and that (2) this property is a direct consequence of the definitions of the curves.

First, a review of these definitions. A parabola is the path traced out by a point which moves (in the plane) so that its distance from a fixed point is always equal to its distance from a fixed line. An ellipse (hyperbola) is the path traced out by a point which moves so that the sum (difference) of its distances from two fixed points is constant.

Consider now the point P tracing out an ellipse (see Figure 1). The length of F_1P increases at exactly the rate at which F_2P decreases, in order to maintain $F_1P+F_2P=$constant. These two equal rates of change are represented by the velocity vectors V_1 and V_2. The resultant velocity V of P itself is in the tangential direction. It can be resolved into any two perpendicular components. Thus V can be resolved into components V_1 and W_1 or it can be resolved into components V_2 and W_2. Here we have two right triangles with a common hypotenuse



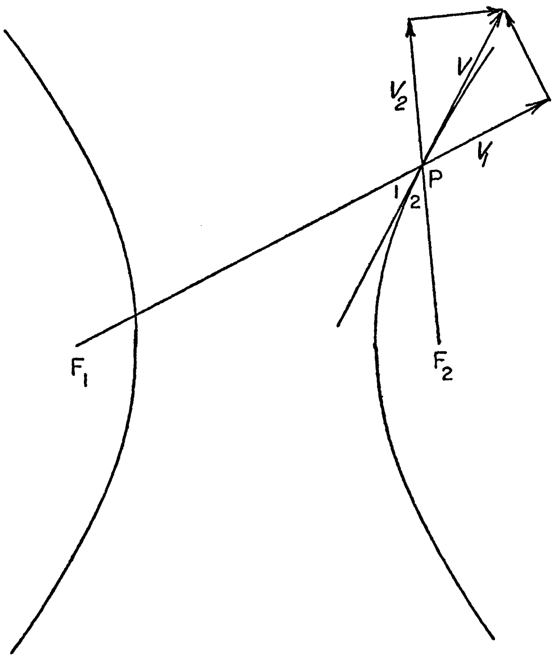
Ellipse

FIG. 1

and two *equal* sides, V_1 and V_2. Hence the right triangles are congruent, $\angle 2 = \angle 3$, and it follows at once that $\angle 1 = \angle 2$, which is what we set out to prove.

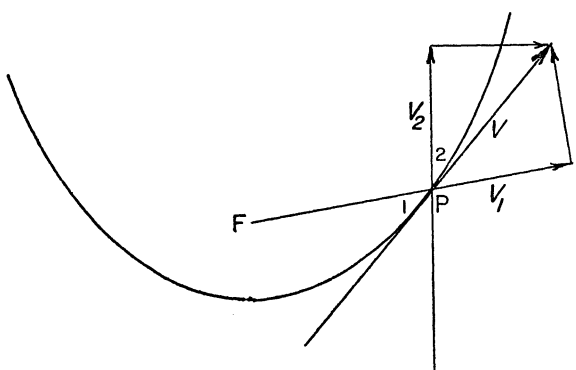
For the hyperbola the proof is almost identical, except that F_1P and F_2P are both increasing (or both decreasing) at the same velocity to maintain $F_1P - F_2P = \text{constant}$. The new "reflection" property is that $\angle 1 = \angle 2$.

For the parabola, it is FP and PD, the distances to the focus and to the directrix, which maintain equality and hence change lengths at the same rate. As before, $\angle 1 = \angle 2$.



Hyperbola

FIG. 2



Parabola

FIG. 3

NEEMIES

PETER KUGEL, Technical Operations, Inc., Burlington, Massachusetts

The advent of the digital computer, with its appetite for binary digits, has made the notion of positional notation important (if not necessarily interesting) for the student who must be weaned from his decimal habits. In this note we will discuss a problem area which involves positional notation and may be of greater interest than most of the problems usually associated with this subject.

Consider the fraction[1] 1288/1449. Read according to the conventions of decimal notation, it represents the rational number[1] 8/9. Read according to the conventions of algebraic notation (on the model of $AB=A\times B$) it represents $1\times2\times8\times8/1\times4\times4\times9$ which is also 8/9. We will be concerned with the question of just how common such "coincidences" are.

Let us call a fraction which can be read as the name of the same rational number according to both conventions a "neemic representation" of the rational number. A rational number will be called a "neemie" if (and only if) there exists a neemic representation of it, and "self neemic" if it is a neemic representation of itself. It is obvious that:

(0) Of the non-negative integers[1] all and only those less than 10 are self neemic.

There is a certain triviality in being neemic only in virtue of being self neemic or of having neemic representations which have only single digits in both numerator and denominator. We shall therefore call such numbers "trivially neemic." A rational number which has a neemic representation with more than two digits will be called "interestingly neemic." Fortunately we have the following:

(1) Every integer, if neemic at all, is interestingly neemic.

Proof. Clearly no non-negative integer larger than 9 can be self neemic (by (0)), so that if it is neemic, then it is interestingly so. The following examples show that the remaining non-negative integers are interestingly neemic. (The case of the negative integers is handled identically.)

$$0 = 0/11,\ 1 = 11/11,\ 2 = 98/49,\ 3 = 153/51,\ 4 = 64/16,\ 5 = 95/19,$$
$$6 = 126/21,\ 7 = 1337/191,\ 8 = 2248/281,\ 9 = 3159/351$$

(2) There are interestingly neemic integers which are not trivially neemic.

Example. $12=168/14=264/22=366/33$, and clearly 12 cannot be trivially neemic.

A number is said to be "perfectly neemic" if every representation of it is neemic, and it is easy to see that there are only two such numbers: 0 and 1. This fact, unlike most of the others we shall consider, is independent of the radix (assuming it to be integral). A number is said to be "infinitely neemic" if it has

[1] For the purposes of this note, a fraction is the name of a rational number and we identify a rational number with its fractional name having the lowest possible numerator. We identify the integers with the rational numbers whose denominator is 1 and write the names of these integers in the usual manner. We shall not distinguish between integers and the numerals which denote them.

an infinite number of neemic representations. Clearly the perfectly neemic numbers are infinitely neemic, but

(3) There are infinitely neemic numbers which are not perfectly neemic.

Example. $1/4 = 861/3444 = 8661/34444 = \cdots = 86 \cdots 1/3 \cdots 444$.

It appears to be rather difficult to discover schemes which mechanically generate or recognize neemies. (The above scheme for (3) merely generates neemic representations.) This difficulty depends, in part, of course, on the chosen radix. The reader should have no difficulty in determining the number of neemies to radix 2, and the truth of the assertion that every integer is neemic to more radixes than it is not.

Let us call numbers which have no neemic representation "aneemic." Again, there are some numbers which will have this characteristic trivially. Thus, it is obvious that:

(4) Every rational number[1] whose numerator or denominator is divisible by 10 is aneemic.

It is easy to derive from this the fact that the set of aneemic numbers in the open interval (1, 0) is dense in itself.

Let us call numbers which are aneemic for any reason other than that indicated in (4) "interestingly aneemic." It turns out that

(5) The set of interestingly aneemic numbers is also dense in itself in the interval (0, 1).

This follows from the following

(5′) Every neemie is expressible in the form: $A^{n_1} \times B^{n_2} \times C^{n_3} / D^{n_4} \times E^{n_5} \times F^{n_6}$. Where A, B, C, D, E, and F range over a finite set of integers, and the n_i are integers of which at least two are 0.

Proof. By the fundamental theorem of arithmetic, every integer can be uniquely expressed as the product of primes. If either the numerator or the denominator of a rational number has a prime factor larger than 10, therefore, it cannot be expressed as the product of (decimal) digits. But this is clearly a necessary (though not sufficient) condition for being neemic. Since there are only four primes smaller than 10, if we exclude 1, it follows that any neemie can be expressed in the form:

$$\frac{A^{n_1} \times B^{n_2} \times C^{n_3} \times D^{n_4}}{A^{n_5} \times B^{n_6} \times C^{n_7} \times D^{n_8}}.$$

Since 2 and 5 cannot both appear in the numerator (denominator) by (4), at least one of the factors in the numerator (denominator) is dispensable. Since 3 and 7 cannot appear in both the numerator and denominator, at least two of the n_i must be 0. (5′) now follows if we let the variables A through F range over the integers 2, 3, 5, 7.

One can then use Euclid's algorithm for generating primes to show that the set of interestingly aneemic numbers is dense in the interval (0, 1), for between any two rationals A/B and C/D, where $A/B < C/D$, there is the following number which is aneemic by (5′):

$$\frac{A((A \times B \times C \times D)! + 1) + 1}{B((A \times B \times C \times D)! + 1)}$$

and clearly this number can be trivially aneemic only if A/B is.

Since (4) rules out 10 as a neemic number, and the proof of (5′) rules out 11, 12 is the smallest neemic integer which is not self neemic. Whether there is a largest neemic integer or whether there are infinitely many neemies are open problems.

AN EXTENDED MEAN VALUE THEOREM

DWIGHT B. GOODNER, The Florida State University

An interesting mean value theorem developed by Flett [1] requires the existence of the derivative. However, since the derivative may fail to exist, it seems desirable to have expressions which may serve us when there is no derivative. The purpose of this note is to extend Flett's mean value theorem to semi-continuous functions which have finite right and left derivates.

DEFINITION. *Let f be a real-valued function of a real variable defined for $\xi \leqq x < \xi + h$. If there exists a sequence $b_1, b_2, b_3, \cdots$ of positive numbers tending to* 0 *and less than h such that*

$$\lim_{n \to \infty} \frac{f(\xi + b_n) - f(\xi)}{b_n}$$

exists, this limit is called a right derivate of f at ξ and is denoted by $D(+)f(\xi)$. Left derivates are analogously defined and denoted.

THEOREM. *Let the function f be upper semi-continuous on the closed interval $a \leqq x \leqq b$, let $f(a) < f(b)$ and let f have at least one finite right derivate $D(+)f(x)$ and at least one finite left derivate $D(-)f(x)$ at each point x of the open interval $a < x < b$. If f has a finite right derivative $f'(a+)$ at a and a finite left derivate $D_o(-)f(b)$ at b and if $f'(a+) = D_o(-)f(b)$, then there exist numbers ξ, p, and q with $a < \xi < b$, $p \geqq 0$, $q \geqq 0$, $p + q = 1$ for which*

$$\frac{f(\xi) - f(a)}{\xi - a} = pD(+)f(\xi) + qD(-)f(\xi).$$

A similar theorem holds if f is lower semi-continuous on $a \leqq x \leqq b$ and $f(a) > f(b)$.

Proof. We may assume without loss of generality that $f'(a+) = D_o(-)f(b) = 0$, since otherwise we would consider the function defined by $f(x) - xf'(a+)$. Let $F(x) = [f(x) - f(a)]/(x - a)$ when $a < x \leqq b$ and let $F(a) = 0$. The function F is upper semi-continuous on the closed interval $a \leqq x \leqq b$ and hence attains its maximum at some point ξ in that interval ([3], p. 102). But ξ is interior to the interval, for $F(a) < F(b)$ while $D_o(-)F(b) = -F(b)/(b-a) < 0$ (the $D_o(-)$ being defined by the same sequence for F as for f). By a proof to be found in many calculus books, all right derivates of F at ξ are $\leqq 0$ and all left derivates of F at ξ are $\geqq 0$. Since f has at least one finite right derivate $D(+)f(\xi)$ and at least one finite left derivate $D(-)f(\xi)$ at ξ, F has a finite right derivate $D(+)F(\xi) = D(+)f(\xi)/(\xi - a) - F(\xi)/(\xi - a)$ and a finite left derivate $D(-)F(\xi) = D(-)f(\xi)/(\xi - a) - F(\xi)/(\xi - a)$ at ξ (the $D(-)$ and the $D(+)$ being defined by the same sequences for F as for f). It follows that $D(+)F(\xi)D(-)F(\xi) \leqq 0$. If $D(+)F(\xi) = 0$, we may choose $p=1$ and $q=0$. If $D(+)F(\xi) \neq 0$, we may choose

$$p = \frac{D(-)F(\xi)}{D(-)F(\xi) - D(+)F(\xi)} \quad \text{and} \quad q = \frac{D(+)F(\xi)}{D(+)F(\xi) - D(-)F(\xi)} .$$

In either case $p \geqq 0$, $q \geqq 0$, $p+q=1$ and $pD(+)F(\xi) + qD(-)F(\xi) = 0$ ([2], p. 852). Substituting from above for $D(+)F(\xi)$ and $D(-)F(\xi)$ and then simplifying gives the desired relation.

References

1. T. M. Flett, "A mean value theorem," *Mathematical Gazette*, **42** (1958) 38–39.
2. D. B. Goodner, "Mean value theorems for functions with finite derivates," *American Mathematical Monthly*, **67** (1960) 852–855.
3. H. P. Thielman, *The theory of functions of real variables*, New York, 1953.

PAPPUSOIDAL

When geometric entities,
Such as a curve (or plate),
About an axis in their plane
Disjointly are rotate,
The distance that the centroid wends
Times given length (or area) tends
To mensurate the odds and ends
With which we terminate.

MARLOW SHOLANDER

ON exp AND log IN ELEMENTARY CALCULUS

A. H. LIGHTSTONE, Carleton University

The trouble with introducing the functions exp and log before the study of the definite integral, is found in the difficulty in establishing that

$$(1) \qquad \lim\left(1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}\right)=\lim\left(\left(1+\frac{1}{n}\right)^n\right).$$

For this reason, many textbook authors, Hardy for example, introduce log in terms of the definite integral and then define exp in terms of log. Since it is difficult to build up a technique of integration without these functions, it follows that the study of the anti-derivative must also be delayed until the definite integral has been presented. Of course, it is possible to live with these organizational restrictions, and many persons may actually prefer to present the definite integral before the anti-derivitive. The purpose of this article is to demonstrate that (1) can be established in a simple manner, and then to compute the derivatives of exp and log. With this accomplished, the above restrictions on the order in which topics are presented, will have disappeared.

In the following discussion, we shall make use of these facts about sequences:

THEOREM 1. $\lim\ (a_n)$ *exists if the sequence* (a_n) *is increasing and bounded above.*

THEOREM 2. $\lim\ (a_n) \leqq \lim\ (b_n)$ *provided that both limits exist and* $\forall_n(a_n \leqq b_n)$.

Turning now to the two sequences appearing in (1), it is easy to show that each sequence is increasing and is bounded above. Therefore, by Theorem 1, each sequence possesses a limit; let us denote these limits by L_1 and L_2, so that

$$\lim\left(1+\frac{1}{1!}+\frac{1}{2!},+\cdots+\frac{1}{n!}\right)=L_1$$

and

$$\lim\left(\left(1+\frac{1}{n}\right)^n\right)=L_2$$

We prove that $L_1=L_2$ by showing that $L_2 \leqq L_1$ and $L_1 \leqq L_2$. There is no difficulty in showing $L_2 \leqq L_1$, in fact the proof is trivial; it is the demonstration that $L_1 \leqq L_2$ that offers difficulty. We shall see that by means of a very simple idea, the difficulty is overcome, and in fact $L_1 \leqq L_2$ can be established in a manner parallel to the proof that $L_2 \leqq L_1$. For this reason, we present the usual proof that $L_2 \leqq L_1$.

By the binomial theorem,

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

$$+\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right)}{n!}$$

$$\leqq 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}$$

for each natural number n.

Therefore, by Theorem 2, $L_2 \leqq L_1$. We wish, now, to prove that $L_1 \leqq L_2$. To this purpose, note that

$$\lim\left(\left(1+\frac{1}{n}\right)^{n+1}\right) = \lim\left(\left(1+\frac{1}{n}\right)^{n}\left(1+\frac{1}{n}\right)\right) = L_2 \lim\left(1+\frac{1}{n}\right) = L_2$$

and

$$\lim\left(\left(1+\frac{1}{n}\right)^{n+2}\right) = \lim\left(\left(1+\frac{1}{n}\right)^{n+1}\left(1+\frac{1}{n}\right)\right) = L_2 \lim\left(1+\frac{1}{n}\right) = L_2,$$

proceeding as indicated, it is easy to prove by mathematical induction that

$$\lim\left(\left(1+\frac{1}{n}\right)^{n+k}\right) = L_2$$

for any natural number k. But, by the binomial theorem

$$\left(1+\frac{1}{n}\right)^{n+k} = 1+\frac{1+\frac{k}{n}}{1!}+\frac{\left(1+\frac{k}{n}\right)\left(1+\frac{k-1}{n}\right)}{2!}+\cdots$$

$$+\frac{\left(1+\frac{k}{n}\right)\cdots\left(1+\frac{k-k}{n}\right)}{(k+1)!}+\cdots$$

$$+\frac{\left(1+\frac{k}{n}\right)\cdots\left(1+\frac{k-k}{n}\right)\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{n-1}{n}\right)}{(k+n)!}$$

$$> 1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{k!}$$

for any natural numbers n and k.

Therefore each term of the constant sequence

$$\left(1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{k!}\right)$$

is less than the corresponding term of the sequence

$$\left(\left(1+\frac{1}{n}\right)^{n+k}\right);$$

hence, by Theorem 2,

$$1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{k!}\leqq L_2$$

for any natural number k. But this states that each term of the sequence

$$\left(1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}\right)$$

is less than or equal to the corresponding term of the constant sequence (L_2), and so by Theorem 2, $L_1 \leqq L_2$. Thus, (1) is established.

It is customary to denote the common limit of the two sequences appearing in (1), by e, and to introduce the function called *exponent* and denoted by exp or by e^x, as follows:

DEFINITION. exp *is* $\{(a, b) \mid a \in R^\# \wedge b = e^a\}$, *where* $R^\#$ *denotes the set of all real numbers*. The power of a positive number, with which this definition is involved, is usually introduced in the discussion of the real number system. We may assume, then, that this concept is known; this means that we may assume the basic properties of powers. In particular we observe that $\exp(1)=e$, $\exp(0)=1$, $\exp(-a)=1/\exp(a)$, and $\exp(a+b)=\exp(a)\cdot\exp(b)$.

What we are after, of course, is to prove that $\exp'=\exp$, where $\exp'$ represents the derivative of exp. To this purpose, we need the following

THEOREM.

$$e^a = \lim\left(1+\frac{a}{1!}+\frac{a^2}{2!}+\cdots+\frac{a^n}{n!}\right).$$

whenever a is a positive real number.

Proof. First, we show that the theorem is true for all positive rational numbers; let r/s be any positive, rational number, so that we may assume $0<s$ and $0<r$, then we shall prove that

$$e^{r/s} = \lim\left(1+\frac{r/s}{1!}+\frac{(r/s)^2}{2!}+\cdots+\frac{(r/s)^n}{n!}\right).$$

It is easy to show that this sequence is increasing and is bounded above; therefore it has a limit, say M. Now,

$$e^{r/s} = \left(\lim\left(\left(1+\frac{1}{n}\right)^n\right)\right)^{r/s} = \lim\left(\left(1+\frac{1}{n}\right)^{nr/s}\right);$$

furthermore, if $\lim (a_n)$ exists, then $\lim (a_n) = \lim (b_n)$ whenever (b_n) is a subsequence of (a_n). Since

$$\left(\left(1+\frac{1}{ns}\right)^{nr}\right) \text{ is a subsequence of } \left(\left(1+\frac{1}{n}\right)^{nr/s}\right),$$

we have

$$e^{r/s} = \lim\left(\left(1+\frac{1}{ns}\right)^{nr}\right).$$

By the binomial theorem

$$\begin{aligned}\left(1+\frac{1}{ns}\right)^{nr} &= 1 + \frac{\frac{r}{s}}{n!} + \frac{\frac{r}{s}\left(\frac{r}{s}-\frac{1}{ns}\right)}{2!} + \cdots + \frac{\frac{r}{s}\cdots\left(\frac{r}{s}-\frac{rn-1}{ns}\right)}{(rn)!}\\ &\leqq 1 + \frac{\frac{r}{s}}{1!} + \frac{\frac{r}{s}}{2!} + \cdots + \frac{\left(\frac{r}{s}\right)^{rn}}{(rn)!}\\ &< M\end{aligned}$$

for each natural number n.

Therefore, each term of the sequence

$$\left(\left(1+\frac{1}{ns}\right)^{nr}\right)$$

is less than the corresponding term of the constant sequence (M); hence, by Theorem 2, $e^{r/s} \leqq M$. We shall now show that $M \leqq e^{r/s}$. Again, it is easy to prove by mathematical induction that

$$\lim\left(\left(1+\frac{1}{ns}\right)^{nr+k}\right) = \lim\left(\left(1+\frac{1}{ns}\right)^{nr}\right) = e^{r/s},$$

for any natural number k. But, by the binomial theorem,

$$\begin{aligned}\left(1+\frac{1}{ns}\right)^{nr+k} &= 1 + \frac{\left(\frac{r}{s}+\frac{k}{ns}\right)}{1!} + \frac{\left(\frac{r}{s}+\frac{k}{ns}\right)\left(\frac{r}{s}+\frac{k-1}{ns}\right)}{2!} + \cdots\\ &\quad + \frac{\left(\frac{r}{s}+\frac{k}{ns}\right)\cdots\left(\frac{r}{s}+\frac{k-k}{ns}\right)}{(k+1)!} + \cdots\end{aligned}$$

$$\frac{+\left(\frac{r}{s}+\frac{k}{ns}\right)\cdots\left(\frac{r}{s}+\frac{k-k}{ns}\right)\left(\frac{r}{s}-\frac{1}{ns}\right)\cdots\left(\frac{r}{s}-\frac{nr-1}{ns}\right)}{(nr+k)!}$$

$$> 1+\frac{\frac{r}{s}}{1!}+\frac{\left(\frac{r}{s}\right)^2}{2!}+\cdots+\frac{\left(\frac{r}{s}\right)^k}{k!}$$

for any natural numbers n and k. Therefore, each term of the constant sequence

$$\left(1+\frac{\frac{r}{s}}{1!}+\frac{\left(\frac{r}{s}\right)^2}{2!}+\cdots+\frac{\left(\frac{r}{s}\right)^k}{\mathrm{k}!}\right)$$

is less than the corresponding term of the sequence

$$\left(\left(1+\frac{1}{ns}\right)^{nr+k}\right);$$

hence, by Theorem 2,

$$1+\frac{r/s}{1!}+\frac{(r/s)^2}{2!}+\cdots+\frac{(r/s)^k}{k!}\leqq e^{r/s}$$

for each natural number k. But this states that each term of the sequence

$$\left(1+\frac{r/s}{1!}+\frac{(r/s)^2}{2!}+\cdots+\frac{(r/s)^n}{n!}\right)$$

is less than or equal to the corresponding term of the constant sequence $(e^{r/s})$, and so, by Theorem 2, $M\leqq e^{r/s}$.

This proves that the theorem is true for each positive, rational number. By recalling the definition of a real number in terms of rational numbers, and by recalling the definition of the limit of a sequence, it can be shown that the theorem is true for any positive real number.

Let us now prove the

THEOREM.

$$\lim\left(\frac{e^{a_n}-1}{a_n}\right)=1,\quad \textit{if}\quad \lim\,(a_n)=0\quad \textit{and}\quad \forall_n(0<a_n).$$

Proof.

$$\frac{e^{a_k}-1}{a_k}=\lim\left(1+\frac{a_k}{2!}+\frac{a_k^{2}}{3!}+\cdots+\frac{a_k^{n-1}}{n!}\right)$$

for any natural number k. Since $0 < a_k$,

$$1+\frac{a_k}{2!}+\frac{a_k^2}{3!}+\cdots+\frac{a_k^{n-1}}{n!}>1,$$

therefore

$$\lim\left(1+\frac{a_k}{2!}+\frac{a_k^2}{3!}+\cdots+\frac{a_k^{n-1}}{n!}\right)\geqq 1, \quad \text{i.e.} \quad \frac{e^{a_k}-1}{a^k}\geqq 1$$

for each natural number k. Also,

$$\begin{aligned}1+\frac{a_k}{2!}+\frac{a_k^2}{3!}+\cdots+\frac{a_k^{n-1}}{n!} &< 1+\frac{a_k}{2}+\frac{a_k^2}{2^2}+\cdots+\frac{a_k^{n-1}}{2^{n-1}}\\ &= \frac{1-(a_k/2)^n}{1-(a_k/2)} \quad \text{provided } a_k<2\\ &< \frac{1}{1-(a_k/2)} \quad \text{whenever } a_k<2.\end{aligned}$$

Therefore, by Theorem 2,

$$\frac{e^{a_k}-1}{a_k}\leqq\frac{1}{1-(a_k/2)}$$

whenever $a_k<2$. Thus, the sequence $((e^{a_n}-1)/a_n)$ possesses a limit and the limit is 1.

COROLLARY. $\lim\,(e^{a_n})=1$ *if* $\lim\,(a_n)=0$ *and* $\forall_n(0<a_n)$.

Proof. $\lim\,(e^{a_n}-1)=0$; therefore, $\lim\,(e^{a_n})=1$.

THEOREM. $\lim\,((e^{a_n}-1)/a_n)=1$ *if* $\lim\,(a_n)=0$ *and* $\forall_n(a_n<0)$.

Proof. if $\lim\,(a_n)=0$ and $\forall_n(a_n<0)$, then

$$\frac{e^{a_n}-1}{a_n}=\frac{(1/e^{-a_n})-1}{a_n}=\frac{1-e^{-a_n}}{a_ne^{-a_n}}=\frac{e^{-a_n}-1}{-a_ne^{-a_n}}\,.$$

Therefore,

$$\begin{aligned}\lim\left(\frac{e^{a_n}-1}{a_n}\right) &= \lim\left(\frac{e^{-a_n}-1}{-a_n}\right)\div\lim(e^{-a_n})\\ &= 1\div\lim(e^{-a_n}),\ \text{by the preceding theorem}\\ &= 1,\ \text{by the preceding corollary.}\end{aligned}$$

Using the two theorems just established, it is easy to prove the

THEOREM.

$$\lim\left(\frac{e^{a_n}-1}{a_n}\right) = 1 \quad \textit{if} \quad \lim(a_n) = 0 \quad \textit{and} \quad \forall_n(a_n \neq 0).$$

Note that this asserts $\exp'(0)=1$.

We are now ready to compute the derivative of exp.

THEOREM. $\exp' = \exp$.

Proof. If $a \in R^{\#}$ then

$$\begin{aligned}\exp'(a) &= \lim_a \frac{\exp - \exp(a)}{x-a} \\ &= \lim\left(\frac{e^{a_n} - e^a}{a_n - a}\right)\end{aligned}$$

whenever (a_n) is a sequence such that $\lim(a_n)=a$ and $\forall_n(a_n \neq a)$. Thus,

$$\exp'(a) = \lim\left(e^a\left(\frac{e^{a_n-a}-1}{a_n-a}\right)\right) = e^a \lim\left(\frac{e^{a_n-a}-1}{a_n-a}\right) = e^a,$$

since $\lim(a_n-a)=0$ and $\forall_n(a_n-a\neq 0)$. Therefore,

$$\exp' = \{(a, b) \mid a \in R^{\#} \wedge b = e^a\} = \exp.$$

Since no two members of exp possess the same second member, we see that the inverse of exp is $\{(a, b) \mid (b, a) \in \exp\}$. This function is denoted by log. Note that the domain of log is $\{a \mid 0<a\}$. Finally, we calculate the derivative of log.

Recalling that f of $g = \{(a, b) \mid b = f(g(a))\}$ we see that exp of log $= \{(a, b) \mid 0<a \wedge b=a\}$. But $(f \text{ of } g)' = (f' \text{ of } g)\cdot g'$; hence

$$\begin{aligned}(\exp' \text{ of } \log)\cdot\log' &= \{(a, b) \mid 0 < a \wedge b = a\}' \\ \text{i.e. } (\exp \text{ of } \log)\cdot\log' &= \{(a, b) \mid 0 < a \wedge b = 1\}\end{aligned}$$

and so

$$\begin{aligned}\log' &= \frac{\{(a, b) \mid 0 < a \wedge b = 1\}}{\{(a, b) \mid 0 < a \wedge b = a\}} \\ &= \left\{(a, b) \mid 0 < a \wedge b = \frac{1}{a}\right\}\end{aligned}$$

References

Hardy, G. H., A course of pure mathematics, Cambridge Univ. Press, 1928.
Jeffery, R. L., Calculus, Univ. of Toronto Press, 1954.
Rudin, W., Principles of Mathematical Analysis, McGraw-Hill, 1953.

TWO CONSTRUCTIONS WITH A TWO-EDGED RULER

DONALD OLIVIER, Carleton College

Given any triangle ABC and one median AH (see Figure 1), select any point D on AH. Then draw BD and extend it until it intersects AC in E. Draw CD and extend it until it intersects AB in F. Prove EF is parallel to BC.

Proof. Points A, D, E, and F are the vertices of a quadrangle, which implies that the intersection of BC and EF is the harmonic conjugate of H with respect to B and C. But since H is the midpoint of $\overline{BC}$, BC and EF must intersect at infinity; that is, EF must be parallel to BC.

It is easy to see that the converse also holds, and this suggests the following construction:

Problem. Given a ruler with two edges (so that at least one line can always be constructed parallel to a given line) but no compass, construct the midpoint of a given segment.

Solution. Let the given segment be $\overline{BC}$. Construct a line l parallel to BC, take any point A not on BC or l, and draw AC intersecting l in E and AB intersecting l in F. Draw CF and BE, and call their intersection D. Draw AD and extend it until it intersects BC at H. Then H is the midpoint of $\overline{BC}$.

We can use this result to make a second construction:

Problem. Given a line m and a point E not on m, construct a line through E parallel to m, using only a two-edged ruler.

Solution. Pick any two points B and C on m and draw EB and EC. Construct H, the midpoint of $\overline{BC}$. Take a point A on EC distinct from C and E and draw AB and AH. Call D the intersection of EB and AH. Draw CD, and extends it until it intersects AB in F. Then EF is parallel to m.

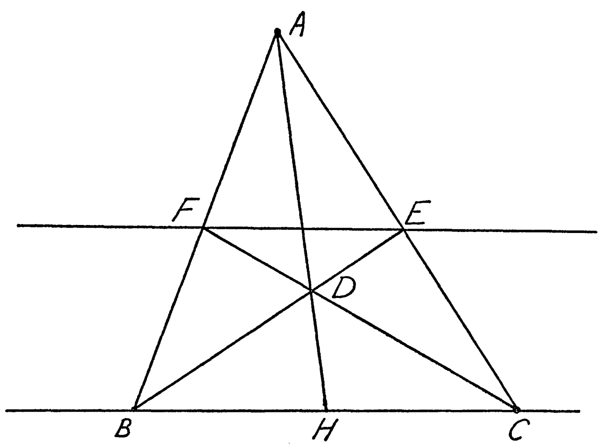


FIG. 1

CONFORMAL TRANSFORMATION CHARTS USED BY ELECTRICAL ENGINEERS

H. F. MATHIS, Ohio State University

Many engineers prefer to solve certain types of problems graphically because they think the visual presentation indicates how various factors interact. If more accuracy is desired, the problem can be reworked using more precise calculations. The most popular chart used for graphical solutions by electrical engineers is the Smith chart [1]. Some engineers find this chart to be almost as useful as their slide rules.

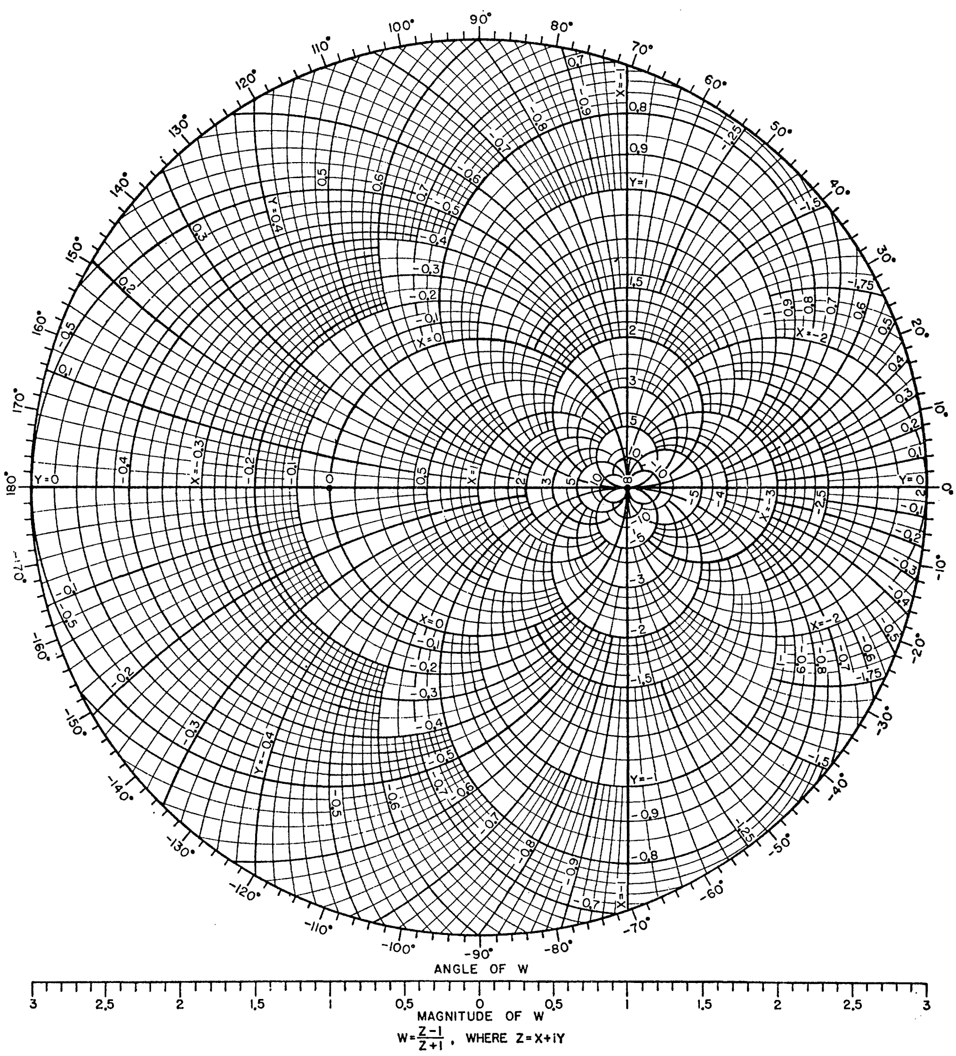


FIG. 1. Circular chart, with rectangular coordinates.

Circular Charts. The Smith chart is simply the map of the right half of the z-plane, i.e., $x \geqq 0$, defined by

$$w = \frac{z-1}{z+1}, \tag{1}$$

where $z = x + iy$. Since electrical engineers work with impedances $Z = R + jX$, admittances $Y = G + jB$, and voltage reflection coefficients

$$\Gamma = \frac{Z-1}{Z+1}, \tag{2}$$

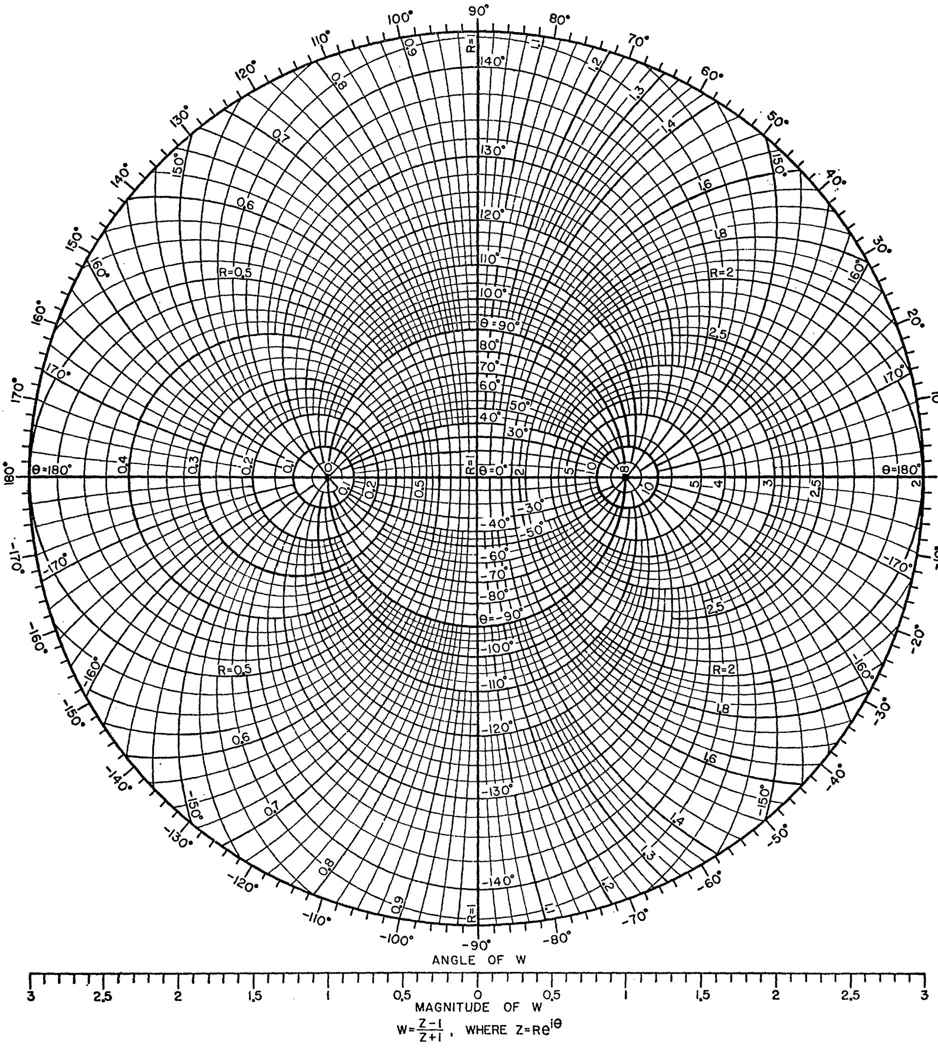


FIG. 2. Circular chart, with polar coordinates.

this nomenclature is found on the Smith chart. The Carter chart [2] is similar to the Smith chart except that z is expressed in polar coordinates. The two charts are called *circular transmission-line charts.*

One reason that engineers like these charts is that the right half of the z-plane is mapped onto the unit circle. Also, infinity is mapped on a single point, which is consistent with their observation that an open circuit, corresponding to $z=\infty$, does not require special attention.

At first these charts were used primarily for solving transmission-line problems. A common problem was to evaluate the function

$$z(d) = \frac{z_L \cos \beta d + i \sin \beta d}{\cos \beta d + i z_L \sin \beta d}, \tag{3}$$

where z_L and β were known. Now these charts are used for many purposes. Hundreds of papers have included some reference to these charts.

For most problems, the restriction that the resistance R be non-negative is not important. However, there are problems where it is desirable to consider negative resistances. For these problems, it may be desirable to use extended circular charts [3], which include portions of the left half of the z-plane.

Conformal transformation charts based on the mapping (1) are shown in Figures 1 and 2. As mentioned above, the charts used by electrical engineers have different notation and several additional scales.

Logarithmic Charts. If β in equation (3) is real, $z(d)$ is a circle when plotted on a circular chart. If β is complex, $z(d)$ is a logarithmic spiral. For this reason, it may be desirable to use a logarithmic chart [4] because $z(d)$ is a straight line in either case. The logarithmic charts are maps defined by

$$w = \log \frac{z-1}{z+1}, \tag{4}$$

where z may be in either rectangular or polar coordinates. Logarithmic charts are shown in Figures 3 and 4.

Construction of Charts. The circular charts are made by drawing circles. An $x=$constant circle has its center at $u=x/(x+1)$, $v=0$ in the w-plane and radius $r=1/|x+1|$. A y-circle has its center at $u=1$, $v=1/y$ and radius $r=1/|y|$. An R-circle has its center at $u=(R^2+1)/(R^2-1)$, $v=0$ and radius $r=2|R/(R^2-1)|$. A θ-circle has its center at $u=0$, $v=-\cot\theta$ and radius $r=|\csc\theta|$.

The logarithmic charts may be drawn by plotting $m=|(z-1)/(z+1)|$ for fixed values of x, y, R, θ as functions of the angle ϕ of $(z-1)/(z+1)$ on semilog paper. The proper substitutions can be made in the above equations for the circles on circular charts to obtain

$$m_x = \frac{1}{x+1}\left(x\cos\phi \pm \sqrt{1-x^2\sin^2\phi}\right),$$

$$m_y = \frac{1}{y}\left(y\cos\phi + \sin\phi \pm \sqrt{(y\cos\phi+\sin\phi)^2 - y^2}\right),$$

$$m_R = \frac{1}{R^2-1}\left[(R^2+1)\cos\phi \pm \sqrt{(R^2-1)^2\cos^2\phi - (R^2-1)^2}\right],$$

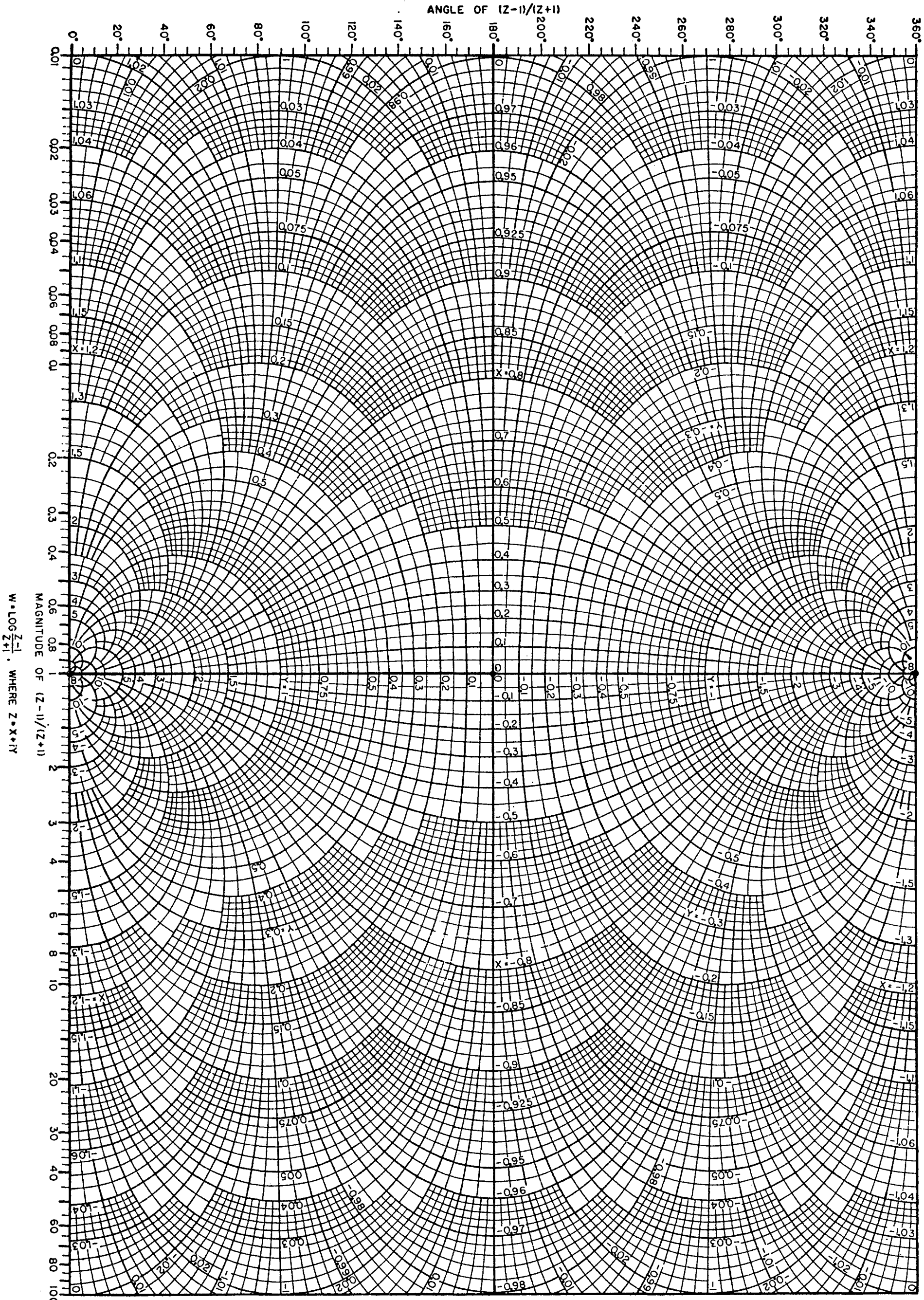


FIG. 3. Logarithmic chart, with rectangular coordinates.

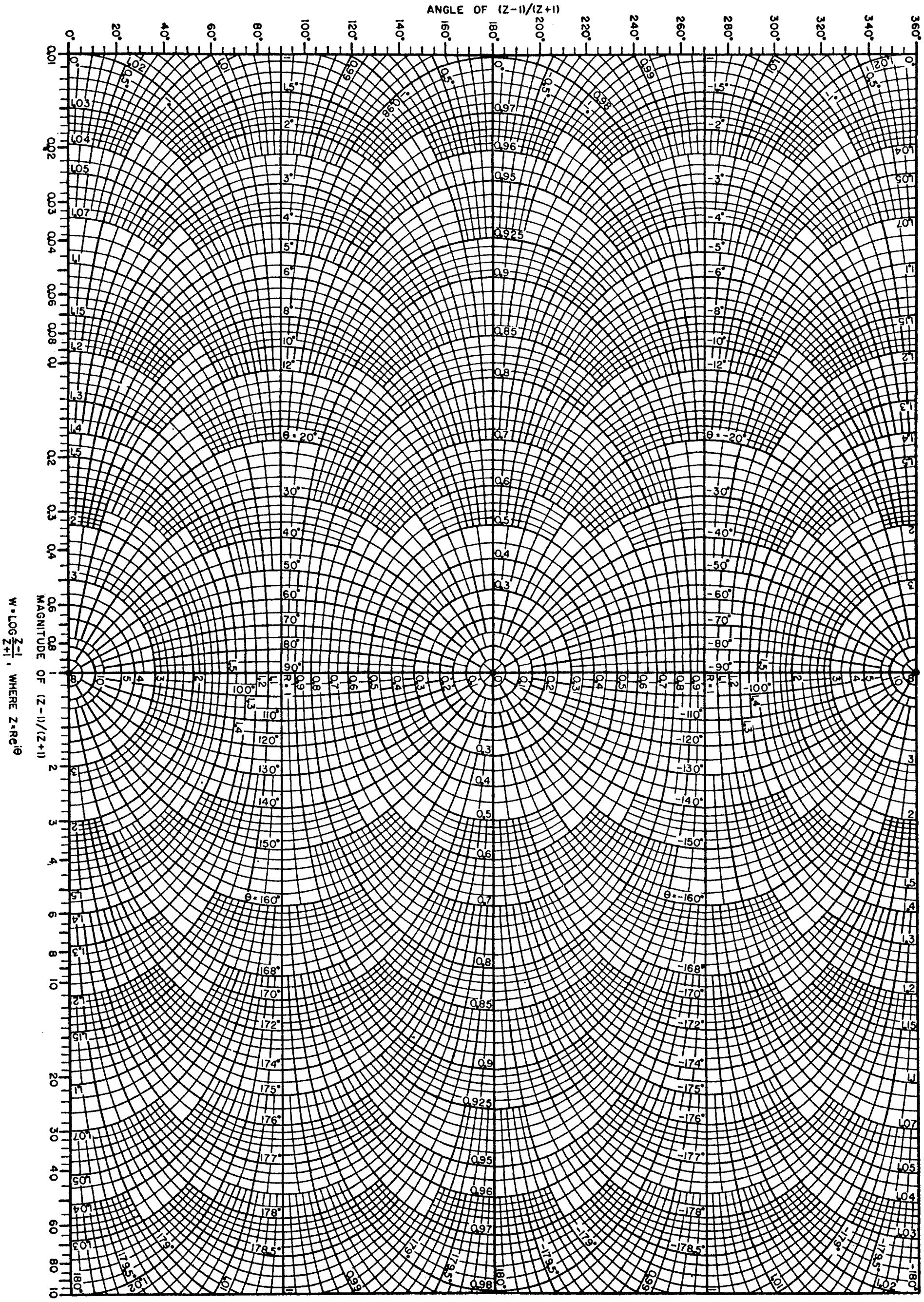


FIG. 4. Logarithmic chart, with polar coordinates.

$$m_\theta = -\cot\theta\sin\phi \pm \sqrt{\cot^2\theta\sin^2\phi + 1}.$$

In order for the curves on the charts to be orthogonal, the distances on the charts corresponding to the ratio e on the log scale and one radian on the linear scale must be equal.

References

1. P. H. Smith, Transmission-line calculator, Electronics, **12** (January, 1939) 29–31. Also, An improved transmission-line calculator, Electronics, **17** (January, 1944) 130–133, 318–325.
2. P. S. Carter, Charts for transmission-line measurements and computations, RCA Review, **3** (January, 1939), 355–368.
3. H. F. Mathis, Extended transmission-line charts, Electronics, **33** (September 23, 1960) 76, 78.
4. ———, Logarithmic transmission-line charts, Electronics, **34** (December 1, 1961) 48–54.

ON THE DERIVATIVE OF THE LOGARITHMIC FUNCTION

F. D. PARKER, University of Alaska

The classical approach to the derivative of log x leaves much to be desired, and has been replaced in some texts by an investigation of the properties of $\int_1^x dt/t$. This, too, has its shortcomings. Accordingly, a third alternative may be in order.

First, we may assume that the derivative of $\log_a x$ exists and is never zero. This cannot be proved for beginning calculus students, but a plausible argument can be made on the basis of the graph of $y = \log_a x$.

Let $u(x)$ and $v(x)$ be differentiable functions of x, each of which takes on all positive values and only positive values, and let $y = uv$. Then $\log_a y = \log_a u + \log_a v$.

Implicit differentiation yields

$$F(y)\frac{dy}{dx} = F(u)\frac{du}{dx} + F(v)\frac{dv}{dx},$$

$$\frac{dy}{dx} = \frac{F(u)}{F(y)}\frac{du}{dx} + \frac{F(v)}{F(y)}\frac{dv}{dx}, \quad \text{and}$$

$$\frac{dy}{dx} = \frac{F(u)}{F(uv)}\frac{du}{dx} + \frac{F(v)}{F(uv)}\frac{dv}{dx}, \quad \text{where}$$

$F(u) = d(\log_a u)/du$. Comparing with the formula for the derivative of a product, we must have $F(u)/F(uv) = v$. Now let $u = 1$, then $F(v) = F(1)/v$, or $d(\log_a v)/dv = F(1)/v$.

Hence the derivative is inversely proportional to the independent variable. Examination of the logarithmic function for various bases suggests that for an appropriate base, the constant is unity. Looking at it in a slightly different way, we can define the number e to be the logarithmic base for which the slope of $\log x$ is unity at $x = 1$.

In order to find the constant of proportionality, we can write $\log_a x = \log_a e \log_e x$, then $d(\log_a x)/dx = 1/x \log_a e$.

AN ELEMENTARY ANALYSIS OF THE FACTORIZATION OF INTEGERS AND THE DETERMINATION OF PRIMES BY THE USE OF INTEGRAL BINARY QUADRATIC FORMS

WILLIAM EDWARD CHRISTILLES, St. Mary's University, Texas

In several previous issues of the MATHEMATICS MAGAZINE, we have discussed the factorization of integers by the use of integral binary quadratic forms. In this paper, we analyze our previous work along with some of the modern advances in number theory in order that we may develop an elementary approach, not only for the factorization of integers, but also as an aid in the determination of primes.

Before continuing, we introduce the following transformation:

Let $M=r_1r_2$, where r_1 and r_2 are any pair of integral factors of M, be an integer properly represented in form,

$$a\alpha^2 + b\alpha\gamma + c\gamma^2, \tag{1}$$

of discriminant $D=4ac-b^2$ where D is not the square of an integer, in integers α and γ with integral coefficients. It follows from Theorem 70 [1] that:

$$r_1 = a_1\alpha_1^2 + b_1\alpha_1\gamma_1 + c_1\gamma_1^2 \tag{2}$$

$$r_2 = a_2\alpha_2^2 + b_2\alpha_2\gamma_2 + c_2\gamma_2^2, \tag{3}$$

where $D=4ac-b^2=4a_1c_1-b_1^2=4a_2c_2-b_2^2$, in integers α_1, α_2, γ_1, and γ_2, with integral coefficients. We can now write:

$$16hM = (u^2+Dv^2)(m^2+Dn^2) = x^2+Dy^2, \tag{4}$$

where $u = (2a_1\alpha_1 + b_1\gamma_1)$, $m = (2a_2\alpha_2 + b_2\gamma_2)$, $v = \gamma_1$, $n = \gamma_2$, $y = (un \mp vm)$, $x=(um \pm Dvn)$, $h=a_1a_2$, $M=r_1r_2$, and D is the discriminant. Thus we have transformed an integer of form (1) into an integer of form

$$x^2 + Dy^2, \tag{5}$$

in integers x and y, of discriminant D, by multiplying M by $16h$.

We now introduce our major theorem:

THEOREM 1 [2]. *If M is an integer, represented properly by form (1), $a\alpha^2 + b\alpha\gamma + c\gamma^2$, of discriminant D, and r_1 and r_2 are any pair of positive integral divisors of M where $M=r_1r_2$, there exist integers x, y, h, w, D, and B, such that:*

$$\text{(6)}\qquad \begin{aligned} &\text{(i)}\quad 16hM = x^2 + Dy^2 \\ &\text{(ii)}\quad y^2 - 4B(DB - x) = w^2. \end{aligned}$$

Moreover, there are integers v, n, a_1, and a_2, where $B=vn$, and $h=a_1a_2$, such that:

$$M = r_1r_2 = \left[\frac{\left(\dfrac{w+y}{2n}\right)^2 + Dv^2}{4a_1}\right]\left[\frac{\left(\dfrac{w-y}{2v}\right)^2 + Dn^2}{4a_2}\right],$$

where

$$\left(\frac{w-y}{2v}\right),\ \left(\frac{w+y}{2n}\right),\ \left[\frac{\left(\frac{w+y}{2n}\right)^2+Dv^2}{4a_1}\right],\quad \textit{and}\quad \left[\frac{\left(\frac{w-y}{2v}\right)^2+Dn^2}{4a_2}\right],$$

are integers.

Now we know from the transformation expressed by (1), (2), (3), and (4), that:

$$16hM = (u^2+Dv^2)(m^2+Dn^2). \tag{7}$$

Thus, from (7), we have:

$$(DB)^2 = 16hM - [(um)^2 + D(vm)^2 + D(un)^2], \tag{8}$$

where $B=vn$. It follows that,

$$(DB)^2 \leqq 16hM, \tag{9}$$

where $D>0$. Consequently,

$$B^2 \leqq \frac{16hM}{D^2} \tag{10}$$

or,

$$-\left|\frac{4\sqrt{hM}}{D}\right| \leqq B \leqq \left|\frac{4\sqrt{hM}}{D}\right|, \tag{11}$$

where $D>0$.

At this point, the author wishes to call attention to an error, and a misleading remark he made in a statement following the proof of his Theorem [3]. The statement, "This is a simple way of proving that if a number, M, be expressed in more than one distinct way in the form, x^2+pqy^2, for a given p and q, then the number is composite and its factors easily determined," is true only if $pq>0$ [4]. If $pq<0$, and not the square of an integer, then we have multiple representations of M in the form, x^2+pqy^2, since the Pell Equation, $u^2-Kv^2=+1$, always has a solution in integers where $K>0$ and not the square of an integer. As an example,

$$13 = (4)^2 - 3(1)^2 = [(4)^2-3(1)^2][(2)^2-3(1)^2] = (5)^2-3(2)^2. \tag{12}$$

Now, as may be seen by examination of the proof of the Theorem [5],

$$B = vn, \tag{13}$$

and hence an integral solution of B always exists if $M=r_1r_2$ is composite. From

$$B = \frac{x \pm \sqrt{x^2-pq(w^2-y^2)}}{2q} = \frac{x\pm z}{2q}, \text{ [6]} \tag{14}$$

there may exist solutions of B which are not integral if we only require that (6), [7] $x^2+pqy^2=z^2+pqw^2$, be satisfied. These values of B will, however, always be rational with the conditions on M and its factors. Hence, to find the factors r_1 and r_2 of M, from (6) [8], we may proceed as follows:

From (13) and (14), we have,

$$B = vn = \frac{x \pm z}{2q} = \frac{t}{2q}, \tag{15}$$

where $t=(x \pm z)$. Let,

$$|vn| = \left(\frac{\sqrt{t}}{\sqrt{2q}}\right)\left(\frac{\sqrt{t}}{\sqrt{2q}}\right), \tag{16}$$

where $|v|=|n|=\sqrt{t}/\sqrt{2q}$. (It is immaterial whether either v or n is positive or negative, since due to (3) of the Theorem [9], whenever either v or n occurs it appears as a square.) Consequently,

$$M = r_1r_2 = \left[p\left(\frac{w-y}{2v}\right)^2 + qn^2\right]\left[p\left(\frac{w+y}{2n}\right)^2 + qv^2\right] \tag{17}$$

$$= \left[\frac{pq(w-y)^2+t^2}{2t}\right]\left[\frac{pq(w+t)^2+t^2}{2t}\right]$$

$$= \frac{[pq(w-y)^2+(x\pm z)^2][pq(w+y)^2+(x\pm z)^2]}{4(x\pm z)^2} \tag{18}$$

where $t=x\pm z$, $B=vn$, and $|v|=|n|=\sqrt{t}/\sqrt{2q}$. To apply the results of (17) through (18) to Theorem 1, we have only to let $p=1$, $q=D$, and replace M by $16hM$.

A similar approach if applied to Theorem 1 [10], where M is represented in more than one distinct way in form

$$px^2 + qy^2, \tag{19}$$

in integers x and y, with integral coefficients, will yield the following results:

$$M = r_1r_2 = \left[p\left(\frac{w-y}{2vp}\right)^2 + qn^2\right]\left[\left(\frac{w+y}{2n}\right)^2 + pqv^2\right]$$

$$= \frac{[q(w-y)^2+pt^2][q(w+y)^2+pt^2]}{4pt^2} \tag{20}$$

$$= \frac{[q(w-y)^2+p(x\pm z)^2][q(w+y)^2+p(x\pm z)^2]}{4p(x\pm z)^2}$$

where $t=x\pm z$, $B=vn$, and $|v|=|n|=(\sqrt{t}/\sqrt{2q})$.

Now, to factor an integer, M, properly represented in form (1), of discriminant D, by use of Theorem 1, we must satisfy both (i) and (ii) of (6). Once (i) is satisfied, we then attempt to satisfy (ii) by successfully substituting, for B, in

(ii) the integers $\pm 1, \pm 2, \pm 3, \cdots$. Thus, in general the rapidity of this method is seen to depend upon the size of the absolute value of B. Since by (11), we can place limits on the range of values of B when $D>0$, we exclude for the remainder of this paper the case where $D \leqq 0$.

We further note, due to Theorem 57 [11], that every integer, M, properly represented in form (1), of discriminant $D>0$, has only a finite number, n, of reduced forms. Let us denote these n reduced forms by f_i, where $i=1, 2, 3, \cdots, n$. Thus M is the product of some two of the n reduced forms, say f_j and f_k, for some j and k where $1 \leqq j \leqq n$ and $1 \leqq k \leqq n$, unless M is a prime. Hence M may be represented as follows:

$$M = r_1 r_2 = f_j f_k, \tag{21}$$

where $r_1 = f_j$ and $r_2 = f_k$ are any pair of integral factors satisfying (21) and each of M, f_j, and f_k, is of form (1), of discriminant $D>0$.

Now, for all possible

$$f_i = a_i \alpha_i^2 + b_i \alpha_i \gamma_i + c_i \gamma_i^2, \tag{22}$$

where $1 \leqq i \leqq n$, of discriminant $D>0$, we will have a finite number of h_q's (since the number of reduced forms is finite), such that,

$$h_q = a_j a_k, \tag{23}$$

where $1 \leqq j \leqq n$, $1 \leqq k \leqq n$, and $1 \leqq q \leqq [C(n, 2)+n]$. Correspondingly, a finite number of B_q's exist, such that,

$$B_q = \gamma_j \gamma_k. \tag{24}$$

Thus, in general (11) may be written as,

$$\left| \frac{4\sqrt{h_q M}}{D} \right| \leqq B_q \leqq \left| \frac{4\sqrt{h_q M}}{D} \right|, \tag{25}$$

where $D>0$. Consequently, if M does not have a solution by Theorem 1, over the range of values of B_q, expressed by (25), then M is a prime integer. By a solution of M is meant the process of finding a B_q which will yield two integral factors $r_1>1$ and $r_2>1$ such that $M=r_1 r_2$.

At this point the author wished to caution the reader concerning the application of Theorem 1. Although the restriction that the discriminant, D, not be positive and not the square of an integer guarantees that none of a_i, b_i, or c_i be zero when dealing with reduced forms, some of the α_i or γ_i may be zero. This would be undesirable from the standpoint of the factorization of a given integer, M, properly represented in form (1), by Theorem 1. Hence, we should first test to see if any of the a_i's or c_i's are factors of M. If one of the a_i or c_i, other than unity, is a factor of M (which will occur if one of the α_i or γ_i is zero) our objective is accomplished; if not, we then proceed with the use of Theorem 1.

In conclusion, we observe from (25) that the factorization of an integer M, properly represented in form (1), of discriminant $D>0$, is most practically obtained by seeking a representation of M in form (1) having as large a discriminant D as possible with relatively few reduced forms.

Example. Let $M=9769$. We find that,

$$M = \alpha^2 + \alpha\gamma + 67\gamma^2 = (98)^2 + (98)(1) + 67(1)^2,$$

where $\alpha=98$, $\gamma=1$, and $D=267$. Now by referring to table 1 [12], we find that,

$$a_1 = 1, \quad b_1 = 1, \quad c_1 = 67, \quad a_2 = 3, \quad b_2 = 3, \quad c_2 = 23.$$

Hence,

$$h_1 = 1, \quad h_2 = 3, \quad \text{and} \quad h_3 = 9.$$

Thus, we have:

$$16h_1M = 4[4\{(98)^2 + (98)(1) + 67(1)^2\}] = 4[(197)^2+267(1)^2] = (394)^2 + 267(2)^2$$

$$16h_3M = 3^2(16M) = (1182)^2 + 267(6)^2.$$

We need not consider $16h_2M$ since it may easily be shown by congruences that $16h_2M$ cannot be represented in form (5) where $D=267$. Now, from (25), we have:

$$-1 \leqq B_1 \leqq 1 \quad \text{and} \quad -4 \leqq B_3 \leqq 4.$$

Thus, either $M=9769$ is a prime or there exists a solution of B_i, where $i=1$ or 3. But by substituting in (ii) of Theorem 1, the integers $\pm 1, \pm 2, \cdots$, over the specified ranges, we find that no such B exists. Hence, 9769 is a prime integer.

References

1. Leonard Eugene Dickson, Introduction to the Theory of Numbers, New York, 1957, p. 95.
2. William Edward Christilles, A Simple Approach to the Factorization of Integers, MATHEMATICS MAGAZINE, **34** (1961) 412–413.
3. William Edward Christilles, A Note on the Factorization of Integers, MATHEMATICS MAGAZINE, **33** (1960) 284–285.
4. Leonard Eugene Dickson, History of the Theory of Numbers, Vol. I, New York, 1952, p. 364.
5. Christilles, *op. cit.*, pp. 284–285.
6. *Ibid.*, p. 285.
7. *Ibid.*
8. *Ibid.*
9. *Ibid.*, p. 284.
10. William Edward Christilles, A Simple Approach to the Factorization of Integers, MATHEMATICS MAGAZINE, **34** (1961) 412–413.
11. Leonard Eugene Dickson, Modern Elementary Theory of Numbers, Chicago, 1943, p. 56.
12. Leonard Eugene Dickson, Introduction to the Theory of Numbers, New York, 1957, p. 85.

3×3 MATRICES FROM KNIGHT'S MOVES

CHARLES W. TRIGG, Los Angeles City College

All of the outside cells of a 3×3 matrix can be entered in a continuous path of Knight's moves. If each cell be numbered according to the order of its entry, two different arrays are obtained depending upon the starting point, be it a corner or the middle cell of a side. The central cell may be numbered 9. The sums of the rows and columns of the arrays are indicated below on the left and bottom of the arrays. The sums of the diagonals and of the broken diagonals are on the top and right. The sum of the elements in the three-celled gnomon at each corner is placed at that corner. The sum of the vertices of the isosceles triangles on each side of the square is placed outside the array near the vertex opposite the base.

			12			
	11	11	15	19	13	
	12	1	4	7	21	
18	17	6	9	2	9	6
	16	3	8	5	15	
	17	10	21	14	15	
			16			

			11			
	10	18	6	21	12	
	11	6	1	4	12	
9	19	3	9	7	16	21
	15	8	5	2	17	
	16	17	15	13	14	
			15			

In the array starting at a corner, the corners are odd and the middle cells are even. The gnomon sums are odd, and the vertex sums are even. In the array starting at a middle cell, the parities are reversed.

Indeed, one array goes into the other by a 45° clockwise rotation. Thus, outside columns and rows go into gnomons, and conversely. Central columns and rows go into diagonals, and conversely. Broken diagonals become vertices of isosceles triangles, and conversely.

It follows that the twenty sums in each array are the same, namely: 6, 9, 10, 11, 11, 12, 12, 13, 14, 15, 15, 15, 16, 16, 17, 17, 18, 19, 21, 21. Furthermore, since each digit except 9 contributes to seven sums and the 9 appears in four, the sum of the sums is $7(36)+4(9)$ or 288.

THE PRODUCT OF TWO EULERIAN POLYNOMIALS

L. CARLITZ, Duke University

The Bernoulli and Euler polynomials can be defined by means of

$$\frac{te^{xt}}{e^t-1}=\sum_{m=0}^{\infty} B_m(x)\frac{t^m}{m!},\qquad \frac{2e^{xt}}{e^t+1}=\sum_{m=0}^{\infty} E_m(x)\frac{t^m}{m!}.$$

The formula

$$B_m(x)B_n(x)=\sum_r\left\{\binom{m}{2r}n+\binom{n}{2r}m\right\}\frac{B_{2r}B_{m+n-2r}(x)}{m+n-2r} \tag{1}$$

is proved in Nielsen's book [3, p. 75]; a different proof occurs in [2]. Nielsen also obtains similar formulas for

$$E_m(x)E_n(x)\quad\text{and}\quad E_m(x)B_n(x).$$

The Eulerian polynomial $H_m(x|\lambda)$ can be defined by means of

$$\frac{(1-\lambda)e^{xt}}{e^t-\lambda}=\sum_{m=0}^{\infty} H_m(x\,|\,\lambda)\frac{t^m}{m!}; \tag{2}$$

for properties of $H_m(x|\lambda)$ see for example [1]. Since

$$H_m(x\,|-1)=E_m(x),$$

it may be of interest to get a formula for the product of two Eulerian polynomials.

We assume that $\alpha\neq 1$, $\beta\neq 1$, $\alpha\beta\neq 1$. It follows from (2) that

$$\begin{aligned}
\sum_{m,n=0}^{\infty} H_m(x\,|\,\alpha)H_n(x\,|\,\beta)\frac{u^mv^n}{m!n!}&=\frac{(1-\alpha)e^{xu}}{e^u-\alpha}\,\frac{(1-\beta)e^{xv}}{e^v-\beta}\\
&=\frac{(1-\alpha)(1-\beta)}{1-\alpha\beta}\,\frac{(1-\alpha\beta)e^{x(u+v)}}{e^{u+v}-\alpha\beta}\,\frac{e^{u+v}-\alpha\beta}{(e^u-\alpha)(e^v-\beta)}\\
&=\frac{(1-\alpha)(1-\beta)}{1-\alpha\beta}\,\frac{(1-\alpha\beta)e^{x(u+v)}}{e^{u+v}-\alpha\beta}\left\{1+\frac{\alpha}{e^u-\alpha}+\frac{\beta}{e^v-\beta}\right\}\\
&=\frac{1}{1-\alpha\beta}\sum_{m,n=0}^{\infty} H_{m+n}(x\,|\,\alpha\beta)\frac{u^mv^n}{m!n!}\\
&\quad\cdot\left\{(1-\alpha)(1-\beta)+\alpha(1-\beta)\sum_{r=0}^{\infty} H_r[\alpha]\frac{u^r}{r!}+\beta(1-\alpha)\sum_{s=0}^{\infty}\mathrm{H}_s[\beta]\frac{v^s}{s!},\right.
\end{aligned}$$

where we have put

$$H_r[\alpha]=H_r(0\,|\,\alpha) \tag{3}$$

the so-called Eulerian number. Comparison of coefficients evidently yields

$$H_m(x\,|\,\alpha)H_n(x\,|\,\beta)=H_{m+n}(x\,|\,\alpha\beta)$$

Supported in part by National Science Foundation grant G16485.

$$(4)\qquad + \frac{\alpha(1-\beta)}{1-\alpha\beta} \sum_{r=1}^{m} \binom{m}{r} H_r[\alpha] H_{m+n-r}(x \mid \alpha\beta) + \frac{\beta(1-\alpha)}{1-\alpha\beta} \sum_{s=1}^{n} \binom{n}{s} H_s[\beta] H_{m+n-s}(x \mid \alpha\beta),$$

provided $\alpha \neq 1$, $\beta \neq 1$, $\alpha\beta \neq 1$.

In the next place we have

$$\sum_{m,n=0}^{\infty} B_m(x) H_n(x \mid \alpha) \frac{u^m v^n}{m!n!} = \frac{ue^{xu}}{e^u - 1} \frac{(1-\alpha)e^{xv}}{e^v - \alpha}$$

$$= u \frac{(1-\alpha)e^{x(u+v)}}{e^{u+v} - \alpha} \frac{e^{u+v} - \alpha}{(e^u - 1)(e^v - \alpha)}$$

$$= \frac{(1-\alpha)e^{x(u+v)}}{e^{u+v} - \alpha} \left\{ u + \frac{u}{e^u - 1} + \frac{\alpha u}{e^v - \alpha} \right\}$$

$$= \sum_{m,n=0}^{\infty} H_{m+n}(x \mid \alpha) \frac{u^m v^n}{m!m!} \cdot \left\{ u + \sum_{r=0}^{\infty} B_r \frac{u^r}{r!} + \frac{u}{1-\alpha} \sum_{s=0}^{\infty} H_s[\alpha] \frac{v^s}{s!} \right\}.$$

It follows that

$$(5)\qquad B_m(x) H_n(x \mid \alpha) = mH_{m+n-1}(x \mid \alpha) + \sum_{r=0}^{m} \binom{m}{r} B_r H_{m+n-r}(x \mid \alpha) + \frac{m\alpha}{1-\alpha} \sum_{s=0}^{n} \binom{n}{s} H_s[\alpha] H_{m+n-s-1}(x \mid \alpha),$$

provided $\alpha \neq 1$.

If $\alpha \neq 1$ but $\alpha\beta = 1$ we take

$$(u+v) \sum_{m,n=0}^{\infty} H_m(x \mid \alpha) H_n(x \mid \alpha^{-1}) \frac{u^m v^n}{m!n!} = (u+v) \frac{(1-\alpha)e^{xu}}{e^u - \alpha} \frac{(1-\alpha^{-1})e^{xv}}{e^v - \alpha^{-1}}$$

$$= (1-\alpha)(1-\alpha^{-1}) \frac{(u+v)e^{x(u+v)}}{e^{u+v} - 1} \left\{ 1 + \frac{\alpha}{e^u - \alpha} + \frac{\alpha^{-1}}{e^v - \alpha^{-1}} \right\}.$$

This implies

$$mH_{m-1}(x \mid \alpha) H_n(x \mid \alpha^{-1}) + nH_m(x \mid \alpha) H_{n-1}(x \mid \alpha^{-1})$$

$$= (1-\alpha)(1-\alpha^{-1}) B_{m+n}(x) - (1-\alpha) \sum_{r=0}^{m} \binom{m}{r} H_r[\alpha] B_{m+n-r}(x)$$

$$- (1-\alpha^{-1}) \sum_{s=0}^{n} \binom{n}{s} H_s[\alpha^{-1}] B_{m+n-s}(x)$$

$$= -(1-\alpha) \sum_{r=1}^{m} \binom{m}{r} H_r[\alpha] B_{m+n-r}(x)$$

$$- (1-\alpha^{-1}) \sum_{s=1}^{n} \binom{n}{s} H_s[\alpha^{-1}] B_{m+n-s}(x).$$

Since

$$\frac{\partial}{\partial x} H_n(x \mid \alpha) = nH_{n-1}(x \mid \alpha),$$

it is clear from (6) that

$$(7)\quad \begin{aligned} H_m(x \mid \alpha) H_n(x \mid \alpha^{-1}) = &- (1-\alpha) \sum_{r=0}^{m-1} \binom{m}{r+1} H_{r+1}[\alpha] \frac{B_{m+n-r}(x)}{m+n-r} \\ &- (1-\alpha^{-1}) \sum_{s=0}^{n-1} \binom{n}{s+1} H_{s+1}[\alpha^{-1}] \frac{B_{m+n-s}(x)}{m+n-s} + C_{m,n}, \end{aligned}$$

where $C_{m,n}$ is independent of x. To determine $C_{m,n}$ we notice first that (6) and (7) imply

$$mC_{m-1,n} + nC_{m,n-1} = 0,$$

so that

$$C_{m,n} = -\frac{n}{m+1} C_{m+1,n-1}.$$

Repeated application of this recursion leads to

$$(8)\quad C_{m,n} = (-1)^n \frac{m!n!}{(m+n)!} C_{m+n,0}.$$

Now if we put $n=0$, $x=0$ in (7) we get

$$\begin{aligned} H_m[\alpha] &= -(1-\alpha) \sum_{r=0}^{m-1} \binom{m}{r+1} H_{r+1}[\alpha] \frac{B_{m-r}}{m-r} + C_{m,0} \\ &= -\frac{1-\alpha}{m+1} \sum_{r=1}^{m} \binom{m+1}{r} H_r[\alpha] B_{m-r+1} + C_{m,0}. \end{aligned}$$

Similarly (5) implies

$$B_{m+1} = (m+1)H_m[\alpha] + \sum_{r=0}^{m+1} \binom{m+1}{r} H_r[\alpha] B_{m-r+1} + \frac{(m+1)\alpha}{1-\alpha} H_m[\alpha],$$

so that

$$(m+1)C_{m,0} = -(1-\alpha)H_{m+1}[\alpha].$$

Therefore by (8)

$$(9)\quad C_{m,n} = (-1)^{n+1} \frac{m!n!}{(m+n+1)!} (1-\alpha) H_{m+n+1}[\alpha].$$

(Since

$$H_n[\alpha^{-1}] = (-1)^n H_n[\alpha],$$

the right member of (9) remains unchanged when we interchange m and n and replace α by α^{-1}).

Combining (7) and (9) we get

$$
\begin{aligned}
H_m(x \mid \alpha)H_n(x \mid \alpha^{-1}) = &- (1-\alpha)\sum_{r=1}^{m}\binom{m}{r}H_r[\alpha]\frac{B_{m+n-r+1}(x)}{m+n-r+1} \\
&- (1-\alpha^{-1})\sum_{s=1}^{n}\binom{n}{s}H_s[\alpha^{-1}]\frac{B_{m+n-s+1}(x)}{m+n-s+1} \\
&+ (-1)^{n+1}\frac{m!n!}{(m+n+1)!}(1-\alpha)H_{m+n+1}[\alpha],
\end{aligned}
\tag{10}
$$

where of course $\alpha \neq 1$.

In particular if we take $\alpha = -1$, (5) and (10) reduce to

$$
\begin{aligned}
B_m(x)E_n(x) = &\; E_{m+n}(x) + \sum_{r=2}^{m}\binom{m}{r}B_rE_{m+n-r}(x) \\
&- \frac{m}{2}\sum_{s=1}^{n}\binom{n}{s}2^{-s}C_sE_{m+n-s-1}(x),
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
E_m(x)E_n(x) = &- 2\sum_{r=1}^{m}\binom{m}{r}2^{-r}C_r\frac{B_{m+n-r+1}(x)}{m+n-r+1} \\
&- 2\sum_{s=1}^{n}\binom{n}{s}2^{-s}C_s\frac{B_{m+n-s+1}(x)}{m+n-s+1} \\
&+ (-1)^{n+1}2^{-m-n}\frac{m!n!}{(m+n+1)!}C_{m+n+1},
\end{aligned}
\tag{12}
$$

where [4, p. 28]

$$C_n = 2^nE_n(0) = (2-2^{-n})\frac{B_{n+1}}{n+1}\cdot$$

The formulas (11) and (12) may be compared with [3, p. 77, formulas (12), (16)].

We note also that since

$$\int_0^1 B_m(x)dx = \frac{B_{m+1}(1) - B_{m+1}(0)}{m+1} = 0 \qquad (m \geqq 1),$$

(10) yields

$$\int_0^1 H_m(x \mid \alpha)H_n(x \mid \alpha^{-1})dx = (-1)^{n+1}\frac{m!n!}{(m+n+1)!}(1-\alpha)H_{m+n+1}[\alpha] \tag{13}$$

$$(m \geqq 1,\ n \geqq 1).$$

Finally we remark that (4), (5) and (10) imply the following special formulas:

$$(14)\qquad H_m(x \mid \alpha) = H_m(x \mid \beta) + \frac{\alpha - \beta}{1 - \beta} \sum_{r=1}^{m} \binom{m}{r} H_r[\alpha] H_{m-r}(x \mid \beta) \qquad (\alpha \neq 1, \beta \neq 1),$$

$$(15)\qquad B_m(x) = mH_{m-1}(x \mid \alpha) + \sum_{r=0}^{m} \binom{m}{r} B_r H_{m-r}(x \mid \alpha) \qquad (\alpha \neq 1),$$

$$(16)\qquad H_m(x \mid \alpha) = -\frac{1 - \alpha}{m + 1} \sum_{r=1}^{m+1} \binom{m+1}{r} H_r[\alpha] B_{m-r+1}(x) \qquad (\alpha \neq 1).$$

It is not difficult to prove these formulas directly. For example (14) follows easily from the identity

$$\frac{(1 - \alpha)e^{xu}}{e^u - \alpha} = \frac{1}{1 - \beta} \left\{ 1 - \alpha + (\alpha - \beta) \frac{1 - \alpha}{e^u - \alpha} \right\} \frac{(1 - \beta)e^{xu}}{e^x - \beta} .$$

References

1. L. Carlitz, Eulerian numbers and polynomials, MATHEMATICS MAGAZINE, **32** (1959) 247–260.
2. L. Carlitz, Note on the integral of the product of several Bernoulli polynomials, Journal of the London Mathematical Society, **34** (1959) 361–363.
3. N. Nielsen, Traité élémentaire des nombres de Bernoulli, Paris, 1923.
4. N. E. Norlund, Vorlesungen über Differenzenrechnung, Berlin, 1924.

NUMBER THEORY

A pump's a composite of handle and spout
That has to be primed, or nothing comes out.
A gun's a composite of barrel and butt
That has to be primed, or nothing will sput.
In the arts, composition is carefully timed
And one doesn't begin till the surface is primed.
You will find composition is easy to do
When you start with a primer and carry it thru.

MARLOW SHOLANDER

A NOTE ON LAURENT EXPANSIONS

MILTIADES S. DEMOS, Drexel Institute of Technology

Most elementary textbooks on function theory give methods for obtaining Laurent series for rational functions, and only very special cases of transcendental functions. For example the series for $1/\sin z$ is obtained only for the region $0<|x|<\pi$. We shall obtain a series for $1/\sin z$ in the region $\pi<|z|<2\pi$; the same method can be used for other functions.

$1/\sin z$ is analytic in $0<|z|<2\pi$ except for poles at π and $-\pi$, with principal parts $-1/(z-\pi)$ and $-1/(z+\pi)$ respectively. We first obtain a Laurent series for

$$f(z) = \frac{1}{\sin z} + \frac{1}{z-\pi} + \frac{1}{z+\pi} \quad \text{valid for} \quad 0 < |z| < 2\pi.$$

$$\left(\text{We define } f(\pi) = \frac{1}{2\pi},\ f(-\pi) = -\frac{1}{2\pi}\right)$$

$$\frac{1}{\sin z} = \frac{1}{z\left(1 - \frac{z^2}{6} + \frac{z^4}{120}\cdots\right)} = S_1 = 1 + \frac{z}{6} + \frac{7z^3}{360}\cdots \qquad 0 < |z| < \pi$$

$$\frac{1}{z-\pi} + \frac{1}{z+\pi} = \frac{2z}{z^2-\pi^2} = -\frac{2z}{\pi^2}\left[\frac{1}{1-\frac{z^2}{\pi^2}}\right] = S_2 = \frac{-2z}{\pi^2}\left(1 + \frac{z^2}{\pi^2} + \frac{z^4}{\pi^4}\cdots\right)$$

$$|z| < \pi$$

$$f(z) = S_1 + S_2 = S_3 = \frac{1}{z} + \left(\frac{1}{6} - \frac{2}{\pi^2}\right)z + \left(\frac{7}{360} - \frac{2}{\pi^4}\right)z^3\cdots$$

$$0 < |z| < \pi$$

But $f(z)$ is analytic for $0<|z|<2\pi$ and has a unique Laurent expansion there, which must be S_3. That is $f(z)=S_3$ is valid for $0<|z|<2\pi$.

Now, we find a Laurent expansion for $2z/(z^2-\pi^2)$ valid for $|z|>\pi$.

$$\frac{2z}{z^2-\pi^2} = \frac{2}{z}\,\frac{1}{1-\frac{\pi^2}{z^2}} = S_4 = \frac{2}{z}\left(1 + \frac{\pi^2}{z^2} + \frac{\pi^4}{z^4} + \cdots\right) \qquad |z| > \pi$$

Then $S_3 - S_4$ represents

$$f(z) - \frac{2z}{z^2-\pi^2} = \frac{1}{\sin z}$$

in the region $\pi<|z|<2\pi$, that is:

$$\frac{1}{\sin z} = \cdots - \frac{2\pi^4}{z^5} - \frac{2\pi^2}{z^3} - \frac{1}{z} + \left(\frac{1}{6} - \frac{2}{\pi^2}\right)z + \left(\frac{7}{360} - \frac{2}{\pi^4}\right)z^3\cdots$$

in the region $\pi<|z|<2\pi$.

HIERARCHIC ALGEBRA

MARCIA BECKMANN PRESTRUD, The RAND Corporation, Santa Monica, California

1. Introduction. Addition, subtraction, multiplication, division, exponentiation and evolution, or taking of roots, form an apparently tidy set of operations, but it is a set with some enticing anomalies.

For one thing, although we seem intuitively satisfied that the relationship between addition and subtraction is analogous to that between multiplication and division, it is clear that something is slightly different about the relationship between exponentiation and evolution. For another thing, addition and multiplication are commutative—why isn't exponentiation?

Such chinks in the mathematical jigsaw are worth mulling over—they sometimes lead to new pieces. In the present case they led to discovery of a widespread family of operations of which the six familiar members mentioned are only a small part.

As Bellman has pointed out in papers on invariant imbedding techniques, an individual phenomenon can often be best understood by being placed in perspective with comparable phenomena. We are hampered, however, in looking clearly at the inter-relationships of this newly extended family of operations by our traditional notion that algebraic operations *connect* two quantities ($a\times b$, for instance). In fact, in the expression $a\times b=c$, the multiplication sign not only says something explicit about a and b, it says something implicit about c (more later about this).

We should really like to be able to deal singly with quantities usually connected by an operation. This requires that we somehow split up the operation so we can parcel it out among the quantities it relates. Such a change of concept and a consequent change of notation will greatly facilitate our investigations.

An illuminating method for our purposes is to consider that each individual quantity has associated with it an operative which we shall call its *hierarchic index*. Operating on quantities with their respective indices can reduce them to a sort of ground state where combining them involves only addition.

For the record let us now make some formal definitions, which we shall then discuss in more detail.

We define the hierarchic index k of the real quantity a as:

$$(1)\qquad \begin{aligned} a^{(k)} &= \log a \text{ iterated } k \text{ times},\\ &\text{that is, } \underbrace{\log\log\cdots\log}_{k \text{ times}} a. \end{aligned}$$

Generally, $a^{(k)}$ is defined only for the region where $a^{(k-1)}\geqq 0$.

We define an inverse relation as:

$$(2)\qquad a^{(-k)} = \text{antilog } a \text{ iterated } k \text{ times}.$$

The general rule of composition for quantities and their associated indices turns out to be:

$$a^{(k)} + b^{(l)} + \cdots + c^{(m)} = d^{(n)}. \tag{3}$$

It will be convenient to refer to the process itself as:

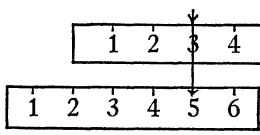
$$(k, l, \cdots, m{:}n). \tag{4}$$

Thus in our new notation, the familiar $a \times b = c$ is written as $a^{(1)} + b^{(1)} = c^{(1)}$, which simply says that adding log a and log b gives log c. The process by itself is (1, 1:1). We now have an inkling of what the multiplication sign says implicitly about c, as well as explicitly about a and b.

So far we have tacitly agreed that hierarchic indices are integers, which implies only discrete members in the set of algebraic operations. But what if we allow non-integral values for the indices? Then the discrete members are seen as part of a continuous range. Addition can now slither rather than jump into multiplication.

To get down to brass tacks, let us take a good hard look at a slide rule, since it has unwittingly been using hierarchic indices all along.

2. Traditional Algebraic Operations. Suppose we begin by looking at addition, multiplication, and exponentiation as measuring operations. The addition of two numbers can be carried out by measuring along one ruler or linear scale with another linear scale, and reading the result as a distance along the first scale.



$$2 + 3 = 5$$

FIG. 1

Similarly, multiplication can be thought of as measuring along one stick with another, but now the sticks are scaled logarithmically instead of linearly, resulting in a slide rule. Exponentiation, interestingly enough, involves two *dissimilar* scales—one loglogarithmic, and the other logarithmic. The exponentiend is found on the loglog scale, the exponent on the log scale, and the result again on the loglog scale. (Subtraction, division and evolution may be thought of as measured backward, or as measured forward on a scale composed of the inverses for the process.)

At this point curiosity sets in. What are the other operations in the set which includes addition, multiplication, and exponentiation, and how are they all related to each other? It turns out that these familiar operations can be thought of as part of Table 1.

We may now introduce the hierarchic index as an indicator for the scale of each number; it is a measure of the "loginess" of the scale. If 0 denotes a linear scale, 1 a log scale, and 2 a loglog scale, Table 1 becomes Table 2.

Since addition involves a linear scale for both addends and also for the result, we can represent it as:

lin, lin: lin, or $a^{(0)} + b^{(0)} = c^{(0)}$, $(0, 0: 0)$;

TABLE 1

	Operation	Scale for Second Number		
		linear	log	loglog
Scale for First Number (and Answer)	linear	addition		
	log		multiplication	
	loglog		exponentiation	

TABLE 2

	Operation	Hierarchic Index for Second Number		
		0	1	2
Hierarchic Index for First Number (and Answer)	0	addition		
	1		multiplication	
	2		exponentiation	

multiplication as:

$$\begin{array}{l} \text{log, log: log, or, } a^{(1)} + b^{(1)} = c^{(1)}, \\ (1\;,\;1\;:\;1\;); \end{array}$$

and exponentiation as:

$$\begin{array}{l} \text{loglog, log: loglog, or, } a^{(2)} + b^{(1)} = c^{(2)}, \\ (\;2\;\;,\;1:\;\;2\;). \end{array}$$

Throughout the following discussion we shall use logarithms to the base 2 for ease in manipulation. Now, as a numerical example of exponentiation in the new notation we can rewrite $4^8 = 65{,}536$ as:

$$4^{(2)} + 8^{(1)} = 65{,}536^{(2)},$$

$$\text{meaning} \left\{ \begin{array}{ccccc} \text{Loglog } 4 & + & \text{Log } 8 & = & \text{Loglog } 65{,}536 \\ 1 & + & 3 & = & 4 \end{array} \right\}.$$

3. New Algebraic Operations. At this point we may happily run riot with new operations to fill in Table 1, perhaps giving them names such as addiplication, etc.

TABLE 3

	linear	log	loglog
linear	addition	addiplication	
log	multidition	multiplication	
loglog		exponentiation	asteration

But the hierarchic index notation is more generally useful, and if we employ it to fill out the matrix of processes consistently we obtain Table 4.

TABLE 4

Hierarchic Index for Second Number

Hierarchic Index for First Number (and Answer)		0	1	2
	0	(addition) lin, lin: lin (0, 0 : 0)	lin, log: lin (0, 1 : 0)	lin, loglog: lin (0, 2 : 0)
	1	log, lin: log (1, 0 : 1)	(multiplication) log, log: log (1, 1 : 1)	log, loglog: log (1, 2 : 1)
	2	loglog, lin: loglog (2, 0 : 2)	(exponentiation) loglog, log: loglog (2, 1 : 2)	loglog, loglog: loglog (2, 2 : 2)

It is now apparent that we have been considering only "measuring-stick" processes—that is, processes whose result is read on the same scale as the first number. (Take another look at Figure 1.) But we needn't restrict ourselves to these; while addition is (lin, lin: lin), we might also have (lin, lin: log), (lin, lin: loglog), etc. As a matter of fact, considering only two numbers to be operated on, instead of just the nine processes in Table 4, there are 27 such processes, the permutations of 3 things 3 at a time with repetitions.

It is also apparent with a moment's thought that processes lying on the principal diagonal will be commutative and all others will not.

This two-dimensional matrix involving hierarchic indices of 0, 1, and 2 can clearly be extended to include higher processes with indices of 3 and above. Before we get carried away in this direction, however, we might recall that an estimate of Gamow's gives the number of atoms in the universe as less than 4^{4^4}—that is, $[4^{(2)}+[4^{(2)}+4^{(1)}]^{(-1)}]^{(-2)}$, where the brackets are used in the customary way to indicate order of operations. (In a later paper we shall investigate the rules of operational order for processes involving more than two quantities.)

And now for a whirl at using one of the new processes! It is the lower right-hand element in Tables 3 and 4, designated as asteration, or (2, 2: 2). On occasion it will be convenient to use the symbol $*$ for this process, to be analogous to $+$ and $\times$.

It is interesting to compare exponentiation, (2, 1: 2), with asteration, (2, 2: 2). For example:

Exponentiation:

$$4^2 = 16, \text{ or } 4^{(2)} + 2^{(1)} = 16^{(2)},$$

$$\text{means} \begin{Bmatrix} \text{Loglog } 4 + \text{Log } 2 = \text{Loglog } 16 \\ 1 \quad + \quad 1 = \quad 2 \end{Bmatrix}.$$

Asteration:

$$4 * 2 = 4, \text{ or } 4^{(2)} + 2^{(2)} = 4^{(2)},$$

$$\text{means} \begin{Bmatrix} \text{Loglog } 4 + \text{Loglog } 2 = \text{Loglog } 4 \\ 1 \quad + \quad 0 \quad = \quad 1 \end{Bmatrix}.$$

An asteration table, unlike a multiplication table, is not constant, but is a function of the log base for reasons which will become clear in the discussion of identities and inverses.

4. Fractional Processes. We are now ready for the consideration of fractional hierarchic indices. If we plot numerical values versus integral hierarchic indices

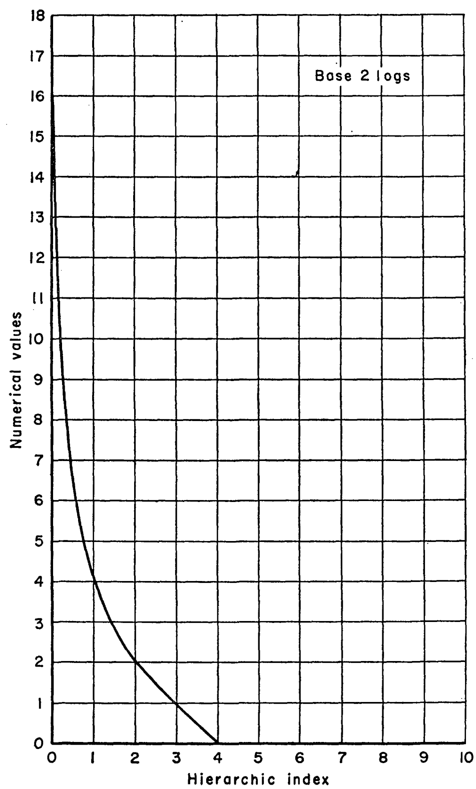


FIG. 2

for, say, the case where the basic number is 16 at index 0 (that is, 16 ready to be used as an addend) and connect these points with a smooth curve, we realize that we may think of processes from a continuous as well as from a discrete point of view (Figure 2).

The problem of ascertaining an equation for such a curve is not a simple one and will be pursued in a later paper. It is connected with the classical iteration problems of Abel and Schröder, to which we might digress for a moment.

Abel proposed linearizing a function $f(x)$ by addition:

$$\phi(f(x)) = \phi(x) + k.$$

Schröder proposed linearizing by multiplication:

$$\phi(f(x)) = c\phi(x),$$

or

$$f(x) = \phi^{-1}(c\phi(x)).$$

We can find iterates of $f(x)$ by means of Schröder's equation, namely:

$$\phi(f(f(x))) = c\phi(f(x)) = c^2\phi(x),$$

and for n in general,

$$f^{(n)}(x) = \phi^{-1}(c^n\phi(x)).$$

This relation can now be used to define $f^{(n)}$ for an arbitrary n, not necessarily integer. It is clear that this generalized iterate satisfies the fundamental semigroup relation

$$f^{(m+n)} = f^{(m)}(f^{(n)}).$$

The problem of determining when a function $\phi(x)$ satisfying Schröder's initial equation exists is a difficult one, which has been discussed at great length in the literature; see [1 and 2].

Returning from our detour, if we now plot more curves of the family of Figure 2 (see Figure 3), we are equipped to think about a process between addition and multiplication, for instance. Let us think of one lying on the matrix diagonal between them, say at an index of 1/2, and compare the three processes in an example.

Addition and multiplication may be approximated by the connecting plots in Figure 11, and for the present the intermediate process *must* be. (It is marked with a dotted line.)

Addition:

$$2 + 4 = 6, \quad \text{or} \quad 2^{(0)} + 4^{(0)} = 6^{(0)} \qquad \text{lin, lin: lin}$$

$$2 + 4 = 6$$

Fractional Process:

$$2^{(1/2)} + 4^{(1/2)} = 7.2^{(1/2)} \qquad \text{“linlog, linlog: linlog”}$$

$$1.5 + 2.8 = 4.3$$

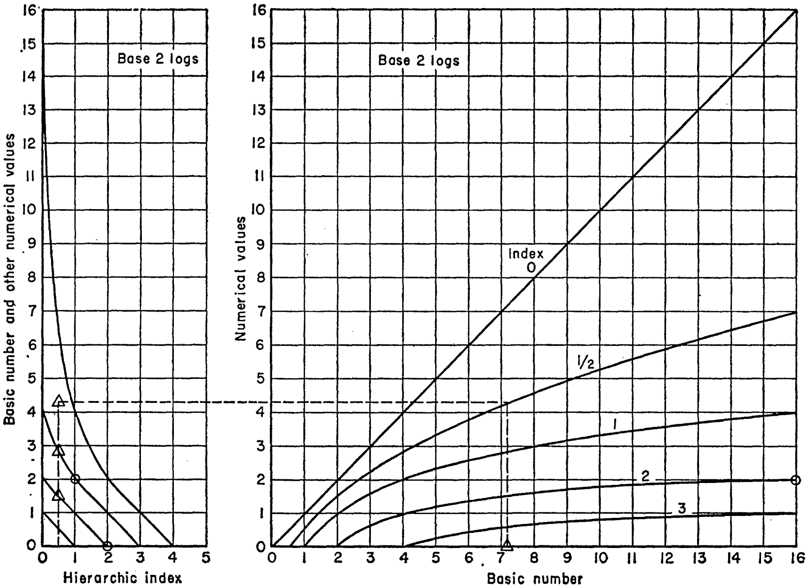


FIG. 3

Multiplication:

$$2 \times 4 = 8, \quad \text{or} \quad 2^{(1)} + 4^{(1)} = 8^{(1)} \qquad \text{Log, Log: Log}$$
$$1 + 2 = 3$$

The operation of exponentiation can also be carried out by means of Figure 3. As an example, take $2^4 = 16$, and look at the points in the figure marked with circles. Other non-commutative processes can be dealt with analogously.

5. Negative Processes. Extrapolation of the curves in Figure 3 into the region of negative hierarchic indices leads to the discovery that an index of -1 indicates the antilog, -2 the antiloglog, etc. (See Figure 4a.) Consideration of various fractional and transcendental numbers enables us to extend curves into the region of negative numerical values as well.

These extensions into the negative regions of Figure 4a are reflected in Figure 4b. As we expect, positive integral indices do not apply to negative basic numbers, and use of the negative integral indices cannot lead to negative numerical values. But most interestingly, these rules do not apply to indices lying between -1 and $+1$. For example, the basic number $-1/2$ does not have a logarithm (index 1), but it does have a "half-logarithm" (index 1/2).

In Figure 4a asymptotes for the curves lie at one index unit beyond their intersection with the abscissa. The asymptotes of Figure 4b may be inferred from those of Figure 4a.

6. Identities. Now a word about identities. In the present context it is more enlightening to think of identities as associated with indices rather than with algebraic operations. For the moment we shall confine ourselves to the prin-

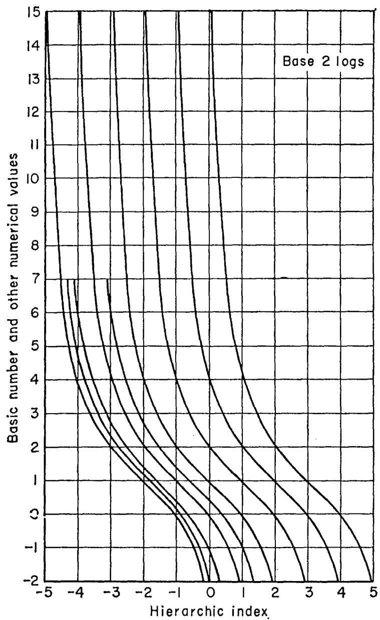


FIG. 4a

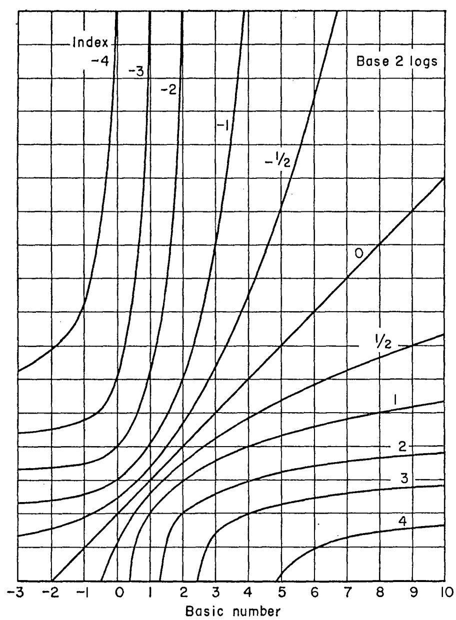


FIG. 4b

cipal-diagonal, commutative processes. Let us define as the identity for an index the basic number which, when operated on by the index, leads to 0. For index 0, it is 0 itself; for index 1 it is 1, since $\log 1=0$. (This is consistent with the identities we associate with addition and multiplication.) For index 2, it is the log base, since loglog (base) $=\log 1=0$. For index 3, it is $\text{base}^{\text{base}}$; for index 4, $\text{base}^{(\text{base}^{\text{base}})}$, etc. In view of Gamow's estimate mentioned previously, we can see that for index 4 with log base 4, the *identity* is roughly the number of atoms in the universe.

As a matter of fact, it is apparent that in Figure 4b the various identities lie along the abscissa, and that the non-integral indices too have their identities. (The identity of index 1/2 is approximately 1/2.)

The negative integral indices have no identities, but surprisingly the fractional indices between 0 and -1 do have. Index -1 approaches an identity at $-\infty$, but indices of greater negativity do not even approach identities.

Once we leave the principal diagonal, we might meditate about identities for exponentiation. 1, the identity for index 1, is suitable as an identity for the exponent, as expected, but the base, identity for index 2, is not suitable as an identity for the exponentiend. As a matter of fact, the suitable "identity" is a variable, $n^{1/n}$, not a constant. (The other "measuring-stick" processes off the main diagonal in Table 4 are similar, and the non-measuring-stick processes, still less well behaved, would appear to have no usefully defined identities at all.)

The identity for a given index may sometimes profitably be defined in usage with higher indices (0 to be used in multiplication, for example). It is worth noting in passing that juggling identities and ∞ has led mathematicians to some interesting gymnastics in the interest of avoiding paradoxes, sometimes to the tiptoe irresolution of indeterminates.

Using the symbol $*$ for process (2, 2: 2) again, by analogy with multiplication we obtain:

$$\begin{array}{lcl} 1 * 1 = 1 & & 0 \times 0 = 0 \\ 1 * 2 = 1 & & 0 \times 1 = 0 \\ 2 * 2 = 2 & , & 1 \times 1 = 1 \\ \infty * 1 = \text{indeterminate} & & \infty \times 0 = \text{indeterminate} \\ \infty * 2 = \infty & & \infty \times 1 = \infty \end{array} \tag{5}$$

It is not immediately clear whether judicious definitions involving 0 are also possible for this process of asteration.

7. Inverses. We arrive at the inverses for the various indices by using the relationship:

$$\text{number}^{(\text{index})} + \text{inverse}^{(\text{index})} = \text{identity}^{(\text{index})}. \tag{6}$$

Since we defined the identities as we did, we must define the inverses accordingly; namely, operating on the inverse with the given index must lead to the negative of the original number (just as operating on the identity with the index must lead to 0).

Using process (2, 2: 2) as an example of principal diagonal processes, we discover the inverses to be $\text{base}^{\text{base}^{(-\log\log n)}}$. As an example:

$$256^{(2)} + (\sqrt[8]{2})^{(2)} = 2^{(2)},$$

that is,

$$[2^{(2^3)}]^{(2)} + [2^{(2^{-3})}]^{(2)} = 2^{(2)}.$$

In the case of exponentiation and other off-diagonal processes, it is not sensible to talk about inverses as defined by the foregoing relationship with identities. Not only have we a decision among three possible identities for exponentiation, but one of them is a variable.

8. Inter-Process Laws. The rule of composition discussed in the introduction takes into account only an additive relationship among indexed quantities. Using the notation in Table 4, however, a generalized relationship for two quantities can be written as follows:

$$[a^{(f)}(p, q\colon r)b^{(g)}]^{(h)} = [a^{(k)}(s, t\colon w)b^{(l)}]^{(m)}, \tag{7}$$

where

$$f + p = k + s,$$
$$g + q = l + t,$$
$$h - r = m - w.$$

This relationship points out that $(p, q\colon r)$ and $(s, t\colon w)$ can be considered to be operators such as multiplication, asteration, etc., as well as addition. For instance, 4×3 can be written in a number of ways. If we write:

$$[4^{(1)}(0,\ 0\colon 0)3^{(1)}]^{(-1)} = [4^{(0)}(1,\ 1\colon 1)3^{(0)}]^{(0)},$$

we are saying that the addition of logarithms (the left) is equal to the multiplication of linear quantities (the right). We may say further that another expression equivalent to the above is:

$$[4^{(-1)}(2,\ 2\colon 2)3^{(-1)}]^{(1)};$$

that is, that asteration of antilogs leads to the same result.

We have so far considered only processes involving two processands. If we add a third, we may think of them as illustrated in Figure 5.

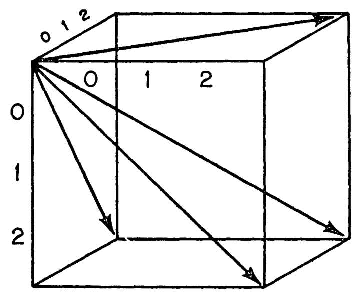


FIG. 5

The diagonals shown on the faces of the cube mark commutative processes, and the principal diagonal through the cube indicates "commutative-associative" processes—processes in which the numbers may be combined in any order. There are also processes of more restricted ordering—for instance, $a * (b^c) = (a * b)^c$, rather surprisingly.

We find too that just as multiplication is distributive over addition, asteration is distributive over multiplication (but not over addition).

In addition to the inter-process relationship stated above, it is likely that other inter-process laws, such as those producing

$$\begin{aligned} a^{(b+c)} &= a^b \times a^c, \\ a^{(b\times c)} &= (a^b)^c, \end{aligned} \tag{8}$$

and

$$(a\times b)^n = a^n \times b^n$$

can be generalized into more complex relationships among the various processes. These in turn will probably lead to an algebra of index manipulation which we shall hope to explore in a later paper.

9. Computer Applications. The viewpoint of processes we have just considered can achieve, in effect, the reduction of any algebraic operations to addition of items from appropriate tables. Since for digital computers addition time is far less than multiplication time, this viewpoint could lead to more efficient computer use. Present computer design makes table search and interpolation a chore, but future design might circumvent this problem. Benefits to analog computer usage from this point of view are perhaps more readily realizable.

10. Arbitrary Functions. In the introduction we gave a definition of the hierarchic index (1) which is closely tied to the exponential function. It is tempting now to speculate on a generalization of this:

$$a^{(k)} = f(a) \text{ iterated } k \text{ times} \tag{9}$$

where f is an arbitrary function. The same rule of composition (3) holds. If k is allowed to assume non-integral values, we again run into the problem of non-integral iterations of an arbitrary function, which we mentioned in Section 4. The most interesting questions from our point of view, however, seem to be associated with the exponential function.

11. Summary. Traditionally, algebraic operations connect numbers. Consider instead that each individual number has an associated process, determined by a real number, the "hierarchic index," k, defined as: $a^{(k)} = \log a$ iterated k times. The inverse relation is $a^{(-k)} = \text{antilog } a$ iterated k times. The rule of composition for numbers and their associated indices is $a^{(k)} + b^{(l)} + \cdots c^{(m)} = d^{(n)}$, abbreviated as $(k, l, \cdots, m: n)$. Exponentiation, $a^b = c$, or $a^{(2)} + b^{(1)} = c^{(2)}$, becomes $(2, 1: 2)$. Addition, multiplication, and exponentiation are thus seen as elements in a family of related operations. Furthermore, operations with integral indices are the discrete elements of a continuous range whose operations have non-integral indices. These can be approximated graphically, or from slide-rule type nomograms. Identities can be indicated graphically for each index, but convenient inverses exist only for certain operations. Ordinary laws of commutation, association, and distribution may later be generalized for this enlarged family. The foregoing definition of the hierarchic index is closely tied to the exponential function but can be generalized to $a^{(k)} = f(a)$ iterated k times, where f is an arbitrary function. The same rule of composition holds. Utilization of such a rule with appropriate tables may prove useful for computers.

I am indebted to several friends at The RAND Corporation for their helpful criticisms of this paper, and particularly to Richard Bellman for suggestions which improved the notation and clarified the rule of composition.

References

1. Hadamard, J., Two works in iteration and related questions, Bull. of the Amer. Math. Soc., **50** (1944) 67–75.

2. Montel, Paul, *Leçons sur les récurrences et leurs applications*, Gauthier-Villars, Paris, 1957.

PARTICULAR SOLUTIONS OF LINEAR DIFFERENCE EQUATIONS WITH CONSTANT COEFFICIENTS

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

This paper develops the operational methods for linear difference equations by means of an anti-difference operator, as can be done for differential equations [3]. It will be assumed that the elementary difference notations and operations are familiar. We shall be primarily concerned with particular solutions and technically should, perhaps, use the notation $Y_{n,p}$ for the results, however with this understanding we shall drop the added encumbrance and merely write Y_n.

Two formulas of considerable importance to which we shall have occasion to refer, and hence for convenience write at once, are

$$Y_{n+k} = (\Delta + 1)^k Y_n = E^k Y_n, \tag{1}$$

n and k once for all integers, and the anti-difference by parts formula,

$$\Delta^{-1} U_n \Delta V_n = U_n V_n - \Delta^{-1} V_{n+1} \Delta U_n. \tag{2}$$

Consider the equation,

$$Y_{n+1} - S Y_n = F_n, \qquad S \neq 1, \tag{3}$$

which by (1) can be written in the form,

$$(\Delta + 1) Y_n - S Y_n = E Y_n - S Y_n = F_n, \tag{4}$$

or in the alternate form,

$$\Delta Y_n - (S - 1) Y_n = F_n. \tag{5}$$

Substitution of $Y_n = B^n$ in the reduced equation,

$$\Delta Y_n - (S - 1) Y_n = 0,$$

yields the result $B = S$ and hence the complementary function $Y_n = A S^n$, wherein A is an arbitrary constant. If we substitute $Y_n = A_n S^n$ in (5), A now regarded as a function of n, we obtain the solution,

$$Y_n = S^n \Delta^{-1} S^{-n} \frac{F_n}{S}, \tag{6}$$

or since

$$S^{-n} = \frac{S}{1 - S} \Delta S^{-n},$$

$$Y_n = -\frac{1}{S - 1} S^n \Delta^{-1} F_n \Delta S^{-n}. \tag{7}$$

By (2) with

$$\Delta V_n = \Delta S^{-n} \qquad V_n = S^{-n}$$
$$U_n = F_n \qquad \Delta U_n = \Delta F_n = F_n',$$

we have

$$\Delta^{-1}F_n\Delta S^{-n} = S^{-n}F_n - \Delta^{-1}S^{-n-1}F_n' = S^{-n}F_n + \frac{1}{S-1}\Delta^{-1}F_n'\Delta S^{-n}$$

hence

$$Y_n = -\frac{1}{S-1}\left[F_n + \frac{1}{S-1}S^n\Delta^{-1}F_n'\Delta S^{-n}\right]. \tag{8}$$

Now we can write (5) in the operational form,

$$[\Delta - (S-1)]Y_n = (E-S)Y_n = F_n,$$

from which we obtain the formal "solution,"

$$Y_n = \frac{1}{\Delta-(S-1)}F_n \quad \text{or} \quad Y_n = \frac{1}{E-S}F_n.$$

In view of (7) we define

$$\frac{1}{\Delta-(S-1)}F_n = \frac{1}{E-S}F_n = -\frac{1}{S-1}S^n\Delta^{-1}F_n\Delta S^{-n} \tag{9}$$

or by (8)

$$\frac{1}{\Delta-(S-1)}F_n = -\frac{1}{S-1}\left[F_n + \frac{1}{S-1}S^n\Delta^{-1}F_n'\Delta S^{-n}\right]. \tag{10}$$

But from the defining relationship, (9), the right side of (10) can be expressed in the operational form,

$$-\frac{1}{S-1}F_n + \frac{1}{S-1}\,\frac{1}{\Delta-(S-1)}F_n'.$$

Therefore we have the fundamental relationship,

$$\frac{1}{\Delta-(S-1)}F_n = -\frac{1}{S-1}F_n + \frac{1}{S-1}\,\frac{1}{\Delta-(S-1)}F_n'. \tag{11}$$

From (11) we have

$$\frac{1}{\Delta-(S-1)}F_n' = -\frac{1}{S-1}F_n' + \frac{1}{S-1}\,\frac{1}{\Delta-(S-1)}F_n'' \tag{12}$$

in which $F_n'' = \Delta^2F_n$, hence substituting (12) in (11) there results

$$\frac{1}{\Delta-(S-1)}F_n = -\frac{1}{S-1}F_n - \frac{1}{(S-1)^2}F_n' + \frac{1}{(S-1)^2}\,\frac{1}{\Delta-(S-1)}F_n''.$$

By repeated applications of (11) we find

$$(13)\qquad \frac{1}{\Delta-(S-1)}F_n = -\frac{1}{S-1}\left[1+\frac{\Delta}{S-1}+\cdots+\frac{\Delta^k}{(S-1)^k}\right]F_n + R$$

where

$$R = \frac{1}{(S-1)^{k+1}}\,\frac{1}{\Delta-(S-1)}\,F_n^{k+1}.$$

The expression

$$-\frac{1}{S-1}\left[1+\frac{\Delta}{S-1}+\cdots+\frac{\Delta^k}{(S-1)^k}\right]F_n$$

is defined as the principle part of (13). If F_n in (3) is of the form

$$F_n = a_0 + a_1 n + a_2 n^2 + \cdots + a_k n^k = P_k(n),$$

a_i, $i=0, 1, \cdots, k$, constants, we expand $P_k(n)$ in a Maclaurin factorial power series,

$$P_k(n) = P(0) + P'(0)n^{(1)} + \frac{P''(0)}{2!}n^{(2)} + \cdots + \frac{P^k(0)}{k!}\,n^{(k)},$$

where $P^k(0)=\Delta^k P_k(n)_{n=0}$, and $n^{(k)}=n(n-1)\cdots(n-k+1)$. Then we may use the principal part of (13) to great advantage in finding a particular solution of (3) for $\Delta n^{(k)}$ has the simple expansion, $kn^{(k-1)}$ while the value of the remaining part of (13), R, is

$$\frac{1}{(S-1)^{k+1}}\,\frac{1}{\Delta-(S-1)}\,F_n^{k+1} = \frac{1}{(S-1)^{k+1}}\left[-\frac{1}{S-1}S^n\Delta^{-1}0\right] = MS^n$$

where M is a constant. This part, however, is already present in the complementary function and thus can be neglected in computing particular solutions. The principal part of (13) can be obtained by formally dividing $\Delta-(S-1)$ into 1 and limiting the expansion to those powers of Δ which yield values other than 0 when operating on F_n.

We now turn to the case, $F_n = V^n$, V a constant, and use equation (3) in the form of equation (4) rather than (5). Hence by (6)

$$Y_n = \frac{1}{E-S}V^n = \frac{1}{(\Delta+1)-S}V^n = \frac{1}{S}S^n\Delta^{-1}S^{-n}V^n = \frac{1}{S}S^n\Delta^{-1}\left(\frac{V}{S}\right)^n$$

$$= \frac{V^n}{V-S}, \qquad V \neq S.$$

Next let F_n be of the form V^nG_n, V a constant. We consider

$$[(\Delta+1)-S]Y_n = V^nG_n.$$

$$Y_n = \frac{1}{E-S}V^nG_n = \frac{1}{(\Delta+1)-S}V^nG_n = S^n\Delta^{-1}S^{-n}\frac{V^nG_n}{S}$$

$$= S^n\Delta^{-1}\left(\frac{S}{V}\right)^{-n}\frac{G_n}{S} = \left(\frac{S}{V}\right)^{n} V^{n-1}\Delta^{-1}\left(\frac{S}{V}\right)^{-n}\frac{G_nV}{S} = V^n\frac{1}{V(\Delta+1)-S}G_n$$

$$= V^n\frac{1}{VE-S}G_n.$$

The three foregoing results may be extended to higher order difference equations upon the following considerations. The first case resulted in a formal division of the operator into 1 and may be symbolized as follows. If

$$[\Delta-(S-1)]Y_n = P_k(n) = Q_k[n^{(k)}],$$

then

$$Y_n = \frac{1}{\Delta-(S-1)}Q_k[n^{(k)}] = p_k(\Delta)Q_k[n^{(k)}]$$

where $p_k(\Delta)$ is a k-th degree polynomial in Δ. Now consider, for example, the case

$$Y_{n+2}-(S+R)Y_{n+1}+SRY_n = Q_k[n^{(k)}],$$

which by the aid of (1) can be expressed in the form,

$$[(\Delta+1)^2-(S+R)(\Delta+1)+SR]Y_n = Q_k[n^{(k)}].$$

Let

$$[\Delta-(S-1)]Y_n = Z_n,$$

then

$$Z_n = p_k(\Delta)Q_k[n^{(k)}]$$

and since the right side remains a factorial polynomial,

$$Y_n = q_k(\Delta)p_k(\Delta)Q_k[n^{(k)}] = r_k(\Delta)Q_k[n^{(k)}]$$

where $r_k(\Delta)$ is that part of the product $q_k(\Delta)p_k(\Delta)$ which is relevant. However, because of the uniqueness of the power series in Δ, formally we might just as well have divided the polynomial $(\Delta+1)^2-(S+R)(\Delta+1)+SR$ into 1 at the outset. The result can be extended to higher order equations. The second two results,

$$\frac{1}{E-S}V^n = \frac{V^n}{V-S}; \qquad V\neq S \quad\text{and}\quad \frac{1}{E-S}V^nG_n = V^n\frac{1}{VE-S}G_n,$$

can be extended by similar observations because the form of the F_n to which these are applicable is preserved by the expansion. Thus we have the two operational results,

$$(14)\qquad \frac{1}{P(E)}V^n = \frac{1}{P(V)}V^n \qquad P(V)\neq 0$$

and

$$\frac{1}{P(E)} V^n G_n = V^n \frac{1}{P(VE)} G_n \tag{15}$$

the second including the case $P(V)=0$, invalid in the first.

We consider now the final case,

$$b_k Y_{n+k} + b_{k-1} Y_{n+k-1} + \cdots + b_1 Y_{n+1} + b_0 Y_n = \cos(na) \text{ or } \sin(na) \tag{16}$$

$$b_i,\ i = 0, 1, \cdots, k \text{ constants, } a \text{ constant.}$$

By (1) we may write (16) in the form

$$P_k(E) Y_n = \cos(na) \quad \text{or} \quad \sin(na)$$

wherein $P_k(E)$ is the k-th degree polynomial in E,

$$b_k E^k + b_{k-1} E^{k-1} + \cdots + b_1 E + b_0, \tag{17}$$

we shall consider the $\cos(na)$, results for the $\sin(na)$ being similar. Either can be worked directly by using e^{ina} and retaining the real or imaginary part of the result as is dictated. However we shall derive a general formula for the $\cos(na)$ leaving other possibilities, such as $(n)\cos(na)$, to be solved as indicated above.

From (14) we have,

$$Y_n = \frac{1}{P_k(E)} \cos(na) = \frac{1}{2P_k(E)} (e^{ina} + e^{-ina}) = \frac{1}{2}\left[\frac{e^{ina}}{P_k(e^{ia})} + \frac{e^{-ina}}{P_k(e^{-ia})}\right].$$

Using the fact that $\overline{e^z} = e^{\bar z}$ and $\overline{P(z)} = P(\bar z)$, $P(z)$ a polynomial in z with real coefficients,

$$Y_n = \frac{e^{ina} P_k(e^{-ia}) + e^{-ina} P_k(e^{ia})}{2\,|P_k(e^{ia})|^2}, \qquad P_k(e^{ia}) \neq 0,$$

where from (17)

$$P_k(e^{ia}) = b_k e^{kia} + b_{k-1} e^{(k-1)ia} + \cdots + b_1 e^{ia} + b_0.$$

From this a computation reveals that

$$\tfrac{1}{2}\left[e^{ina} P_k(e^{-ia}) + e^{-ina} P_k(e^{ia})\right] = \sum_{i=0}^{k} b_i \cos(n-i)a$$

while $|P_k(e^{ia})|^2 = \sum_{i=0}^{k} c_i \cos(ia)$, in which

$$c_i = M_i \sum_{j=0}^{k-i} b_j b_{j+1}, \quad M_0 = 1, \quad M_i = 2 \qquad \text{for } i = 1, 2, \cdots, k. \tag{18}$$

Thus

$$Y_n = \frac{\sum_{i=0}^{k} b_i \cos(n-1)a}{\sum_{i=0}^{k} c_i \cos(ia)} \tag{19}$$

with c_i given in (18). If the right side of (16) had been $\sin(na)$ then the numerator sum in (19) would have contained $\sin(n-1)a$ rather than $\cos(n-1)a$. If $P_k(e^{ia})=0$, we must use the shift formula (15) on e^{ina} and then select the real or imaginary part of the end result as indicated.

Example. Let,

(20) $Y_{n+2}+3Y_{n+1}+2Y_n = n^2 2^n$ whose reduced equation is,

(21) $Y_{n+2}+3Y_{n+1}+2Y_n = 0.$ Substitution of $Y_n = B^n$ in (21) yields

$B=-1$ and $B=-2$ from which we have the complementary function, $Y_{n,c} = C_1(-1)^n+C_2(-2)^n$, in which C_1 and C_2 are arbitrary constants. For $Y_{n,p}$ we have,

$$Y_{n,p} = \frac{1}{E^2+3E+2}n^2 2^n = 2^n \frac{1}{4E^2+6E+2}\left[n^{(2)}+n^{(1)}\right]$$

$$= \frac{2^n}{432}(36-42\Delta+37\Delta^2)\left[n^{(2)}+n^{(1)}\right] = \frac{2^n}{108}(9n^2-21n+8).$$

Thus,

$$Y_n = C_1(-1)^n + C_2(-2)^n + \frac{2^n}{108}(9n^2-21n+8)$$

is the general solution of (20).

References

1. Jordan, Charles, *Calculus of Finite Differences*, Chelsea, N. Y., 1947.
2. Boole, George, *Finite Differences*, Chelsea, N. Y.
3. Grimm, C. A., Particular Solutions of Linear Differential Equations, MATHEMATICS MAGAZINE, Volume 32, Number 1, Sept.–Oct. 1958, pp. 25–30.

A RADICAL SUGGESTION

ROY J. DOWLING, University of Manitoba

$\sqrt{10}$ is a useful number for illustrative purposes when one is discussing irrational numbers at an elementary level. It is the hypotenuse of a right triangle whose other sides are 1 and 3, and it can be shown to be irrational by a most unsophisticated argument.

When the non-zero integer p is squared, the resulting integer has exactly twice as many terminal zero digits as does p (even if p has none). Thus, if p is a non-zero integer, p^2 ends with an even number of zeros, and, if q is a non-zero integer, $10q^2$ ends with an odd number of zeros. It follows that p^2 cannot equal $10q^2$.

THE PERIMETRIC BISECTION OF TRIANGLES

DOV AVISHALOM, University of Minnesota

I. The Cleavers. If we drop from the middle M of a circular arc AMB a perpendicular MN on the chord AB, then $AN=NB$ (Fig. 1).

If we drop from the middle M of a circular arc AMB a perpendicular MN on the segment AC of the broken line ACB, then $AN=NC+CB$ (Fig. 2).

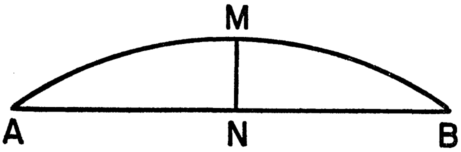


FIG. 1

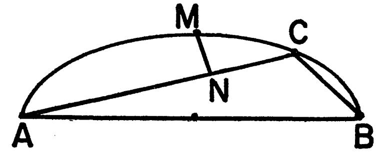


FIG. 2

Point M (Fig. 3) is the middle of the circular arc AMB; $AC \perp MN$; therefore $AN=NC+CB$. Let $AD=DB$, then $AN+AD=NC+CB+BD=p$ ($2p=$perimeter of $\triangle ABC$).

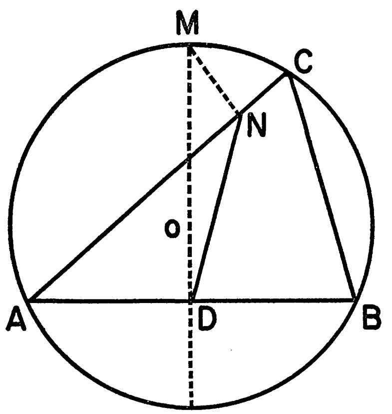


FIG. 3

The line ND divides the perimeter of $\triangle ABC$ in two halves. The point N is called the cleavance-point. The segment ND is called the cleaver. Every triangle has three cleavers.

THEOREM 1. *Every cleaver in a triangle is parallel to the bisector of the angle opposite the side of the middle where it passes.*

Proof. (Fig. 4) Let BN bisect the angle ABC; then $CN/NA=b/c$; $CN/a-CN=b/c$; hence $CN=ab/b+c$; but $CM=a/2$; $CQ=b+c/2$; therefore $CQ/CB=CM/CN$, that means $QM \| BN$ (Thales' theorem),

THEOREM 2. *The three cleavers of a triangle are concurrent in a point called the cleavance-center.*

Proof. (a) In an isosceles triangle, the cleavers will meet on the altitude of the triangle (symmetry).

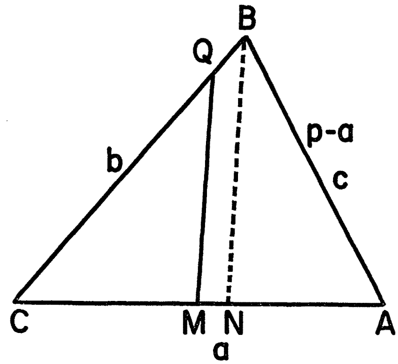


FIG. 4

(b) For any oblique triangle, let us first prove a Lemma of Dov Jarden [2, p. 50]:

The parallels through middles of sides of a triangle to three cevians are concurrent.

The proof of this lemma is simple. The medial triangle is homologous to the given triangle (the triangles are similar and sides are parallel correspondingly). Cevians are concurrent lines in a triangle through the vertices. The bisectors of interior angles in a triangle are cevians; therefore by Theorem 1 and Jarden's lemma, Theorem 2 is proved.

Remarks. (a) The cleavance-center is the incenter of the medial triangle, called the center of the Spieker circle [3, p. 226].

(b) The cleavance-center is identical with the perimetric gravity-center of our triangle.

Problem. Is there in the plane an affine-transformation which leaves invariant the cleavers and the center of cleavance? The answer is given in:

THEOREM 3. *The similarity-transformation leaves the cleavers and cleavance-center invariant.*

Proof. Every affine-transformation transforms the middle of a segment into middle of a segment, parallel lines into parallels; the similarity transformation preserves the angles and their bisectors. Therefore, by Theorem 1, the similarity transformation preserves the cleaver and cleavance-center.

II. The Splitters. Through each vertex of the triangle passes a line bisecting the perimeter of the triangle; p.e.: $BC+CE_2=BA+AE_2=p$; (Fig. 5). The point E_2 is called a splitting point. The segment BE_2 is called a splitter. There are in a triangle three splitters: AE_1, BE_2, CE_3.

THEOREM 4. *The three splitters in a triangle are concurrent in a point called the splitting-center.*

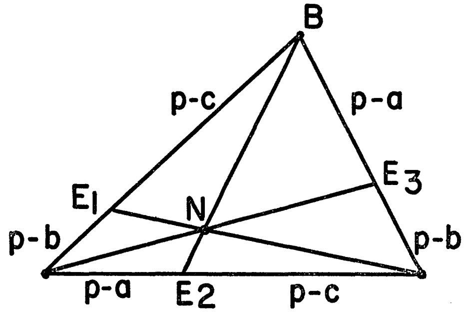


FIG. 5

Proof.

$$\frac{CE_2}{AE_2}\cdot\frac{AE_3}{BE_3}\cdot\frac{BE_1}{CE_1}=\frac{(p-a)(p-b)(p-c)}{(p-c)(p-a)(p-b)}=1;$$

According to the Ceva-theorem, the splitters are concurrent, q.e.d.

Remarks. (a) The lines from the vertices of a triangle to the internal points of contact of the respective escribed circles meet at a point called the Nagel-point [3, p. 184].

The splitting center of the triangle is identical with the Nagel-point [3, p. 149].

(b) The gravity-center G, the incenter I, the splitting-center N, are collinear, and $GN=2IG$ [3, p. 225].

(c) The incenter I, the center of Spieker circle S, and the Nagel-point are collinear, and $IS=SN$ [3, pp. 225–228].

(d) Generally, the splitting center and the cleavance-center are 2 different points [remark c]. Only in an equilateral triangle they are identical.

(e) In fig. 3 it is easy to see that $NC=b-a/2$ and arc $MC=\beta-\alpha$; Prof. H. Guggenheimer kindly pointed out to me, that these properties give easy ways for solving construction problems. Example: Construct a triangle given:

$$R,\ a-b,\ \alpha-\beta\ (R=\text{radius of circumcircle}).$$

(f) In [1, p. 46] I gave an analytical proof of theorem 2, which is proven here synthetically.

References

1. Dov Avishalom, Perimeter-Bisectors in a triangle, Riveon Lematematika, vol. 13, 1959.

2. Dov Jarden, Synthetical proof for the theorem on the center of perimeter-bisectors in a triangle, Riveon Lematematika, vol. 13, 1959.

3. Roger A. Johnson, Advanced Euclidean Geometry, Dover, 1960.

TEACHING OF MATHEMATICS

EDITED BY ROTHWELL STEPHENS, Knox College

This department is devoted to the teaching of mathematics. Thus, articles of methodology, exposition, curriculum, tests and measurements, and any other topic related to teaching, are invited. Papers on any subject in which you, as a teacher, are interested, or questions which you would like others to discuss, should be sent to Rothwell Stephens, Mathematics Department, Knox College, Galesburg, Illinois.

DIFFERENCE QUOTIENTS AND THE TEACHING OF THE DERIVATIVE

WILLIAM ZLOT, Yeshiva University

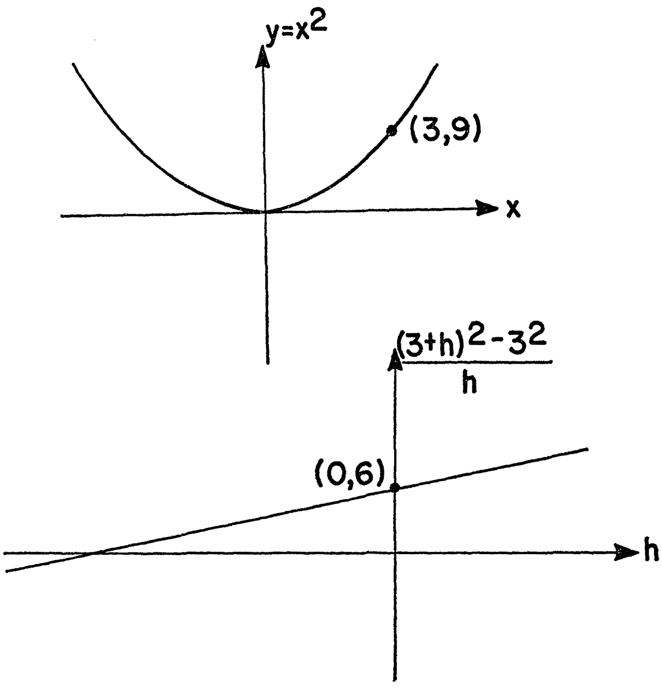
This note suggests a technique that can be used to help clarify the definition of the derivative of a real-valued function of a real variable. The essence of this technique is a graphical study of the function whose limit is being considered, namely, the difference quotient. Let us illustrate the procedure by considering the derivative of $y=x^2$ at $x=3$.

The following are the steps in the method:

(a) Sketch the function $y=x^2$.

(b) On a different set of axes, sketch the difference quotient $[(3+h)^2-3^2]/h$ as a function of h, being careful (for suggestive purposes) to place the vertical axis directly below the point (3, 9) on the graph of $y=x^2$.

The graphs look as follows:



Difference quotient for $y=x^2$ at (3, 9).

Of course, we were aided in our sketching by our ability to perform an algebraic transformation of the expression $[(3+h)^2-3^2]/h$ in order to obtain a "simple" equivalent expression $6+h(h\neq 0)$.* The graph of $[(3+h)^2-3^2]/h$ is thus a straight line punctured at the point (0, 6) and with "slope" equal to 1.

The number 6 is, of course, defined to be the derivative of $y=x^2$ at $x=3$, and from the graph of $[(3+h)^2-3^2]/h$, it is intuitively clear that although 6 is not achieved, it can be approximated as closely as one desires by the difference quotient for all values of h that are "small" in absolute value. The defining of the instantaneous rate of change to be 6 can thus be made plausible by appeal to the graph of the difference quotient.

Naturally, this graphical approach in no way constitutes a formal mathematical proof that $\lim_{h\to 0}[(3+h)^2-3^2]/h=6$ and it may be mentioned in the classroom that there is an analytical definition of the *limit of a function* that can be used as the basis for a formal proof.**

In conclusion, it may be pointed out that the device described in this article can be used to advantage in the study of functions that are not differentiable as well as in the study of functions that only have either a right-hand or a left-hand derivative at a point.

---

* Cases in which such a simplifying reduction is not available can also be treated in the manner advocated, but then the sketching of the difference quotient can, of course, become a cumbersome problem.

** If a study of the definition of the limit of a function is part of the syllabus, then the instructor has an opportunity to motivate the study of the limit of a function that is discontinuous at the point in question.

GEOMETRIC INTERPRETATION OF THE IMPLICIT FUNCTION THEOREM

KURT KREITH, University of California, Davis

Most texts on advanced calculus give a geometric motivation for the implicit function theorem in the simplest case (x and t are real variables, $f(x, t)=0$; when can we solve for $x=g(t)$?) and then proceed to prove a more general version of the theorem without giving further geometric interpretation. The purpose of this note is to point out that that simple geometric interpretations are readily available in more general cases.

We begin by giving a general statement of the implicit function theorem as found in Apostol [1]:

Implicit Function Theorem. Let $f=(f_1, \cdots, f_n)$ be a vector valued function of $n+k$ variables defined on an open set S in E_{n+k}. Suppose $f\in C^1$ on S. Let $(x_0; t_0)=(x_{1,0}, \cdots, x_{n,0}; t_{1,0}, \cdots, t_{k,0})$ be a point in S for which $f(x_0; t_0)=0$ and for which the $n\times n$ determinant $J_f(x_0; t_0)\equiv\det[D_jf_i(x_0; t_0)]\neq 0$. Then there exists a k-dimensional neighborhood T_0 of t_0 and one, and only one, vector-valued function g, defined on T_0 and having values in E_n such that

(i) $g\in C^1$ on T_0.

(ii) $g(t_0)=x_0$.

(iii) $f(g(t); t)=0$ for every t in T_0.

Case 1. $n=k=1$. This is the "simplest case" referred to above.

Case 2. $n=1$, $k=2$. In this case the graph of $f(x; s, t)=0$ is a smooth surface

G in E_3 which contains $(x_0; s_0, t_0)$. The normal vector to G at $(x_0; s_0, t_0)$ is given by $N=(f_x, f_s, f_t)_0$, so that the condition $J_f(x_0; s_0, t_0)\equiv(f_x)_0\neq 0$ implies that N is not perpendicular to the x-axis. Thus, near $(x_0; s_0, t_0)$ the coordinates (s, t) can be used to describe the surface G by means of the equation $x=g(s, t)$.

Case 3. $n=2$, $k=1$. In this case the graphs of the equations

$$f_i(x, y; t) = 0; \qquad i = 1, 2$$

are smooth surfaces G_i which contain the point $(x_0, y_0; t_0)$. There are two possibilities:

(a) G_1 and G_2 intersect in a curve C near $(x_0, y_0; t_0)$.

(b) G_1 and G_2 have the same tangent plane at $(x_0, y_0; t_0)$.

In case (b) the normals $N_i=(f_{i,x}, f_{i,y}, f_{i,t})_0$ to G_i at $(x_0, y_0; t_0)$ are parallel and therefore

$$J_f(x_0, y_0; t_0) \equiv \left|\begin{matrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_2}{\partial x} \\ \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_2}{\partial y} \end{matrix}\right|_0 = 0.$$

Thus the condition $J_f(x_0, y_0; t_0)\neq 0$ implies that the surfaces intersect in a curve C near $(x_0, y_0; t_0)$. The non-vanishing of the Jacobian has further implications; the tangent to C at $(x_0, y_0; t_0)$ is given by $T=N_1\times N_2$, and the third component of T is just $J_f(x_0, y_0; t_0)$. Therefore the condition $J_f(x_0, y_0; t_0)\neq 0$ implies that T is not perpendicular to the t-axis and that near $(x_0, y_0; t_0)$ the coordinate t can be used as a parameter to describe C by means of the equations $x=g_1(t)$, $y=g_2(t)$.

Bibliography

1. Apostol, T. *Mathematical Analysis*, Addison-Wesley (1957).

SIX EQUAL INSCRIBED CIRCLES

LEON BANKOFF, Los Angeles, California

The six cross hatched circles are equal. This is shown in the separate diagrams where circles (A) and (B) are unit circles. The radii of (X), (M), (E), and (D) are r_1, r_2, r_3, and r_4, respectively.

$$(XY)^2 = (XZ)^2 + (ZY)^2$$
$$(r_1 + \tfrac{1}{4})^2 = (\tfrac{1}{2} - r_1)^2 + (\tfrac{1}{4})^2$$
$$r_1 = \tfrac{1}{6}$$

$$(AE)^2 - (EF)^2 = (EB)^2 - (GB)^2$$
$$(1 - r_3)^2 - r_3^2 = (1 + r_3)^2 - (1 - r_3)^2$$
$$r_3 = \tfrac{1}{6}$$

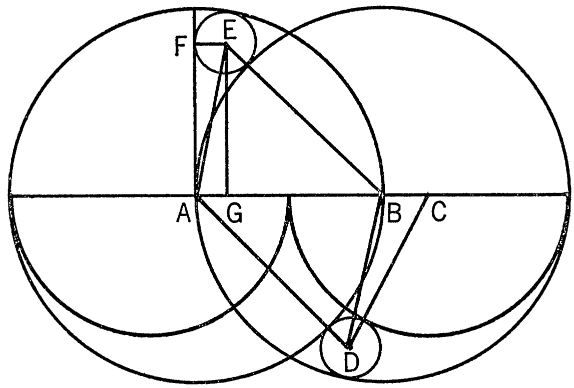
Also,

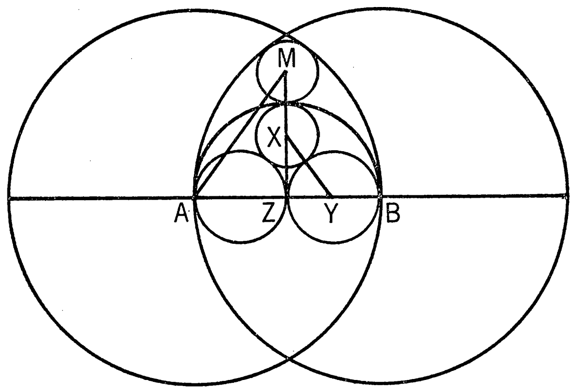
$$(AM)^2 = (AZ)^2 + (MZ)^2$$
$$(1 - r_2)^2 = (\tfrac{1}{2})^2 + (r_2 + \tfrac{1}{2})^2$$
$$r_2 = \tfrac{1}{6}$$

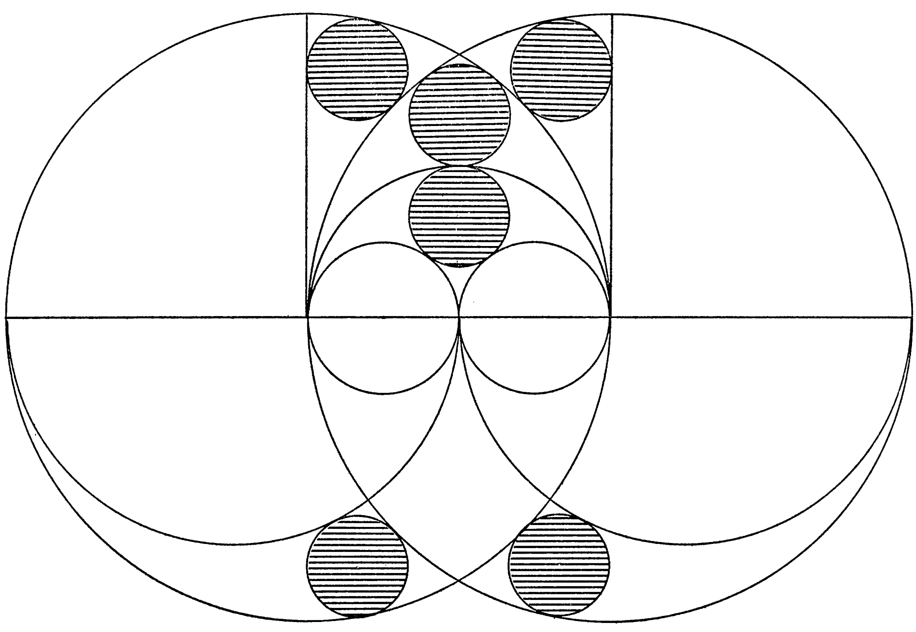
Also, by Stewart's Theorem,

$$(CD)^2\,(AB) + (AD)^2(BC) = (BD)^2(AC) + (AB)(BC)(AC)$$
$$(r_4 + 3/4)^2(1) + (r_4 + 1)^2(\tfrac{1}{4}) = (1 - r_4)^2(5/4) + (1)(\tfrac{1}{4})(5/4)$$
$$r_4 = \tfrac{1}{6}$$

Therefore, $r_1=r_2=r_3=r_4$.







Six equal inscribed circles.

COMMENTS ON PAPERS AND BOOKS

EDITED BY HOLBROOK M. MACNEILLE, Case Institute of Technology

This department will present comments on papers published in the MATHEMATICS MAGAZINE, lists of new books, and 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 to Holbrook M. MacNeille, Department of Mathematics, Case Institute of Technology, Cleveland 6, Ohio.

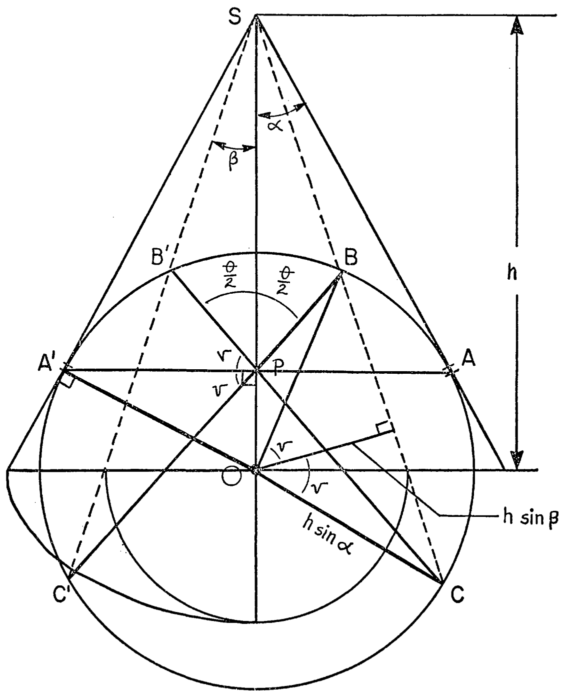
COMMENTS ON A PAPER ON THE CONE

JOSEF ANDERSSON, Vaxholm, Sweden

In the article "Some Old Slants and New Twists to the Cone" (May–June 1961, p. 296), the author maintains that the equation

$$y^2 - z^2 \cdot \frac{b^2}{h^2} \cdot \frac{a^2 + h^2}{a^2 - b^2} + \frac{2a^2b^2z}{h(a^2 - b^2)} = 0$$

represents two planes. In reality this is the equation of a hyperbolic cylinder passing through the curve γ common to the cone and sphere. It is the asymptotic planes of this cylinder, parallel to the planes of the circular sections of the cone, which were used and consequently led to the correct result. To obtain the equation of two symmetric circular sections, we should set $r^2 = a^2h^2 : (a^2 + h^2)$ and not



equal to a^2, in order that the sphere O be tangent to the cone at double points of γ situated on the longest generatrices of the cone.

With regard to the superiority of analysis over geometry (page 295) let us consider a solution of the problem in question by geometry.

1. $SB \cdot SC = \overline{AS}^2 = SP \cdot SO$, therefore O, P, B, C are concyclic.

2. BP, $B'P$ are harmonic conjugates with respect to orthogonal straight lines PA, PS. From this we show the equality of the angles designated by v and then $\sin \theta/2 = \cos v = \sin \beta : \sin \alpha$.

BOOK REVIEWS

Lectures on Tensor Calculus and Differential Geometry. By Johan C. H. Gerretsen. P. Noordhoff N. V. Groningen, The Netherlands, 1962, lx+202 pp.

The first three chapters of this book are devoted to a quick review of a vector space over the real field, linear transformations, and quadratic forms. The approach is geometrical. Chapter four consists of tensor calculus. Thus far the tools are developed. A discussion of point spaces and manifolds is given in chapter five. The properties of tangent space and correspondences between manifolds are also given in this chapter. Chapter six contains a treatment of curves and their osculating spaces. Then Frenet's equations are obtained. This chapter has quite an extensive treatment of curves and related ideas. Chapter seven deals with the geodesic derivative, geodesic curves, the Christoffel symbols, and geodesic and conformal correspondences. Chapter eight is about hypersurfaces and related subjects. The theory of curvature of general manifolds is treated in chapter nine. This chapter is quite interesting and contains covariant derivatives, the Riemannian curvature, geodesic mapping, and other ideas.

The book ends with chapter ten which is the theory of integrability. This chapter contains the conditions of integrability, geodesic and conformal mappings, and Bonnet's problem.

This book is an excellent text for seniors and beginning graduate students in tensor analysis and its application to differential geometry.

ALI R. AMIR-MOEZ
University of Florida

The Scientific American Book of Mathematical Puzzles and Diversions. By Martin Gardner. Simon and Schuster, New York, 1959, xi+178 pp., hard cover, $3.95.

The 2nd Scientific American Book of Mathematical Puzzles and Diversions. By Martin Gardner. Simon and Schuster, New York, 1961, 253 pp., hard cover, $3.95.

These two volumes consist of selections from the author's feature, *Mathematical Games*, which has been appearing regularly in the Scientific American since January 1958, augmented by comments from some of his multitudinous readers. Together they constitute a recent Selection of The Library of Science. The text is directed toward those interested in recreational mathematics, but should prove stimulating for the serious student as well as the professional mathematician.

The 16 chapters in the first book and 20 chapters in the second book deal with topics that range from origami to hexaflexagons, from ticktacktoe to Eleusis, from polyominoes to Soma cubes, from magic squares to digital roots, from the puzzles of Dudeney and Loyd to mathematical card tricks. Two chapters in each book are devoted to Nine Problems, nicely chosen and challenging.

The material is well-organized and is written in a flowing, understandable style. The large print and many excellent figures facilitate reading. Chapters which contain problems also have the answers at the ends of the chapters. References for further reading on the various topics are appended to each book. These are not books to be read and discarded, but are certain to become treasured occupants of the permanent shelves.

CHARLES W. TRIGG
Los Angeles City College

BOOKS RECEIVED FOR REVIEW

Axiomatics. By R. Blanche, The Free Press, New York, 1962, v+65 pages, $1.25 (paper).

The Development of Mathematical Logic. By P. H. Nidditch, The Free Press, New York, 1962, viii+88 pages, $1.25 (paper).

Logic and Boolean Algebra. By B. H. Arnold, Prentice-Hall, Englewood Cliffs, N. J., 1962, 144 pages, $6.75.

Sets, Relations, and Functions. By James F. Gray, Holt, Rinehart & Winston, New York, 1962, ix+143 pages, $2.50 (paper).

An Introduction to Mathematical Machine Theory. By Seymour Ginsburg, Addison-Wesley, Reading, Mass., 1962, ix+148 pages, $8.75.

High-Speed Analog Computers. By R. Tomovic and W. J. Karplus, Wylie, New York, 1962, xi+255 pages, $9.95.

Advanced Calculus for Applications. By F. B. Hildebrand, Prentice-Hall, Englewood Cliffs, N. J., 1962, ix+646 pages, $13.00.

Calculus and Analytic Geometry. By John F. Randolph, Wadsworth, Belmont, Calif., 1961, xi+618 pages, $8.50.

Analytic Geometry. By William L. Schaaf, Doubleday, New York, 1962, 378 pages, $1.95 (paper).

Statistical Theory. By B. W. Lindgren, Macmillan, New York, 1962, xii+427 pages, $7.95.

Statistics An Intuitive Approach. By G. H. Weinberg and J. A. Schumaker, Wadsworth, Belmont, Calif., 1962, xii+338 pages, $6.50.

Geometry, Algebra and Trigonometry by Vector Methods. By A. H. Copeland, Sr., Macmillan, New York, 1962, x+298 pages, $6.25.

Algebra II. By C. F. Brumfiel, R. E. Eicholz and M. E. Shanks, Addison-Wesley, Reading, Mass., 1962, xiii+466 pages, $5.28.

Intermediate Algebra. By William Wootan and Irving Drooyan, Wadsworth, Belmont, Calif., 1962, xi+334 pages, $4.95.

Elementary Technical Mathematics. By F. L. Juszli and C. A. Rodgers, Prentice-Hall, Englewood Cliffs, N. J., 1962, xii+522 pages, $7.95.

Television Production Handbook. By H. Zettl, Wadsworth, Belmont, Calif., 1961, xi+450 pages, $6.95.

Play Mathematics. By Harry Langman, Hafner, New York, 1962, 216 pages, $4.95.

Foundations of Geometry and Trigonometry. By Howard Levi. Prentice-Hall, Inc., Englewood Cliffs, N. J., xiv+347 pages, $7.95.

PROBLEMS AND SOLUTIONS

EDITED BY ROBERT E. HORTON, Los Angeles City College

Readers of this department are invited to submit for solution problems believed to be new that may arise in study, in research, or in extra-academic situations. Proposals should be accompanied by solutions, when available, and by any information that will assist the editor. Ordinarily, problems in well-known textbooks should not be submitted. Solutions should be submitted on separate, signed sheets. Figures should be drawn in India ink and exactly the size desired for reproduction. Send all communications for this department to Robert E. Horton, Los Angeles City College, 855 North Vermont Avenue, Los Angeles 29, California.

PROPOSALS

502. *Proposed by Raphael T. Coffman, Richland, Washington.*

Using only plane geometry for construction and proof, show how to draw a tangent to an ellipse at any given point on the ellipse.

503. *Proposed by Brother U. Alfred, St. Mary's College, California.*

Consider the series $a, a^2, a^3, a^5, a^8, \cdots$ where every term beginning with the third is the product of the two previous terms. As the number of the terms approaches infinity, what is the limiting relationship that is approached by two successive terms?

504. *Proposed by M. S. Demos, Drexel Institute of Technology.*

The orbit of the earth about the sun is an ellipse with the sun at the focus. Astronomy textbooks say that the mean distance of the sun from the earth is the semi-major axis a.

Show that the correct mean distance with respect to time is $(1+e^2/2)a$, where e is the eccentricity.

505. *Proposed by Leonard Carlitz, Duke University.*

Show that for $n \geqq 1$ $\left(x^2 \dfrac{d}{dx}\right)^n = \sum_{r=1}^{n} a_{nr}x^n \left(x \dfrac{d}{dx}\right)^r$

where the a_{nr} are independent of x. Also identify the a_{nr}.

506. *Proposed by Leon Bankoff, Los Angeles, California.*

Let X, Y, Z be the intersections of the adjacent internal angle trisectors of a triangle ABC, forming the triangles ABZ, BCX, and CAY. If R, S, T are the feet of the internal angle bisectors XR, YS, and ZT, show that AR, BS, and CT are concurrent.

507. *Proposed by D. Rameswar Rao, Secundrabad, India.*

Prove that the power of a power always can be expressed as the geometric mean of two powers.

508. *Proposed by David L. Silverman, Beverly Hills, California.*

Prove that every polyhedron has at least two faces with the same number of edges.

SOLUTIONS

Late Solutions

467, 470. *Josef Andersson, Vaxholm, Sweden.*

472, 479 (two solutions). *Jose Gallego-Diaz, Universidad del Zulia, Maracaibo, Venezuela.*

Polygonal Numbers

481. [May 1962]. *Proposed by Daniel I. A. Cohen, Central High School, Philadelphia, Pennsylvania.*

Prove that the r-th m-gonal number can never be the m-th r-gonal number unless $r=m$.

Solution by Jerry L. Pietenpol, Columbia University.

The r-th m-gonal number is

$$\frac{(m-2)r^2-(m-4)r}{2}.$$

Equating this to the m-th r-gonal number and simplifying yields

$$(r-m)(r-2)(m-2)=0.$$

Thus if $r\neq m$ this cannot be satisfied, unless we admit the degenerate case $m=2$.

Comment by Charles W. Trigg, Los Angeles City College.

If the $(r+k)$th m-gonal number equals the $(m+k)$th r-gonal number, then

$$\tfrac{1}{2}(r+k)[2+(r+k-1)(m-2)]=\tfrac{1}{2}(m+k)[2+(m+k-1)(r-2)]$$

or

$$(r-m)[(m-2)r-(k^2+2m+3k-4)]=0,$$

so

$$r=m \text{ or } r=2+k(k+3)/(m-2), \text{ where } r, m, k \text{ are integers.}$$

Some solutions for other than the trivial $r=m$, are:

$$k=1,\quad (r,m)=(3,6),(4,4),(6,3);$$
$$k=2,\quad (r,m)=(3,12),(4,7),(7,4),(12,3);$$
$$k=3,\quad (r,m)=(3,20),(4,11),(5,8),(8,5),(11,4),(20,3).$$

Also solved by Brother U. Alfred, St. Mary's College, California; J. L. Brown, Jr., Pennsylvania State University; Frederick Carty, Hofstra College; Gilbert Labelle, University of Montreal; Francis D. Parker, University of Alaska; C. F. Pinzka, University of Cincinnati; and the proposer.

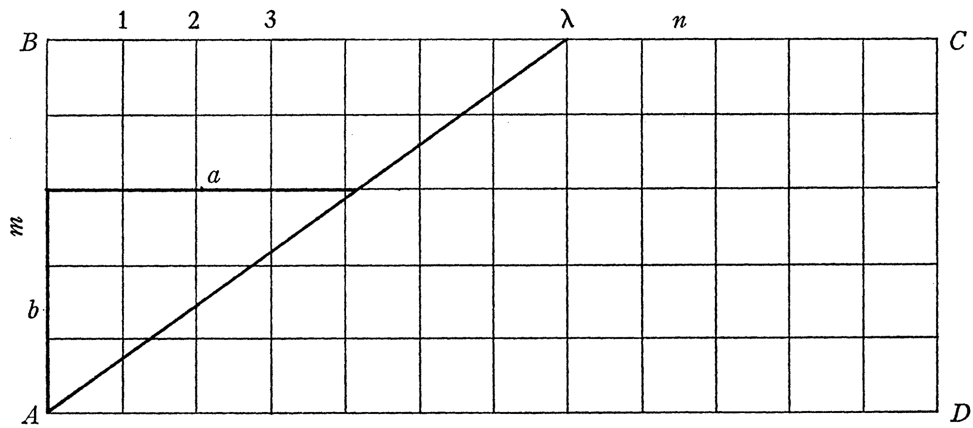
Collinear Lattice Points

482. [May 1962] *Proposed by Brother U. Alfred, St. Mary's College, California.*

Given a rectangular array of equally spaced points with $n+1$ points along one side and $m+1$ points along the other. If $m=\prod p_i^{\alpha_i}$ in terms of prime factors and $n=\prod q_j^{\beta_j}$ in terms of prime factors, determine an expression for the number of lines passing through at least three points of the array, one point being a corner and a second one of the other points on the periphery.

Solution by the proposer.

Let there be $m+1$ points along AB and $n+1$ along BC. Count the points along BC as shown in the figure. Connecting the λth point to A it can be seen



that if a third point is to lie on the line, there would have to be a ratio $\lambda/m=a/b$, so that λ and m would have a common factor. Hence, there will be a third point on the line for those points which have a value λ not prime to m. The number of such points is given by:

$$\sum\left[\frac{n}{p_i}\right]-\sum\left[\frac{n}{p_ip_j}\right]+\sum\left[\frac{n}{p_ip_jp_k}\right]-\cdots$$

Similarly, taking points along CD, there will be a third point on the line connecting them to A if the quantity λ measured from D is not prime to n. Hence, the number of such points would be:

$$\sum\left[\frac{m}{q_i}\right]-\sum\left[\frac{m}{q_iq_j}\right]+\sum\left[\frac{m}{q_iq_jq_k}\right]-\cdots$$

There are also lines AB and AD. Furthermore, we have taken AC twice in the summations. Hence the net result is:

$$\sum\left[\frac{n}{p_i}\right]-\sum\left[\frac{n}{p_ip_j}\right]+\sum\left[\frac{n}{p_ip_jp_k}\right]-\cdots$$
$$+\sum\left|\frac{m}{q_i}\right|-\sum\left|\frac{m}{q_iq_j}\right|+\sum\left|\frac{m}{q_iq_jq_k}\right|-\cdots$$
$$+1.$$

Non-Zero Jacobians

483. [May 1962] *Proposed by Dermott A. Breault, Sylvania Electric Products, Waltham, Massachusetts.*

Given that $w=f(z)$ is conformal in a region R, and $z=f^{-1}(w)$ is conformal in R'. Set $w=\phi(x, y)+i\psi(x, y)$ and $z=x(\phi, \psi)+iy(\phi, \psi)$. Show that the Jacobians $\partial(\phi, \psi)/\partial(x, y)$ and $\partial(x, y)/\partial(\phi, \psi)$ are not zero in R and R' respectively.

Solution by Jose Gallego-Diaz, Universidad del Zulia, Maracaibo, Venezuela.

By the conditions of Cauchy-Riemann:

$$\phi_x' = \psi_y'$$
$$\phi_y' = -\psi_x'$$

Therefore the first Jacobian is:

$$J_1 = \begin{vmatrix} \phi_x' & \psi_x' \\ \phi_y' & \psi_y' \end{vmatrix} = \begin{vmatrix} \phi_x' & \phi_y' \\ \phi_y' & \phi_x' \end{vmatrix} = [\phi_x']^2 + [\phi_y']^2 \neq 0$$

and the same reason holds for the other one.

Equal Triangle Segments

484. [May 1962] *Proposed by Leon Bankoff, Los Angeles, California.*

A square $ADEB$ is constructed externally on the hypotenuse AB of a right triangle ABC, $(CB>CA)$. CE cuts AB at S, and a perpendicular to AB at S cuts CB at Q. T is the foot of the bisector of angle BCA, and P is the foot of the perpendicular from T on CB. TP cuts QS at R. Show that $QR=TS$.

I. *Solution by Francis D. Parker, University of Alaska.*

This is an example of a problem which yields to analytic geometry methods. If the coordinates of B and A are $(b, 0)$ and $(0, a)$, respectively, then the coordinates of T are

$$\left(\frac{ab}{a+b}, \frac{ab}{a+b}\right)$$

and those of S are

$$\left(\frac{ab(a+b)}{a^2+ab+b^2}, \frac{ab^2}{a^2+ab+b^2}\right).$$

With a moderate amount of calculation, it appears that the coordinates of Q are

$$\left(\frac{ab^2}{a^2+ab+b^2}, 0\right)$$

and those of R are

$$\left(\frac{ab}{a+b}, \frac{a^2b^2}{(a+b)(a^2+ab+b^2)}\right).$$

Then $(ST)^2$ and $(QR)^2$ have a common value,

$$\frac{a^4b^2(a^2+b^2)}{(a+b)(a^2+ab+b^2)}.$$

II. *Solution by Jose Gallego-Diaz, Universidad del Zulia, Maracaibo, Venezuela.*

Ang. $CEB = A - \alpha$; $\alpha = ECB$; $\tan ECB = a/(a+b)$; $\tan A = a/b$

Tan $CEB = \tan(A-\alpha) = a^2/(a^2+b^2+ab) = SB/c$

$SB = a^2c/(a^2+b^2+ab)$.

$SA = c - SB = bc(a+b)/(a^2+b^2+ab)$, $TA = bc/(a+b)$;

$ST = SA - TA = ab^2c/(a+b)(a^2+b^2+ab)$

$QS = BS \tan B = abc/(a^2+b^2+ab)$

$RS = TS \cot B = a^2bc/(a+b)(a^2+b^2+ab)$

$RQ = QS - RS = abc(a+b)/(a+b)(a^2+b^2+ab) - a^2bc/(a+b)(a^2+b^2+ab)$

$\quad = ab^2/c(a+b)(a^2+b^2+ab)$. Q.E.D.

III. *Solution by the proposer.*

Let L be the intersection of CD and AB. Extend CT, cutting DE at V. The center of the square $ADEB$ lies on the circumcircle of triangle ABC and also on the line CV. So $AT=VE$ and $DV=TB$.

Now $AC/CB=AT/TB=VE/DV=TS/LT$. Also, in the similar right triangles ABC and RTS, we have $AC/CB=TS/RS$. Hence $RS=LT$.

But $DE/LS=CE/CS=BE/QS=DE/QS$. So $LS=QS$ and $TS=QR$.

Note: The condition $CB>CA$ is unnecessary.

Also solved by Dermott A. Breault, Sylvania Electric Products, Inc., Waltham, Massachusetts; Frederick Carty, Hofstra College; Win Myint, Watervliet Arsenal, New York and Stephen Meskin, New York University (Two solutions, jointly); Hazel S. Wilson, Jacksonville University, Florida; and the proposer (Two solutions).

Divergent P-Series

485. [May 1962] *Proposed by R. P. Steinkirk, University of Missouri.*

$$\sum_{n=1}^{\infty}\left(\frac{1}{n}\right)^{1+(1/p)}$$

converges for all p greater than zero. What about

$$\sum_{n=1}^{\infty}\left(\frac{1}{n}\right)^{1+(1/n)}?$$

Solution by David L. Silverman, Beverly Hills, California.

If $\sum a_n$ and $\sum b_n$ are two series of positive terms and

$$0 < \lim_{n\to\infty} \frac{a_n}{b_n} < \infty,$$

then the two series have the same convergence character. Comparing the series in question with the harmonic series, we have

$$\lim_{n\to\infty} \frac{\left(\frac{1}{n}\right)^{1+1/n}}{\frac{1}{n}} = \lim_{n\to\infty} \left(\frac{1}{n}\right)^{1/n} = 1,$$

by applying L'Hospital's Rule. Since the harmonic series diverges, so does

$$\sum\left(\frac{1}{n}\right)^{1+1/n}.$$

Also solved by Dermott A. Breault, Sylvania Electric Products, Inc., Waltham, Massachusetts; R. Bumcrot, University of Missouri; Frederick Carty, Hofstra College; Kenneth J. Herty, Philadelphia, Pennsylvania; J. M. Howell, Los Angeles City College; Glenn D. James, Los Angeles City College; J. P. Nelsinger, Northeast Missouri State Teachers College; Jerry L. Pietenpol, Columbia University; C. F. Pinzka, University of Cincinnati; Jean Richard, Lachenaie, P.Q., Canada; and the proposer. One incorrect solution was received.

M. S. Klamkin pointed out that Problem 485 is a special case of Problem 97, this magazine, Vol. 27, No. 4, March, 1954, p. 218. Frederick Carty found a solution to the problem in Brand's *Advanced Calculus*, page 53.

A Partitioned Determinant

486. [May 1962] *Proposed by M. Rumney, London, England.*

Given three different positive integers N_1, N_2, and N_3. Find a partition of N_1 into three different positive integers a_{11}, a_{12}, a_{13}, N_2 into three different positive integers a_{21}, a_{22}, a_{23}, and N_3 into three different positive integers a_{31}, a_{32}, a_{33} so that the determinant $|a_{ij}| = 0$, $i = 1, 2, 3$ and $j = 1, 2, 3$.

Solution by Brother U. Alfred, St. Mary's College, California.

We can take

$$a_{11} = a_{21} = a_{31} = 1$$

$$a_{21} = a_{22} = a_{32} = 2$$

$$a_{31} = N_1 - 3, \qquad a_{32} = N_2 - 3, \qquad a_{33} = N_3 - 3$$

there being no statement about ALL the quantities a_{ij} being different.

If this latter requirement were intended it would not be possible to fulfill in some cases, as for example, $N_1 = 6$, $N_2 = 7$, $N_3 = 8$.

Also solved by Daniel I. A. Cohen, Central High School, Philadelphia, Pennsylvania; and David L. Silverman, Beverly Hills, California.

The Square Root of a Matrix

487. *Proposed by Huseyin Demir, Kandilli, Eregli, Kdz., Turkey.*

Find the square root of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Solution by Maurice Brisebois, Universitê de Sherbrooke, Canada.

Let X, U be arbitrary square matrices of order n, A a given non-singular matrix, $\tilde{A}$ a matrix similar to A, $X_{\tilde{A}}$ an arbitrary nonsingular matrix permutable with $\tilde{A}$, (λi) the set of all characteristic values of A; $i=1, \cdots, n$ (they need not be all distinct), E_{p_i} the identity matrix of order p_i with $\sum_i p_i=n$, H_{p_i} the matrix with 1's in the superdiagonal and 0's elsewhere. Let $(\sqrt[n]{\lambda_1 E_{p1}+H_{p1}} \cdots, \sqrt[n]{\lambda_n E_{pn}+H_{pn}})$ be a matrix built with square matrices along the diagonal, matrices of order p_i of the type

$$\begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix} \quad ; i=1, \cdots, n$$

and having 0's elsewhere.

Then all solutions of the matrix equation

$$X^m = A$$

are given by the formula:

$$X = UX_{\tilde{A}}(\sqrt[m]{\lambda_1 E_{p_1} + H_{p_1}}, \cdots, \sqrt[m]{\lambda_n E_{p_n} + H_{p_n}}) \cdot X_{\tilde{A}}^{-1} U^{-1}.$$

The particular case $m=n=2$ yields:

$$X = UX_{\tilde{A}} \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{pmatrix} X_{\tilde{A}}^{-1} U^{-1}.$$

If the matrix A is singular, a more elaborate study is needed and the existence of the mth roots of A is bound with the existence of a system of admissible elementary divisors for X_2 a matrix such that (X_1, X_2) is a matrix similar to X. (We call a system of elementary divisors for X_2 "admissible" if, after raising X_2 to the mth power, these elementary divisors split and generate the system of elementary divisors for A_2 where $A=(A_1, A_2)$ with A_1 and A_2 similar to X_1 and X_2 respectively.)

Remarks. 1. In the general case, the solutions of $X^m=A(|A| \neq 0)$ are not expressible as polynomials in A unless all λ_i are distinct.

2. The solutions of $X^m=A$ are parametric in nature and the number of parameters present in $X_{\tilde{A}}$ is given by the number N of linearly independent

matrices commuting with A, where $N=\sum_{i+1}^{t}(2i-1)n_i$; $(t\leqq n)$, n_i being the degrees of the non-constant invariant polynomials of A.

3. For some results along this line, see Lusternik-Sobolen, "Elem. of Functional Analysis," p. 283, Dunford-Schwartz, "Linear Operations" (Part I), problem 31 on page 583 and Bellman "Introd. to Matrix Analysis," problems 1–3 on page 93.

Also solved by Brother U. Alfred, St. Mary's College, California; J. A. H. Hunter, Toronto, Canada; Francis D. Parker, University of Alaska; Gilbert Labelle, University of Montreal, Canada; C. F. Pinzka, University of Cincinnati; J. L. Stearn, Washington, D. C.; and the proposer.

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 306. Show that

$$\frac{1}{11}+\frac{1}{111}+\frac{1}{1111}+\cdots+\frac{1}{10}=\frac{1}{1100}+\frac{1}{111000}+\cdots$$

[*Submitted by Murray S. Klamkin.*]

Q 307. The nine positive digits can be arranged into 3×3 arrays in 9! ways. Find the sum of the determinants of these arrays. [*Submitted by C. W. Trigg.*]

Q 308. It is ubiquitously stated that the ingenious Gauss astonished his primary school teacher by mentally obtaining the sum of the first hundred natural numbers as 5050. Now show that the sum of the cubes of these hundred numbers just as easily can be given mentally as 25502500. [*Submitted by Dewey Duncan.*]

Q 309. Determine

$$\prod_{n=2}^{\infty}\left[1-\frac{2}{1+n^3}\right].$$

[*Submitted by Murray S. Klamkin.*]

Quickening a Quickie

Q 237. [Jan. 1959] Show that

$$\begin{vmatrix} x-2 & x-3 & x-4 \\ x+1 & x-1 & x-3 \\ x-4 & x-7 & x-10 \end{vmatrix}=0.$$

A 237. *Alternate solutions by C. W. Trigg.*

1. When the elements of the second column are subtracted from the elements of the first column, and those of the third column from those of the second column, the elements of the first and second columns of the derived determinant are identical. Hence the determinant vanishes.

2. When the elements of any column are subtracted from the elements of each of the other two, the elements in two of the columns of the derived determinant are proportional, so the determinant vanishes. That is,

$$\begin{vmatrix} 2 & 1 & x-4 \\ 4 & 2 & x-3 \\ 6 & 3 & x-10 \end{vmatrix} = \begin{vmatrix} 1 & x-3 & -1 \\ 2 & x-1 & -2 \\ 3 & x-7 & -3 \end{vmatrix} = \begin{vmatrix} x-2 & -1 & -2 \\ x+1 & -2 & -4 \\ x-4 & -3 & -6 \end{vmatrix} = 0.$$

3. If $x=1$, the first and third rows are proportional, so the determinant vanishes. If $x=5$, the first and second rows are proportional and the determinant vanishes. Thus the polynomial has two distinct zeros, and since (from method 2) its degree is not greater than 1, it must be identically zero.

Answers

A 306. Since

$$\frac{1}{11}\left(1-\frac{1}{100}\right)+\frac{1}{111}\left(1-\frac{1}{100}\right)+\cdots = 1/9\left[\frac{1}{100}+\frac{1}{1000}+\cdots\right]=\frac{1}{10},$$

the result follows immediately.

A 307. When two rows of a determinant are interchanged, the sign of the determinant is changed. When the rows of a three-by-three determinant are permuted, 3 positive and 3 negative determinants equal in absolute value are obtained. Hence the 9! determinants fall into 9!/6 groups each of which sums to zero.

A 308. $(1^3+2^3+\cdots+100^3)=(1+2+\cdots+100)^2=(5050)^2$
$=[(51\times 50)100+5^2]100.$

A 309.

$$\prod\left[1-\frac{2}{1+n^3}\right]=\prod\left(\frac{n-1}{n+1}\right)\cdot\prod\left(\frac{n^2+n+1}{n^2-n+1}\right)=(2)\left(\frac{1}{3}\right)=\frac{2}{3}\cdot$$

EDITOR'S NOTE

Beginning with this issue the Miscellaneous Notes department is being discontinued. Articles formerly intended for this department will be included among the general articles. Professor Roy Dubisch, who formerly edited Miscellaneous Notes, will continue to edit the algebra and number theory papers appearing as general articles.

THE MATHEMATICAL ASSOCIATION OF AMERICA

The Association is a national organization of persons interested in mathematics at the college level. It was organized at Columbus, Ohio, in December 1915 with 1045 individual charter members and was incorporated in the State of Illinois on September 8, 1920. Its present membership is over 12,500, including more than 500 members residing in foreign countries.

Any person interested in the field of mathematics is eligible for election to membership. Annual dues of $5.00 include a subscription to the American Mathematical Monthly. Members are also entitled to reduced rates for purchases of the Carus Mathematical Monographs, MAA Studies, and for subscriptions to several journals.

Further information about the Association, its publications and its activities may be obtained by writing to:

HARRY M. GEHMAN, *Executive Director*
Mathematical Association of America
SUNY at Buffalo (University of Buffalo)
Buffalo 14, New York

ANALYTIC GEOMETRY

Third Edition

R. S. Underwood and Fred W. Sparks

This new version of a clearly presented and readable text retains the well-tested and widely accepted features of the former edition yet includes a completely new set of problems (except for a few well-tried examples which merit continued use) and a new section on the "Silhouette Method."

About 300 pages *A Late Spring 1963 Publication*

MATHEMATICS FOR BUSINESS AND ECONOMICS

Robert Cissell and Thomas Bruggeman

This timely text recognizes the trend toward the application of mathematical methods to problems in business, economics, and the social sciences in general. An excellent preparation for courses in mathematics of finance, business statistics, linear programing, econometrics, and in the newer management techniques. Instructor's Manual available.

229 pages *1962* *$4.75*

DIFFERENTIAL AND INTEGRAL CALCULUS

James R. F. Kent

The author uses a modern approach to the traditional topics of calculus to make his text readily understandable to the student without the aid of an interpreter, yet rigorous enough to answer his more searching questions. A generous amount of carefully planned exercises insures a good understanding of each topic.

511 pages *1960* *$6.75*