THE BELL SYSTEM
TECHNICAL JOURNAL

A JOURNAL DEVOTED TO THE
SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL
COMMUNICATION

TABLE OF CONTENTS
AND
INDEX

VOLUME XX

1941

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

PRINTED IN U. S. A.

THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XX, 1941

Table of Contents

JANUARY, 1941

APRIL, 1941

JULY, 1941

October, 1941

Index to Volume XX

A

B

C

D

F

H

I

L

V. 20

VOLUME XX JANUARY, 1941 NUMBER 1

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

50c per copy *$1.50 per Year*

THE BELL SYSTEM TECHNICAL JOURNAL

Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.

SUBSCRIPTIONS

Subscriptions are accepted at $1.50 per year. Single copies are 50 cents each. The foreign postage is 35 cents per year or 9 cents per copy.

The Bell System Technical Journal

Vol. XX *January, 1941* *No. 1*

Single Sampling and Double Sampling Inspection Tables

By H. F. DODGE and H. G. ROMIG

Introduction

A CONSIDERABLE amount of attention has been given to the application of statistical methods to problems of inspection with emphasis on means for securing certain definite advantages such as reduction in the cost of inspection, reduction in the cost of production by minimizing rejections, and the attainment of uniform quality of manufactured products.[1,2,3,4] This paper presents four sets of sampling inspection tables that have contributed in a notable way to important reductions in such costs and to substantial improvements in control of quality for many characteristics of products used in the Bell System.

Whether sampling may be employed to advantage in place of 100% inspection usually depends, of course, on the purpose for which inspection is made. The sampling tables here presented provide definite procedures for conducting inspections that have certain immediate purposes which are described in some detail. Through their provision for instituting a "screening" inspection whenever quality falls below an acceptable level, the procedures have been found in practice to enforce a program of controlling quality in process as the alternative to high inspection costs.

General Field of Application

The sampling tables presented herewith have been developed for use in consumer or producer inspections of products composed of similar individual articles or pieces, where it is desired to have assurance of a definite degree of conformance to specification requirements with a minimum of expense.

The following paragraphs indicate the general conditions under which the tables are applicable, as well as some of the assumptions involved in their development.

Acceptance Inspection of Lots—The tables are intended for application in inspections whose immediate purpose is to determine the acceptability of individual lots of product.

By a lot will be meant a collection of individual pieces from a common source, possessing a common set of quality characteristics, and offered as a group for inspection and acceptance at one time. These pieces may be parts, partial assemblies or finished units of product. For purposes of inspection, it is desirable that a lot be composed of pieces all of which have been produced under what are judged to be the same essential conditions. To this end, an attempt should be made to avoid grouping together batches of product that are likely to differ from one another in quality, because of differences in the raw materials used, or differences in manufacturing methods or conditions. For inspections made in a manufacturing plant, particularly where production is continuous as with conveyor systems, the time element may often be the deciding factor in fixing the size of lot, and such items as convenience in handling, and stocking or shipping facilities may make it desirable to take an hour's, a half-day's, or a day's production as the quantity to be considered as a lot for inspection purposes.

Quantity Production—Maximum advantage in the use of the tables may be expected for products produced more or less continuously on a quantity basis as distinguished from those produced intermittently on a small scale.

Inspection by "Method of Attributes"—Inspection by the "method of attributes"[5] is assumed. That is, each piece inspected is examined, gauged, or tested to determine whether it does or does not conform to the requirements imposed by specification.

For some characteristics, the requirements may be expressed as numerical limits to be met by the piece, such as maximum and minimum tolerance limits for a dimension, or the minimum tolerance limit for the illumination of a lamp. For others, the requirements may be expressed in less precise terms, and inspection may consist in observing whether the piece does or does not conform to the finish, appearance, color, etc., of say a standard sample, or to the grade of workmanship commonly understood by the phrase "accepted standards of good workmanship."

Nondestructive Inspection—The tables are applicable primarily to quality characteristics that may be inspected by nondestructive means, so that at any time it is entirely practicable to inspect every piece in the lot.

This limitation is a consequence of the inspection procedure adopted in the development of the tables, wherein complete inspection of individual lots is prescribed under certain conditions.

Quality Measured by "Fraction Defective"—The yardstick of quality used in the tables is "fraction defective" (or fraction nonconforming), that is, the ratio of the number of pieces that fail to conform to a specified requirement to the total number of pieces under consideration.

A piece of product that fails to meet the requirement for a characteristic is classed as nonconforming with respect to that characteristic, and for convenience is referred to as defective. Thus, a deviation from a specified requirement or from accepted standards of good workmanship is termed a "defect." If, in the inspection of the "end illumination" of 1000 lamps, it were found that 10 of the lamps had illumination less than the minimum value specified, and the remaining 990 had illumination equal to or greater than the minimum value, we would say that 10 defects were observed, and the lot of 1000 was 1% defective (fraction defective, $p = 0.01$).

Sampling Inspection—The tables are applicable where, under normal conditions, it will be satisfactory to inspect only a portion of the pieces in the lot and to accept the lot if the inspection results for this sample of pieces meet certain criteria. This, in effect, imposes the condition that it is not the purpose of this inspection to make sure that each piece in the lot conforms to the requirements for the characteristic inspected.

Such a situation is common, for example, in the process inspection of component parts of product units, where it may be the purpose of inspection to make reasonably certain that the quality passing on to the next stage is such that no extraordinary effort will be expended on defective parts. This situation is also common for various characteristics of finished units of product, such as some adjustment and dimensional items, items of condition, finish and workmanship that can be covered by a "surface" inspection, as well as items for which 100% inspections or tests have been made previously during process or are to be made in subsequent operations before delivery to the ultimate consumer. Characteristics, whose conformance to specified requirements is of vital importance to the functional quality of the product, and for which 100% inspection is feasible, may not of course be candidates for sampling inspection.

Acceptance Based on Observed Number of Defects—The acceptance criterion used in the tables is a stated allowable number of defects in a sample of stated size.

If only one defect is allowed in a sample of n pieces selected from a lot, then the "Allowable Defect Number" is 1 (referred to as the "Acceptance Number" in an earlier paper[3]). The criterion for the acceptance of a lot is the finding of a number of defects equal to or less than the Allowable Defect Number.

Random Samples—The theory used in the development of the tables assumes that each sample drawn from a lot is a random sample.

A random sample is one selected by a random operation,[6] such as would obtain if a number of physically similar chips, numbered to correspond to the pieces of product under consideration, were thoroughly mixed in a mixing bowl, and a number of them, equal to the desired sample size, were withdrawn to identify which pieces of product should be included in the inspection sample. When, in practice, there are indications that individual lots may be stratified in quality, it is of course best to select a "representative" sample, one such that each stratum or subportion of the lot is proportionately represented by a subsample that is selected by a random operation.

Inspection Procedures

Two distinct methods of inspection are employed—single sampling and double sampling. In single sampling, only one sample is permitted before a decision is reached regarding the disposition of the lot, and the acceptance criterion is expressed as an allowable defect number, c. In double sampling, a second sample is permitted if the first fails, and two allowable defect numbers are used—the first, c_1, applying to the observed number of defects for the first sample alone, and the second, c_2, applying to the observed num-

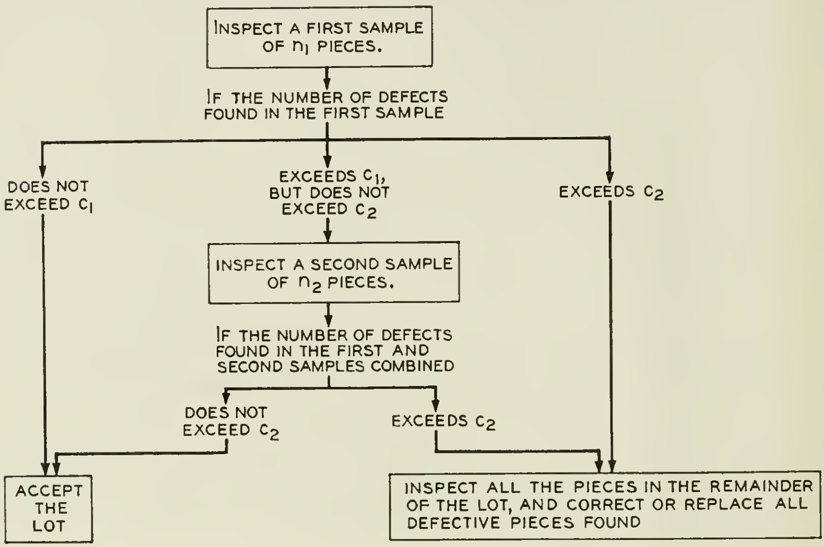


Fig. 1—Double sampling inspection procedure

ber of defects for the first and second samples combined. The specific procedures assumed in the development of the tables are as follows:

Single Sampling Inspection Procedure

(a) Inspect a sample of n pieces.
(b) If the number of defects found in the sample does not exceed c, accept the lot.
(c) If the number of defects found in the sample exceeds c, inspect all the pieces in the remainder of the lot.
(d) Correct or replace all defective pieces found.

Double Sampling Inspection Procedure

(a) Inspect a first sample of n_1 pieces.
(b) If the number of defects found in the first sample does not exceed c_1, accept the lot.

(c) If the number of defects found in the first sample exceeds c_2, inspect all the pieces in the remainder of the lot.
(d) If the number of defects found in the first sample exceeds c_1 but does not exceed c_2, inspect a second sample of n_2 pieces.
(e) If the total number of defects found in the first and second samples combined does not exceed c_2, accept the lot.
(f) If the total number of defects found in the first and second samples combined exceeds c_2, inspect all the pieces in the remainder of the lot.
(g) Correct or replace all defective pieces found.

The double sampling procedure can, perhaps, be visualized more easily by reference to Fig. 1.

The theoretical development assumes that the inspection operation itself never overlooks a defect and that all defective pieces found, whether in samples or in the remainders of those lots that are inspected completely, will be corrected or replaced by conforming pieces.* Thus, lots that fail to be accepted by sample are assumed to be completely cleared of defects.

Protection and Economy Features

When a consumer† adopts sampling inspection in place of 100 per cent inspection, he forgoes the opportunity of assuring himself that each piece of product will conform to requirements, and must choose a sampling plan that will provide a degree of protection against defective material that is consistent with his needs. This choice may be narrowed down by choosing some value of allowable per cent defective, and by deciding whether this allowable value should apply to a limited quantity of product such as a lot, or to the general output comprising a more or less steady flow of lots.

Two Kinds of Consumer Protection

For both the single sampling and double sampling procedures outlined above, tables are developed for each of the following two kinds of consumer protection:

(*a*) *Lot Quality Protection*—in which there is prescribed (1) some chosen value of allowable per cent defective in a lot (Lot Tolerance Per Cent Defective), and also (2) some chosen value for the probability of accepting

* While the mathematical solution assumes correction or replacement of defective pieces, it may be expedient practically to reject defective pieces and not replace them. The effect of following this, rather than the assumed procedure, involves differences in results too small to be of any practical consequence for the small values of per cent defective covered by the tables.

† The term "consumer" is used in the general sense of the recipient of the product after the inspection has been completed. This may, of course, be the ultimate consumer or his agent. However, in a manufacturing unit, if one department produces parts for use by a subsequent assembly department, the first department may be considered as the producer and the second, the consumer.

a submitted lot that has a per cent defective equal to the Lot Tolerance Per Cent Defective. This probability is termed the Consumer's Risk.

(*b*) *Average Quality Protection*—in which there is prescribed some chosen value of *average* per cent defective in the product *after inspection* (Average Outgoing Quality Limit, *AOQL*), that shall not be exceeded no matter what may be the level of per cent defective in the product submitted to the inspector.

Single sampling plans employing the first of these two types of protection were developed in an earlier paper.[3] An extension of the underlying theory as applied to double sampling will be given here. Sampling plans employing the second type of protection will likewise be covered for both the single sampling and double sampling procedures.*

The development of the second concept (*AOQL*) in 1927 was the result of a practical need in certain types of manufacturing process inspections, following considerable experience in the application of inspection procedures based on the first concept (Lot Tolerance and Consumer's Risk) which had been developed in 1924. Both have since been used extensively.

Minimum Amount of Inspection

For all of the four inspection plans covered, certain general principles, given in the earlier paper,[3] are used.

For each plan two requirements are imposed—first, that the plan shall provide a specified degree of protection (as covered by (a) or (b) above), and second, that the amount of inspection shall be a minimum for product of *expected* quality, subject to the degree of protection imposed by the first requirement.

The first requirement can be satisfied by a large number of different combinations of sample sizes and allowable defect numbers. The second requirement dictates which one of these combinations shall be chosen, and requires a determination of the value of per cent defective to be normally expected in product submitted to the inspector. This expected value is referred to as the "process average" per cent defective.

For the inspection procedures here adopted, the amount of inspection that will be done *in the long run* is made up of two parts: (1) the number of pieces inspected in the samples and (2) the number of pieces inspected in the remainder of those lots that fail to be accepted by sample. We are

* An adaptation of these concepts to inspection by the method of variables, using the arithmetic mean as an acceptance criterion, is given in a doctorate thesis (Columbia University) by H. G. Romig, "Allowable Average in Sampling Inspection," March 1939, for the case of a normally distributed characteristic that is statistically controlled with respect to the standard deviation.

to find a solution that will minimize the amount of inspection for uniform product* of process average quality.

In single sampling, for each combination of sample size and allowable defect number, there will be a definite probability of exceeding the allowable defect number for a sample drawn from uniform product of process average quality. This probability is termed the Producer's Risk. It represents the chance of not accepting a lot on the basis of the sample findings under these postulated conditions, and for the adopted inspection procedure is thus the chance of inspecting the remainder of the pieces in the lot. The average (expected) amount of inspection per lot then equals the number inspected in the sampled portion plus the product of the Producer's Risk and the number of pieces in the remainder of the lot. This average value can be found for each combination, and the desired solution is obtained by choosing that combination of sample size and allowable defect number for which the average amount of inspection is smallest.

In double sampling, an entirely similar procedure is followed. Here, of course, we must consider the probability of taking a second sample when the first sample fails, and then the probability of failure for the second sample. The overall chance of failure constitutes the Producer's Risk for the complete double sampling plan.

No distinction is made here as to who actually inspects the remainders of those lots that fail to be accepted by sample. Whether the consumer does this inspection, or rejects such lots and thus in effect requires the producer to do it, will be considered immaterial. Interest will be centered only on the total amount of inspection done, recognizing that no matter which agency performs this service the cost will probably be reflected in the overall cost to the consumer.

It should be noted that, in the theoretical developments, the number of defects observed in a sample is not used to "estimate" the quality of the lot. Instead, it serves to indicate what action should be taken—whether the lot should be accepted, subjected to further sampling, or inspected completely—the entire process constituting a set of operations which when repeated over and over again produce a desired end result.

Single Sampling—Lot Quality Protection

The solution for this plan was given in the earlier paper,[3] but will be reviewed briefly since certain of the principles and terms employed will be extended to the other three inspection plans.

* By "uniform product" is meant one produced under statistically controlled conditions such that the probability of producing a defective piece remains constant at some definite value p. The solution thus provides for a minimum of inspection *if* quality is statistically controlled at a per cent defective level equal to the process average per cent defective.

Protection is defined by specifying values of,

(a) Lot Tolerance Per Cent Defective, the allowable per cent defective in a lot.

(b) Consumer's Risk, the probability of accepting a lot of tolerance quality.

If the allowable defect number is c, then the Consumer's Risk is the probability of finding c or less defects in a random sample of n pieces drawn from a lot of N pieces in which the per cent defective is equal to the lot tolerance per cent defective. The tables presented are based on a Consumer's Risk of 0.10, a value found most useful in practice. For this choice, the chances of accepting a lot of worse than tolerance quality are less than 1 in 10.

TABLE 1

SOLUTION FOR A PARTICULAR CASE—SINGLE SAMPLING, LOT QUALITY PROTECTION

n and c Combinations for Lot Size, 1000; Lot Tol. % Def., 3%; Cons. Risk, 0.10		Application to Product having Proc. Av. % Def. = 0.45%				
Sample Size n	Allowable Defect Number, c	Prob. of Acceptance by Sample	Prob. of Inspecting Remainder of Lot (Producer's Risk)	Av. No. of Pieces Inspected per Lot: In Sample	In Remainder of Lot	Total
75	0	.713	.287	75	265	340
125	1	.891	.109	125	95	220
*170	*2	.*958	*.042	*170	*35	*205
210	3	.984	.016	210	13	223
250	4	.994	.006	250	5	255
290	5	.998	.002	290	1	291
325	6	.999+	.000+	325	0	325

* Plan involving minimum amount of inspection.

For each value of c, such as 0, 1, 2, etc., there is a unique value of sample size n, such that the probability of finding c or less defects is 0.10. Any of these combinations of n and c will thus provide the desired consumer protection.

Now, for a given value of process average per cent defective, one of these combinations involves a smaller total amount of inspection than any of the others, as illustrated in Table 1. This combination of n and c, which provides the desired solution, gives the most efficient adjustment between the Consumer's Risk and Producer's Risk from the standpoint of minimizing inspection effort. Fig. 2 shows the relationship between these two risks for the conditions given in Table 1.

Curves providing a basis for solutions, such as that given in Table 1, have been published[3] for a Consumer's Risk of 0.10. The appended SL tables (Single Sampling Lot Quality Protection) provide for practical use a complete set of such solutions for lot tolerance values from 0.5% to 10%. Each table is based on a particular value of lot tolerance per cent defective, and each solution, comprising a sample size, n, and allowable defect number, c, covers a range of lot sizes and a range of process average values.* The value of n given in the tables is based on the largest lot size for each lot size range, and the value of c corresponds to the mean lot size in each lot size range and to the mean value of process average in each

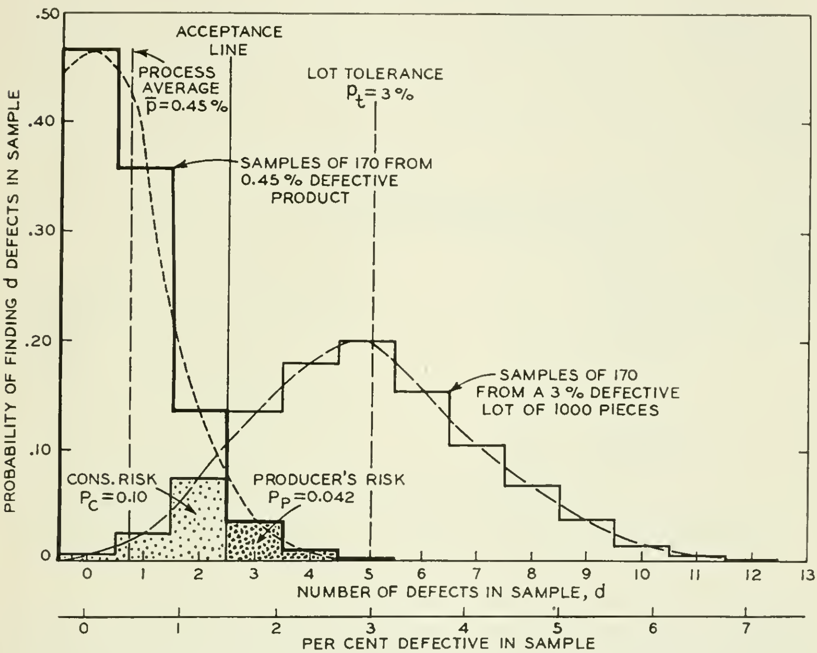


Fig. 2—Relation between consumer's risk and producer's risk

process average range, as indicated in Fig. 3. This procedure is followed for all of the sampling tables presented with this paper.

For the lot quality protection tables for both single and double sampling (SL and DL Tables), these choices are made to insure that, for the lot size range covered, the risk will not exceed the specified value (0.10) and to give *on the average*, for the process average range covered, the most economical plan. For reasons found advantageous in practice, sample sizes for samples of over 50 pieces are given to the nearest 5 units. For

* The extremely small process average range in the first column of each table has been specifically provided for those cases, increasingly common with long continued use of these inspection procedures, where the process average per cent defective is for all practical purposes zero.

extremely large samples, the size is given to the nearest 10. This basis of rounding sample sizes is followed for all of the sampling tables presented with this paper.

On each table are listed values of AOQL to indicate the upper bound to the long term average per cent defective in product after inspection that may be reached under the most adverse conditions.

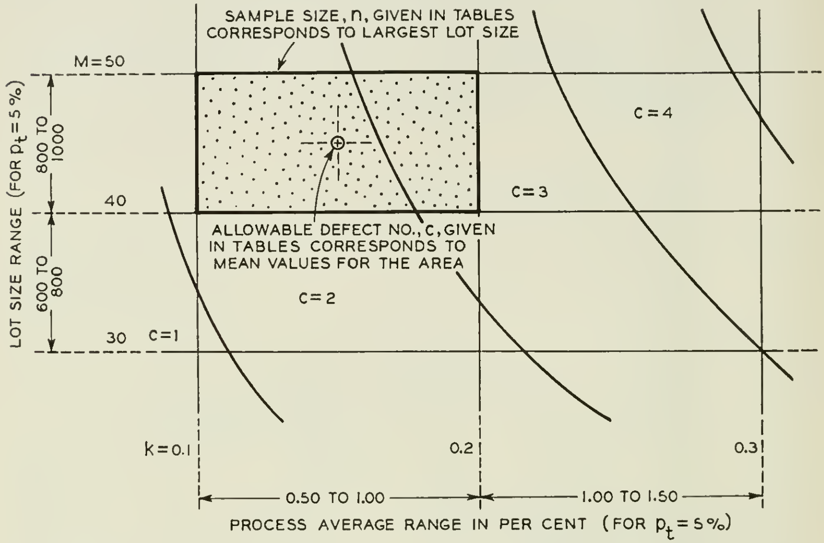


Fig. 3—Basis of choosing the n and c values given in the sampling tables

Double Sampling—Lot Quality Protection

The solution for this plan is carried out in substantially the same way as for single sampling. Protection is defined, as before, by specifying values of lot tolerance per cent defective and Consumer's Risk. As for the single sampling procedure, a Consumer's Risk value of 0.10 is adopted. In double sampling, a lot is given a second chance of acceptance if the first sample results are unfavorable, so that the Consumer's Risk is the sum of two parts: (a) the probability of accepting a lot of tolerance quality for the first sample, and (b) the probability of its acceptance for the second sample, if the first fails. For example, if the two allowable defect numbers, c_1 and c_2, are 1 and 7, respectively, the Consumer's Risk is the sum of the probabilities for all of the following possible ways in which these criteria may be met, as shown in Table 2.

As in the case of single sampling, for any given process average value there are a large number of acceptance criteria—pairs of c_1 and c_2 in this case—for each of which sample sizes may be selected so as to give the desired Consumer's Risk of 0.10, but we wish to choose the combinations of n_1, n_2, c_1 and c_2 that will involve a minimum amount of inspection for product of process average quality. Furthermore, there are an unlimited number of ways of apportioning the Consumer's Risk between the first and second samples for each process average value. This latter factor introduces one more variable factor than will permit of a ready solution by other than trial and error methods, and accordingly an empirical choice has been made on the basis of a complete investigation of the relative practical advantages of several possible choices. Specifically, the solutions are based on an apportionment such that the risk for the first sample is equal to the risk for an independent sample equal in size to the first and second samples

TABLE 2

COMPUTATION OF CONSUMER'S RISK—DOUBLE SAMPLING

No. of Defects		Probability for $n_1 = 88$, $n_2 = 154$ 5% Defective Lot of 1000 pieces	
In 1st Sample	In 2nd Sample		
0		.010	Accepted by 1st Sample
1		.048	
2	0, 1, 2, 3, 4 or 5	.018	Accepted by 2nd Sample
3	0, 1, 2, 3 or 4	.015	
4	0, 1, 2 or 3	.007	
5	0, 1 or 2	.002	
6	0 or 1	.000	
7	0	.000	
Total		.100	Consumer's Risk

combined. The use of an 0.06 risk in determining n_1 and $n_1 + n_2$ for given values of c_1 and c_2 provides a Consumer's Risk of almost exactly 0.10 over a considerable portion of the field covered by the tables, though in some areas a value as low as 0.056 is necessary. The "minimum" solutions for double sampling are, of course, conditioned by this choice.*

As shown in the Appendix, paired values of c_1 and c_2 that satisfy the condition of minimum inspection depend on (1) the tolerance number of defects for a lot, and (2) the ratio of the process average to the lot tolerance

* Study of the effect of different apportionments of the Consumer's Risk on the average amount of inspection for product of process average quality indicates that considerably more than half of the 0.10 risk should be taken for small process average values and that less than half should be taken for large process average values. The single choice that was made provides a solution that closely approximates the true minimum over a large portion of the tables, and was considered justified by the great saving in computation effort. With this choice, the average amount of inspection per lot does not in general exceed the true minimum by more than 3 to 5% although for extremely low process average values the excess may be as much as 15%.

per cent defective. These values have been determined by trial and error and form the basis of the c_1 c_2 zones given in Fig. 7 of the Appendix.

The appended DL tables (Double Sampling Lot Quality Protection) provide a complete set of solutions using paired values of c_1 and c_2 determined from Fig. 7. These tables are constructed on the same principles as the single sampling tables described above.

Single Sampling—Average Quality Protection

The solution for this plan considers the degree to which the entire inspection procedure screens out defects in the product submitted to the inspector. Lots accepted by sample undergo a partial screening through the elimination of defects found in samples. Lots that fail to be accepted

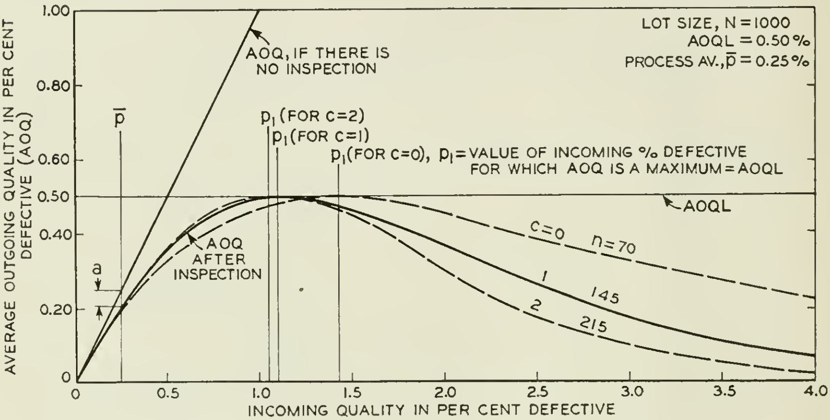


Fig. 4—Relationship between incoming quality, outgoing quality and *AOQL*

by sample are completely cleared of defects. The overall result is some average per cent defective in the product as it leaves the inspector, termed the "average outgoing quality," which depends on the level of per cent defective for incoming product and the proportion of total defects that are screened out.

The solid curve of Fig. 4 shows how the average outgoing quality varies for different values of incoming quality for a lot size of $N = 1000$, a sample size of $n = 145$ and an allowable defect number of $c = 1$. The curve is based on the concept of incoming product of uniform quality treated mathematically as an homogeneous universe. As the level of incoming per cent defective gets higher and higher, more and more lots are completely inspected. In turn, the average outgoing per cent defective increases, reaches a maximum value (0.50%, in Fig. 4), and then falls off as a result of rapid increase in the amount of screening. This maximum value is termed the average outgoing quality limit (AOQL).

For this plan, protection is defined by specifying a definite value of AOQL. For each possible value of c such as 0, 1, 2, etc. there is a unique value of sample size that will give the specified value of AOQL. This is illustrated in Fig. 4. Any of these combinations of n and c provide the desired protection, and as for the lot quality protection plans, we choose that combination of n and c that gives a minimum amount of inspection for uniform product of process average quality.

In the Appendix it is shown that the allowable defect number satisfying the condition of minimum inspection is dependent on two factors (1) the number of defects per lot for process average quality, and (2) the ratio of the process average per cent defective to the AOQL value. Fig. 9 of the Appendix defines zones of allowable defect numbers for which the average amount of inspection is a minimum.

The appended SA tables (Single Sampling Average Quality Protection) provide a complete set of minimum inspection solutions for AOQL values from 0.1% to 10%. The choice of n and c for each solution in the tables is based on the procedure of Fig. 3 (using c zones given by Fig. 9), to insure that the AOQL value over the area in question will not exceed the specified value and to give on the average for this area the most economical plan.

On each table are given values of lot tolerance per cent defective for a Consumer's Risk of 10%. These values are found useful in practice since it is often desirable to know the degree of protection afforded to individual lots.

Double Sampling—Average Quality Protection

The solution for double sampling differs from that for single sampling in that no simple relation has been found that gives directly the sample sizes that will result in a specified value of AOQL for a given lot size. This, together with the lack of simple relations for determining the choice of allowable defect numbers (c_1 and c_2) that provide a minimum solution, has necessitated an empirical choice, the consequence of which is much the same as for the similar action taken in the solution of the problem of double sampling for lot quality protection.* Specifically, the interrelationship between n_1, n_2, c_1 and c_2 used in the latter case for a 10% Consumer's Risk is used again here and the solutions given are consequently minima that are contingent on this choice. An extensive trial and error investigation, using the underlying theoretical relations, leads to the conclusion that the degree to which the solutions given in these tables approach the true minima, is of the same order of magnitude as for the double sampling tables for lot quality protection.

The method of solution is essentially that illustrated by example in the

* See footnote page 11.

Appendix. The pairs of values of c_1 and c_2 used in the solution are confined to those given in Fig. 7 of the Appendix. For each of these pairs of c_1 and c_2, sample sizes are determined, using the above mentioned relationship to a 10% Consumer's Risk, that will give the desired AOQL value. Of these several sets of c_1, c_2, n_1 and n_2, that one is selected which involves the least amount of inspection.

The appended DA tables (Double Sampling Average Quality Protection) provide a complete set of such minimum inspection solutions for AOQL values from 0.1% to 10%. The choice of n_1, n_2, c_1 and c_2 for each solution in the tables is based on the general procedure of Fig. 3 (using the zones given in Fig. 7) to insure that the AOQL value over the area in question will not exceed the specified value and to give on the average for this area the most economical plan.

As for the single sampling AOQL tables there are listed values of lot tolerance per cent defective for a Consumer's Risk of 10%. In this case, these values have entered directly into the solution as explained above.

Application of Sampling Tables

In the above description of the sampling tables, attention has been confined to the inspection of a single characteristic. The tables are, however, equally applicable to a group of characteristics considered collectively provided defects with respect to these characteristics are of essentially the same seriousness and may, therefore, be considered additive. When such application is made, the per cent defective values given in the tables embrace all such defects collectively, and since more than one defect may occur on a single piece of product, any allowable defect number listed in the tables should, by agreement, be considered either as a "number of defective pieces" or as a "number of defects."

The sampling tables based on lot quality protection (Tables SL and DL) are perhaps best adapted to conditions where interest centers on each lot separately—for example, where the individual lot tends to retain its identity either from a shipment or a service standpoint. They have been found particularly useful in inspections made by the ultimate consumer or his purchasing agent for lots or shipments purchased more or less intermittently.

The sampling tables based on average quality protection (Tables SA and DA) are especially adapted for use where interest centers on the *average* quality of product after inspection rather than on the quality of each individual lot and where inspection is, therefore, intended to serve, if necessary, as a partial screen for defective pieces. The latter point of view has been found particularly helpful, for example, in consumer inspections of continuing purchases of large quantities of a product, and in manufacturing process inspections of parts where the inspection lots tend to lose their identity by merger in a common storeroom from which quantities are withdrawn on order as needed.

Other things being equal the average amount of inspection for double sampling is less than for single sampling. Fig. 5* gives a direct comparison for the lot protection tables (SL and DL). The saving obtained by using double instead of single sampling is greatest for large lot sizes and low process averages. Over the area of the tables found most useful in practice (per-

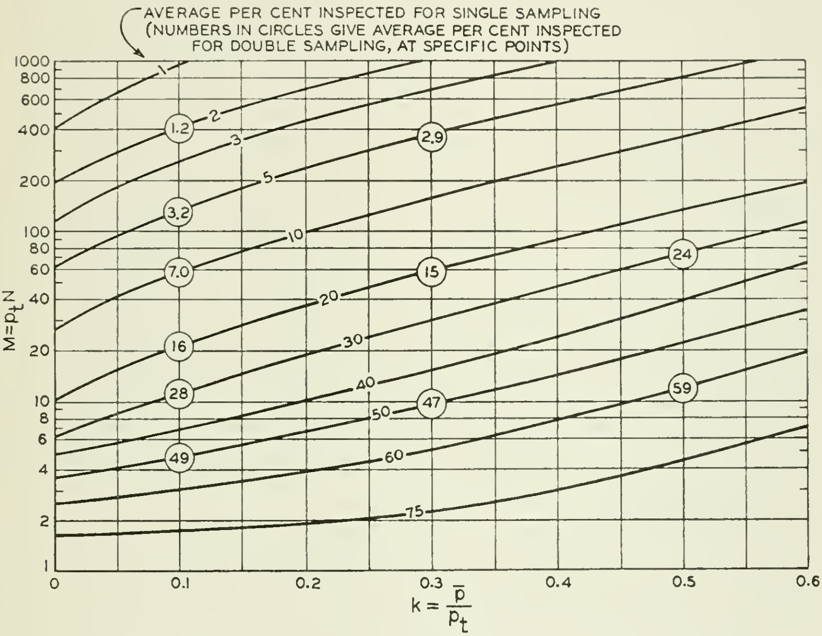


Fig. 5—Relative amount of inspection, double and single sampling

centage inspection less than 25 or 30%), the saving generally exceeds 10% and may be as great as 50%. The saving that results from using the double sampling instead of the single sampling AOQL tables (SA and DA) is of the same order of magnitude and may be estimated roughly from Fig. 5 by using the associated lot tolerance values listed in the AOQL tables, for a chosen set of AOQL, lot size, and process average values. While the amount of inspection is a major cost item, other costs associated with double

* The curves and figures on this chart should be regarded as approximate. The mathematical relations involved are such that there exist unique values to be plotted on the M–k plane when certain approximate probability equations, referred to in the appendix, are employed in the solution, but not when exact equations are employed.

sampling frequently throw the advantage to single sampling. Among the added costs are those associated with interruption of work, extra handling of product, etc. incidental to the selection of an independent second sample. Aside from these considerations, it is common to find a psychological preference for double sampling. This appears to be associated with the tendency to look with favor on any plan that permits a "second chance" to make good, particularly when an initial failure is of a marginal character.

Given a specific problem of replacing 100% screening inspection by a sampling inspection, the first step is to decide on the type of protection desired, to select the desired limit of per cent defective—lot tolerance or AOQL value—for that type of protection, and to choose between single and double sampling. This results in the selection of one of the appended tables. The second step is to determine whether the quality of product is good enough to warrant the introduction of sampling. The economies of sampling will be realized, of course, only insofar as the per cent defective in submitted product is such that the acceptance criteria of the selected sampling plan will be met. A statistical analysis of past inspection results should first be made, therefore, in order to determine existing levels and fluctuations in the per cent defective for the characteristic or the group of characteristics under consideration. This provides information with respect to the degree of control of quality as well as the usual level of per cent defective to be expected under existing conditions. From this and other information is to be determined a value for the "process average" per cent defective that should be used in applying the selected sampling table, if sampling is to be introduced.

The determination of the process average per cent defective is an engineering problem, essentially one of prediction, in which use is made of all available information—knowledge of manufacturing conditions past and anticipated, judgment as to what periods of the past, if any, may be taken as representative of the future, results of analyses showing uniformity and level of per cent defective for such past periods, etc. The application of "control chart" analysis[1,7] to past data is especially recommended.* If

* The following procedure has been used with general success. Tabulate the observed values of fraction defective, p, for at least 25 immediately preceding lots (or groups of lots, say by days or weeks, if p is very small), excluding lots that are nonrepresentative for known reasons, and apply the control chart test to the observed values of p. If the data show statistical control, and if there are grounds for believing that future manufacturing conditions will be essentially the same as those of the past, use the average of the observed values of p as the process average value, $\bar{p}$. If lack of statistical control is shown, replace values of p that are beyond $\pm 3\sigma$ control limits[1, 7] by values corresponding to $\pm 2\sigma$ control limits (where $\sigma = \sqrt{\bar{p}\,(1 - \bar{p})/n}$). Compute a corrected average value of p, in which the individually corrected values are used in place of the corresponding observed values. Unless other conflicting evidence predominates, use this corrected value as a tentative process average value, until such time as a revision appears warranted on the basis of new evidence.

the process average value thus determined is well within the range of process average values listed in the selected sampling table then sampling can advantageously be introduced. If it is beyond this range, it would be quite satisfactory from a protection standpoint to use the last process average column of the selected table but the sampling plan itself would force rejection or a screening inspection of such a large proportion of the lots that the introduction of sampling probably would not pay. If the process average value is but poorly estimated, the amount of inspection will be somewhat larger than need be but the specified degree of protection will still be realized. Where there is uncertainty it is better to overestimate than to underestimate the process average value since, for a given magnitude of error, a lesser amount of excess inspection will thereby be incurred.

It should be especially noted that the tables may be safely applied whether quality is well controlled or not. If, for example, the usual level of per cent defective is well within the range of process average values listed in the selected table but individual lots are frequently well outside this range, the sampling plan will usually permit acceptance by sampling while quality is good but force 100% inspection when it is bad.

Experience with the tables indicates that where the procedures are used by a manufacturer within his own organization or by a consumer who rejects lots that are not accepted by sample, the general plan forces corrective action whenever quality becomes poorer than normally expected. The attendant increase in overall inspection costs provides a compelling argument, in a language well understood by all, for determining the cause of trouble in the manufacturing process and for instituting measures for eliminating it as speedily as possible. Thus, while the inspection procedures have as their immediate purpose the provision of a curative technique whereby product already made is cleared of abnormal proportions of defects, they are found by experience to enforce the adoption of a preventive technique—one that exerts economic pressure to track down and remove causes of abnormal quality variations, thus enforcing control of quality in the process and assuring better health in the product of tomorrow. Because of these factors the long term average outgoing per cent defective may rarely be expected to exceed half the AOQL value associated with the inspection plan in use.

Quality control is achieved most efficiently, of course, not by the inspection operation itself but by getting at causes.[6] It may be expedited by carrying out regular statistical control analyses of the cumulative results of sampling inspection—preparing quality control charts[1,7] for "per cent defective" with subgrouping of results on a lot-by-lot, a day-by-day, or a week-by-week basis—and making the findings available to those directly responsible for manufacturing processes.

Where a steady supply of product is offered for acceptance on a lot-by-lot basis, the use of these sampling procedures and tables, together with continuing control chart analyses of the inspection results obtained therefrom, have been found to provide a balanced and economical inspection program.

Acknowledgment

Work underlying the development and application of these tables has been contributed by many individuals in the Bell Telephone Laboratories and the Western Electric Company. The authors here express their indebtedness to these associates, particularly to those in the Western Electric Company who cooperated in the early development of the technical features of the plans and worked out shop procedures for use in their application. The laborious work of computing and preparing the tables in their final form was carried out by Miss Mary N. Torrey and Miss Ruth A. Bender—we wish to express our appeciation to them for their efforts to make the tables as free from error as possible.

References

1. W. A. Shewhart, "Economic Control of Quality of Manufactured Product," (New York: D. Van Nostrand Co., 1931).
2. E. S. Pearson, "The Application of Statistical Methods to Industrial Standardisation and Quality Control" (British Standards Institution, London, 1935).
3. H. F. Dodge and H. G. Romig, "A Method of Sampling Inspection," *The Bell System Technical Journal*, Vol. VIII, October 1929, pp. 613–631.
4. H. F. Dodge, "Acceptance-Rejection Requirements in Specifications," *Proc. A.S.T.M.* Vol. 34, Part II, 1934, pp. 877–890.
5. G. Udney Yule and M. G. Kendall, "An Introduction to the Theory of Statistics," (11th Ed.; London: Charles Griffin & Co., 1937).
6. W. A. Shewhart, "Statistical Method from the Viewpoint of Quality Control," (The Graduate School, U. S. Dept. of Agriculture, Washington, 1939).
7. *A.S.T.M. Manual on Presentation of Data*, Supplement B (Philadelphia: A.S.T.M., 1940).
8. Karl Pearson, "Tables for Statisticians and Biometricians," (London: Cambridge University Press, 1914), Table LI.
9. G. A. Campbell, "Probability Curves Showing Poisson's Exponential Summation," *The Bell System Technical Journal*, Vol. II, January 1923, pp. 95–113.
10. Frances Thorndike, "Applications of Poisson's Probability Summation," *The Bell System Technical Journal*, Vol. V, October 1926, pp. 604–624.

MATHEMATICAL APPENDIX

Fundamental Probability Formulas

The mathematical probabilities used in the solutions are based on equations corresponding to one or the other of the following two sets of conditions:

(a) Sampling from a finite universe.

(b) Sampling from an infinite universe.

In relations involving the determination of the Consumer's Risk, the sample is considered as a sample from a lot of a finite number of pieces and probabilities are correspondingly based on (a). For all other relations in the solutions—involving the determination of the Producer's Risk, the determination of the average number of pieces inspected per lot, etc.— the sample is considered as a sample from the general output of product—a source of supply—and probabilities are correspondingly based on (b).

Finite Universe

The probability of finding m defects in a random sample of n units drawn from a finite universe (lot) of N pieces in which the number of defects is $M = pN$, is given exactly by

$$P_{m,n,N,M} = \frac{1}{C_n^N} C_{n-m}^{N-M} C_m^M . \tag{1}$$

When $p < 0.10$, a good approximation to (1) is given by the $m + 1$st term of the expansion of the binomial, $\left[\left(1 - \frac{n}{N}\right) + \frac{n}{N}\right]^M$,

$$P_{m,n,N,M} \approx P_{m,\frac{n}{N},M} = C_m^M \left(1 - \frac{n}{N}\right)^{M-m} \left(\frac{n}{N}\right)^m . \tag{1'}$$

When $p < 0.10$ and when $\frac{n}{N} < 0.10$, a good approximation to (1) is given by the $m + 1$st term of the Poisson exponential distribution,

$$P_{m,n,N,M} \approx P_{m,pn} = \frac{e^{-pn}(pn)^m}{m!} . \tag{1''}$$

These are general equations applicable for any fraction defective, p, but are used in this paper only for the specific case where $p = p_t$, the lot tolerance fraction defective, and where in turn $M = p_tN$.

The Consumer's Risk P_C, is the probability of meeting the acceptance criteria—c, for single sampling, and c_1 and c_2, for double sampling—in samples drawn from a lot of N pieces containing exactly the tolerance number of defects $M = p_tN$.

For single sampling,

$$P_C = \sum_{m=0}^{m=c} P_{m,n,N,M} \quad (\text{when } p = p_t). \tag{2}$$

For double sampling,

$$\begin{aligned} P_C = \sum_{m=0}^{m=c_1} P_{m,n_1,N,M} &+ P_{c_1+1,n_1,N,M} \sum_{m=0}^{m=c_2-c_1-1} P_{m,n_2,N-n_1,M-c_1-1} \\ &+ P_{c_1+2,n_1,N,M} \sum_{m=0}^{m=c_2-c_1-2} P_{m,n_2,N-n_1,M-c_1-2} + \cdots \\ &+ P_{c_2,n_1,N,M} P_{0,n_2,N-n_1,M-c_2} \quad (\text{when } p = p_t). \end{aligned} \tag{3}$$

Values of P_C in equations (2) and (3) are given approximately by substituting $P_{m,\frac{n}{N},M}$ or $P_{m,pn}$ for $P_{m,n,N,M}$ throughout, in accordance with equations (1′) and (1″), using $p = p_t$. The resulting equations will be referred to as (2′), (2″), (3′) and (3″), respectively. (2′) (2″) (3′) (3″)

Infinite Universe

The probability of finding m defects in a random sample of n pieces drawn from an infinite universe (general output of uniform product) in which the fraction defective is p, is given exactly by the $m + 1$st term of the expansion of the binomial, $[(1 - p) + p]^n$,

$$P_{m,n,p} = C_m^n(1 - p)^{n-m} p^m. \tag{4}$$

When $p < 0.10$, a good approximation to (4) is given by the $m + 1$st term of the Poisson exponential distribution,

$$P_{m,n,p} \approx P_{m,pn} = \frac{e^{-pn}(pn)^m}{m!}. \tag{4'}$$

The probability of meeting the acceptance criteria—c, for single sampling, and c_1 and c_2 for double sampling—in samples drawn from submitted product having a fraction defective of p, is termed the probability of acceptance, P_a. For single sampling,

$$P_a = \sum_{m=0}^{m=c} P_{m,n,p}. \tag{5}$$

For double sampling,

$$P_a = \sum_{m=0}^{m=c_1} P_{m,n_1,p} + P_{c_1+1,n_1,p} \sum_{m=0}^{m=c_2-c_1-1} P_{m,n_2,p} + P_{c_1+2,n_1,p} \sum_{m=0}^{m=c_2-c_1-2} P_{m,n_2,p}$$
$$+ \cdots + P_{c_2,n_1,p} P_{0,n_2,p}. \tag{6}$$

Values of P_a in equations (5) and (6) are given approximately by substituting Poisson exponential probabilities, $P_{m,pn}$, for $P_{m,n,p}$ throughout in accordance with equation (4′). The resulting equations will be referred to as equations (5′) and (6′), respectively. (5′) (6′)

The Poisson exponential approximation is used in subsequent paragraphs wherever probabilities in sampling from an infinite universe apply. Tables[8] and charts[9, 10] are available from which these probability values (single term values, or cumulative values for "c or less defects") may be read directly.* Figure 6 gives a cumulative probability chart for the Poisson exponential distribution, which is widely useful in the solutions involved.

The Producer's Risk, P_P, is the probability of failing to meet the acceptance criteria in samples drawn from product of process average ($\bar{p}$) quality. Using $p = \bar{p}$ in equations (5) and (6),

$$P_P = 1 - P_a \text{ (when } p = \bar{p}). \tag{7}$$

Lot Quality Protection

Single Sampling

Given: Lot Size (N), lot tolerance fraction defective (p_t), Consumer's Risk ($P_C = 0.10$), process average fraction defective ($\bar{p}$).

* In this work use was made of more complete tables, giving cumulative probabilities for pn values up to 100, prepared by Office of the Switching Theory Engineer, Bell Telephone Laboratories.

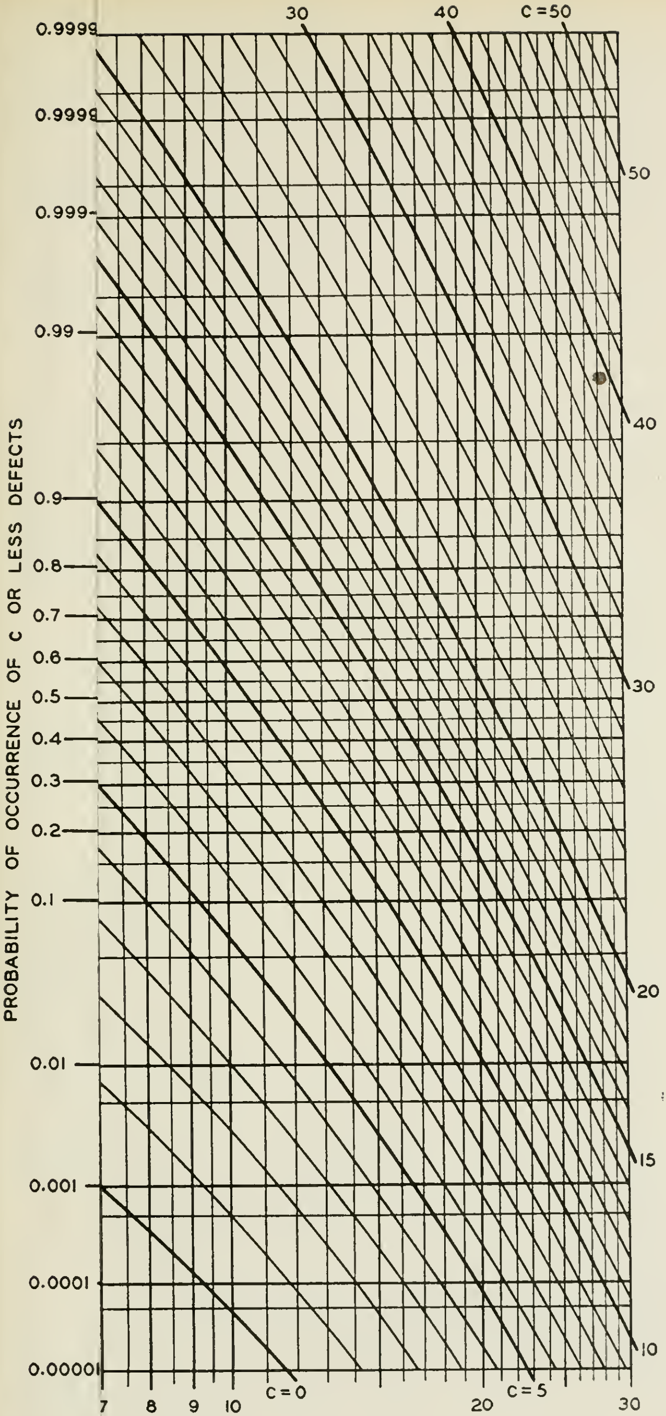


Fig. 6—universe in which the fraction defective is p (A modi-

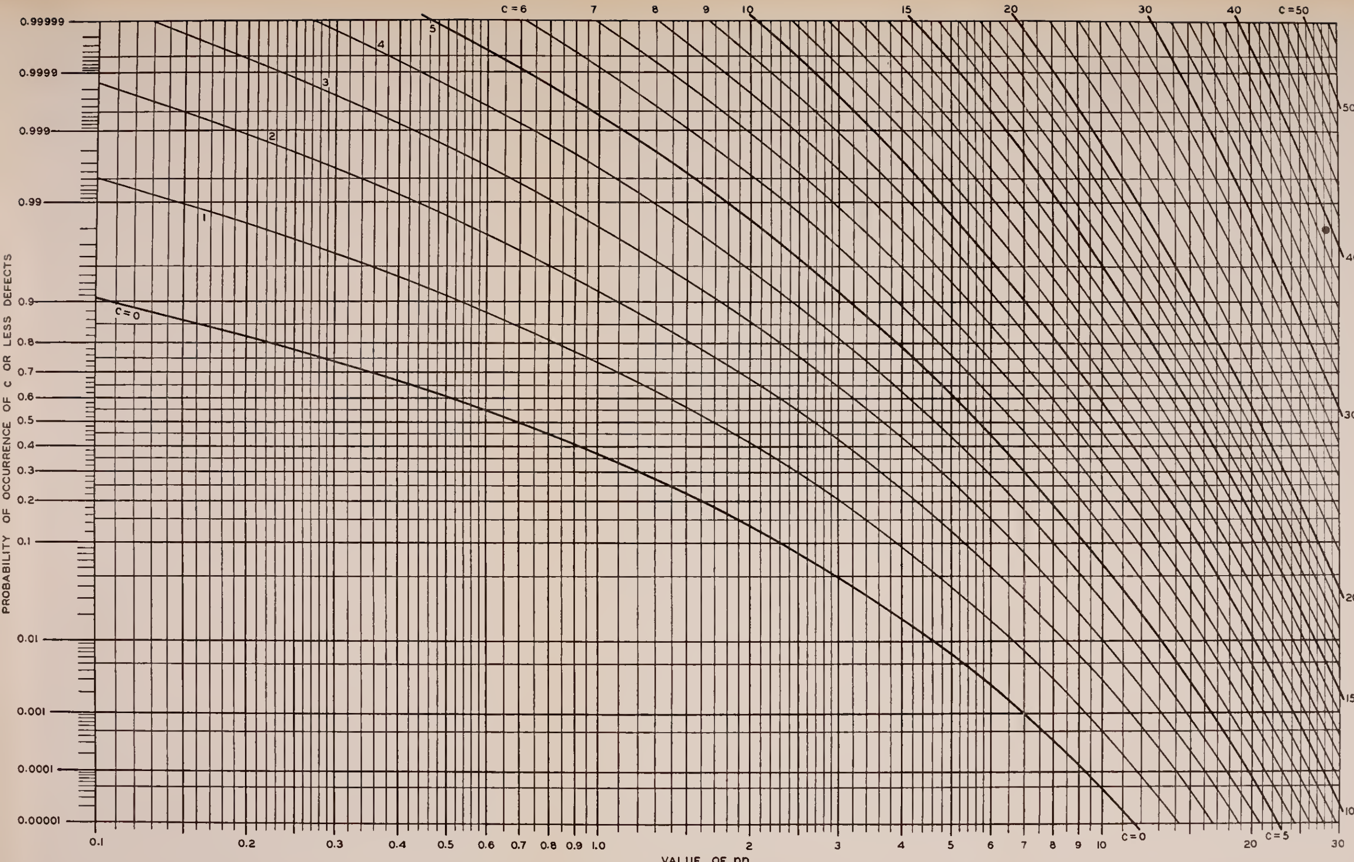


Fig. 6—Cumulative probability curves—Poisson exponential. For determining probability of occurrence of c or less defects in a sample of n pieces selected from an infinite universe in which the fraction defective is p (A modification of chart given by Miss F. Thorndike *B. S. T. J.*, October 1926).

To find: Values of n and c that will minimize $\bar{I}$, the average number of pieces inspected per lot for product of process average ($\bar{p}$) quality.

The average number of pieces inspected per lot (I) for product of p quality is given by

$$I = n + (N - n)(1 - P_a), \tag{8}$$

where P_a is given by equation (5). Substituting the approximation of equation (5′) gives

$$I = n + (N - n)\left(1 - \sum_{m=0}^{m=c} P_{m,pn}\right). \tag{8′}$$

$\bar{I}$ is a specific value of I and is obtained from equation (8′) by using $p = \bar{p}$. The value of c that makes $\bar{I}$ a minimum may be read from the chart of Fig. 2 of the previous paper,[3] which uses coordinates of $M = p_t N$ and $k = \dfrac{\bar{p}}{p_t}$ and is based on $P_C = 0.10$. The corresponding sample size n may be read from Fig. 3 of the previous paper[3] (based on equation (2′)), from Fig. 6 if appropriate, or by direct computation from equation (2), (2′), or (2″), using $P_C = 0.10$.

Double Sampling

Given: Lot size (N), lot tolerance fraction defective (p_t), Consumer's Risk ($P_C = 0.10$), process average fraction defective ($\bar{p}$).

To find: Values of n_1, n_2, c_1, c_2 that will minimize $\bar{I}$.

The average number of pieces inspected per lot (I) for product of p quality is given by

$$I = n_1 + n_2\left(1 - \sum_{m=0}^{m=c_1} P_{m,pn_1}\right) + (N - n_1 - n_2)(1 - P_a), \tag{9}$$

where P_a is determined from equation (6′).

$\bar{I}$ is a specific value of I and is obtained from equation (9) by using $p = \bar{p}$. As outlined on page 11, the pair of values of c_1 and c_2 that makes $\bar{I}$ a minimum is determined by trial and error, conditioned by the choice that the Consumer's Risk of 0.10 be divided between the first and second samples so that the "initial risk" for the first sample is 0.06. Figure 7 gives such pairs of c_1, c_2 values, corresponding to values $M = p_t N$ and $k = \dfrac{\bar{p}}{p_t}$.

For the selected apportionment of Consumer's Risk, the sample sizes n_1 and n_2 may be determined approximately from the following equations, which are based on equation (1′),

$$\left.\begin{aligned} 0.06 &= \sum_{m=0}^{m=c_1} C_m^M \left(1 - \frac{n_1}{N}\right)^{M-m} \left(\frac{n_1}{N}\right)^m, \\ 0.06 &= \sum_{m=0}^{m=c_2} C_m^M \left(1 - \frac{n_1 + n_2}{N}\right)^{M-m} \left(\frac{n_1 + n_2}{N}\right)^m. \end{aligned}\right\} \tag{10}$$

Figure 8 based on these equations gives $p_t n_1$ and $p_t(n_1 + n_2)$ values associated with c_1 and c_2 for a given value of $M = p_t N$, and thus provides the desired values of n_1 and n_2.

To find: Values of n and c that will minimize $\bar{I}$, the average number of pieces inspected per lot for product of process average ($\bar{p}$) quality.

The average number of pieces inspected per lot (I) for product of p quality is given by

$$I = n + (N - n)\,(1 - P_a), \tag{8}$$

where P_a is given by equation (5). Substituting the approximation of equation (5′) gives

$$I = n + (N - n)\left(1 - \sum_{m=0}^{m=c} P_{m,pn}\right). \tag{8'}$$

$\bar{I}$ is a specific value of I and is obtained from equation (8′) by using $p = \bar{p}$. The value of c that makes $\bar{I}$ a minimum may be read from the chart of Fig. 2 of the previous paper,[3] which uses coordinates of $M = p_t N$ and $k = \dfrac{\bar{p}}{p_t}$ and is based on $P_C = 0.10$. The corresponding sample size n may be read from Fig. 3 of the previous paper[3] (based on equation (2′)), from Fig. 6 if appropriate, or by direct computation from equation (2), (2′), or (2″), using $P_C = 0.10$.

Double Sampling

Given: Lot size (N), lot tolerance fraction defective (p_t), Consumer's Risk ($P_C = 0.10$), process average fraction defective ($\bar{p}$).

To find: Values of n_1, n_2, c_1, c_2 that will minimize $\bar{I}$.

The average number of pieces inspected per lot (I) for product of p quality is given by

$$I = n_1 + n_2\left(1 - \sum_{m=0}^{m=c_1} P_{m,pn_1}\right) + (N - n_1 - n_2)(1 - P_a), \tag{9}$$

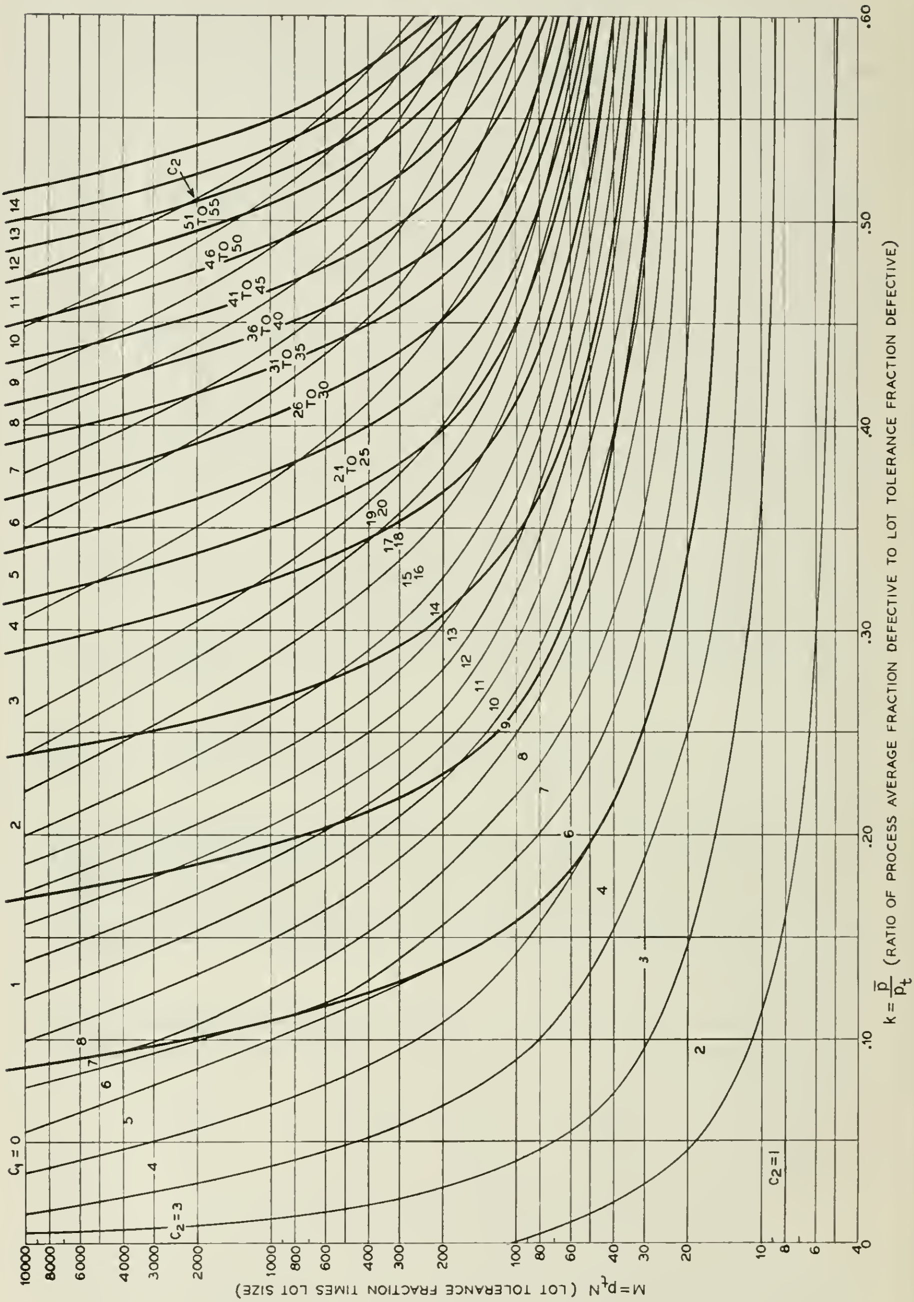
where P_a is determined from equation (6′).

$\bar{I}$ is a specific value of I and is obtained from equation (9) by using $p = \bar{p}$. As outlined on page 11, the pair of values of c_1 and c_2 that makes $\bar{I}$ a minimum is determined by trial and error, conditioned by the choice that the Consumer's Risk of 0.10 be divided between the first and second samples so that the "initial risk" for the first sample is 0.06. Figure 7 gives such pairs of c_1, c_2 values, corresponding to values $M = p_t N$ and $k = \dfrac{\bar{p}}{p_t}$.

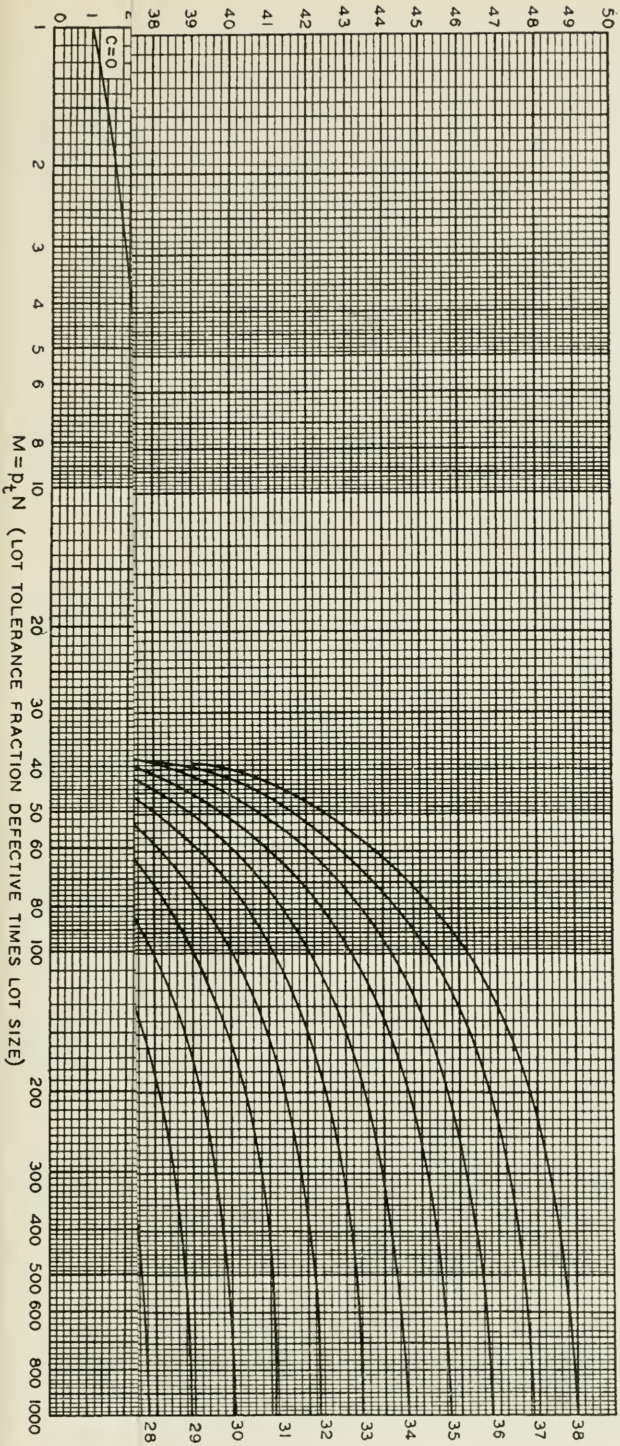
For the selected apportionment of Consumer's Risk, the sample sizes n_1 and n_2 may be determined approximately from the following equations, which are based on equation (1′),

$$\left.\begin{aligned} 0.06 &= \sum_{m=0}^{m=c_1} C_m^M \left(1 - \frac{n_1}{N}\right)^{M-m} \left(\frac{n_1}{N}\right)^m, \\ 0.06 &= \sum_{m=0}^{m=c_2} C_m^M \left(1 - \frac{n_1 + n_2}{N}\right)^{M-m} \left(\frac{n_1 + n_2}{N}\right)^m. \end{aligned}\right\} \tag{10}$$

Figure 8 based on these equations gives $p_t n_1$ and $p_t(n_1 + n_2)$ values associated with c_1 and c_2 for a given value of $M = p_t N$, and thus provides the desired values of n_1 and n_2.



$k = \frac{\bar{p}}{p_t}$ (RATIO OF PROCESS AVERAGE FRACTION DEFECTIVE TO LOT TOLERANCE FRACTION DEFECTIVE)
$M = p_t N$ (LOT TOLERANCE FRACTION TIMES LOT SIZE)
$c_1 = 0$
$c_2 = 3$
$c_2 = 1$
c_2
51 TO 55
46 TO 50
41 TO 45
36 TO 40
31 TO 35
26 TO 30
21 TO 25
19 20
17 18
15 16



ll as
roxi-
(3)

(11)

ieses

ality

(12)

re-

(12′)

In
rom

will
sub-
rs is

(13)

tion

(14)

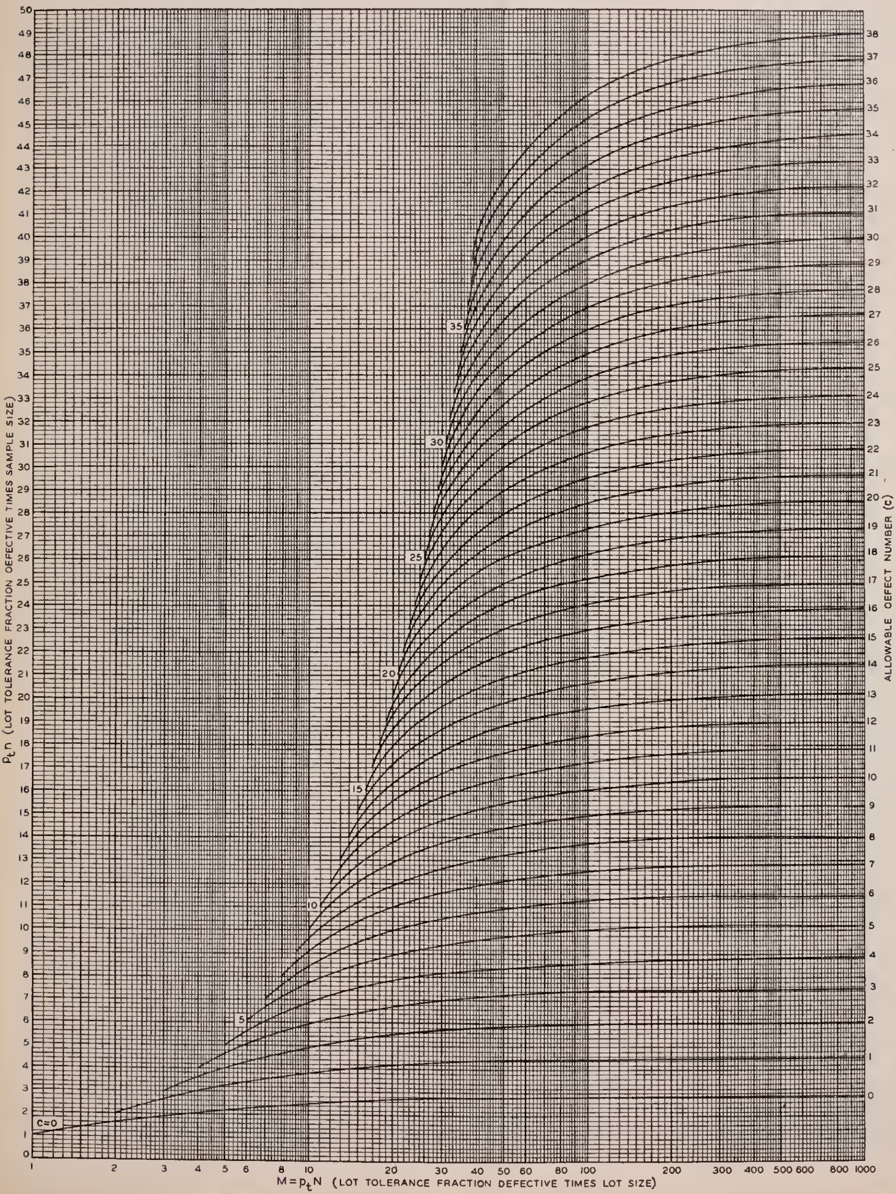


Fig. 8—Curves for determining sample sizes, n_1 and n_2—lot tolerance protection, Consumer's Risk, 0.10.

The use of $P = 0.06$ for determining $n_1 + n_2$ corresponding to c_2 as well as for determining n_1 corresponding to c_1 results in a Consumer's Risk of approximately 0.10, as may be checked by writing the Consumer's Risk equation (3) as follows:

$$P_C = \sum_{m=0}^{m=c_1} P_{m,n_1,N,M} + \sum_{m=0}^{m=c_2} P_{m,n_1+n_2,N,M} - \Big(P_{0,n_1,N,M} \sum_{m=0}^{m=c_2} P_{m,n_2,N-n_1,M}$$
$$+ P_{1,n_1,N,M} \sum_{m=0}^{m=c_2-1} P_{m,n_2,N-n_1,M-1} + \cdots$$
$$+ P_{c_1,n_1,N,M} \sum_{m=0}^{m=c_2-c_1} P_{m,n_2,N-n_1,M-c_1}\Big). \qquad (11)$$

The sum of the first two terms is 0.12 and the sum of the terms in parentheses is of the order of 0.02.

Average Quality Protection

General Relations

When the fraction defective in submitted product is p, the average quality after inspection (p_A) is given by

$$p_A = p\frac{N - I}{N} \qquad (12)$$

when all defective pieces found are replaced. If defective pieces found are removed but not replaced,

$$p_A = p\frac{N - I}{N - pI}, \qquad (12')$$

the factor pI representing the average number of defective pieces removed. In deriving the tables, equation (12) has been used. The error in p_A resulting from the use of (12) rather than (12′) is $\frac{pI}{N}$, which is generally small.

The average outgoing quality limit (p_L) is the maximum value of p_A that will result under any sampling plan, considering all possible values of p in the submitted product. The value of p for which this maximum value of p_A occurs is designated as p_1, hence

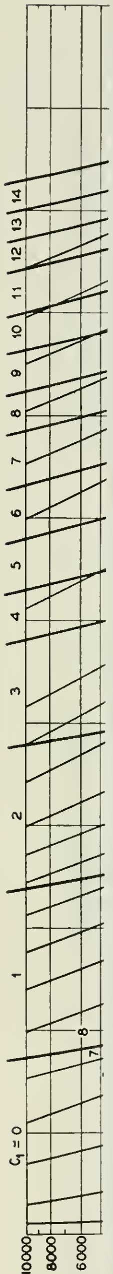
$$p_L = p_1\frac{N - I}{N}. \qquad (13)$$

The value of p_1 for which $p_A = p_L$ may be determined by differentiating equation (12) with respect to p, equating to 0, and solving for p, that is

$$\frac{dp_A}{dp} = \frac{N - I}{N} - \frac{p}{N}\frac{dI}{dp} = 0. \qquad (14)$$

Single Sampling

Given: Lot size (N), *AOQL* (p_L), process average fraction defective ($\bar{p}$).
To find: Values of n and c that will minimize $\bar{I}$.


10000
8000
6000
$c_1 = 0$
1
2
3
4
5
6
7
8
9
10
11
12
13
14
8
7

The use of $P = 0.06$ for determining $n_1 + n_2$ corresponding to c_2 as well as for determining n_1 corresponding to c_1 results in a Consumer's Risk of approximately 0.10, as may be checked by writing the Consumer's Risk equation (3) as follows:

$$P_C = \sum_{m=0}^{m=c_1} P_{m,n_1,N,M} + \sum_{m=0}^{m=c_2} P_{m,n_1+n_2,N,M} - \Bigg(P_{0,n_1,N,M} \sum_{m=0}^{m=c_2} P_{m,n_2,N-n_1,M} + P_{1,n_1,N,M} \sum_{m=0}^{m=c_2-1} P_{m,n_2,N-n_1,M-1} + \cdots + P_{c_1,n_1,N,M} \sum_{m=0}^{m=c_2-c_1} P_{m,n_2,N-n_1,M-c_1} \Bigg). \quad (11)$$

The sum of the first two terms is 0.12 and the sum of the terms in parentheses is of the order of 0.02.

Average Quality Protection

General Relations

When the fraction defective in submitted product is p, the average quality after inspection (p_A) is given by

$$p_A = p \frac{N - I}{N} \quad (12)$$

when all defective pieces found are replaced. If defective pieces found are removed but not replaced,

$$p_A = p \frac{N - I}{N - pI}, \quad (12')$$

the factor pI representing the average number of defective pieces removed. In deriving the tables, equation (12) has been used. The error in p_A resulting from the use of (12) rather than (12′) is $\frac{pI}{N}$, which is generally small.

The average outgoing quality limit (p_L) is the maximum value of p_A that will result under any sampling plan, considering all possible values of p in the submitted product. The value of p for which this maximum value of p_A occurs is designated as p_1, hence

$$p_L = p_1 \frac{N - I}{N}. \quad (13)$$

The value of p_1 for which $p_A = p_L$ may be determined by differentiating equation (12) with respect to p, equating to 0, and solving for p, that is

$$\frac{dp_A}{dp} = \frac{N - I}{N} - \frac{p}{N}\frac{dI}{dp} = 0. \quad (14)$$

Single Sampling

Given: Lot size (N), AOQL (p_L), process average fraction defective ($\bar{p}$).
To find: Values of n and c that will minimize $\bar{I}$.

The average quality after inspection (p_A), after substituting in equation (12) the value of I given in equation (8′), is obtained from the relation

$$p_A = p\frac{(N-n)}{N}\sum_{m=0}^{m=c}\frac{e^{-pn}(pn)^m}{m!}. \tag{15}$$

Differentiating with respect to p in accordance with equation (14) gives,

$$\frac{dp_A}{dp} = \frac{(N-n)}{N}\left[\sum_{m=0}^{m=c}\frac{e^{-pn}(pn)^m}{m!} - \frac{e^{-pn}(pn)^{c+1}}{c!}\right]. \tag{16}$$

Equating to zero and solving for p, gives the value of $p = p_1$ that makes p_A a maximum; i.e., $p_A = p_L$.

Let $p_1 n = x$; the particular case covered by equation (15) where $p = p_1$, and $p_A = p_L$ may then be expressed as

$$p_L = \frac{N-n}{Nn}\,x\sum_{m=0}^{m=c}\frac{e^{-x}x^m}{m!}, \tag{17}$$

or

$$p_L = y\left(\frac{1}{n} - \frac{1}{N}\right), \tag{18}$$

where

$$y = x\sum_{m=0}^{m=c}\frac{e^{-x}x^m}{m!}. \tag{19}$$

Similarly, equation (16) equated to zero becomes, after substituting $p_1 n = x$ and simplifying,

$$\sum_{m=0}^{m=c}\frac{e^{-x}x^m}{m!} - \frac{e^{-x}x^{c+1}}{c!} = 0. \tag{20}$$

Substituting in equation (19) the second term of equation (20) for the summation term gives

$$y = \frac{e^{-x}x^{c+2}}{c!}. \tag{21}$$

These relations* provide a basis for determining the values of x and y, corresponding to specific values of c, listed in Table A. The values of x for $c = 0$ to 30 were determined from equation (20) using Newton's Method of Approximation. The values of x for $c = 31$ to 40 were estimated on the basis of successive differences. The listed values of y are averages of the two values determined from equations (19) and (21), which differ slightly because values of x were determined to only two decimal places.

* Reduction of the mathematical relations to this simplified form and the determination of several x and y values, were contributed by Dr. Walter Bartky of the University of Chicago (when he was associated with the Western Electric Co.) shortly after the development of the *AOQL* concept and the preparation of preliminary *AOQL* double sampling tables. The methods and work of computing the values in Table A were contributed by Mr. George C. Campbell, formerly of the Bell Telephone Laboratories.

TABLE A

VALUES OF x AND y FOR GIVEN VALUES OF c

Used in equation (18) for determining p_L when N, n and c are given, or in equation (22) for determining n when N, c and p_L are given

c =	0	1	2	3	4	5	6	7	8	9	10
x =	1.00	1.62	2.27	2.95	3.64	4.35	5.07	5.80	6.55	7.30	8.06
y =	0.3679	0.8408	1.372	1.946	2.544	3.172	3.810	4.465	5.150	5.836	6.535
c =	11	12	13	14	15	16	17	18	19	20	
x =	8.82	9.59	10.37	11.15	11.93	12.72	13.52	14.32	15.12	15.92	
y =	7.234	7.948	8.677	9.404	10.12	10.87	11.63	12.38	13.14	13.88	
c =	21	22	23	24	25	26	27	28	29	30	
x =	16.73	17.54	18.35	19.17	19.98	20.81	21.63	22.46	23.29	24.13	
y =	14.66	15.42	16.18	16.97	17.73	18.54	19.30	20.11	20.91	21.75	
c =	31	32	33	34	35	36	37	38	39	40	
x =	24.96	25.81	26.65	27.50	28.35	29.21	30.06	30.93	31.79	32.66	
y =	22.54	23.40	24.22	25.08	25.94	26.83	27.68	28.62	29.50	30.44	

The value of c that minimizes $\bar{I}$ (equation (8′), using $p = \bar{p}$), is given directly by Fig. 9, which uses coordinates of $\bar{M} = \bar{p}N$ and $\bar{k} = \bar{p}/p_L$. The curves bounding the c zones on Fig. 9 were obtained directly from relations between equations (18) and (8′), using $p = \bar{p}$, that define values of $\bar{M}$ and $\bar{k}$ such that $\bar{I}$ is the same for c and $c + 1$.

The value of n, corresponding to the value of c given on Fig. 9, is determined from equation (18), expressed as

$$n = \frac{yN}{p_L N + y}. \tag{22}$$

Example: Given: $N = 750$, $p_L = 0.01$, $\bar{p} = 0.004$.
To Find: n and c.

Solution: $\bar{M} = \bar{p}N = (0.004)\ (750) = 3$; $\bar{k} = \dfrac{\bar{p}}{p_L} = \dfrac{0.004}{0.01} = 0.4.$

Consulting Fig. 9, for $\bar{M} = 3$ and $\bar{k} = 0.4$, read $c = 1$.
From Table A, for $c = 1$, read $y = 0.8408$.
From equation (22), $n = \dfrac{(0.8408)\ (750)}{(0.01)\ (750) + 0.8408} = 75.6.$
Sampling Plan: $n = 76$, $c = 1$.

Double Sampling

Given: Lot Size (N), *AOQL* (p_L), process average fraction defective ($\bar{p}$).
To find: Values of n_1, n_2, c_1 and c_2 that will minimize $\bar{I}$.

The average quality after inspection (p_A) is found by substituting in equation (12), the value of I given in equation (9).

$$p_A = \frac{p}{N}\left[(N - n_1)\sum_{m=0}^{m=c_1} P_{m,pn_1} + (N - n_1 - n_2)\left(P_{c_1+1,pn_1}\sum_{m=0}^{m=c_2-c_1-1} P_{m,pn_2} + \cdots + P_{c_2,pn_1}P_{0,pn_2}\right)\right]. \tag{23}$$

Differentiating equation (23) with respect to p and equating to 0, in accordance with equation (14), and solving for p, gives the value of $p = p_1$ that makes p_A a maximum; i.e., $p_A = p_L$. The resulting equation is not reproduced here since

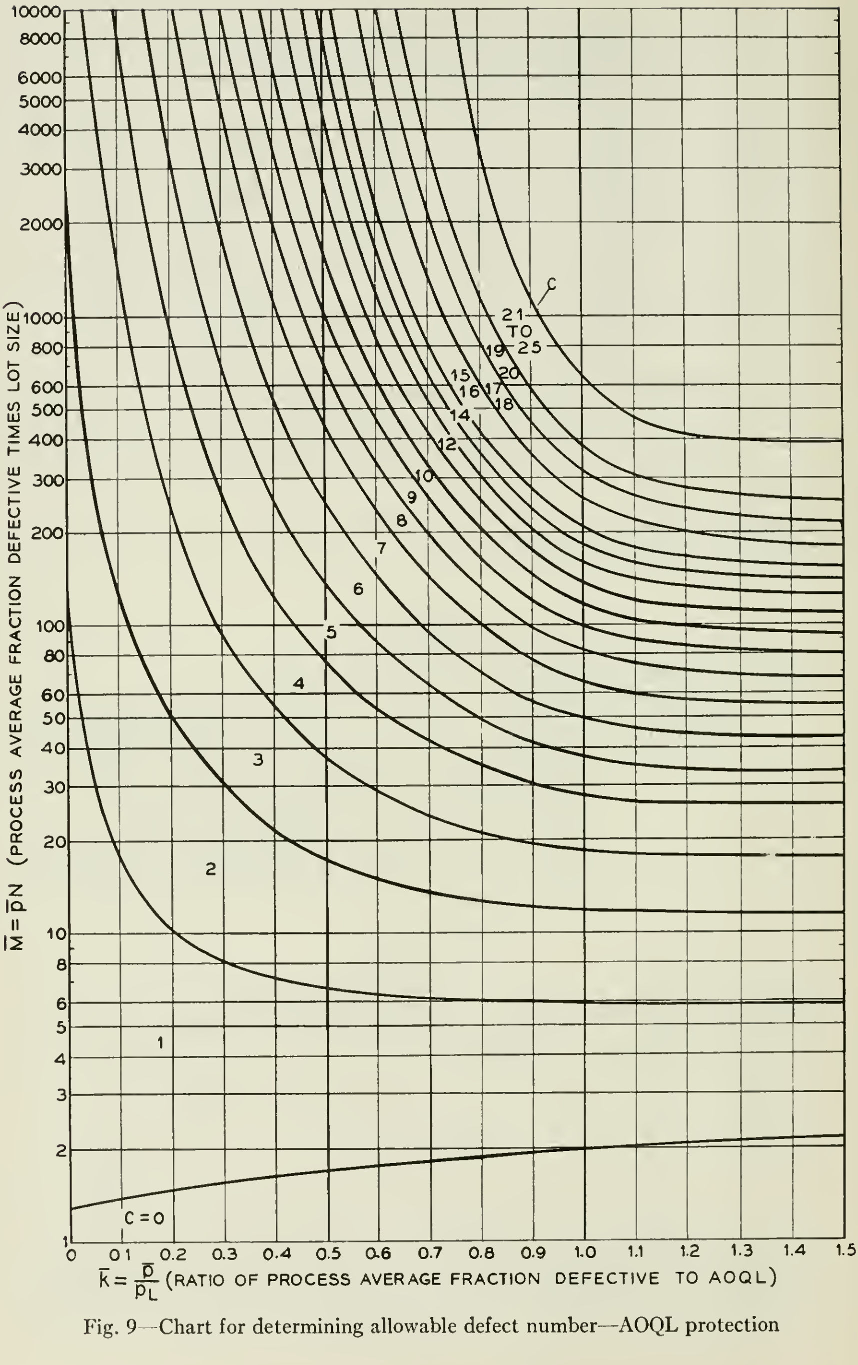


Fig. 9—Chart for determining allowable defect number—AOQL protection

it can be readily solved only for small values of c_1 and c_2. It is usually easier, particularly for the larger values of c_1 and c_2, to determine the maximum value of p_A (i.e., p_L) by trial and error, using work charts for estimating the region in which p_1 will be found.

The procedure used in preparing the tables and in finding the solution for a specific set of conditions is probably best illustrated by working out an actual example. In this procedure, use is made of known relationships between p_t and p_L values as given by the DL tables, where an initial risk of 0.06 and a Consumer's Risk of 0.10 are associated with p_t as outlined on page 11. For a given lot size, a work chart is prepared on which points corresponding to associated p_L and p_t values are plotted for each pair of c_1, c_2 values given in Fig. 7. A line drawn through all points for a single pair, such as $c_1 = 0$, $c_2 = 1$, indicates what p_t value should be associated with any p_L value specified. Fig. 10 indicates the nature of the work charts and the following example illustrates its use.

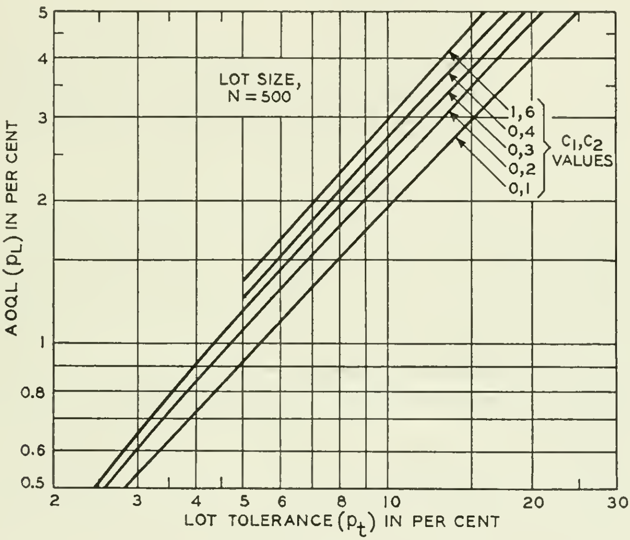


Fig. 10—Work chart giving p_t values corresponding to p_L values for given pairs of c_1, c_2 values—lot size, $N = 500$

Example: Given: $N = 500$, $p_L = .01$, $\bar{p} = .004$.

To find: n_1, n_2, c_1 and c_2 that will minimize average amount of inspection per lot. (Condition: For the associated lot tolerance value, p_t, the initial risk is 0.06 and the Consumer's Risk $P_C = 0.10$).

Solution: *Step 1*—Consult work chart, Fig. 10 for $N = 500$. Try $c_1 = 0$, $c_2 = 1$, and corresponding to $p_L = .01$, read $p_t = .054$.

Step 2—To determine if first choice of c_1, c_2 was the best.

$$M = p_t N = 0.054\ (500) = 27;\ k = \frac{\bar{p}}{p_t} = \frac{0.004}{0.054} = 0.074.$$

Consult Fig. 7, giving best c_1, c_2 values for given M and

k values. Corresponding to $M = 27$, $k = 0.074$, read $c_1 = 0$, $c_2 = 2$. Hence the first choice was not the best.

Step 3—Similar to Step 1. Consult work chart, Fig. 10. For $c_1 = 0$, $c_2 = 2$, corresponding to $p_L = 0.01$, read $p_t = .047$.

Step 4—Similar to Step 2. $M = p_t N = .047\ (500) = 23.5$; $k = \dfrac{\bar{p}}{p_t} = 0.085$. Consult Fig. 7 and corresponding to $M = 23.5$, $k = 0.085$, read $c_1 = 0, c_2 = 2$. This agrees with the choice in Step 3 and gives desired solution.

Step 5—To determine n_1 and n_2 for $c_1 = 0$, $c_2 = 2$. On Fig. 8, corresponding to $M = 23.5$, for $c_1 = 0$, read $p_t n_1 = 2.67$ and for $c_2 = 2$, read $p_t\ (n_1 + n_2) = 5.60$. Since per Step 3, $p_t = .047, n_1 = 57, n_1 + n_2 = 119$ and $n_2 = 62$.

Sampling Plan. $n_1 = 57, n_2 = 62, c_1 = 0, c_2 = 2$. (Rounding these values of n to the nearest 5 in accordance with the practice used in preparing the tables, gives $n_1 = 55$, $n_1 + n_2 = 120$, $n_2 = 65$, the values shown in Table DA-1 for $N = 401$–500, $\bar{p} = 0.21$–0.40%.)

Nature and Magnitude of Errors

Each sampling plan (combination of n and c values for single sampling, and of n_1, n_2, c_1 and c_2 values for double sampling) in the tables constitutes a solution for a range of process average values and a range of lot sizes. The following paragraphs give information regarding the magnitude of errors, associated with these solutions, that may be present because of the following two factors:

(1) Approximate equations and curves derived therefrom were used in place of exact equations over most areas of the tables, in order to minimize computative effort.
(2) The sample sizes, n_1 and $n_1 + n_2$, listed in the tables represent computed values rounded to the nearest unit for $n = 50$ or less, rounded to the nearest 5 for $50 < n < 1000$, and rounded to the nearest 10 for $n > 1000$.

Effect of Approximations—The percentage error in the Consumer's Risk value of 0.10, corresponding to lot tolerance values listed in the tables, attributable to the use of approximate equations and curves derived therefrom, is on the average about 3% and should not exceed 7%. The percentage error in the *AOQL* values, listed in the tables, attributable to the use of approximate relations involving the Poisson exponential rather than the binomial distribution, is on the average about 4% and should not exceed 12%. In a large number of exploratory checks for both single and double sampling, it was found in every instance that the Consumer's Risk and the *AOQL* values derived from approximate equations were larger than the corresponding exact values. The largest error observed in the Consumer's Risk for single sampling occurred when, instead of 0.10, the exact relation gave a value of 0.0937. Similarly the largest error in the *AOQL* occurred in single sampling when, instead of 0.0883, the exact relation gave a value of 0.0786. The observed errors in double sampling were of the same order of magnitude.

Effect of Rounding—The use of rounded values of n, n_1 and n_2 gives values of Consumer's Risk other than exactly 0.10. However each sampling plan lists sample sizes based on the largest lot size in the corresponding lot size range. As a result, the Consumer's Risk associated with the p_t value designated at the top of the Lot Tolerance tables does not exceed 0.10 except in a few isolated cases, where the risk may be as high as 0.12 for the largest lot size. Likewise, the *AOQL* value for any sampling plan in the *AOQL* tables does not exceed the value designated at the top of each table except in a few isolated cases, where the error due to rounding may be as much as 10% of the designated value for the largest lot size.

The Consumer's Risk value of 0.10 and the *AOQL* values listed in the tables, are therefore with few exceptions, upper bounds that will not be exceeded in the application of the tables.

NOMENCLATURE

N Number of pieces in lot.

n Number of pieces in sample.

n_1 Number of pieces in first sample.

n_2 Number of pieces in second sample.

c Allowable defect number.

c_1 Allowable defect number for first sample, n_1.

c_2 Allowable defect number for first and second samples combined, $n_1 + n_2$.

p_t Lot tolerance fraction defective.

p Fraction defective; also used specifically to denote fraction defective in submitted product.

$\bar{p}$ Process average (expected) fraction defective in submitted product.

p_A Average fraction defective in product after inspection—Average Outgoing Quality (*AOQ*).

p_L Maximum value of average fraction defective in product after inspection—Average Outgoing Quality Limit (*AOQL*).

p_1 Specific value of p in submitted product, for which $p_A = p_L$.

P_C Consumer's Risk.

P_a Probability of acceptance.

P_P Producer's Risk.

I Average number of pieces inspected per lot for submitted product of p quality.

$\bar{I}$ Specific value of I when p in submitted product $= \bar{p}$.

$\bar{I}_{min}$ Minimum value of $\bar{I}$.

$M = p_t N$ Number of defects in lot of tolerance (p_t) quality.

$\bar{M} = \bar{p}N$ Number of defects in a lot of process average ($\bar{p}$) quality.

$k = \dfrac{\bar{p}}{p_t}$ Ratio of process average fraction defective to lot tolerance fraction defective.

$\bar{k} = \dfrac{\bar{p}}{p_L}$ Ratio of process average fraction defective to *AOQL*.

m Number of defects found in sample.

$C_n^N = \dfrac{N!}{(N - n)!\, n!}$ Number of combinations of N things taken n at a time.

TABLE I: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE SL-0.5
Lot Tolerance Per Cent Defective = 0.5%

Process Average %	0-.005			.006-.050			.051-.100			.101-.150			.151-.200			.201-.250		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1-180	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
181-210	180	0	.02	180	0	.02	180	0	.02	180	0	.02	180	0	.02	180	0	.02
211-250	210	0	.03	210	0	.03	210	0	.03	210	0	.03	210	0	.03	210	0	.03
251-300	240	0	.03	240	0	.03	240	0	.03	240	0	.03	240	0	.03	240	0	.03
301-400	275	0	.04	275	0	.04	275	0	.04	275	0	.04	275	0	.04	275	0	.04
401-500	300	0	.05	300	0	.05	300	0	.05	300	0	.05	300	0	.05	300	0	.05
501-600	320	0	.05	320	0	.05	320	0	.05	320	0	.05	320	0	.05	320	0	.05
601-800	350	0	.06	350	0	.06	350	0	.06	350	0	.06	350	0	.06	350	0	.06
801-1000	365	0	.06	365	0	.06	365	0	.06	365	0	.06	365	0	.06	365	0	.06
1001-2000	410	0	.07	410	0	.07	410	0	.07	670	1	.08	670	1	.08	670	1	.08
2001-3000	430	0	.07	430	0	.07	705	1	.09	705	1	.09	955	2	.10	955	2	.10
3001-4000	440	0	.07	440	0	.07	730	1	.09	985	2	.10	1230	3	.11	1230	3	.11
4001-5000	445	0	.08	740	1	.10	1000	2	.11	1000	2	.11	1250	3	.12	1480	4	.12
5001-7000	450	0	.08	750	1	.10	1020	2	.12	1280	3	.12	1510	4	.13	1760	5	.14
7001-10,000	455	0	.08	760	1	.10	1040	2	.12	1530	4	.14	1790	5	.14	2240	7	.16
10,001-20,000	460	0	.08	775	1	.10	1330	3	.14	1820	5	.16	2300	7	.17	2780	9	.18
20,001-50,000	775	1	.11	1050	2	.13	1600	4	.15	2080	6	.18	3060	10	.20	4200	15	.22
50,001-100,000	780	1	.11	1060	2	.13	1840	5	.17	2590	8	.19	3780	13	.22	5140	19	.24

TABLE SL-1
Lot Tolerance Per Cent Defective = 1.0%

Process Average %	0-.010			.011-.10			.11-.20			.21-.30			.31-.40			.41-.50		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1-120	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
121-150	120	0	.06	120	0	.06	120	0	.06	120	0	.06	120	0	.06	120	0	.06
151-200	140	0	.08	140	0	.08	140	0	.08	140	0	.08	140	0	.08	140	0	.08
201-300	165	0	.10	165	0	.10	165	0	.10	165	0	.10	165	0	.10	165	0	.10
301-400	175	0	.12	175	0	.12	175	0	.12	175	0	.12	175	0	.12	175	0	.12
401-500	180	0	.13	180	0	.13	180	0	.13	180	0	.13	180	0	.13	180	0	.13
501-600	190	0	.13	190	0	.13	190	0	.13	190	0	.13	190	0	13	305	1	.14
601-800	200	0	.14	200	0	.14	200	0	.14	330	1	.15	330	1	.15	330	1	.15
801-1000	205	0	.14	205	0	.14	205	0	.14	335	1	.17	335	1	.17	335	1	.17
1001-2000	220	0	.15	220	0	.15	360	1	.19	490	2	.21	490	2	.21	610	3	.22
2001-3000	220	0	.15	375	1	.20	505	2	.23	630	3	.24	745	4	.26	870	5	.26
3001-4000	225	0	.15	380	1	.20	510	2	.24	645	3	.25	880	5	.28	1000	6	.29
4001-5000	225	0	.16	380	1	.20	520	2	.24	770	4	.28	895	5	.29	1120	7	.31
5001-7000	230	0	.15	385	1	.21	655	3	.27	780	4	.29	1020	6	.32	1260	8	.34
7001-10,000	230	0	.16	520	2	.25	660	3	.28	910	5	.32	1150	7	.34	1500	10	.37
10,001-20,000	390	1	.21	525	2	.26	785	4	.31	1040	6	.35	1400	9	.39	1980	14	.43
20,001-50,000	390	1	.21	530	2	.26	920	5	.34	1300	8	.39	1890	13	.44	2570	19	.48
50,001-100,000	390	1	.21	670	3	.29	1040	6	.36	1420	9	.41	2120	15	.47	3150	23	.50

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
AOQL = Average Outgoing Quality Limit.

TABLE I CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE SL-2
Lot Tolerance Per Cent Defective = 2.0%

Process Average %	0–.02			.03–.20			.21–.40			.41–.60			.61–.80			.81–1.00		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1–75	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
76–100	70	0	.16	70	0	.16	70	0	.16	70	0	.16	70	0	.16	70	0	.16
101–200	85	0	.25	85	0	.25	85	0	.25	85	0	.25	85	0	.25	85	0	.25
201–300	95	0	.26	95	0	.26	95	0	.26	95	0	.26	95	0	.26	95	0	.26
301–400	100	0	.28	100	0	.28	100	0	.28	160	1	.32	160	1	.32	160	1	.32
401–500	105	0	.28	105	0	.28	105	0	.28	165	1	.34	165	1	.34	165	1	.34
501–600	105	0	.29	105	0	.29	175	1	.34	175	1	.34	175	1	.34	235	2	.36
601–800	110	0	.29	110	0	.29	180	1	.36	240	2	.40	240	2	.40	300	3	.41
801–1000	115	0	.28	115	0	.28	185	1	.37	245	2	.42	305	3	.44	305	3	.44
1001–2000	115	0	.30	190	1	.40	255	2	.47	325	3	.50	380	4	.54	440	5	.56
2001–3000	115	0	.31	190	1	.41	260	2	.48	385	4	.58	450	5	.60	565	7	.64
3001–4000	115	0	.31	195	1	.41	330	3	.54	450	5	.63	510	6	.65	690	9	.70
4001–5000	195	1	.41	260	2	.50	335	3	.54	455	5	.63	575	7	.69	750	10	.74
5001–7000	195	1	.42	265	2	.50	335	3	.55	515	6	.69	640	8	.73	870	12	.80
7001–10,000	195	1	.42	265	2	.50	395	4	.62	520	6	.69	760	10	.79	1050	15	.86
10,001–20,000	200	1	.42	265	2	.51	460	5	.67	650	8	.77	885	12	.86	1230	18	.94
20,001–50,000	200	1	.42	335	3	.58	520	6	.73	710	9	.81	1060	15	.93	1520	23	1.0
50,001–100,000	200	1	.42	335	3	.58	585	7	.76	770	10	.84	1180	17	.97	1690	26	1.1

TABLE SL-3
Lot Tolerance Per Cent Defective = 3.0%

Process Average %	0–.03			.04–.30			.31–.60			.61–.90			.91–1.20			1.21–1.50		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1–40	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
41–55	40	0	.18	40	0	.18	40	0	.18	40	0	.18	40	0	.18	40	0	.18
56–100	55	0	.30	55	0	.30	55	0	.30	55	0	.30	55	0	.30	55	0	.30
101–200	65	0	.38	65	0	.38	65	0	.38	65	0	.38	65	0	.38	65	0	.38
201–300	70	0	.40	70	0	.40	70	0	.40	110	1	.48	110	1	.48	110	1	.48
301–400	70	0	.43	70	0	.43	115	1	.52	115	1	.52	115	1	.52	155	2	.54
401–500	70	0	.45	70	0	.45	120	1	.53	120	1	.53	160	2	.58	160	2	.58
501–600	75	0	.43	75	0	.43	120	1	.56	160	2	.63	160	2	.63	200	3	.65
601–800	75	0	.44	125	1	.57	125	1	.57	165	2	.66	205	3	.71	240	4	.74
801–1000	75	0	.45	125	1	.59	170	2	.67	210	3	.73	250	4	.76	290	5	.78
1001–2000	75	0	.47	130	1	.60	175	2	.72	260	4	.85	300	5	.90	380	7	.95
2001–3000	75	0	.48	130	1	.62	220	3	.82	300	5	.95	385	7	1.0	460	9	1.1
3001–4000	130	1	.63	175	2	.75	220	3	.84	305	5	.96	425	8	1.1	540	11	1.2
4001–5000	130	1	.63	175	2	.76	260	4	.91	345	6	1.0	465	9	1.1	620	13	1.2
5001–7000	130	1	.63	175	2	.76	265	4	.92	390	7	1.1	505	10	1.2	700	15	1.3
7001–10,000	130	1	.64	175	2	.77	265	4	.93	390	7	1.1	550	11	1.2	775	17	1.4
10,001–20,000	130	1	.64	175	2	.78	305	5	1.0	430	8	1.2	630	13	1.3	900	20	1.5
20,001–50,000	130	1	.65	225	3	.86	350	6	1.1	520	10	1.2	750	16	1.4	1090	25	1.6
50,001–100,000	130	1	.65	265	4	.96	390	7	1.1	590	12	1.3	830	18	1.5	1215	28	1.6

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
AOQL = Average Outgoing Quality Limit.

TABLE I CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE SL-4

Lot Tolerance Per Cent Defective = 4.0%

Process Average %	0–.04			.05–.40			.41–.80			.81–1.20			1.21–1.60			1.61–2.00		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1–35	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
36–50	34	0	.35	34	0	.35	34	0	.35	34	0	.35	34	0	.35	34	0	.35
51–100	44	0	.47	44	0	.47	44	0	.47	44	0	.47	44	0	.47	44	0	.47
101–200	50	0	.55	50	0	.55	50	0	.55	50	0	.55	50	0	.55	50	0	.55
201–300	55	0	.57	55	0	.57	85	1	.71	85	1	.71	85	1	.71	85	1	.71
301–400	55	0	.58	55	0	.58	90	1	.72	120	2	.80	120	2	.80	145	3	.86
401–500	55	0	.60	55	0	.60	90	1	.77	120	2	.87	150	3	.91	150	3	.91
501–600	55	0	.61	95	1	.76	125	2	.87	125	2	.87	155	3	.93	185	4	.95
601–800	55	0	.62	95	1	.78	125	2	.93	160	3	.97	190	4	1.0	220	5	1.0
801–1000	55	0	.63	95	1	.80	130	2	.92	165	3	.98	220	5	1.1	255	6	1.1
1001–2000	55	0	.65	95	1	.84	165	3	1.1	195	4	1.2	255	6	1.3	315	8	1.4
2001–3000	95	1	.86	130	2	1.0	165	3	1.1	230	5	1.3	320	8	1.4	405	11	1.6
3001–4000	95	1	.86	130	2	1.0	195	4	1.2	260	6	1.4	350	9	1.5	465	13	1.6
4001–5000	95	1	.87	130	2	1.0	195	4	1.2	290	7	1.4	380	10	1.6	520	15	1.7
5001–7000	95	1	.87	130	2	1.0	200	4	1.2	290	7	1.5	410	11	1.7	575	17	1.9
7001–10,000	95	1	.88	130	2	1.1	230	5	1.4	325	8	1.5	440	12	1.7	645	19	1.9
10,001–20,000	95	1	.88	165	3	1.2	265	6	1.4	355	9	1.6	500	14	1.8	730	22	2.0
20,001–50,000	95	1	.88	165	3	1.2	295	7	1.5	380	10	1.7	590	17	2.0	870	26	2.1
50,001–100,000	95	1	.88	200	4	1.3	325	8	1.6	410	11	1.8	620	18	2.0	925	29	2.2

TABLE SL-5

Lot Tolerance Per Cent Defective = 5.0%

Process Average %	0–.05			.06–.50			.51–1.00			1.01–1.50			1.51–2.00			2.01–2.50		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1–30	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
31–50	30	0	.49	30	0	.49	30	0	.49	30	0	.49	30	0	.49	30	0	.49
51–100	37	0	.63	37	0	.63	37	0	.63	37	0	.63	37	0	.63	37	0	.63
101–200	40	0	.74	40	0	.74	40	0	.74	40	0	.74	40	0	.74	40	0	.74
201–300	43	0	.74	43	0	.74	70	1	.92	70	1	.92	95	2	.99	95	2	.99
301–400	44	0	.74	44	0	.74	70	1	.99	100	2	1.0	120	3	1.1	145	4	1.1
401–500	45	0	.75	75	1	.95	100	2	1.1	100	2	1.1	125	3	1.2	150	4	1.2
501–600	45	0	.76	75	1	.98	100	2	1.1	125	3	1.2	150	4	1.3	175	5	1.3
601–800	45	0	.77	75	1	1.0	100	2	1.2	130	3	1.2	175	5	1.4	200	6	1.4
801–1000	45	0	.78	75	1	1.0	105	2	1.2	155	4	1.4	180	5	1.4	225	7	1.5
1001–2000	45	0	.80	75	1	1.0	130	3	1.4	180	5	1.6	230	7	1.7	280	9	1.8
2001–3000	75	1	1.1	105	2	1.3	135	3	1.4	210	6	1.7	280	9	1.9	370	13	2.1
3001–4000	75	1	1.1	105	2	1.3	160	4	1.5	210	6	1.7	305	10	2.0	420	15	2.2
4001–5000	75	1	1.1	105	2	1.3	160	4	1.5	235	7	1.8	330	11	2.0	440	16	2.2
5001–7000	75	1	1.1	105	2	1.3	185	5	1.7	260	8	1.9	350	12	2.2	490	18	2.4
7001–10,000	75	1	1.1	105	2	1.3	185	5	1.7	260	8	1.9	380	13	2.2	535	20	2.5
10,001–20,000	75	1	1.1	135	3	1.4	210	6	1.8	285	9	2.0	425	15	2.3	610	23	2.6
20,001–50,000	75	1	1.1	135	3	1.4	235	7	1.9	305	10	2.1	470	17	2.4	700	27	2.7
50,001–100,000	75	1	1.1	160	4	1.6	235	7	1.9	355	12	2.2	515	19	2.5	770	30	2.8

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
AOQL = Average Outgoing Quality Limit.

TABLE I CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE SL-7

Lot Tolerance Per Cent Defective = 7.0%

Process Average %	0–.07			.08–.70			.71–1.40			1.41–2.10			2.11–2.80			2.81–3.50		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1–25	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
26–50	24	0	.80	24	0	.80	24	0	.80	24	0	.80	24	0	.80	24	0	.80
51–100	28	0	.95	28	0	.95	28	0	.95	28	0	.95	28	0	.95	28	0	.95
101–200	30	0	1.0	30	0	1.0	49	1	1.3	49	1	1.3	49	1	1.3	65	2	1.4
201–300	31	0	1.1	31	0	1.1	50	1	1.4	70	2	1.5	85	3	1.6	85	3	1.6
301–400	32	0	1.1	55	1	1.4	70	2	1.6	90	3	1.7	105	4	1.8	125	5	1.8
401–500	32	0	1.1	55	1	1.4	75	2	1.6	90	3	1.8	110	4	1.9	140	6	2.0
501–600	32	0	1.1	55	1	1.4	75	2	1.7	95	3	1.8	125	5	2.0	145	6	2.1
601–800	32	0	1.1	55	1	1.4	75	2	1.7	110	4	2.0	130	5	2.1	160	7	2.2
801–1000	33	0	1.1	55	1	1.4	95	3	1.9	110	4	2.1	145	6	2.2	180	8	2.4
1001–2000	55	1	1.5	75	2	1.8	95	3	2.0	130	5	2.3	185	8	2.5	230	11	2.8
2001–3000	55	1	1.5	75	2	1.8	115	4	2.1	150	6	2.4	215	10	2.8	300	15	3.0
3001–4000	55	1	1.5	75	2	1.8	115	4	2.2	165	7	2.6	235	11	2.9	330	17	3.2
4001–5000	55	1	1.5	75	2	1.8	130	5	2.4	185	8	2.7	250	12	3.0	350	18	3.3
5001–7000	55	1	1.5	75	2	1.8	130	5	2.4	185	8	2.7	270	13	3.1	385	20	3.4
7001–10,000	55	1	1.5	95	3	2.0	150	6	2.5	200	9	2.9	285	14	3.2	415	22	3.6
10,001–20,000	55	1	1.5	95	3	2.0	150	6	2.5	220	10	2.9	320	16	3.3	470	25	3.7
20,001–50,000	55	1	1.5	115	4	2.2	170	7	2.6	235	11	3.1	355	18	3.5	530	29	3.9
50,001–100,000	55	1	1.5	115	4	2.2	185	8	2.7	270	13	3.1	370	19	3.5	530	29	3.9

TABLE SL-10

Lot Tolerance Per Cent Defective = 10.0%

Process Average %	0–.10			.11–1.00			1.01–2.00			2.01–3.00			3.01–4.00			4.01–5.00		
Lot Size	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %	n	c	AOQL %
1–20	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0	All	0	0
21–50	17	0	1.3	17	0	1.3	17	0	1.3	17	0	1.3	17	0	1.3	17	0	1.3
51–100	20	0	1.5	20	0	1.5	20	0	1.5	33	1	1.7	33	1	1.7	33	1	1.7
101–200	22	0	1.5	22	0	1.5	35	1	2.0	48	2	2.2	48	2	2.2	60	3	2.4
201–300	23	0	1.5	38	1	1.9	50	2	2.3	65	3	2.4	75	4	2.6	85	5	2.7
301–400	23	0	1.5	38	1	2.0	50	2	2.4	65	3	2.5	90	5	2.7	100	6	2.9
401–500	23	0	1.5	38	1	2.0	50	2	2.5	75	4	2.8	90	5	2.9	110	7	3.2
501–600	23	0	1.5	38	1	2.1	65	3	2.7	80	4	3.0	100	6	3.2	125	8	3.3
601–800	23	0	1.6	38	1	2.1	65	3	2.8	90	5	3.1	100	6	3.3	140	9	3.4
801–1000	39	1	2.1	50	2	2.6	65	3	2.8	90	5	3.2	115	7	3.4	150	10	3.7
1001–2000	39	1	2.1	50	2	2.6	80	4	3.1	105	6	3.4	140	9	3.9	195	14	4.4
2001–3000	39	1	2.1	50	2	2.6	80	4	3.1	115	7	3.7	165	11	4.1	230	17	4.7
3001–4000	39	1	2.1	50	2	2.6	90	5	3.4	130	8	3.8	190	13	4.4	255	19	4.8
4001–5000	39	1	2.1	50	2	2.6	90	5	3.5	130	8	3.9	200	14	4.5	270	20	4.9
5001–7000	39	1	2.1	65	3	3.0	105	6	3.6	140	9	4.1	200	14	4.6	295	22	5.0
7001–10,000	39	1	2.2	65	3	3.0	105	6	3.6	150	10	4.2	210	15	4.7	315	24	5.2
10,001–20,000	39	1	2.2	65	3	3.0	120	7	3.7	150	10	4.3	240	17	4.8	340	26	5.4
20,001–50,000	39	1	2.2	80	4	3.2	120	7	3.7	165	11	4.4	260	19	5.0	380	30	5.7
50,001–100,000	39	1	2.2	95	5	3.3	130	8	4.0	180	12	4.4	270	20	5.1	380	30	5.7

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
AOQL = Average Outgoing Quality Limit.

TABLE II: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE DL-0.5

Lot Tolerance Per Cent Defective = 0.5%

Process Average %	0–.005						.006–.050						.051–.100					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–180	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
181–210	180	0	–	–	–	.02	180	0	–	–	–	.02	180	0	–	–	–	.02
211–250	210	0	–	–	–	.03	210	0	–	–	–	.03	210	0	–	–	–	.03
251–300	240	0	–	–	–	.03	240	0	–	–	–	.03	240	0	–	–	–	.03
301–400	275	0	–	–	–	.04	275	0	–	–	–	.04	275	0	–	–	–	.04
401–450	290	0	–	–	–	.04	290	0	–	–	–	.04	290	0	–	–	–	.04
451–500	340	0	110	450	1	.04	340	0	110	450	1	.04	340	0	110	450	1	.04
501–550	350	0	130	480	1	.05	350	0	130	480	1	.05	350	0	130	480	1	.05
551–600	360	0	150	510	1	.05	360	0	150	510	1	.05	360	0	150	510	1	.05
601–800	400	0	185	585	1	.06	400	0	185	585	1	.06	400	0	185	585	1	.06
801–1000	430	0	200	630	1	.07	430	0	200	630	1	.07	430	0	200	630	1	.07
1001–2000	490	0	265	755	1	.08	490	0	265	755	1	.08	490	0	265	755	1	.08
2001–3000	520	0	290	810	1	.09	520	0	290	810	1	.09	520	0	530	1050	2	.10
3001–4000	530	0	310	840	1	.09	530	0	570	1100	2	.11	530	0	570	1100	2	.11
4001–5000	540	0	305	845	1	.09	540	0	580	1120	2	.11	540	0	830	1370	3	.12
5001–7000	545	0	315	860	1	.10	545	0	615	1160	2	.11	545	0	865	1410	3	.12
7001–10,000	550	0	330	880	1	.10	550	0	620	1170	2	.12	550	0	1130	1680	4	.14
10,001–20,000	555	0	345	900	1	.10	555	0	925	1480	3	.13	555	0	1185	1740	4	.15
20,001–50,000	560	0	650	1210	2	.12	560	0	940	1500	3	.14	900	1	1400	2300	6	.16
50,001–100,000	560	0	650	1210	2	.12	560	0	1210	1770	4	.15	905	1	1655	2560	7	.17

Process Average %	.101–.150						.151–.200						.201–.250					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–180	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
181–210	180	0	–	–	–	.02	180	0	–	–	–	.02	180	0	–	–	–	.02
211–250	210	0	–	–	–	.03	210	0	–	–	–	.03	210	0	–	–	–	.03
251–300	240	0	–	–	–	.03	240	0	–	–	–	.03	240	0	–	–	–	.03
301–400	275	0	–	–	–	.04	275	0	–	–	–	.04	275	0	–	–	–	.04
401–450	290	0	–	–	–	.04	290	0	–	–	–	.04	290	0	–	–	–	.04
451–500	340	0	110	450	1	.04	340	0	110	450	1	.04	340	0	110	450	1	.04
501–550	350	0	130	480	1	.05	350	0	130	480	1	.05	350	0	130	480	1	.05
551–600	360	0	150	510	1	.05	360	0	150	510	1	.05	360	0	150	510	1	.05
601–800	400	0	185	585	1	.06	400	0	185	585	1	.06	400	0	185	585	1	.06
801–1000	430	0	200	630	1	.07	430	0	200	630	1	.07	430	0	200	630	1	.07
1001–2000	490	0	500	990	2	.09	490	0	500	990	2	.09	490	0	500	990	2	.09
2001–3000	520	0	530	1050	2	.10	520	0	760	1280	3	.11	520	0	980	1500	4	.11
3001–4000	530	0	810	1340	3	.11	530	0	1030	1560	4	.12	840	1	1160	2000	6	.13
4001–5000	540	0	1060	1600	4	.13	845	1	1205	2050	6	.14	845	1	1425	2270	7	.14
5001–7000	545	0	1105	1650	4	.13	860	1	1490	2350	7	.15	860	1	1700	2560	8	.16
7001–10,000	880	1	1300	2180	6	.15	880	1	1770	2650	8	.16	1170	2	2160	3330	11	.17
10,001–20,000	900	1	1840	2740	8	.18	1200	2	2250	3450	11	.19	1740	4	2620	4360	15	.21
20,001–50,000	1210	2	2330	3540	11	.20	1500	3	2980	4480	15	.22	2300	6	4240	6540	24	.24
50,001–100,000	1210	2	2590	3800	12	.21	1770	4	3690	5460	19	.23	2560	7	5420	7980	30	.26

TABLE DL-1

LOT TOLERANCE PER CENT DEFECTIVE = 1.0%

Process Average %	0–.010						.011–.10						.11–.20					
Lot Size	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–120	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
121–150	120	0	–	–	–	.06	120	0	–	–	–	.06	120	0	–	–	–	.06
151–200	140	0	–	–	–	.08	140	0	–	–	–	.08	140	0	–	–	–	.08
201–260	165	0	–	–	–	.10	165	0	–	–	–	.10	165	0	–	–	–	.10
261–300	180	0	75	255	1	.10	180	0	75	255	1	.10	180	0	75	255	1	.10
301–400	200	0	90	290	1	.12	200	0	90	290	1	.12	200	0	90	290	1	.12
401–500	215	0	100	315	1	.14	215	0	100	315	1	.14	215	0	100	315	1	.14
501–600	225	0	115	340	1	.15	225	0	115	340	1	.15	225	0	115	340	1	.15
601–800	235	0	125	360	1	.16	235	0	125	360	1	.16	235	0	125	360	1	.16
801–1000	245	0	135	380	1	.17	245	0	135	380	1	.17	245	0	250	495	2	.19
1001–2000	265	0	155	420	1	.18	265	0	155	420	1	.18	265	0	285	550	2	.21
2001–3000	270	0	160	430	1	.19	270	0	300	570	2	.22	270	0	420	690	3	.25
3001–4000	275	0	160	435	1	.19	275	0	305	580	2	.22	275	0	435	710	3	.25
4001–5000	275	0	165	440	1	.19	275	0	310	585	2	.23	275	0	565	840	4	.28
5001–7000	275	0	170	445	1	.20	275	0	315	590	2	.23	275	0	580	855	4	.29
7001–10,000	280	0	320	600	2	.24	280	0	460	740	3	.26	280	0	590	870	4	.30
10,001–20,000	280	0	325	605	2	.24	280	0	465	745	3	.27	450	1	700	1150	6	.33
20,001–50,000	280	0	325	605	2	.25	280	0	605	885	4	.30	450	1	830	1280	7	.36
50,001–100,000	280	0	325	605	2	.25	280	0	605	885	4	.30	450	1	960	1410	8	.38

Process Average %	.21–.30						.31–.40						.41–.50					
Lot Size	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–120	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
121–150	120	0	–	–	–	.06	120	0	–	–	–	.06	120	0	–	–	–	.06
151–200	140	0	–	–	–	.08	140	0	–	–	–	.08	140	0	–	–	–	.08
201–260	165	0	–	–	–	.10	165	0	–	–	–	.10	165	0	–	–	–	.10
261–300	180	0	75	255	1	.10	180	0	75	255	1	.10	180	0	75	255	1	.10
301–400	200	0	90	290	1	.12	200	0	90	290	1	.12	200	0	90	290	1	.12
401–500	215	0	100	315	1	.14	215	0	100	315	1	.14	215	0	100	315	1	.14
501–600	225	0	115	340	1	.15	225	0	115	340	1	.15	225	0	205	430	2	.16
601–800	235	0	230	465	2	.18	235	0	230	465	2	.18	235	0	230	465	2	.18
801–1000	245	0	250	495	2	.19	245	0	250	495	2	.19	245	0	250	495	2	.19
1001–2000	265	0	405	670	3	.23	265	0	515	780	4	.24	265	0	515	780	4	.24
2001–3000	270	0	545	815	4	.26	430	1	620	1050	6	.28	430	1	830	1260	8	.30
3001–4000	435	1	645	1080	6	.29	435	1	865	1300	8	.30	580	2	940	1520	10	.33
4001–5000	440	1	660	1100	6	.30	440	1	1000	1440	9	.33	585	2	1075	1660	11	.35
5001–7000	445	1	785	1230	7	.33	590	2	990	1580	10	.36	730	3	1190	1920	13	.38
7001–10,000	450	1	920	1370	8	.35	600	2	1240	1840	12	.39	870	4	1540	2410	17	.41
10,001–20,000	605	2	1035	1640	10	.39	745	3	1485	2230	15	.43	1150	6	1990	3140	23	.44
20,001–50,000	605	2	1295	1900	12	.42	885	4	1845	2730	19	.47	1280	7	2600	3880	29	.52
50,001–100,000	605	2	1545	2150	14	.44	885	4	2085	2970	21	.49	1410	8	3280	4690	36	.55

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
AOQL = Average Outgoing Quality Limit.

TABLE II CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE DL-2

LOT TOLERANCE PER CENT DEFECTIVE = 2.0%

Process Average %	0–.02						.03–.20						.21–.40					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–75	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
76–100	70	0	–	–	–	.16	70	0	–	–	–	.16	70	0	–	–	–	.16
101–200	85	0	–	–	–	.25	85	0	–	–	–	.25	85	0	–	–	–	.25
201–300	115	0	50	165	1	.29	115	0	50	165	1	.29	115	0	50	165	1	.29
301–400	120	0	60	180	1	.32	120	0	60	180	1	.32	120	0	60	180	1	.32
401–500	125	0	65	190	1	.33	125	0	65	190	1	.33	125	0	120	245	2	.37
501–600	125	0	70	195	1	.34	125	0	70	195	1	.34	125	0	130	255	2	.39
601–800	130	0	75	205	1	.35	130	0	75	205	1	.35	130	0	135	265	2	.41
801–1000	135	0	75	210	1	.36	135	0	140	275	2	.42	135	0	140	275	2	.42
1001–2000	135	0	85	220	1	.38	135	0	155	290	2	.45	135	0	220	355	3	.50
2001–3000	140	0	85	225	1	.39	140	0	155	295	2	.46	140	0	285	425	4	.56
3001–4000	140	0	85	225	1	.40	140	0	225	365	3	.52	140	0	290	430	4	.57
4001–5000	140	0	160	300	2	.47	140	0	230	370	3	.53	140	0	360	500	5	.61
5001–7000	140	0	160	300	2	.48	140	0	230	370	3	.54	140	0	365	505	5	.62
7001–10,000	140	0	160	300	2	.48	140	0	235	375	3	.54	225	1	350	575	6	.66
10,001–20,000	140	0	165	305	2	.49	140	0	235	375	3	.54	225	1	415	640	7	.71
20,001–50,000	140	0	165	305	2	.49	140	0	305	445	4	.59	225	1	480	705	8	.75
50,001–100,000	140	0	165	305	2	.49	140	0	305	445	4	.60	225	1	545	770	9	.78

Process Average %	.41–.60						.61–.80						.81–1.00					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–75	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
76–100	70	0	–	–	–	.16	70	0	–	–	–	.16	70	0	–	–	–	.16
101–200	85	0	–	–	–	.25	85	0	–	–	–	.25	85	0	–	–	–	.25
201–300	115	0	50	165	1	.29	115	0	50	165	1	.29	115	0	50	165	1	.29
301–400	120	0	115	235	2	.34	120	0	115	235	2	.34	120	0	115	235	2	.34
401–500	125	0	120	245	2	.37	125	0	120	245	2	.37	125	0	120	245	2	.37
501–600	125	0	130	255	2	.39	125	0	185	310	3	.41	125	0	185	310	3	.41
601–800	130	0	195	325	3	.44	130	0	250	380	4	.45	130	0	250	380	4	.45
801–1000	135	0	200	335	3	.46	135	0	255	390	4	.48	210	1	290	500	6	.54
1001–2000	135	0	285	420	4	.54	220	1	375	595	7	.62	220	1	485	705	9	.65
2001–3000	225	1	385	610	7	.65	295	2	435	730	9	.69	360	3	535	895	12	.72
3001–4000	225	1	455	680	8	.69	295	2	555	850	11	.74	365	3	715	1080	15	.77
4001–5000	225	1	460	685	8	.70	300	2	620	920	12	.77	435	4	775	1210	17	.81
5001–7000	300	2	450	750	9	.74	370	3	680	1050	14	.82	505	5	935	1440	21	.89
7001–10,000	300	2	520	820	10	.77	375	3	735	1110	15	.85	575	6	1055	1630	24	.95
10,001–20,000	305	2	645	950	12	.83	375	3	935	1310	18	.92	640	7	1240	1880	28	1.0
20,001–50,000	305	2	715	1020	13	.86	445	4	1045	1490	21	.98	705	8	1635	2340	36	1.1
50,001–100,000	305	2	830	1135	15	.90	510	5	1150	1660	24	1.0	770	9	1850	2620	41	1.2

TABLE DL-3

LOT TOLERANCE PER CENT DEFECTIVE = 3.0%

Process Average %	0–.03						.04–.30					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–40	All	0	–	–	–	0	All	0	–	–	–	0
41–55	40	0	–	–	–	.18	40	0	–	–	–	.18
56–100	55	0	–	–	–	.30	55	0	–	–	–	.30
101–150	70	0	30	100	1	.37	70	0	30	100	1	.37
151–200	75	0	40	115	1	.45	75	0	40	115	1	.45
201–300	75	0	40	115	1	.50	75	0	40	115	1	.50
301–400	80	0	45	125	1	.52	80	0	45	125	1	.52
401–500	85	0	50	135	1	.53	85	0	50	135	1	.53
501–600	85	0	50	135	1	.54	85	0	50	135	1	.54
601–800	90	0	50	140	1	.55	90	0	95	185	2	.64
801–1000	90	0	55	145	1	.56	90	0	100	190	2	.66
1001–2000	90	0	60	150	1	.58	90	0	105	195	2	.70
2001–3000	90	0	60	150	1	.59	90	0	155	245	3	.80
3001–4000	95	0	105	200	2	.72	95	0	150	245	3	.80
4001–5000	95	0	105	200	2	.73	95	0	155	250	3	.81
5001–7000	95	0	105	200	2	.73	95	0	155	250	3	.81
7001–10,000	95	0	105	200	2	.73	95	0	155	250	3	.81
10,001–20,000	95	0	105	200	2	.74	95	0	200	295	4	.92
20,001–50,000	95	0	105	200	2	.74	95	0	200	295	4	.93
50,001–100,000	95	0	105	200	2	.75	95	0	245	340	5	1.0

Process Average %	.31–.60						.61–.90					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–40	All	0	–	–	–	0	All	0	–	–	–	0
41–55	40	0	–	–	–	.18	40	0	–	–	–	.18
56–100	55	0	–	–	–	.30	55	0	–	–	–	.30
101–150	70	0	30	100	1	.37	70	0	30	100	1	.37
151–200	75	0	40	115	1	.45	75	0	40	115	1	.45
201–300	75	0	40	115	1	.50	75	0	80	155	2	.54
301–400	80	0	85	165	2	.57	80	0	85	165	2	.57
401–500	85	0	90	175	2	.60	85	0	125	210	3	.64
501–600	85	0	95	180	2	.62	85	0	130	215	3	.67
601–800	90	0	135	225	3	.70	90	0	170	260	4	.74
801–1000	90	0	140	230	3	.72	90	0	180	270	4	.77
1001–2000	90	0	190	280	4	.84	150	1	210	360	6	.90
2001–3000	90	0	200	290	4	.86	150	1	300	450	8	1.0
3001–4000	95	0	235	330	5	.92	150	1	350	500	9	1.1
4001–5000	150	1	230	380	6	.98	200	2	340	540	10	1.2
5001–7000	150	1	230	380	6	1.0	200	2	385	585	11	1.2
7001–10,000	150	1	275	425	7	1.0	200	2	425	625	12	1.2
10,001–20,000	150	1	320	470	8	1.1	200	2	475	675	13	1.3
20,001–50,000	150	1	365	515	9	1.2	200	2	515	715	14	1.3
50,001–100,000	150	1	405	555	10	1.2	200	2	555	755	15	1.3

Process Average %	.91–1.20						1.21–1.50					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–40	All	0	–	–	–	0	All	0	–	–	–	0
41–55	40	0	–	–	–	.18	40	0	–	–	–	.18
56–100	55	0	–	–	–	.30	55	0	–	–	–	.30
101–150	70	0	30	100	1	.37	70	0	30	100	1	.37
151–200	75	0	65	140	2	.47	75	0	65	140	2	.47
201–300	75	0	80	155	2	.54	75	0	80	155	2	.54
301–400	80	0	120	200	3	.62	80	0	120	200	3	.62
401–500	85	0	125	210	3	.64	85	0	160	245	4	.69
501–600	85	0	170	255	4	.72	135	1	185	320	6	.76
601–800	140	1	195	335	6	.79	140	1	210	350	7	.81
801–1000	145	1	235	380	7	.85	145	1	270	415	8	.86
1001–2000	150	1	325	475	9	1.0	195	2	350	545	11	1.1
2001–3000	200	2	365	565	11	1.1	290	4	470	760	16	1.2
3001–4000	245	3	405	650	13	1.2	330	5	545	875	19	1.2
4001–5000	250	3	445	695	14	1.2	380	6	620	1000	22	1.3
5001–7000	250	3	530	780	16	1.3	380	6	700	1080	24	1.4
7001–10,000	250	3	575	825	17	1.3	425	7	785	1210	27	1.5
10,001–20,000	295	4	655	950	20	1.4	470	8	900	1370	31	1.6
20,001–50,000	295	4	755	1050	23	1.5	515	9	1165	1680	39	1.7
50,001–100,000	340	5	840	1180	26	1.6	515	9	1315	1830	43	1.8

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
AOQL = Average Outgoing Quality Limit.

TABLE II CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE DL-4

Lot Tolerance Per Cent Defective = 4.0%

Process Average %	0–.04						.05–.40					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–35	All	0	–	–	–	0	All	0	–	–	–	0
36–50	34	0	–	–	–	.35	34	0	–	–	–	.35
51–75	40	0	–	–	–	.43	40	0	–	–	–	.43
76–100	50	0	25	75	1	.46	50	0	25	75	1	.46
101–150	55	0	30	85	1	.55	55	0	30	85	1	.55
151–200	60	0	30	90	1	.64	60	0	30	90	1	.64
201–300	60	0	35	95	1	.70	60	0	35	95	1	.70
301–400	65	0	35	100	1	.71	65	0	35	100	1	.71
401–500	65	0	40	105	1	.73	65	0	70	135	2	.83
501–600	65	0	40	105	1	.74	65	0	75	140	2	.85
601–800	65	0	40	105	1	.75	65	0	75	140	2	.87
801–1000	70	0	40	110	1	.76	70	0	75	145	2	.90
1001–2000	70	0	40	110	1	.78	70	0	80	150	2	.94
2001–3000	70	0	80	150	2	.95	70	0	115	185	3	1.1
3001–4000	70	0	80	150	2	.96	70	0	115	185	3	1.1
4001–5000	70	0	80	150	2	.97	70	0	115	185	3	1.1
5001–7000	70	0	80	150	2	.98	70	0	115	185	3	1.1
7001–10,000	70	0	80	150	2	.98	70	0	150	220	4	1.2
10,001–20,000	70	0	80	150	2	.98	70	0	150	220	4	1.2
20,001–50,000	70	0	80	150	2	.99	70	0	150	220	4	1.2
50,001–100,000	70	0	80	150	2	.99	70	0	185	255	5	1.3

Process Average %	.41–.80						.81–1.20					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–35	All	0	–	–	–	0	All	0	–	–	–	0
36–50	34	0	–	–	–	.35	34	0	–	–	–	.35
51–75	40	0	–	–	–	.43	40	0	–	–	–	.43
76–100	50	0	25	75	1	.46	50	0	25	75	1	.46
101–150	55	0	30	85	1	.55	55	0	30	85	1	.55
151–200	60	0	30	90	1	.64	60	0	55	115	2	.68
201–300	60	0	65	125	2	.75	60	0	65	125	2	.75
301–400	65	0	65	130	2	.80	65	0	95	160	3	.86
401–500	65	0	70	135	2	.83	65	0	100	165	3	.92
501–600	65	0	100	165	3	.93	65	0	135	200	4	1.0
601–800	65	0	110	175	3	.97	65	0	140	205	4	1.0
801–1000	70	0	105	175	3	.98	110	1	155	265	6	1.2
1001–2000	70	0	145	215	4	1.1	110	1	195	305	7	1.3
2001–3000	70	0	180	250	5	1.2	110	1	260	370	9	1.4
3001–4000	110	1	175	285	6	1.3	150	2	255	405	10	1.5
4001–5000	115	1	170	285	6	1.3	150	2	285	435	11	1.6
5001–7000	115	1	205	320	7	1.4	150	2	320	470	12	1.6
7001–10,000	115	1	205	320	7	1.4	150	2	325	475	12	1.7
10,001–20,000	115	1	235	350	8	1.5	150	2	355	505	13	1.7
20,001–50,000	115	1	270	385	9	1.6	150	2	420	570	15	1.7
50,001–100,000	115	1	300	415	10	1.7	150	2	450	600	16	1.8

Process Average %	1.21–1.60						1.61–2.00					
	Trial 1		Trial 2			AOQL in %	Trial 1		Trial 2			AOQL in %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–35	All	0	–	–	–	0	All	0	–	–	–	0
36–50	34	0	–	–	–	.35	34	0	–	–	–	.35
51–75	40	0	–	–	–	.43	40	0	–	–	–	.43
76–100	50	0	25	75	1	.46	50	0	25	75	1	.46
101–150	55	0	30	85	1	.55	55	0	30	85	1	.55
151–200	60	0	55	115	2	.68	60	0	55	115	2	.68
201–300	60	0	90	150	3	.84	60	0	90	150	3	.84
301–400	65	0	95	160	3	.86	65	0	120	185	4	.92
401–500	65	0	130	195	4	.96	105	1	140	245	6	1.0
501–600	105	1	145	250	6	1.1	105	1	175	280	7	1.1
601–800	105	1	185	290	7	1.2	105	1	210	315	8	1.2
801–1000	110	1	210	320	8	1.2	145	2	230	375	10	1.3
1001–2000	150	2	240	390	10	1.5	180	3	295	475	13	1.6
2001–3000	185	3	305	490	13	1.6	220	4	410	630	18	1.7
3001–4000	185	3	340	525	14	1.6	285	6	465	750	22	1.8
4001–5000	185	3	395	580	16	1.7	285	6	520	805	24	1.9
5001–7000	185	3	435	620	17	1.7	320	7	585	905	27	2.0
7001–10,000	220	4	460	680	19	1.9	320	7	645	965	29	2.1
10,001–20,000	220	4	495	715	20	1.9	350	8	790	1140	35	2.2
20,001–50,000	255	5	575	830	24	2.0	385	9	895	1280	40	2.3
50,001–100,000	255	5	665	920	27	2.1	415	10	985	1400	44	2.4

TABLE DL-5

Lot Tolerance Per Cent Defective = 5.0%

Process Average %	0–.05						.06–.50						.51–1.00					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–30	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
31–50	30	0	–	–	–	.49	30	0	–	–	–	.49	30	0	–	–	–	.49
51–75	38	0	–	–	–	.59	38	0	–	–	–	.59	38	0	–	–	–	.59
76–100	44	0	21	65	1	.64	44	0	21	65	1	.64	44	0	21	65	1	.64
101–200	49	0	26	75	1	.84	49	0	26	75	1	.84	49	0	26	75	1	.84
201–300	50	0	30	80	1	.91	50	0	30	80	1	.91	50	0	55	105	2	1.0
301–400	55	0	30	85	1	.92	55	0	55	110	2	1.1	55	0	55	110	2	1.1
401–500	55	0	30	85	1	.93	55	0	55	110	2	1.1	55	0	80	135	3	1.2
501–600	55	0	30	85	1	.94	55	0	60	115	2	1.1	55	0	85	140	3	1.2
601–800	55	0	35	90	1	.95	55	0	65	120	2	1.1	55	0	85	140	3	1.3
801–1000	55	0	35	90	1	.96	55	0	65	120	2	1.1	55	0	115	170	4	1.4
1001–2000	55	0	35	90	1	.98	55	0	95	150	3	1.3	55	0	120	175	4	1.4
2001–3000	55	0	65	120	2	1.2	55	0	95	150	3	1.3	55	0	150	205	5	1.5
3001–4000	55	0	65	120	2	1.2	55	0	95	150	3	1.3	90	1	140	230	6	1.6
4001–5000	55	0	65	120	2	1.2	55	0	95	150	3	1.4	90	1	165	255	7	1.8
5001–7000	55	0	65	120	2	1.2	55	0	95	150	3	1.4	90	1	165	255	7	1.8
7001–10,000	55	0	65	120	2	1.2	55	0	120	175	4	1.5	90	1	190	280	8	1.9
10,001–20,000	55	0	65	120	2	1.2	55	0	120	175	4	1.5	90	1	190	280	8	1.9
20,001–50,000	55	0	65	120	2	1.2	55	0	150	205	5	1.7	90	1	215	305	9	2.0
50,001–100,000	55	0	65	120	2	1.2	55	0	150	205	5	1.7	90	1	240	330	10	2.1

Process Average %	1.01–1.50						1.51–2.00						2.01–2.50					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–30	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
31–50	30	0	–	–	–	.49	30	0	–	–	–	.49	30	0	–	–	–	.49
51–75	38	0	–	–	–	.59	38	0	–	–	–	.59	38	0	–	–	–	.59
76–100	44	0	21	65	1	.64	44	0	21	65	1	.64	44	0	21	65	1	.64
101–200	49	0	51	100	2	.91	49	0	51	100	2	.91	49	0	51	100	2	.91
201–300	50	0	55	105	2	1.0	50	0	80	130	3	1.1	50	0	100	150	4	1.1
301–400	55	0	80	135	3	1.1	55	0	100	155	4	1.2	85	1	105	190	6	1.3
401–500	55	0	105	160	4	1.3	85	1	120	205	6	1.4	85	1	140	225	7	1.4
501–600	55	0	110	165	4	1.3	85	1	145	230	7	1.4	85	1	165	250	8	1.5
601–800	90	1	125	215	6	1.5	90	1	170	260	8	1.5	120	2	185	305	10	1.6
801–1000	90	1	150	240	7	1.5	90	1	200	290	9	1.6	120	2	210	330	11	1.7
1001–2000	90	1	185	275	8	1.7	120	2	225	345	11	1.9	175	4	260	435	15	2.0
2001–3000	120	2	180	300	9	1.9	150	3	270	420	14	2.1	205	5	375	580	21	2.3
3001–4000	120	2	210	330	10	2.0	150	3	295	445	15	2.3	230	6	420	650	24	2.4
4001–5000	120	2	255	375	12	2.1	150	3	345	495	17	2.3	255	7	445	700	26	2.5
5001–7000	120	2	260	380	12	2.1	150	3	370	520	18	2.3	255	7	495	750	28	2.6
7001–10,000	120	2	285	405	13	2.1	175	4	370	545	19	2.4	280	8	540	820	31	2.7
10,001–20,000	120	2	310	430	14	2.2	175	4	420	595	21	2.4	280	8	660	940	36	2.8
20,001–50,000	120	2	335	455	15	2.2	205	5	485	690	25	2.5	305	9	745	1050	41	2.9
50,001–100,000	120	2	360	480	16	2.3	205	5	555	760	28	2.6	330	10	810	1140	45	3.0

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
AOQL = Average Outgoing Quality Limit.

TABLE II CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "LOT TOLERANCE PER CENT DEFECTIVE" AND CONSUMER'S RISK = 0.10

TABLE DL-7

LOT TOLERANCE PER CENT DEFECTIVE = 7.0%

Process Average %	0–07						.08–.70					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–25	All	0	–	–	–	0	All	0	–	–	–	0
26–50	24	0	–	–	–	.80	24	0	–	–	–	.80
51–75	31	0	15	46	1	.90	31	0	15	46	1	.90
76–110	34	0	16	50	1	1.1	34	0	16	50	1	1.1
111–200	36	0	19	55	1	1.2	36	0	19	55	1	1.2
201–300	37	0	23	60	1	1.3	37	0	23	60	1	1.3
301–400	38	0	22	60	1	1.3	38	0	42	80	2	1.5
401–500	39	0	21	60	1	1.3	39	0	41	80	2	1.5
501–600	39	0	26	65	1	1.3	39	0	46	85	2	1.6
601–800	39	0	26	65	1	1.4	39	0	46	85	2	1.6
801–1000	39	0	26	65	1	1.4	39	0	46	85	2	1.6
1001–2000	40	0	45	85	2	1.7	40	0	65	105	3	1.9
2001–3000	40	0	45	85	2	1.7	40	0	65	105	3	1.9
3001–4000	40	0	45	85	2	1.7	40	0	65	105	3	1.9
4001–5000	40	0	45	85	2	1.7	40	0	85	125	4	2.1
5001–7000	40	0	45	85	2	1.7	40	0	85	125	4	2.1
7001–10,000	40	0	45	85	2	1.7	40	0	85	125	4	2.1
10,001–20,000	40	0	45	85	2	1.7	40	0	85	125	4	2.1
20,001–50,000	40	0	45	85	2	1.7	40	0	105	145	5	2.3
50,001–100,000	40	0	45	85	2	1.7	40	0	105	145	5	2.3

Process Average %	.71–1.40						1.41–2.10					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–25	All	0	–	–	–	0	All	0	–	–	–	0
26–50	24	0	–	–	–	.80	24	0	–	–	–	.80
51–75	31	0	15	46	1	.90	31	0	15	46	1	.90
76–110	34	0	16	50	1	1.1	34	0	31	65	2	1.2
111–200	36	0	39	75	2	1.4	36	0	54	90	3	1.5
201–300	37	0	38	75	2	1.5	37	0	58	95	3	1.6
301–400	38	0	57	95	3	1.7	38	0	77	115	4	1.8
401–500	39	0	61	100	3	1.7	39	0	76	115	4	1.8
501–600	39	0	61	100	3	1.7	65	1	90	155	6	2.1
601–800	39	0	81	120	4	1.9	65	1	105	170	7	2.2
801–1000	39	0	86	125	4	2.0	65	1	110	175	7	2.2
1001–2000	40	0	105	145	5	2.2	65	1	150	215	9	2.5
2001–3000	65	1	100	165	6	2.3	85	2	165	250	11	2.7
3001–4000	65	1	115	180	7	2.4	85	2	185	270	12	2.8
4001–5000	65	1	115	180	7	2.5	85	2	185	270	12	2.9
5001–7000	65	1	135	200	8	2.5	85	2	205	290	13	3.0
7001–10,000	65	1	135	200	8	2.6	85	2	205	290	13	3.1
10,001–20,000	65	1	155	220	9	2.7	85	2	220	305	14	3.2
20,001–50,000	65	1	170	235	10	2.8	85	2	240	325	15	3.3
50,001–100,000	65	1	170	235	10	2.9	105	3	270	375	18	3.4

Process Average %	2.11–2.80						2.81–3.50					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1–25	All	0	–	–	–	0	All	0	–	–	–	0
26–50	24	0	–	–	–	.80	24	0	–	–	–	.80
51–75	31	0	15	46	1	.90	31	0	15	46	1	.90
76–110	34	0	31	65	2	1.2	34	0	31	65	2	1.2
111–200	36	0	54	90	3	1.5	36	0	69	105	4	1.5
201–300	37	0	73	110	4	1.7	60	1	80	140	6	1.9
301–400	60	1	85	145	6	1.9	60	1	100	160	7	2.0
401–500	60	1	105	165	7	2.1	60	1	135	195	9	2.2
501–600	65	1	120	185	8	2.2	85	2	130	215	10	2.3
601–800	65	1	140	205	9	2.3	85	2	165	250	12	2.5
801–1000	85	2	140	225	10	2.5	105	3	180	285	14	2.7
1001–2000	105	3	175	280	13	2.8	145	5	230	375	19	3.1
2001–3000	105	3	210	315	15	3.0	165	6	300	465	24	3.3
3001–4000	105	3	250	355	17	3.1	180	7	335	515	27	3.4
4001–5000	125	4	245	370	18	3.2	180	7	370	550	29	3.6
5001–7000	125	4	265	390	19	3.3	200	8	385	585	31	3.8
7001–10,000	125	4	300	425	21	3.4	200	8	450	650	35	3.9
10,001–20,000	125	4	335	460	23	3.5	220	9	485	705	38	4.0
20,001–50,000	145	5	360	505	26	3.5	220	9	565	785	43	4.1
50,001–100,000	165	6	390	555	29	3.7	235	10	610	845	47	4.2

TABLE DL-10

LOT TOLERANCE PER CENT DEFECTIVE = 10.0%

Process Average %	0-.10					.11-1.00					1.01-2.00					2.01-3.00					3.01-4.00					4.01-5.00				
	Trial 1		Trial 2				Trial 1		Trial 2			Trial 1		Trial 2			Trial 1		Trial 2			Trial 1		Trial 2			Trial 1		Trial 2		
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %	n_1	c_1	n_2	n_1+n_2	c_2	AOQL in %
1-20	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0	All	0	–	–	–	0
21-50	17	0	–	–	–	1.3	17	0	–	–	–	1.3	17	0	–	–	–	1.3	17	0	–	–	–	1.3	17	0	–	–	–	1.3	17	0	–	–	–	1.3
51-100	25	0	13	38	1	1.6	25	0	13	38	1	1.6	25	0	13	38	1	1.6	25	0	24	49	2	1.8	25	0	24	49	2	1.8	25	0	24	49	2	1.8
101-200	27	0	15	42	1	1.8	27	0	15	42	1	1.8	27	0	28	55	2	2.1	27	0	38	65	3	2.3	27	0	53	80	4	2.4	27	0	53	80	4	2.4
201-300	27	0	16	43	1	1.9	27	0	30	57	2	2.2	27	0	43	70	3	2.4	27	0	53	80	4	2.7	43	1	62	105	6	2.8	43	1	82	125	8	3.0
301-400	27	0	17	44	1	1.9	27	0	33	60	2	2.2	27	0	43	70	3	2.5	44	1	66	110	6	2.9	44	1	86	130	8	3.1	60	2	90	150	10	3.2
401-500	28	0	16	44	1	1.9	28	0	32	60	2	2.3	28	0	57	85	4	2.7	44	1	76	120	7	3.1	44	1	101	145	9	3.3	60	2	105	165	11	3.4
501-600	28	0	17	45	1	1.9	28	0	32	60	2	2.3	28	0	57	85	4	2.8	45	1	75	120	7	3.3	60	2	100	160	10	3.4	75	3	115	190	13	3.6
601-800	28	0	17	45	1	2.0	28	0	47	75	3	2.6	28	0	57	85	4	2.9	45	1	90	135	8	3.5	60	2	110	170	11	3.7	75	3	140	215	15	3.9
801-1000	28	0	32	60	2	2.3	28	0	47	75	3	2.6	28	0	72	100	5	3.0	45	1	90	135	8	3.5	60	2	125	185	12	3.9	90	4	150	240	17	4.1
1001-2000	28	0	32	60	2	2.4	28	0	47	75	3	2.7	45	1	70	115	6	3.3	60	2	105	165	10	3.9	75	3	150	225	15	4.3	115	6	200	315	23	4.8
2001-3000	28	0	32	60	2	2.4	28	0	47	75	3	2.7	45	1	85	130	7	3.5	60	2	130	190	12	4.1	75	3	175	250	17	4.4	130	7	235	365	27	5.0
3001-4000	28	0	32	60	2	2.4	28	0	62	90	4	2.9	45	1	85	130	7	3.5	60	2	130	190	12	4.2	90	4	170	260	18	4.6	130	7	255	385	29	5.1
4001-5000	28	0	32	60	2	2.4	28	0	62	90	4	3.0	45	1	95	140	8	3.7	60	2	140	200	13	4.3	90	4	180	270	19	4.7	140	8	270	410	31	5.2
5001-7000	28	0	32	60	2	2.4	28	0	62	90	4	3.0	45	1	95	140	8	3.8	60	2	140	200	13	4.4	90	4	205	295	21	4.9	140	8	315	455	35	5.3
7001-10,000	28	0	32	60	2	2.5	28	0	62	90	4	3.0	45	1	95	140	8	3.8	60	2	155	215	14	4.4	90	4	220	310	22	5.0	140	8	340	480	37	5.4
10,001-20,000	28	0	32	60	2	2.5	28	0	62	90	4	3.0	45	1	110	155	9	3.9	60	2	165	225	15	4.4	100	5	230	330	24	5.1	155	9	370	525	41	5.6
20,001-50,000	28	0	32	60	2	2.5	28	0	72	100	5	3.3	45	1	120	165	10	3.9	75	3	165	240	16	4.5	100	5	280	380	28	5.2	165	10	405	570	45	5.7
50,001-100,000	28	0	32	60	2	2.5	28	0	72	100	5	3.3	45	1	135	180	11	4.2	75	3	200	275	19	4.8	115	6	285	400	30	5.3	165	10	440	605	48	6.2

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Com bined.
AOQL = Average Outgoing Quality Limit.

TABLE III: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-0.1
Average Outgoing Quality Limit = 0.1%

Process Average %	0–.002			.003–.020			.021–.040			.041–.060			.061–.080			.081–.100		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1–75	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
76–95	75	0	1.5	75	0	1.5	75	0	1.5	75	0	1.5	75	0	1.5	75	0	1.5
96–130	95	0	1.4	95	0	1.4	95	0	1.4	95	0	1.4	95	0	1.4	95	0	1.4
131–200	130	0	1.2	130	0	1.2	130	0	1.2	130	0	1.2	130	0	1.2	130	0	1.2
201–300	165	0	1.1	165	0	1.1	165	0	1.1	165	0	1.1	165	0	1.1	165	0	1.1
301–400	190	0	.96	190	0	.96	190	0	.96	190	0	.96	190	0	.96	190	0	.96
401–500	210	0	.91	210	0	.91	210	0	.91	210	0	.91	210	0	.91	210	0	.91
501–600	230	0	.86	230	0	.86	230	0	.86	230	0	.86	230	0	.86	230	0	.86
601–800	250	0	.81	250	0	.81	250	0	.81	250	0	.81	250	0	.81	250	0	.81
801–1000	270	0	.76	270	0	.76	270	0	.76	270	0	.76	270	0	.76	270	0	.76
1001–2000	310	0	.71	310	0	.71	310	0	.71	310	0	.71	310	0	.71	310	0	.71
2001–3000	330	0	.67	330	0	.67	330	0	.67	330	0	.67	330	0	.67	655	1	.64
3001–4000	340	0	.64	340	0	.64	340	0	.64	695	1	.59	695	1	.59	695	1	.59
4001–5000	345	0	.62	345	0	.62	345	0	.62	720	1	.54	720	1	.54	720	1	.54
5001–7000	350	0	.61	350	0	.61	750	1	.51	750	1	.51	750	1	.51	750	1	.51
7001–10,000	355	0	.60	355	0	.60	775	1	.49	775	1	.49	775	1	.49	1210	2	.44
10,001–20,000	360	0	.59	810	1	.48	810	1	.48	1280	2	.42	1280	2	.42	1770	3	.38
20,001–50,000	365	0	.58	830	1	.47	1330	2	.41	1870	3	.37	2420	4	.34	2980	5	.33
50,001–100,000	365	0	.58	835	1	.46	1350	2	.40	2480	4	.33	3070	5	.32	4270	7	.30

TABLE SA-0.25
Average Outgoing Quality Limit = 0.25%

Process Average %	0–.005			.006–.050			.051–.100			.101–.150			.151–.200			.201–.250		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1–60	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
61–100	60	0	2.5	60	0	2.5	60	0	2.5	60	0	2.5	60	0	2.5	60	0	2.5
101–200	85	0	2.1	85	0	2.1	85	0	2.1	85	0	2.1	85	0	2.1	85	0	2.1
201–300	100	0	1.9	100	0	1.9	100	0	1.9	100	0	1.9	100	0	1.9	100	0	1.9
301–400	110	0	1.8	110	0	1.8	110	0	1.8	110	0	1.8	110	0	1.8	110	0	1.8
401–500	115	0	1.8	115	0	1.8	115	0	1.8	115	0	1.8	115	0	1.8	115	0	1.8
501–600	120	0	1.7	120	0	1.7	120	0	1.7	120	0	1.7	120	0	1.7	120	0	1.7
601 800	125	0	1.7	125	0	1.7	125	0	1.7	125	0	1.7	125	0	1.7	125	0	1.7
801–1000	130	0	1.7	130	0	1.7	130	0	1.7	130	0	1.7	130	0	1.7	250	1	1.4
1001–2000	135	0	1.6	135	0	1.6	135	0	1.6	290	1	1.3	290	1	1.3	290	1	1.3
2001 3000	140	0	1.6	140	0	1.6	300	1	1.3	300	1	1.3	300	1	1.3	300	1	1.3
3001–4000	140	0	1.6	140	0	1.6	310	1	1.3	310	1	1.3	310	1	1.3	485	2	1.1
4001–5000	145	0	1.6	145	0	1.6	315	1	1.2	315	1	1.2	495	2	1.1	495	2	1.1
5001–7000	145	0	1.6	320	1	1.2	320	1	1.2	510	2	1.0	510	2	1.0	700	3	.94
7001–10,000	145	0	1.6	325	1	1.2	325	1	1.2	520	2	1.0	720	3	.91	720	3	.91
10,001–20,000	145	0	1.6	330	1	1.2	535	2	1.0	750	3	.89	970	4	.81	1190	5	.75
20,001–50,000	145	0	1.6	335	1	1.2	545	2	1.0	995	4	.80	1240	5	.74	1980	8	.66
50,001–100,000	335	1	1.2	545	2	1.0	775	3	.87	1250	5	.73	1750	7	.67	2810	11	.62

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE III CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-0.5
Average Outgoing Quality Limit = 0.5%

Process Average %	0-.010			.011-.10			.11-.20			.21-.30			.31-.40			41.-.50		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1-30	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-
31-50	30	0	5.0	30	0	5.0	30	0	5.0	30	0	5.0	30	0	5.0	30	0	5.0
51-100	42	0	4.2	42	0	4.2	42	0	4.2	42	0	4.2	42	0	4.2	42	0	4.2
101-200	55	0	3.6	55	0	3.6	55	0	3.6	55	0	3.6	55	0	3.6	55	0	3.6
201-300	60	0	3.4	60	0	3.4	60	0	3.4	60	0	3.4	60	0	3.4	60	0	3.4
301-400	60	0	3.5	60	0	3.5	60	0	3.5	60	0	3.5	60	0	3.5	60	0	3.5
401-500	65	0	3.3	65	0	3.3	65	0	3.3	65	0	3.3	65	0	3.3	125	1	2.9
501-600	65	0	3.3	65	0	3.3	65	0	3.3	65	0	3.3	130	1	2.7	130	1	2.7
601-800	65	0	3.4	65	0	3.4	65	0	3.4	140	1	2.6	140	1	2.6	140	1	2.6
801-1000	70	0	3.2	70	0	3.2	70	0	3.2	145	1	2.6	145	1	2.6	145	1	2.6
1001-2000	70	0	3.2	70	0	3.2	155	1	2.5	155	1	2.5	155	1	2.5	240	2	2.2
2001-3000	70	0	3.3	70	0	3.3	160	1	2.4	160	1	2.4	250	2	2.1	250	2	2.1
3001-4000	70	0	3.3	160	1	2.4	160	1	2.4	255	2	2.1	255	2	2.1	355	3	1.9
4001-5000	75	0	3.0	165	1	2.4	165	1	2.4	260	2	2.0	360	3	1.9	460	4	1.7
5001-7000	75	0	3.0	165	1	2.4	265	2	2.0	265	2	2.0	370	3	1.8	475	4	1.7
7001-10,000	75	0	3.1	165	1	2.4	265	2	2.0	375	3	1.8	485	4	1.7	595	5	1.6
10,001-20,000	75	0	3.1	165	1	2.4	270	2	1.9	380	3	1.7	615	5	1.5	855	7	1.4
20,001-50,000	170	1	2.3	275	2	1.9	390	3	1.7	625	5	1.5	875	7	1.3	1410	11	1.2
50,001-100,000	170	1	2.3	275	2	1.9	510	4	1.6	755	6	1.4	1290	10	1.2	2130	16	1.1

TABLE SA-0.75
Average Outgoing Quality Limit = 0.75%

Process Average %	0-.015			.016-.15			.16-.30			.31-.45			.46-.60			.61-.75		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1-25	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-
26-50	25	0	6.4	25	0	6.4	25	0	6.4	25	0	6.4	25	0	6.4	25	0	6.4
51-100	33	0	5.6	33	0	5.6	33	0	5.6	33	0	5.6	33	0	5.6	33	0	5.6
101-200	39	0	5.2	39	0	5.2	39	0	5.2	39	0	5.2	39	0	5.2	39	0	5.2
201-300	42	0	5.0	42	0	5.0	42	0	5.0	42	0	5.0	42	0	5.0	42	0	5.0
301-400	44	0	4.9	44	0	4.9	44	0	4.9	44	0	4.9	90	1	4.0	90	1	4.0
401-500	45	0	4.8	45	0	4.8	45	0	4.8	90	1	4.1	90	1	4.1	90	1	4.1
501-600	45	0	4.9	45	0	4.9	45	0	4.9	95	1	3.9	95	1	3.9	95	1	3.9
601-800	46	0	4.9	46	0	4.9	100	1	3.8	100	1	3.8	100	1	3.8	100	1	3.8
801-1000	47	0	4.8	47	0	4.8	100	1	3.8	100	1	3.8	100	1	3.8	155	2	3.2
1001-2000	48	0	4.7	48	0	4.7	105	1	3.7	105	1	3.7	170	2	3.1	170	2	3.1
2001-3000	48	0	4.7	110	1	3.5	110	1	3.5	170	2	3.1	170	2	3.1	240	3	2.8
3001-4000	48	0	4.7	110	1	3.5	110	1	3.5	175	2	3.1	245	3	2.7	315	4	2.5
4001-5000	49	0	4.6	110	1	3.6	175	2	3.1	175	2	3.1	245	3	2.7	320	4	2.5
5001-7000	49	0	4.6	110	1	3.6	180	2	3.0	250	3	2.7	325	4	2.5	400	5	2.3
7001-10,000	49	0	4.6	110	1	3.7	180	2	3.0	255	3	2.6	405	5	2.3	560	7	2.1
10,001-20,000	49	0	4.6	110	1	3.7	255	3	2.6	335	4	2.4	495	6	2.1	750	9	1.9
20,001-50,000	110	1	3.7	180	2	3.0	260	3	2.6	420	5	2.2	675	8	1.9	1130	13	1.6
50,001-100,000	110	1	3.7	185	2	2.9	335	4	2.4	590	7	2.0	955	11	1.7	1720	19	1.5

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE III CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-1.0
Average Outgoing Quality Limit = 1.0%

Process Average %	0–.02			.03–.20			.21–.40			.41–.60			.61–.80			.81–1.00		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1–25	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
26–50	22	0	7.7	22	0	7.7	22	0	7.7	22	0	7.7	22	0	7.7	22	0	7.7
51–100	27	0	7.1	27	0	7.1	27	0	7.1	27	0	7.1	27	0	7.1	27	0	7.1
101–200	32	0	6.4	32	0	6.4	32	0	6.4	32	0	6.4	32	0	6.4	32	0	6.4
201–300	33	0	6.3	33	0	6.3	33	0	6.3	33	0	6.3	33	0	6 3	65	1	5.0
301–400	34	0	6.1	34	0	6.1	34	0	6.1	70	1	4.6	70	1	4.6	70	1	4.6
401–500	35	0	6.1	35	0	6.1	35	0	6.1	70	1	4.7	70	1	4.7	70	1	4.7
501–600	35	0	6.1	35	0	6.1	75	1	4.4	75	1	4.4	75	1	4.4	75	1	4.4
601–800	35	0	6.2	35	0	6.2	75	1	4.4	75	1	4.4	75	1	4.4	120	2	4.2
801–1000	35	0	6.3	35	0	6.3	80	1	4.4	80	1	4.4	120	2	4.3	120	2	4.3
1001–2000	36	0	6.2	80	1	4.5	80	1	4.5	130	2	4.0	130	2	4.0	180	3	3.7
2001–3000	36	0	6.2	80	1	4.6	80	1	4.6	130	2	4.0	185	3	3.6	235	4	3.3
3001–4000	36	0	6.2	80	1	4.7	135	2	3.9	135	2	3.9	185	3	3.6	295	5	3.1
4001–5000	36	0	6.2	85	1	4.6	135	2	3.9	190	3	3.5	245	4	3.2	300	5	3.1
5001–7000	37	0	6.1	85	1	4.6	135	2	3.9	190	3	3.5	305	5	3.0	420	7	2.8
7001–10,000	37	0	6.2	85	1	4.6	135	2	3.9	245	4	3.2	310	5	3.0	430	7	2.7
10,001–20,000	85	1	4.6	135	2	3.9	195	3	3.4	250	4	3.2	435	7	2.7	635	10	2.4
20,001–50,000	85	1	4.6	135	2	3.9	255	4	3.1	380	6	2.8	575	9	2.5	990	15	2.1
50,001–100,000	85	1	4.6	135	2	3.9	255	4	3.1	445	7	2.6	790	12	2.3	1520	22	1.9

TABLE SA-1.5
Average Outgoing Quality Limit = 1.5%

Process Average %	0–.03			.04–.30			.31–.60			.61–.90			.91–1.20			1.21–1.50		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1–15	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
16–50	16	0	11.6	16	0	11.6	16	0	11.6	16	0	11.6	16	0	11.6	16	0	11.6
51–100	20	0	9.8	20	0	9.8	20	0	9.8	20	0	9.8	20	0	9.8	20	0	9.8
101–200	22	0	9.5	22	0	9.5	22	0	9.5	22	0	9.5	22	0	9.5	44	1	8.2
201–300	23	0	9.2	23	0	9.2	23	0	9.2	47	1	7.9	47	1	7.9	47	1	7.9
301–400	23	0	9.3	23	0	9.3	49	1	7.8	49	1	7.8	49	1	7.8	49	1	7.8
401–500	23	0	9.4	23	0	9.4	50	1	7.7	50	1	7.7	50	1	7.7	50	1	7.7
501–600	24	0	9.0	24	0	9.0	50	1	7.7	50	1	7.7	50	1	7.7	50	1	7.7
601–800	24	0	9.1	24	0	9.1	50	1	7.8	50	1	7.8	80	2	6.4	80	2	6.4
801–1000	24	0	9.1	55	1	7.0	55	1	7.0	85	2	6.2	85	2	6.2	85	2	6.2
1001–2000	24	0	9.1	55	1	7.0	55	1	7.0	85	2	6.2	120	3	5.4	155	4	5.0
2001–3000	24	0	9.2	55	1	7.1	90	2	5.9	125	3	5.3	160	4	4.9	200	5	4.6
3001–4000	24	0	9.2	55	1	7.1	90	2	5.9	125	3	5.3	165	4	4.8	240	6	4.4
4001–5000	24	0	9.2	55	1	7.1	90	2	5.9	125	3	5.3	205	5	4.6	280	7	4.2
5001–7000	24	0	9.2	55	1	7.1	90	2	5.9	165	4	4.8	205	5	4.6	325	8	4.0
7001–10,000	24	0	9.2	55	1	7.1	130	3	5.2	165	4	4.8	250	6	4.2	375	9	3.8
10,001–20,000	55	1	7.1	90	2	5.9	130	3	5.2	210	5	4.4	340	8	3.8	515	12	3.4
20,001–50,000	55	1	7.1	90	2	5.9	170	4	4.7	295	7	4.0	480	11	3.5	860	19	3.0
50,001–100,000	55	1	7.1	130	3	5.2	210	5	4.4	340	8	3.8	625	14	3.3	1120	24	2.8

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE III CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-2.0
Average Outgoing Quality Limit = 2.0%

Process Average %	0-.04			.05-.40			.41-.80			.81-1.20			1.21-1.60			1.61-2.00		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1-15	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-
16-50	14	0	13.6	14	0	13.6	14	0	13.6	14	0	13.6	14	0	13.6	14	0	13.6
51-100	16	0	12.4	16	0	12.4	16	0	12.4	16	0	12.4	16	0	12.4	16	0	12.4
101-200	17	0	12.2	17	0	12.2	17	0	12.2	17	0	12.2	35	1	10.5	35	1	10.5
201-300	17	0	12.3	17	0	12.3	17	0	12.3	37	1	10.2	37	1	10.2	37	1	10.2
301-400	18	0	11.8	18	0	11.8	38	1	10.0	38	1	10.0	38	1	10.0	60	2	8.5
401-500	18	0	11.9	18	0	11.9	39	1	9.8	39	1	9.8	60	2	8.6	60	2	8.6
501-600	18	0	11.9	18	0	11.9	39	1	9.8	39	1	9.8	60	2	8.6	60	2	8.6
601-800	18	0	11.9	40	1	9.6	40	1	9.6	65	2	8.0	65	2	8.0	85	3	7.5
801-1000	18	0	12.0	40	1	9.6	40	1	9.6	65	2	8.1	65	2	8.1	90	3	7.4
1001-2000	18	0	12.0	41	1	9.4	65	2	8.2	65	2	8.2	95	3	7.0	120	4	6.5
2001-3000	18	0	12.0	41	1	9.4	65	2	8.2	95	3	7.0	120	4	6.5	180	6	5.8
3001-4000	18	0	12.0	42	1	9.3	65	2	8.2	95	3	7.0	155	5	6.0	210	7	5.5
4001-5000	18	0	12.0	42	1	9.3	70	2	7.5	125	4	6.4	155	5	6.0	245	8	5.3
5001-7000	18	0	12.0	42	1	9.3	95	3	7.0	125	4	6.4	185	6	5.6	280	9	5.1
7001-10,000	42	1	9.3	70	2	7.5	95	3	7.0	155	5	6.0	220	7	5.4	350	11	4.8
10,001-20,000	42	1	9.3	70	2	7.6	95	3	7.0	190	6	5.6	290	9	4.9	460	14	4.4
20,001-50,000	42	1	9.3	70	2	7.6	125	4	6.4	220	7	5.4	395	12	4.5	720	21	3.9
50,001-100,000	42	1	9.3	95	3	7.0	160	5	5.9	290	9	4.9	505	15	4.2	955	27	3.7

TABLE SA-2.5
Average Outgoing Quality Limit = 2.5%

Process Average %	0-.05			.06-.50			.51-1.00			1.01-1.50			1.51-2.00			2.01-2.50		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1-10	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-	All	0	-
11-50	11	0	17.6	11	0	17.6	11	0	17.6	11	0	17.6	11	0	17.6	11	0	17.6
51-100	13	0	15.3	13	0	15.3	13	0	15.3	13	0	15.3	13	0	15.3	13	0	15.3
101-200	14	0	14.7	14	0	14.7	14	0	14.7	29	1	12.9	29	1	12.9	29	1	12.9
201-300	14	0	14.9	14	0	14.9	30	1	12.7	30	1	12.7	30	1	12.7	30	1	12.7
301-400	14	0	15.0	14	0	15.0	31	1	12.3	31	1	12.3	31	1	12.3	48	2	10.7
401-500	14	0	15.0	14	0	15.0	32	1	12.0	32	1	12.0	49	2	10.6	49	2	10.6
501-600	14	0	15.1	32	1	12.0	32	1	12.0	50	2	10.4	50	2	10.4	70	3	9.3
601-800	14	0	15.1	32	1	12.0	32	1	12.0	50	2	10.5	50	2	10.5	70	3	9.4
801-1000	15	0	14.2	33	1	11.7	33	1	11.7	50	2	10.6	70	3	9.4	90	4	8.5
1001-2000	15	0	14.2	33	1	11.7	55	2	9.3	75	3	8.8	95	4	8.0	120	5	7.6
2001-3000	15	0	14.2	33	1	11.8	55	2	9.4	75	3	8.8	120	5	7.6	145	6	7.2
3001-4000	15	0	14.3	33	1	11.8	55	2	9.5	100	4	7.9	125	5	7.4	195	8	6.6
4001-5000	15	0	14.3	33	1	11.8	75	3	8.9	100	4	7.9	150	6	7.0	225	9	6.3
5001-7000	33	1	11.8	55	2	9.7	75	3	8.9	125	5	7.4	175	7	6.7	250	10	6.1
7001-10,000	34	1	11.4	55	2	9.7	75	3	8.9	125	5	7.4	200	8	6.4	310	12	5.8
10,001-20,000	34	1	11.4	55	2	9.7	100	4	8.0	150	6	7.0	260	10	6.0	425	16	5.3
20,001-50,000	34	1	11.4	55	2	9.7	100	4	8.0	180	7	6.7	345	13	5.5	640	23	4.8
50,001-100,000	34	1	11.4	80	3	8.4	125	5	7.4	235	9	6.1	435	16	5.2	800	28	4.5

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE III CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-3.0
Average Outgoing Quality Limit = 3.0%

Process Average %	0–.06			.07–.60			.61–1.20			1.21–1.80			1.81–2.40			2.41–3.00		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1–10	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
11–50	10	0	19.0	10	0	19.0	10	0	19.0	10	0	19.0	10	0	19.0	10	0	19.0
51–100	11	0	18.0	11	0	18.0	11	0	18.0	11	0	18.0	11	0	18.0	22	1	16.4
101–200	12	0	17.0	12	0	17.0	12	0	17.0	25	1	15.1	25	1	15.1	25	1	15.1
201–300	12	0	17.0	12	0	17.0	26	1	14.6	26	1	14.6	26	1	14.6	40	2	12.8
301–400	12	0	17.1	12	0	17.1	26	1	14.7	26	1	14.7	41	2	12.7	41	2	12.7
401–500	12	0	17.2	27	1	14.1	27	1	14.1	42	2	12.4	42	2	12.4	42	2	12.4
501–600	12	0	17.3	27	1	14.2	27	1	14.2	42	2	12.4	42	2	12.4	60	3	10.8
601–800	12	0	17.3	27	1	14.2	27	1	14.2	43	2	12.1	60	3	10.9	60	3	10.9
801–1000	12	0	17.4	27	1	14.2	44	2	11.8	44	2	11.8	60	3	11.0	80	4	9.8
1001–2000	12	0	17.5	28	1	13.8	45	2	11.7	65	3	10.2	80	4	9.8	100	5	9.1
2001–3000	12	0	17.5	28	1	13.8	45	2	11.7	65	3	10.2	100	5	9.1	140	7	8.2
3001–4000	12	0	17.5	28	1	13.8	65	3	10.3	85	4	9.5	125	6	8.4	165	8	7.8
4001–5000	28	1	13.8	28	1	13.8	65	3	10.3	85	4	9.5	125	6	8.4	210	10	7.4
5001–7000	28	1	13.8	45	2	11.8	65	3	10.3	105	5	8.8	145	7	8.1	235	11	7.1
7001–10,000	28	1	13.9	46	2	11.6	65	3	10.3	105	5	8.8	170	8	7.6	280	13	6.8
10,001–20,000	28	1	13.9	46	2	11.7	85	4	9.5	125	6	8.4	215	10	7.2	380	17	6.2
20,001–50,000	28	1	13.9	65	3	10.3	105	5	8.8	170	8	7.6	310	14	6.5	560	24	5.7
50,001–100,000	28	1	13.9	65	3	10.3	125	6	8.4	215	10	7.2	385	17	6.2	690	29	5.4

TABLE SA-4.0
Average Outgoing Quality Limit = 4.0%

Process Average %	0–.08			.09–.80			.81–1.60			1.61–2.40			2.41–3.20			3.21–4.00		
Lot Size	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%	n	c	pt%
1–10	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
11–50	8	0	23.0	8	0	23.0	8	0	23.0	8	0	23.0	8	0	23.0	8	0	23.0
51–100	8	0	24.0	8	0	24.0	8	0	24.0	8	0	24.0	17	1	21.5	17	1	21.5
101–200	9	0	22.0	9	0	22.0	19	1	20.0	19	1	20.0	19	1	20.0	19	1	20.0
201–300	9	0	22.5	9	0	22.5	20	1	19.0	20	1	19.0	31	2	16.8	31	2	16.8
301–400	9	0	22.5	20	1	19.1	20	1	19.1	32	2	16.2	32	2	16.2	43	3	15.2
401–500	9	0	22.5	20	1	19.1	20	1	19.1	32	2	16.3	32	2	16.3	44	3	14.9
501–600	9	0	22.5	20	1	19.2	20	1	19.2	32	2	16.3	45	3	14.6	60	4	12.9
601–800	9	0	22.5	20	1	19.2	33	2	15.9	33	2	15.9	46	3	14.3	60	4	13.0
801–1000	9	0	22.5	21	1	18.3	33	2	16.0	46	3	14.3	60	4	13.0	75	5	12.2
1001–2000	9	0	22.5	21	1	18.4	34	2	15.6	47	3	14.1	75	5	12.2	105	7	11.0
2001–3000	9	0	22.5	21	1	18.4	34	2	15.6	60	4	13.2	90	6	11.3	125	8	10.4
3001–4000	21	1	18.4	21	1	18.4	48	3	13.8	65	4	12.2	110	7	10.7	155	10	9.8
4001–5000	21	1	18.5	34	2	15.7	48	3	13.9	80	5	11.6	110	7	10.8	175	11	9.5
5001–7000	21	1	18.5	34	2	15 7	48	3	13.9	80	5	11.6	125	8	10.4	210	13	9.0
7001–10,000	21	1	18.5	34	2	15.7	65	4	12.3	95	6	11.1	145	9	9.8	245	15	8.6
10,001–20,000	21	1	18.5	34	2	15.7	65	4	12.3	110	7	10.8	195	12	9.0	340	20	7.9
20,001–50,000	21	1	18.5	49	3	13.6	80	5	11.6	145	9	9.8	250	15	8.5	460	26	7.4
50,001–100,000	21	1	18.5	49	3	13.6	95	6	11.1	165	10	9.6	310	18	8.0	540	30	7.1

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE III CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-5.0

Average Outgoing Quality Limit = 5.0%

Process Average %	0-.10			.11-1.00			1.01-2.00			2.01-3.00			3.01-4.00			4.01-5.00		
Lot Size	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%
1-5	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
6-50	6	0	30.5	6	0	30.5	6	0	30.5	6	0	30.5	6	0	30.5	6	0	30.5
51-100	7	0	27.0	7	0	27.0	7	0	27.0	14	1	26.5	14	1	26.5	14	1	26.5
101-200	7	0	27.5	7	0	27.5	16	1	24.0	16	1	24.0	16	1	24.0	24	2	21.5
201-300	7	0	27.5	16	1	24.0	16	1	24.0	16	1	24.0	25	2	21.0	25	2	21.0
301-400	7	0	27.5	16	1	24.0	16	1	24.0	26	2	20.0	26	2	20.0	35	3	18.8
401-500	7	0	27.5	16	1	24.0	16	1	24.0	26	2	20.0	36	3	18.3	46	4	17.0
501-600	7	0	28.0	16	1	24.0	26	2	20.0	26	2	20.0	37	3	17.9	47	4	16.6
601-800	7	0	28.0	16	1	24.0	27	2	19.4	37	3	17.9	48	4	16.3	60	5	15.2
801-1000	7	0	28.0	17	1	22.5	27	2	19.5	37	3	17.9	48	4	16.3	70	6	14.3
1001-2000	7	0	28.0	17	1	23.0	27	2	19.6	38	3	17.6	60	5	15.3	85	7	13.7
2001-3000	7	0	28.0	17	1	23.0	38	3	17.6	50	4	15.8	75	6	13.9	125	10	12.3
3001-4000	17	1	23.0	27	2	19.6	39	3	17.0	60	5	15.4	85	7	13.8	140	11	11.8
4001-5000	17	1	23.0	27	2	19.6	39	3	17.0	65	5	14.2	100	8	12.9	155	12	11.6
5001-7000	17	1	23.0	27	2	19.7	39	3	17.1	75	6	13.9	115	9	12.3	185	14	11.0
7001-10,000	17	1	23.0	27	2	19.7	50	4	15.9	75	6	14.0	130	10	12.0	225	17	10.4
10,001-20,000	17	1	23.0	27	2	19.7	50	4	15.9	90	7	13.1	170	13	11.0	305	22	9.6
20,001-50,000	17	1	23.0	39	3	17.1	65	5	14.3	115	9	12.3	215	16	10.4	400	28	9.0
50,001-100,000	17	1	23.0	39	3	17.1	75	6	14.0	145	11	11.6	275	20	9.8	450	31	8.8

TABLE SA-7.0

Average Outgoing Quality Limit = 7.0%

Process Average %	0-.14			.15-1.40			1.41-2.80			2.81-4.20			4.21-5.60			5.61-7.00		
Lot Size	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%
1-5	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
6-50	5	0	35.5	5	0	35.5	5	0	35.5	5	0	35.5	5	0	35.5	5	0	35.5
51-100	5	0	36.0	5	0	36.0	5	0	36.0	11	1	28.5	11	1	28.5	11	1	28.5
101-200	5	0	36.5	5	0	36.5	11	1	30.5	11	1	30.5	18	2	26.5	18	2	26.5
201-300	5	0	36.5	12	1	28.5	12	1	28.5	18	2	26.5	18	2	26.5	25	3	26.0
301-400	5	0	37.0	12	1	28.5	12	1	28.5	19	2	25.5	26	3	25.0	33	4	23.5
401-500	5	0	37.0	12	1	28.5	19	2	25.5	19	2	25.5	26	3	25.0	34	4	23.0
501-600	5	0	37.0	12	1	28.5	19	2	25.5	27	3	24.5	34	4	23.0	42	5	21.5
601-800	5	0	37.0	12	1	29.0	19	2	25.5	27	3	24.5	35	4	22.5	50	6	20.5
801-1000	5	0	37.0	12	1	29.0	19	2	25.5	27	3	24.5	43	5	21.5	60	7	19.3
1001-2000	5	0	37.0	12	1	29.0	27	3	24.5	36	4	22.0	50	6	21.0	70	8	17.7
2001-3000	12	1	29.0	19	2	25.5	28	3	23.5	45	5	20.5	60	7	19.6	100	11	16.5
3001-4000	12	1	29.0	20	2	24.5	28	3	24.0	45	5	20.5	70	8	18.1	120	13	15.8
4001-5000	12	1	29.0	20	2	24.5	36	4	22.0	55	6	19.0	80	9	17.3	140	15	15.1
5001-7000	12	1	29.0	20	2	24.5	36	4	22.0	55	6	19.1	90	10	16.8	160	17	14.6
7001-10,000	12	1	29.0	20	2	24.5	36	4	22.0	65	7	18.4	110	12	15.9	195	20	13.9
10,001-20,000	12	1	29.0	28	3	24.0	45	5	20.5	75	8	17.8	135	14	15.2	240	24	13.2
20,001-50,000	12	1	29.0	28	3	24.0	55	6	19.2	95	10	16.6	175	18	14.1	310	30	12.4
50,001-100,000	12	1	29.0	28	3	24.0	55	6	19.2	115	12	15.8	210	21	13.4	355	34	12.1

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE III CONT'D: SINGLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE SA-10.0

Average Outgoing Quality Limit = 10.0%

Process Average %	0–.20			.21–2.00			2.01–4.00			4.01–6.00			6.01–8.00			8.01–10.00		
Lot Size	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%	n	c	p_t%
1-3	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–	All	0	–
4-50	3	0	52.5	3	0	52.5	3	0	52.5	3	0	52.5	3	0	52.5	7	1	43.5
51-100	4	0	43.0	4	0	43.0	8	1	40.0	8	1	40.0	8	1	40.0	12	2	37.5
101-200	4	0	43.5	8	1	40.0	8	1	40.0	13	2	35.5	13	2	35.5	18	3	33.0
201-300	4	0	43.5	8	1	40.5	8	1	40.5	13	2	35.5	18	3	33.0	23	4	32.0
301-400	4	0	43.5	8	1	40.5	13	2	35.5	13	2	35.5	24	4	30.0	29	5	30.0
401-500	4	0	43.5	8	1	40.5	13	2	36.0	19	3	31.5	24	4	30.0	30	5	29.5
501-600	4	0	43.5	8	1	40.5	13	2	36.0	19	3	31.5	24	4	30.5	36	6	28.5
601-800	4	0	43.5	8	1	40.5	13	2	36.0	19	3	31.5	31	5	29.5	42	7	27.5
801-1000	4	0	44.0	8	1	40.5	14	2	33.5	25	4	30.0	37	6	28.0	49	8	26.5
1001-2000	8	1	40.5	14	2	33.5	19	3	32.0	31	5	30.0	44	7	26.5	65	10	23.5
2001-3000	8	1	40.5	14	2	33.5	19	3	32.0	31	5	30.0	50	8	26.0	85	13	22.5
3001-4000	8	1	40.5	14	2	33.5	25	4	30.0	38	6	27.5	65	10	24.0	100	15	21.5
4001-5000	8	1	40.5	14	2	33.5	25	4	30.0	38	6	27.5	65	10	24.0	120	18	20.5
5001-7000	8	1	40.5	14	2	33.5	25	4	30.0	44	7	27.0	80	12	22.5	135	20	19.8
7001-10,000	8	1	40.5	14	2	33.5	32	5	29.0	50	8	26.0	85	13	22.5	160	23	19.2
10,001-20,000	8	1	40.5	19	3	32.0	32	5	29.0	60	9	24.5	110	16	21.0	190	27	18.3
20,001-50,000	8	1	40.5	19	3	32.0	38	6	27.5	70	11	23.0	130	19	19.7	225	31	17.5
50,001-100,000	14	2	33.5	19	3	32.0	44	7	27.0	80	12	22.5	155	22	19.0	260	35	16.9

n = Size of Sample; entry of "All" indicates that each piece in lot is to be inspected.
c = Allowable Defect Number for Sample.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE IV: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-0.1

Average Outgoing Quality Limit = 0.1%

Process Average %	0–.002						.003–.020						.021–.040					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–75	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
76–95	75	0	–	–	–	1.5	75	0	–	–	–	1.5	75	0	–	–	–	1.5
96–130	95	0	–	–	–	1.4	95	0	–	–	–	1.4	95	0	–	–	–	1.4
131–200	130	0	–	–	–	1.2	130	0	–	–	–	1.2	130	0	–	–	–	1.2
201–300	165	0	–	–	–	1.1	165	0	–	–	–	1.1	165	0	–	–	–	1.1
301–350	190	0	–	–	–	.96	190	0	–	–	–	.96	190	0	–	–	–	.96
351–400	225	0	95	320	1	.86	225	0	95	320	1	.86	225	0	95	320	1	.86
401–500	250	0	120	370	1	.80	250	0	120	370	1	.80	250	0	120	370	1	.80
501–600	275	0	130	405	1	.77	275	0	130	405	1	.77	275	0	130	405	1	.77
601–800	310	0	155	465	1	.71	310	0	155	465	1	.71	310	0	155	465	1	.71
801–1000	350	0	185	535	1	.66	350	0	185	535	1	.66	350	0	185	535	1	.66
1001–2000	430	0	240	670	1	.58	430	0	240	670	1	.58	430	0	240	670	1	.58
2001–3000	465	0	265	730	1	.56	465	0	265	730	1	.56	465	0	265	730	1	.56
3001–4000	495	0	285	780	1	.54	495	0	285	780	1	.54	540	0	570	1110	2	.49
4001–5000	505	0	295	800	1	.53	505	0	295	800	1	.53	555	0	615	1170	2	.48
5001–7000	520	0	320	840	1	.52	520	0	320	840	1	.52	590	0	660	1250	2	.46
7001–10,000	540	0	335	875	1	.51	625	0	715	1340	2	.44	625	0	715	1340	2	.44
10,001–20,000	555	0	345	900	1	.50	650	0	750	1400	2	.43	720	0	1150	1870	3	.38
20,001–50,000	660	0	760	1420	2	.42	660	0	760	1420	2	.42	770	0	1700	2470	4	.36
50,001–100,000	670	0	770	1440	2	.42	740	0	1230	1970	3	.38	805	0	1725	2530	4	.35

Process Average %	.041–.060						.061–.080						.081–.100					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–75	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
76–95	75	0	–	–	–	1.5	75	0	–	–	–	1.5	75	0	–	–	–	1.5
96–130	95	0	–	–	–	1.4	95	0	–	–	–	1.4	95	0	–	–	–	1.4
131–200	130	0	–	–	–	1.2	130	0	–	–	–	1.2	130	0	–	–	–	1.2
201–300	165	0	–	–	–	1.1	165	0	–	–	–	1.1	165	0	–	–	–	1.1
301–350	190	0	–	–	–	.96	190	0	–	–	–	.96	190	0	–	–	–	.96
351–400	225	0	95	320	1	.86	225	0	95	320	1	.86	225	0	95	320	1	.86
401–500	250	0	120	370	1	.80	250	0	120	370	1	.80	250	0	120	370	1	.80
501–600	275	0	130	405	1	.77	275	0	130	405	1	.77	275	0	130	405	1	.77
601–800	310	0	155	465	1	.71	310	0	155	465	1	.71	310	0	155	465	1	.71
801–1000	350	0	185	535	1	.66	350	0	185	535	1	.66	350	0	185	535	1	.66
1001–2000	430	0	240	670	1	.58	430	0	240	670	1	.58	475	0	450	925	2	.54
2001–3000	520	0	530	1050	2	.50	520	0	530	1050	2	.50	520	0	530	1050	2	.50
3001–4000	540	0	570	1110	2	.49	540	0	570	1110	2	.49	585	0	885	1470	3	.45
4001–5000	555	0	615	1170	2	.48	605	0	935	1540	3	.44	605	0	935	1540	3	.44
5001–7000	655	0	1045	1700	3	.41	655	0	1045	1700	3	.41	680	0	1360	2040	4	.40
7001–10,000	700	0	1120	1820	3	.39	700	0	1120	1820	3	.39	1250	1	1780	3030	6	.35
10,001–20,000	740	0	1530	2270	4	.37	1350	1	2020	3370	6	.33	1400	1	2420	3820	7	.32
20,001–50,000	1400	1	2170	3570	6	.32	1450	1	3030	4480	8	.31	2230	2	4650	6880	12	.27
50,001–100,000	1460	1	3060	4520	8	.31	2330	2	4910	7240	12	.26	3690	4	7580	11270	19	.24

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE IV CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-0.25

Average Outgoing Quality Limit = 0.25%

Process Average %	0–.005						.006–.050					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–60	All	0	–	–	–	–	All	0	–	–	–	–
61–100	60	0	–	–	–	2.5	60	0	–	–	–	2.5
101–200	85	0	–	–	–	2.1	85	0	–	–	–	2.1
201–300	120	0	65	185	1	1.8	120	0	65	185	1	1.8
301–400	135	0	70	205	1	1.7	135	0	70	205	1	1.7
401–500	145	0	80	225	1	1.6	145	0	80	225	1	1.6
501–600	160	0	90	250	1	1.5	160	0	90	250	1	1.5
601–800	165	0	95	260	1	1.5	165	0	95	260	1	1.5
801–1000	180	0	105	285	1	1.4	180	0	105	285	1	1.4
1001–2000	205	0	120	325	1	1.3	205	0	120	325	1	1.3
2001–3000	210	0	125	335	1	1.3	210	0	125	335	1	1.3
3001–4000	210	0	130	340	1	1.3	210	0	130	340	1	1.3
4001–5000	215	0	130	345	1	1.3	245	0	280	525	2	1.1
5001–7000	215	0	135	350	1	1.3	250	0	285	535	2	1.1
7001–10,000	255	0	290	545	2	1.1	255	0	290	545	2	1.1
10,001–20,000	260	0	295	555	2	1.1	260	0	295	555	2	1.1
20,001–50,000	265	0	300	565	2	1.0	305	0	505	810	3	.91
50,001–100,000	270	0	305	575	2	1.0	310	0	510	820	3	.91

Process Average %	.051–.100						.101–.150					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–60	All	0	–	–	–	–	All	0	–	–	–	–
61–100	60	0	–	–	–	2.5	60	0	–	–	–	2.5
101–200	85	0	–	–	–	2.1	85	0	–	–	–	2.1
201–300	120	0	65	185	1	1.8	120	0	65	185	1	1.8
301–400	135	0	70	205	1	1.7	135	0	70	205	1	1.7
401–500	145	0	80	225	1	1.6	145	0	80	225	1	1.6
501–600	160	0	90	250	1	1.5	160	0	90	250	1	1.5
601–800	165	0	95	260	1	1.5	165	0	95	260	1	1.5
801–1000	180	0	105	285	1	1.4	200	0	195	395	2	1.3
1001–2000	220	0	245	465	2	1.2	220	0	245	465	2	1.2
2001–3000	235	0	275	510	2	1.1	260	0	435	695	3	1.0
3001–4000	240	0	280	520	2	1.1	270	0	440	710	3	1.0
4001–5000	275	0	445	720	3	1.0	275	0	445	720	3	1.0
5001–7000	290	0	475	765	3	.95	315	0	660	975	4	.88
7001–10,000	295	0	490	785	3	.94	325	0	705	1030	4	.84
10,001–20,000	330	0	730	1060	4	.83	590	1	910	1500	6	.76
20,001–50,000	335	0	735	1070	4	.83	645	1	1355	2000	8	.70
50,001–100,000	350	0	930	1280	5	.80	665	1	1615	2280	9	.68

Process Average %	.151–.200						.201–.250					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–60	All	0	–	–	–	–	All	0	–	–	–	–
61–100	60	0	–	–	–	2.5	60	0	–	–	–	2.5
101–200	85	0	–	–	–	2.1	85	0	–	–	–	2.1
201–300	120	0	65	185	1	1.8	120	0	65	185	1	1.8
301–400	135	0	70	205	1	1.7	135	0	70	205	1	1.7
401–500	145	0	80	225	1	1.6	145	0	80	225	1	1.6
501–600	160	0	90	250	1	1.5	175	0	160	335	2	1.4
601–800	195	0	185	380	2	1.3	195	0	185	380	2	1.3
801–1000	200	0	195	395	2	1.3	200	0	195	395	2	1.3
1001–2000	220	0	245	465	2	1.2	240	0	375	615	3	1.1
2001–3000	260	0	435	695	3	1.0	280	0	570	850	4	.96
3001–4000	290	0	600	890	4	.94	290	0	600	890	4	.94
4001–5000	300	0	615	915	4	.92	520	1	750	1270	6	.85
5001–7000	545	1	795	1340	6	.82	555	1	965	1520	7	.80
7001–10,000	585	1	1045	1630	7	.76	620	1	1420	2040	9	.72
10,001–20,000	660	1	1530	2190	9	.68	1220	3	2000	3220	13	.61
20,001–50,000	990	2	2340	3330	13	.61	1850	5	3500	5350	21	.55
50,001–100,000	1340	3	2910	4250	16	.56	2930	8	6090	9020	33	.48

TABLE DA-0.5

Average Outgoing Quality Limit = 0.5%

Process Average %	0–.010						.011–.10						.11–.20					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–30	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
31–50	30	0	–	–	–	5.0	30	0	–	–	–	5.0	30	0	–	–	–	5.0
51–75	40	0	–	–	–	4.6	40	0	–	–	–	4.6	40	0	–	–	–	4.6
76–100	47	0	23	70	1	4.4	47	0	23	70	1	4.4	47	0	23	70	1	4.4
101–150	60	0	30	90	1	3.8	60	0	30	90	1	3.8	60	0	30	90	1	3.8
151–200	70	0	35	105	1	3.3	70	0	35	105	1	3.3	70	0	35	105	1	3.3
201–300	80	0	45	125	1	3.0	80	0	45	125	1	3.0	80	0	45	125	1	3.0
301–400	85	0	50	135	1	2.9	85	0	50	135	1	2.9	85	0	50	135	1	2.9
401–500	90	0	55	145	1	2.8	90	0	55	145	1	2.8	90	0	55	145	1	2.8
501–600	95	0	55	150	1	2.8	95	0	55	150	1	2.8	95	0	55	150	1	2.8
601–800	100	0	55	155	1	2.7	100	0	55	155	1	2.7	110	0	115	225	2	2.4
801–1000	100	0	60	160	1	2.7	100	0	60	160	1	2.7	115	0	125	240	2	2.3
1001–2000	105	0	60	165	1	2.6	125	0	135	260	2	2.2	125	0	135	260	2	2.2
2001–3000	110	0	60	170	1	2.6	130	0	145	275	2	2.1	145	0	235	380	3	1.9
3001–4000	110	0	65	175	1	2.5	130	0	155	285	2	2.1	145	0	240	385	3	1.9
4001–5000	135	0	150	285	2	2.1	135	0	150	285	2	2.1	150	0	240	390	3	1.9
5001–7000	135	0	155	290	2	2.1	135	0	155	290	2	2.1	155	0	245	400	3	1.8
7001–10,000	135	0	160	295	2	2.1	135	0	160	295	2	2.1	175	0	375	550	4	1.6
10,001–20,000	140	0	160	300	2	2.0	155	0	250	405	3	1.8	185	0	500	685	5	1.5
20,001–50,000	140	0	165	305	2	2.0	155	0	255	410	3	1.8	185	0	505	690	5	1.5
50,001–100,000	140	0	170	310	2	2.0	155	0	260	415	3	1.8	325	1	495	820	6	1.4

Process Average %	.21–.30						.31–.40						.41–.50					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–30	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
31–50	30	0	–	–	–	5.0	30	0	–	–	–	5.0	30	0	–	–	–	5.0
51–75	40	0	–	–	–	4.6	40	0	–	–	–	4.6	40	0	–	–	–	4.6
76–100	47	0	23	70	1	4.4	47	0	23	70	1	4.4	47	0	23	70	1	4.4
101–150	60	0	30	90	1	3.8	60	0	30	90	1	3.8	60	0	30	90	1	3.8
151–200	70	0	35	105	1	3.3	70	0	35	105	1	3.3	70	0	35	105	1	3.3
201–300	80	0	45	125	1	3.0	80	0	45	125	1	3.0	80	0	45	125	1	3.0
301–400	85	0	50	135	1	2.9	95	0	90	185	2	2.7	95	0	90	185	2	2.7
401–500	100	0	100	200	2	2.6	100	0	100	200	2	2.6	100	0	100	200	2	2.6
501–600	105	0	105	210	2	2.5	105	0	105	210	2	2.5	105	0	105	210	2	2.5
601–800	110	0	115	225	2	2.4	110	0	115	225	2	2.4	120	0	180	300	3	2.2
801–1000	115	0	125	240	2	2.3	125	0	185	310	3	2.2	125	0	185	310	3	2.2
1001–2000	135	0	220	355	3	2.0	135	0	220	355	3	2.0	145	0	295	440	4	1.9
2001–3000	145	0	235	380	3	1.9	150	0	320	470	4	1.8	275	1	475	750	7	1.6
3001–4000	155	0	325	480	4	1.8	280	1	415	695	6	1.6	295	1	600	895	8	1.5
4001–5000	165	0	345	510	4	1.7	300	1	525	825	7	1.5	430	2	700	1130	10	1.4
5001–7000	175	0	455	630	5	1.6	310	1	670	980	8	1.4	460	2	860	1320	11	1.3
7001–10,000	300	1	460	760	6	1.5	465	2	785	1250	10	1.3	620	3	1120	1740	14	1.2
10,001–20,000	320	1	680	1000	8	1.4	495	2	1175	1670	13	1.2	740	4	1420	2160	18	1.2
20,001–50,000	350	1	930	1280	10	1.3	680	3	1490	2170	16	1.1	925	5	2085	3010	24	1.1
50,001–100,000	505	2	1075	1580	12	1.2	680	3	1810	2490	19	1.1	1550	9	3410	4960	38	.99

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10

TABLE IV CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-0.75

AVERAGE OUTGOING QUALITY LIMIT = 0.75%

Process Average %	0–.015						.016–.15						.16–.30					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–25	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
26–50	25	0	–	–	–	6.4	25	0	–	–	–	6.4	25	0	–	–	–	6.4
51–75	35	0	15	50	1	6.0	35	0	15	50	1	6.0	35	0	15	50	1	6.0
76–100	39	0	21	60	1	5.6	39	0	21	60	1	5.6	39	0	21	60	1	5.6
101–200	55	0	30	85	1	4.5	55	0	30	85	1	4.5	55	0	30	85	1	4.5
201–300	60	0	30	90	1	4.3	60	0	30	90	1	4.3	60	0	30	90	1	4.3
301–400	60	0	35	95	1	4.2	60	0	35	95	1	4.2	60	0	35	95	1	4.2
401–500	65	0	35	100	1	4.1	65	0	35	100	1	4.1	70	0	75	145	2	3.7
501–600	65	0	40	105	1	4.0	65	0	40	105	1	4.0	75	0	80	155	2	3.6
601–800	70	0	40	110	1	3.9	70	0	40	110	1	3.9	75	0	85	160	2	3.5
801–1000	70	0	40	110	1	3.9	70	0	40	110	1	3.9	80	0	90	170	2	3.4
1001–2000	75	0	40	115	1	3.8	85	0	90	175	2	3.3	95	0	155	250	3	2.9
2001–3000	75	0	40	115	1	3.8	85	0	100	185	2	3.2	100	0	160	260	3	2.8
3001–4000	90	0	105	195	2	3.1	90	0	105	195	2	3.1	100	0	160	260	3	2.8
4001–5000	90	0	105	195	2	3.1	90	0	105	195	2	3.1	105	0	235	340	4	2.6
5001–7000	90	0	105	195	2	3.1	90	0	105	195	2	3.1	110	0	245	355	4	2.5
7001–10,000	90	0	110	200	2	3.0	100	0	165	265	3	2.8	110	0	245	355	4	2.5
10,001–20,000	90	0	110	200	2	3.0	100	0	165	265	3	2.8	120	0	325	445	5	2.3
20.001–50,000	90	0	110	200	2	3.0	105	0	165	270	3	2.7	130	0	335	465	5	2.2
50,001–100,000	95	0	110	205	2	3.0	110	0	250	360	4	2.5	215	1	395	610	7	2.1

Process Average %	.31–.45						.46–.60						.61–.75					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–25	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
26–50	25	0	–	–	–	6.4	25	0	–	–	–	6.4	25	0	–	–	–	6.4
51–75	35	0	15	50	1	6.0	35	0	15	50	1	6.0	35	0	15	50	1	6.0
76–100	39	0	21	60	1	5.6	39	0	21	60	1	5.6	39	0	21	60	1	5.6
101–200	55	0	30	85	1	4.5	55	0	30	85	1	4.5	55	0	30	85	1	4.5
201–300	60	0	30	90	1	4.3	60	0	30	90	1	4.3	60	0	30	90	1	4.3
301–400	70	0	70	140	2	3.8	70	0	70	140	2	3.8	70	0	70	140	2	3.8
401–500	70	0	75	145	2	3.7	70	0	75	145	2	3.7	70	0	75	145	2	3.7
501–600	75	0	80	155	2	3.6	75	0	80	155	2	3.6	80	0	120	200	3	3.3
601–800	75	0	85	160	2	3.5	85	0	130	215	3	3.2	85	0	130	215	3	3.2
801–1000	85	0	140	225	3	3.1	85	0	140	225	3	3.1	95	0	185	280	4	2.9
1001–2000	95	0	155	250	3	2.9	100	0	215	315	4	2.7	175	1	265	440	6	2.5
2001–3000	105	0	230	335	4	2.6	185	1	280	465	6	2.4	205	1	405	610	8	2.2
3001–4000	110	0	290	400	5	2.5	195	1	345	540	7	2.3	285	2	475	760	10	2.1
4001–5000	195	1	300	495	6	2.3	210	1	415	625	8	2.2	300	2	560	860	11	2.0
5001–7000	200	1	375	575	7	2.2	300	2	510	810	10	2.0	390	3	710	1100	14	1.9
7001–10,000	205	1	375	580	7	2.2	320	2	665	985	12	1.9	415	3	955	1370	17	1.8
10,001–20,000	215	1	455	670	8	2.1	335	2	855	1190	14	1.8	600	5	1200	1800	22	1.7
20.001–50,000	225	1	605	830	10	2.0	440	3	1030	1470	17	1.7	800	7	1640	2440	29	1.6
50,001–100,000	335	2	725	1060	12	1.8	555	4	1305	1860	21	1.6	1025	9	2405	3430	40	1.5

TABLE DA-1

AVERAGE OUTGOING QUALITY LIMIT = 1.0%

Process Average %	0–.02						.03–.20						.21–.40					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–25	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
26–50	22	0	–	–	–	7.7	22	0	–	–	–	7.7	22	0	–	–	–	7.7
51–100	33	0	17	50	1	6.9	33	0	17	50	1	6.9	33	0	17	50	1	6.9
101–200	43	0	22	65	1	5.8	43	0	22	65	1	5.8	43	0	22	65	1	5.8
201–300	47	0	28	75	1	5.5	47	0	28	75	1	5.5	47	0	28	75	1	5.5
301–400	49	0	31	80	1	5.4	49	0	31	80	1	5.4	55	0	60	115	2	4.8
401–500	50	0	30	80	1	5.4	50	0	30	80	1	5.4	55	0	65	120	2	4.7
501–600	50	0	30	80	1	5.4	50	0	30	80	1	5.4	60	0	65	125	2	4.6
601–800	50	0	35	85	1	5.3	60	0	70	130	2	4.5	60	0	70	130	2	4.5
801–1000	55	0	30	85	1	5.2	60	0	75	135	2	4.4	60	0	75	135	2	4.4
1001–2000	55	0	35	90	1	5.1	65	0	75	140	2	4.3	75	0	120	195	3	3.8
2001–3000	65	0	80	145	2	4.2	65	0	80	145	2	4.2	75	0	125	200	3	3.7
3001–4000	70	0	80	150	2	4.1	70	0	80	150	2	4.1	80	0	175	255	4	3.5
4001–5000	70	0	80	150	2	4.1	70	0	80	150	2	4.1	80	0	180	260	4	3.4
5001–7000	70	0	80	150	2	4.1	75	0	125	200	3	3.7	80	0	180	260	4	3.4
7001–10,000	70	0	80	150	2	4.1	80	0	125	205	3	3.6	85	0	180	265	4	3.3
10,001–20,000	70	0	80	150	2	4.1	80	0	130	210	3	3.6	90	0	230	320	5	3.2
20,001–50,000	75	0	80	155	2	4.0	80	0	135	215	3	3.6	95	0	300	395	6	2.9
50,001–100,000	75	0	80	155	2	4.0	85	0	180	265	4	3.3	170	1	380	550	8	2.6

Process Average %	.41–.60						.61–.80						.81–1.00					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–25	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
26–50	22	0	–	–	–	7.7	22	0	–	–	–	7.7	22	0	–	–	–	7.7
51–100	33	0	17	50	1	6.9	33	0	17	50	1	6.9	33	0	17	50	1	6.9
101–200	43	0	22	65	1	5.8	43	0	22	65	1	5.8	47	0	43	90	2	5.4
201–300	55	0	50	105	2	4.9	55	0	50	105	2	4.9	55	0	50	105	2	4.9
301–400	55	0	60	115	2	4.8	55	0	60	115	2	4.8	60	0	80	140	3	4.5
401–500	55	0	65	120	2	4.7	60	0	95	155	3	4.3	60	0	95	155	3	4.3
501–600	60	0	65	125	2	4.6	65	0	100	165	3	4.2	65	0	100	165	3	4.2
601–800	65	0	105	170	3	4.1	65	0	105	170	3	4.1	70	0	140	210	4	3.9
801–1000	65	0	110	175	3	4.0	70	0	150	220	4	3.8	125	1	180	305	6	3.5
1001–2000	80	0	165	245	4	3.7	135	1	200	335	6	3.3	140	1	245	385	7	3.2
2001–3000	80	0	170	250	4	3.6	150	1	265	415	7	3.0	215	2	355	570	10	2.8
3001–4000	85	0	220	305	5	3.3	160	1	330	490	8	2.8	225	2	455	680	12	2.7
4001–5000	145	1	225	370	6	3.1	225	2	375	600	10	2.7	240	2	595	835	14	2.5
5001–7000	155	1	285	440	7	2.9	235	2	440	675	11	2.6	310	3	665	975	16	2.4
7001–10,000	165	1	355	520	8	2.7	250	2	585	835	13	2.4	385	4	785	1170	19	2.3
10,001–20,000	175	1	415	590	9	2.6	325	3	655	980	15	2.3	520	6	980	1500	24	2.2
20,001–50,000	250	2	490	740	11	2.4	340	3	910	1250	19	2.2	610	7	1410	2020	32	2.1
50,001–100,000	275	2	700	975	14	2.2	420	4	1050	1470	22	2.1	770	9	1850	2620	41	2.0

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE IV CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-1.5

AVERAGE OUTGOING QUALITY LIMIT = 1.5%

Process Average %	0–.03						.04–.30					
Lot Size	Trial 1 n_1	Trial 1 c_1	Trial 2 n_2	Trial 2 n_1+n_2	Trial 2 c_2	p_t %	Trial 1 n_1	Trial 1 c_1	Trial 2 n_2	Trial 2 n_1+n_2	Trial 2 c_2	p_t %
1–15	All	0	–	–	–	–	All	0	–	–	–	–
16–50	16	0	–	–	–	11.6	16	0	–	–	–	11.6
51–75	23	0	11	34	1	10.5	23	0	11	34	1	10.5
76–100	26	0	14	40	1	9.4	26	0	14	40	1	9.4
101–200	31	0	18	49	1	8.4	31	0	18	49	1	8.4
201–300	33	0	22	55	1	8.0	33	0	22	55	1	8.0
301–400	34	0	21	55	1	7.9	34	0	21	55	1	7.9
401–500	35	0	20	55	1	7.8	35	0	20	55	1	7.8
501–600	35	0	20	55	1	7.8	40	0	45	85	2	6.8
601–800	35	0	20	55	1	7.8	41	0	49	90	2	6.7
801–1000	36	0	19	55	1	7.8	42	0	48	90	2	6.5
1001–2000	44	0	51	95	2	6.3	44	0	51	95	2	6.3
2001–3000	45	0	50	95	2	6.2	45	0	50	95	2	6.2
3001–4000	45	0	50	95	2	6.2	50	0	85	135	3	5.5
4001–5000	45	0	50	95	2	6.2	50	0	85	135	3	5.5
5001–7000	46	0	54	100	2	6.1	50	0	90	140	3	5.4
7001–10,000	46	0	54	100	2	6.1	50	0	90	140	3	5.4
10,001–20,000	46	0	54	100	2	6.1	50	0	90	140	3	5.4
20,001–50,000	47	0	53	100	2	6.1	55	0	125	180	4	5.0
50,001–100,000	47	0	53	100	2	6.1	55	0	130	185	4	4.9

Process Average %	.31–.60						.61–.90					
Lot Size	Trial 1 n_1	Trial 1 c_1	Trial 2 n_2	Trial 2 n_1+n_2	Trial 2 c_2	p_t %	Trial 1 n_1	Trial 1 c_1	Trial 2 n_2	Trial 2 n_1+n_2	Trial 2 c_2	p_t %
1–15	All	0	–	–	–	–	All	0	–	–	–	–
16–50	16	0	–	–	–	11.6	16	0	–	–	–	11.6
51–75	23	0	11	34	1	10.5	23	0	11	34	1	10.5
76–100	26	0	14	40	1	9.4	26	0	14	40	1	9.4
101–200	31	0	18	49	1	8.4	35	0	35	70	2	7.5
201–300	38	0	37	75	2	7.0	38	0	37	75	2	7.0
301–400	39	0	41	80	2	6.9	39	0	41	80	2	6.9
401–500	39	0	46	85	2	6.9	44	0	71	115	3	6.1
501–600	40	0	45	85	2	6.8	44	0	71	115	3	6.1
601–800	46	0	74	120	3	6.0	46	0	74	120	3	6.0
801–1000	47	0	78	125	3	5.9	47	0	78	125	3	5.9
1001–2000	49	0	81	130	3	5.7	55	0	105	160	4	5.4
2001–3000	55	0	110	165	4	5.3	100	1	145	245	6	4.6
3001–4000	55	0	115	170	4	5.2	105	1	190	295	7	4.3
4001–5000	55	0	120	175	4	5.1	105	1	190	295	7	4.3
5001–7000	60	0	155	215	5	4.7	110	1	225	335	8	4.2
7001–10,000	60	0	160	220	5	4.6	115	1	280	395	9	3.9
10,001–20,000	60	0	165	225	5	4.5	120	1	315	435	10	3.8
20,001–50,000	65	0	195	260	6	4.4	165	2	350	515	12	3.7
50,001–100,000	115	1	235	350	8	4.0	175	2	440	615	14	3.5

Process Average %	.91–1.20						1.21–1.50					
Lot Size	Trial 1 n_1	Trial 1 c_1	Trial 2 n_2	Trial 2 n_1+n_2	Trial 2 c_2	p_t %	Trial 1 n_1	Trial 1 c_1	Trial 2 n_2	Trial 2 n_1+n_2	Trial 2 c_2	p_t %
1–15	All	0	–	–	–	–	All	0	–	–	–	–
16–50	16	0	–	–	–	11.6	16	0	–	–	–	11.6
51–75	23	0	11	34	1	10.5	23	0	11	34	1	10.5
76–100	26	0	14	40	1	9.4	26	0	14	40	1	9.4
101–200	35	0	35	70	2	7.5	35	0	35	70	2	7.5
201–300	38	0	37	75	2	7.0	40	0	60	100	3	6.6
301–400	42	0	63	105	3	6.3	42	0	63	105	3	6.3
401–500	44	0	71	115	3	6.1	46	0	94	140	4	5.8
501–600	44	0	71	115	3	6.1	47	0	98	145	4	5.7
601–800	49	0	101	150	4	5.6	85	1	125	210	6	5.0
801–1000	50	0	105	155	4	5.5	90	1	125	215	6	4.9
1001–2000	95	1	175	270	7	4.6	100	1	230	330	9	4.4
2001–3000	110	1	255	365	9	4.1	155	2	345	500	13	3.9
3001–4000	160	2	300	460	11	3.8	205	3	405	610	15	3.6
4001–5000	165	2	340	505	12	3.7	250	4	480	730	18	3.5
5001–7000	165	2	375	540	13	3.7	310	5	610	920	22	3.3
7001–10,000	170	2	420	590	14	3.6	360	6	660	1020	24	3.2
10,001–20,000	210	3	420	630	15	3.6	415	7	835	1250	29	3.1
20,001–50,000	225	3	640	865	20	3.3	510	9	1130	1640	38	3.0
50,001–100,000	275	4	725	1000	23	3.2	570	10	1400	1970	45	2.9

TABLE DA-2

AVERAGE OUTGOING QUALITY LIMIT = 2.0%

Process Average %	0–.04						.05–.40					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–15	All	0	–	–	–	–	All	0	–	–	–	–
16–50	14	0	–	–	–	13.6	14	0	–	–	–	13.6
51–100	21	0	12	33	1	11.7	21	0	12	33	1	11.7
101–200	24	0	13	37	1	11.0	24	0	13	37	1	11.0
201–300	26	0	15	41	1	10.4	26	0	15	41	1	10.4
301–400	26	0	16	42	1	10.3	26	0	16	42	1	10.3
401–500	27	0	16	43	1	10.3	30	0	35	65	2	9.0
501–600	27	0	16	43	1	10.3	31	0	34	65	2	8.9
601–800	27	0	17	44	1	10.2	31	0	39	70	2	8.8
801–1000	27	0	17	44	1	10.2	32	0	38	70	2	8.7
1001–2000	33	0	37	70	2	8.5	33	0	37	70	2	8.5
2001–3000	34	0	41	75	2	8.2	34	0	41	75	2	8.2
3001–4000	34	0	41	75	2	8.2	38	0	62	100	3	7.3
4001–5000	34	0	41	75	2	8.2	38	0	62	100	3	7.3
5001–7000	35	0	40	75	2	8.1	38	0	62	100	3	7.3
7001–10,000	35	0	40	75	2	8.1	38	0	62	100	3	7.3
10,001–20,000	35	0	40	75	2	8.1	39	0	66	105	3	7.2
20,001–50,000	35	0	40	75	2	8.1	43	0	92	135	4	6.6
50,001–100,000	35	0	45	80	2	8.0	43	0	92	135	4	6.6

Process Average %	.41–.80						.81–1.20					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–15	All	0	–	–	–	–	All	0	–	–	–	–
16–50	14	0	–	–	–	13.6	14	0	–	–	–	13.6
51–100	21	0	12	33	1	11.7	21	0	12	33	1	11.7
101–200	24	0	13	37	1	11.0	27	0	28	55	2	9.6
201–300	29	0	31	60	2	9.1	29	0	31	60	2	9.1
301–400	30	0	35	65	2	9.0	33	0	52	85	3	8.2
401–500	30	0	35	65	2	9.0	34	0	56	90	3	7.9
501–600	35	0	55	90	3	7.9	35	0	55	90	3	7.9
601–800	35	0	60	95	3	7.7	38	0	82	120	4	7.3
801–1000	36	0	59	95	3	7.6	38	0	87	125	4	7.2
1001–2000	37	0	63	100	3	7.5	43	0	112	155	5	6.5
2001–3000	41	0	84	125	4	7.0	75	1	115	190	6	6.1
3001–4000	41	0	89	130	4	6.9	80	1	140	220	7	5.8
4001–5000	42	0	88	130	4	6.9	80	1	175	255	8	5.5
5001–7000	44	0	116	160	5	6.4	85	1	205	290	9	5.3
7001–10,000	45	0	115	160	5	6.3	85	1	210	295	9	5.2
10,001–20,000	45	0	115	160	5	6.3	90	1	260	350	11	5.1
20,001–50,000	47	0	148	195	6	6.0	130	2	300	430	13	4.7
50,001–100,000	85	1	185	270	8	5.2	135	2	345	480	14	4.5

Process Average %	1.21–1.60						1.61–2.00					
	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–15	All	0	–	–	–	–	All	0	–	–	–	–
16–50	14	0	–	–	–	13.6	14	0	–	–	–	13.6
51–100	21	0	12	33	1	11.7	23	0	23	46	2	10.9
101–200	27	0	28	55	2	9.6	27	0	28	55	2	9.6
201–300	32	0	48	80	3	8.4	32	0	48	80	3	8.4
301–400	33	0	52	85	3	8.2	36	0	69	105	4	7.6
401–500	36	0	74	110	4	7.5	60	1	90	150	6	7.0
501–600	37	0	78	115	4	7.4	65	1	95	160	6	6.8
601–800	38	0	82	120	4	7.3	70	1	120	190	7	6.4
801–1000	70	1	100	170	6	6.5	70	1	145	215	8	6.2
1001–2000	80	1	160	240	8	5.8	110	2	205	315	11	5.5
2001–3000	115	2	195	310	10	5.3	160	3	310	470	15	4.7
3001–4000	120	2	255	375	12	5.0	235	5	415	650	20	4.3
4001–5000	125	2	285	410	13	4.9	275	6	475	750	23	4.2
5001–7000	125	2	320	445	14	4.8	280	6	575	855	26	4.1
7001–10,000	165	3	335	500	15	4.5	320	7	645	965	29	4.0
10,001–20,000	170	3	425	595	18	4.4	395	9	835	1230	37	3.9
20,001–50,000	205	4	515	720	22	4.3	480	11	1090	1570	46	3.7
50,001–100,000	250	5	615	865	26	4.1	580	13	1460	2040	58	3.5

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE IV CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-2.5

AVERAGE OUTGOING QUALITY LIMIT = 2.5%

Process Average %	0–.05						.06–.50						.51–1.00					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–10	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
11–50	11	0	–	–	–	17.6	11	0	–	–	–	17.6	11	0	–	–	–	17.6
51–100	18	0	10	28	1	14.1	18	0	10	28	1	14.1	18	0	10	28	1	14.1
101–200	20	0	11	31	1	13.7	20	0	11	31	1	13.7	23	0	25	48	2	11.7
201–300	21	0	13	34	1	13.0	21	0	13	34	1	13.0	24	0	25	49	2	11.4
301–400	21	0	14	35	1	12.8	24	0	26	50	2	11.3	24	0	26	50	2	11.3
401–500	22	0	13	35	1	12.7	25	0	25	50	2	11.1	28	0	47	75	3	9.8
501–600	22	0	14	36	1	12.5	25	0	30	55	2	10.9	28	0	47	75	3	9.8
601–800	22	0	14	36	1	12.5	26	0	29	55	2	10.8	28	0	47	75	3	9.8
801–1000	26	0	29	55	2	10.8	26	0	29	55	2	10.8	29	0	46	75	3	9.6
1001–2000	27	0	33	60	2	10.5	27	0	33	60	2	10.5	33	0	72	105	4	8.3
2001–3000	27	0	33	60	2	10.5	30	0	50	80	3	9.3	33	0	72	105	4	8.3
3001–4000	27	0	33	60	2	10.5	31	0	49	80	3	9.1	33	0	77	110	4	8.2
4001–5000	27	0	33	60	2	10.5	31	0	49	80	3	9.1	36	0	94	130	5	7.6
5001–7000	28	0	32	60	2	10.3	31	0	49	80	3	9.1	36	0	94	130	5	7.7
7001–10,000	28	0	32	60	2	10.3	31	0	49	80	3	9.2	36	0	94	130	5	7.7
10,001–20,000	28	0	32	60	2	10.3	31	0	49	80	3	9.2	36	0	94	130	5	7.8
20,001–50,000	28	0	32	60	2	10.3	33	0	87	120	4	7.7	70	1	145	215	8	6.6
50,001–100,000	28	0	37	65	2	10.2	33	0	92	125	4	7.6	70	1	170	240	9	6.4

Process Average %	1.01–1.50						1.51–2.00						2.01–2.50					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–10	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
11–50	11	0	–	–	–	17.6	11	0	–	–	–	17.6	11	0	–	–	–	17.6
51–100	18	0	10	28	1	14.1	20	0	20	40	2	13.0	20	0	20	40	2	13.0
101–200	23	0	25	48	2	11.7	23	0	25	48	2	11.7	25	0	35	60	3	10.8
201–300	26	0	44	70	3	10.3	26	0	44	70	3	10.3	28	0	57	85	4	9.5
301–400	27	0	43	70	3	9.9	29	0	61	90	4	9.3	49	1	71	120	6	8.8
401–500	28	0	47	75	3	9.8	30	0	60	90	4	9.2	50	1	80	130	6	8.4
501–600	30	0	65	95	4	9.1	30	0	65	95	4	9.1	55	1	95	150	7	8.0
601–800	31	0	69	100	4	8.8	55	1	85	140	6	8.0	60	1	115	175	8	7.6
801–1000	32	0	68	100	4	8.7	60	1	100	160	7	7.8	85	2	120	205	9	7.2
1001–2000	60	1	90	150	6	7.6	65	1	150	215	9	7.0	95	2	210	305	13	6.5
2001–3000	65	1	115	180	7	7.2	90	2	170	260	11	6.8	125	3	265	390	16	6.0
3001–4000	65	1	140	205	8	6.8	95	2	205	300	12	6.4	185	5	350	535	21	5.5
4001–5000	70	1	160	230	9	6.5	100	2	255	355	14	6.0	220	6	410	630	24	5.2
5001–7000	75	1	190	265	10	6.2	130	3	265	395	15	5.7	255	7	495	750	28	5.0
7001–10,000	100	2	195	295	11	6.0	140	3	355	495	18	5.3	325	9	665	990	36	4.7
10,001–20,000	105	2	215	320	12	5.9	170	4	380	550	20	5.2	360	10	830	1190	43	4.6
20,001–50,000	105	2	245	350	13	5.8	205	5	485	690	25	5.0	415	11	1145	1560	54	4.3
50,001–100,000	110	2	295	405	15	5.6	245	6	610	855	30	4.7	510	14	1370	1880	65	4.2

TABLE DA-3

AVERAGE OUTGOING QUALITY LIMIT = 3.0%

Process Average %	0–.06						.07–.60						.61–1.20					
Lot Size	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–10	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
11–50	10	0	–	–	–	19.0	10	0	–	–	–	19.0	10	0	–	–	–	19.0
51–100	16	0	9	25	1	16.4	16	0	9	25	1	16.4	16	0	9	25	1	16.4
101–200	17	0	9	26	1	16.0	17	0	9	26	1	16.0	17	0	9	26	1	16.0
201–300	18	0	10	28	1	15.5	18	0	10	28	1	15.5	21	0	23	44	2	13.3
301–400	18	0	11	29	1	15.2	21	0	24	45	2	13.2	23	0	37	60	3	12.0
401–500	18	0	11	29	1	15.2	21	0	25	46	2	13.0	24	0	36	60	3	11.7
501–600	18	0	12	30	1	15.0	21	0	25	46	2	13.0	24	0	41	65	3	11.5
601–800	21	0	25	46	2	13.0	21	0	25	46	2	13.0	24	0	41	65	3	11.5
801–1000	21	0	26	47	2	12.8	21	0	26	47	2	12.8	25	0	40	65	3	11.4
1001–2000	22	0	26	48	2	12.6	22	0	26	48	2	12.6	27	0	58	85	4	10.3
2001–3000	22	0	26	48	2	12.6	25	0	40	65	3	11.4	28	0	62	90	4	10.0
3001–4000	23	0	26	49	2	12.4	25	0	45	70	3	11.0	29	0	76	105	5	9.6
4001–5000	23	0	26	49	2	12.4	26	0	44	70	3	11.0	30	0	75	105	5	9.5
5001–7000	23	0	27	50	2	12.2	26	0	44	70	3	11.0	30	0	80	110	5	9.4
7001–10,000	23	0	27	50	2	12.2	27	0	43	70	3	11.0	30	0	80	110	5	9.4
10,001–20,000	23	0	27	50	2	12.2	27	0	43	70	3	11.0	31	0	94	125	6	9.2
20,001–50,000	23	0	27	50	2	12.2	28	0	67	95	4	9.7	55	1	120	175	8	8.0
50,001–100,000	23	0	27	50	2	12.2	31	0	84	115	5	9.0	60	1	140	200	9	7.6

Process Average %	1.21–1.80						1.81–2.40						2.41–3.00					
Lot Size	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %	Trial 1		Trial 2			p_t %
	n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2		n_1	c_1	n_2	n_1+n_2	c_2	
1–10	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
11–50	10	0	–	–	–	19.0	10	0	–	–	–	19.0	10	0	–	–	–	19.0
51–100	17	0	17	34	2	15.8	17	0	17	34	2	15.8	17	0	17	34	2	15.8
101–200	20	0	21	41	2	13.7	22	0	33	55	3	12.4	22	0	33	55	3	12.4
201–300	23	0	37	60	3	12.0	23	0	37	60	3	12.0	24	0	51	75	4	11.1
301–400	23	0	37	60	3	12.0	25	0	55	80	4	10.8	42	1	63	105	6	10.4
401–500	24	0	36	60	3	11.7	25	0	55	80	4	10.8	46	1	79	125	7	9.7
501–600	26	0	54	80	4	10.7	46	1	69	115	6	9.7	48	1	97	145	8	9.2
601–800	26	0	54	80	4	10.7	49	1	81	130	7	9.4	50	1	115	165	9	8.9
801–1000	27	0	58	85	4	10.3	49	1	86	135	7	9.2	70	2	120	190	10	8.4
1001–2000	49	1	76	125	6	9.1	50	1	150	200	10	8.0	100	3	180	280	14	7.5
2001–3000	50	1	95	145	7	8.7	80	2	165	245	12	7.6	130	4	260	390	19	6.9
3001–4000	55	1	110	165	8	8.5	105	3	200	305	14	7.0	155	5	330	485	23	6.5
4001–5000	60	1	135	195	9	7.8	110	3	225	335	15	6.7	215	7	390	605	27	6.0
5001–7000	60	1	165	225	10	7.3	110	3	250	360	16	6.6	270	9	505	775	34	5.7
7001–10,000	85	2	160	245	11	7.2	115	3	290	405	18	6.5	285	9	680	965	41	5.4
10,001–20,000	85	2	180	265	12	7.2	140	4	315	455	20	6.3	315	10	805	1120	47	5.3
20,001–50,000	85	2	205	290	13	7.0	170	5	420	590	26	6.0	390	13	940	1330	56	5.2
50,001–100,000	90	2	245	335	15	6.8	200	6	505	705	30	5.7	445	15	1105	1550	65	5.1

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

TABLE IV CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-4

Average Outgoing Quality Limit = 4.0%

Process Average %	0–.08						.09–.80						.81–1.60					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–10	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
11–50	8	0	–	–	–	23.0	8	0	–	–	–	23.0	8	0	–	–	–	23.0
51–100	12	0	7	19	1	22.0	12	0	7	19	1	22.0	12	0	7	19	1	22.0
101–200	13	0	8	21	1	21.0	13	0	8	21	1	21.0	15	0	17	32	2	18.0
201–300	13	0	9	22	1	20.5	16	0	18	34	2	17.4	16	0	18	34	2	17.4
301–400	14	0	8	22	1	20.0	16	0	19	35	2	17.0	18	0	28	46	3	15.5
401–500	14	0	8	22	1	20.0	16	0	19	35	2	17.0	19	0	28	47	3	15.3
501–600	16	0	19	35	2	17.0	16	0	19	35	2	17.0	19	0	29	48	3	15.1
601–800	16	0	20	36	2	16.7	16	0	20	36	2	16.7	19	0	30	49	3	14.9
801–1000	16	0	20	36	2	16.7	16	0	20	36	2	16.7	20	0	45	65	4	13.8
1001–2000	17	0	19	36	2	16.6	19	0	31	50	3	14.8	21	0	44	65	4	13.6
2001–3000	17	0	19	36	2	16.6	19	0	31	50	3	14.8	21	0	44	65	4	13.6
3001–4000	17	0	20	37	2	16.5	19	0	31	50	3	14.8	22	0	58	80	5	13.0
4001–5000	17	0	20	37	2	16.5	19	0	31	50	3	14.8	22	0	58	80	5	13.0
5001–7000	17	0	20	37	2	16.5	19	0	31	50	3	14.8	22	0	58	80	5	13.0
7001–10,000	17	0	20	37	2	16.5	19	0	36	55	3	14.6	23	0	57	80	5	12.7
10,001–20,000	17	0	20	37	2	16.5	21	0	44	65	4	13.6	23	0	72	95	6	12.0
20,001–50,000	17	0	20	37	2	16.5	21	0	44	65	4	13.6	43	1	92	135	8	10.6
50,001–100,000	17	0	20	37	2	16.5	23	0	62	85	5	12.5	44	1	106	150	9	10.3

Process Average %	1.61–2.40						2.41–3.20						3.21–4.00					
	Trial 1		Trial 2				Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–10	All	0	–	–	–	–	All	0	–	–	–	–	All	0	–	–	–	–
11–50	8	0	–	–	–	23.0	8	0	–	–	–	23.0	8	0	–	–	–	23.0
51–100	13	0	14	27	2	20.5	13	0	14	27	2	20.5	13	0	14	27	2	20.5
101–200	16	0	26	42	3	16.5	16	0	26	42	3	16.5	16	0	26	42	3	16.5
201–300	17	0	28	45	3	16.0	18	0	37	55	4	15.0	33	1	47	80	6	13.2
301–400	19	0	41	60	4	14.3	19	0	41	60	4	14.3	35	1	60	95	7	12.8
401–500	20	0	40	60	4	14.0	34	1	51	85	6	13.0	36	1	74	110	8	12.2
501–600	20	0	45	65	4	13.8	37	1	63	100	7	12.2	50	2	75	125	9	11.6
601–800	22	0	58	80	5	13.0	39	1	81	120	8	11.6	55	2	105	160	11	10.8
801–1000	37	1	58	95	6	12.2	41	1	94	135	9	11.1	55	2	120	175	12	10.5
1001–2000	39	1	71	110	7	11.5	55	2	110	165	11	10.6	80	3	165	245	16	9.5
2001–3000	41	1	89	130	8	11.0	60	2	145	205	13	9.8	95	4	210	305	20	9.2
3001–4000	43	1	102	145	9	10.5	80	3	160	240	15	9.4	115	5	250	365	23	8.8
4001–5000	45	1	120	165	10	10.0	85	3	180	265	16	8.9	160	7	305	465	28	8.1
5001–7000	65	2	120	185	11	9.6	85	3	200	285	17	8.7	210	9	450	660	38	7.4
7001–10,000	65	2	140	205	12	9.3	90	3	230	320	19	8.5	235	10	550	785	44	7.1
10,001–20,000	65	2	160	225	13	9.0	105	4	265	370	22	8.3	270	12	625	895	50	7.0
20,001–50,000	70	2	175	245	14	8.8	125	5	315	440	26	8.1	295	13	725	1020	57	6.9
50,001–100,000	70	2	205	275	16	8.7	150	6	385	535	31	7.7	335	15	845	1180	66	6.8

TABLE DA-5

AVERAGE OUTGOING QUALITY LIMIT = 5.0%

Process Average %	0–.10						.11–1.00					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–5	All	0	–	–	–	–	All	0	–	–	–	–
6–50	6	0	–	–	–	30.5	6	0	–	–	–	30.5
51–100	10	0	6	16	1	26.5	10	0	6	16	1	26.5
101–200	11	0	6	17	1	26.0	12	0	15	27	2	22.0
201–300	11	0	7	18	1	25.0	13	0	15	28	2	21.0
301–400	11	0	8	19	1	25.0	13	0	15	28	2	21.0
401–500	13	0	15	28	2	21.0	13	0	15	28	2	21.0
501–600	13	0	15	28	2	21.0	13	0	15	28	2	21.0
601–800	13	0	16	29	2	20.5	13	0	16	29	2	20.5
801–1000	13	0	16	29	2	20.5	13	0	16	29	2	20.5
1001–2000	13	0	16	29	2	20.5	15	0	25	40	3	18.7
2001–3000	13	0	16	29	2	21.0	15	0	26	41	3	18.4
3001–4000	14	0	15	29	2	21.0	15	0	26	41	3	18.4
4001–5000	14	0	16	30	2	20.5	16	0	25	41	3	18.0
5001–7000	14	0	16	30	2	20.5	16	0	26	42	3	18.0
7001–10,000	14	0	16	30	2	20.5	16	0	26	42	3	18.0
10,001–20,000	14	0	17	31	2	20.5	17	0	38	55	4	16.4
20,001–50,000	14	0	17	31	2	20.5	17	0	38	55	4	16.4
50,001–100,000	14	0	18	32	2	20.5	18	0	47	65	5	15.6

Process Average %	1.01–2.00						2.01–3.00					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–5	All	0	–	–	–	–	All	0	–	–	–	–
6–50	6	0	–	–	–	30.5	6	0	–	–	–	30.5
51–100	11	0	11	22	2	25.0	11	0	11	22	2	25.0
101–200	12	0	15	27	2	22.0	14	0	22	36	3	19.8
201–300	14	0	24	38	3	19.3	14	0	24	38	3	19.3
301–400	15	0	24	39	3	19.0	16	0	33	49	4	17.5
401–500	15	0	24	39	3	19.0	16	0	34	50	4	17.1
501–600	15	0	25	40	3	18.7	16	0	34	50	4	17.1
601–800	16	0	34	50	4	17.1	17	0	33	60	5	16.2
801–1000	16	0	34	50	4	17.1	30	1	45	75	6	15.0
1001–2000	17	0	33	50	4	17.1	31	1	59	90	7	14.5
2001–3000	17	0	48	65	5	15.5	32	1	68	100	8	14.0
3001–4000	18	0	47	65	5	15.5	34	1	81	115	9	13.5
4001–5000	18	0	47	65	5	15.5	35	1	95	130	10	13.0
5001–7000	18	0	47	65	5	15.5	50	2	90	140	11	12.5
7001–10,000	19	0	56	75	6	15.0	50	2	105	155	12	12.1
10,001–20,000	19	0	56	75	6	15.0	50	2	125	175	13	11.7
20,001–50,000	33	1	72	105	8	13.5	50	2	135	185	14	11.3
50,001–100,000	34	1	86	120	9	13.1	55	2	160	215	16	11.0

Process Average %	3.01–4.00						4.01–5.00					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–5	All	0	–	–	–	–	All	0	–	–	–	–
6–50	6	0	–	–	–	30.5	6	0	–	–	–	30.5
51–100	11	0	11	22	2	25.0	12	0	18	30	3	23.0
101–200	14	0	22	36	3	19.8	14	0	30	44	4	19.0
201–300	15	0	32	47	4	18.0	27	1	48	75	7	16.3
301–400	27	1	38	65	6	16.6	29	1	56	85	8	15.5
401–500	29	1	51	80	7	15.5	30	1	70	100	9	14.9
501–600	31	1	64	95	8	14.3	43	2	72	115	10	13.9
601–800	32	1	78	110	9	13.9	45	2	90	135	12	13.5
801–1000	45	2	75	120	10	13.3	60	3	110	170	14	12.4
1001–2000	50	2	100	150	12	12.7	75	4	160	235	19	11.5
2001–3000	50	2	130	180	14	12.0	95	5	185	280	22	11.0
3001–4000	65	3	135	200	15	11.3	95	5	255	350	27	10.5
4001–5000	70	3	155	225	17	11.0	130	7	260	390	29	10.0
5001–7000	70	3	185	255	19	10.7	160	9	355	515	38	9.5
7001–10,000	85	4	200	285	21	10.4	180	10	430	610	44	9.2
10,001–20,000	100	5	220	320	23	10.0	215	12	490	705	50	8.9
20,001–50,000	120	6	290	410	29	9.5	230	13	605	835	59	8.7
50,001–100,000	140	7	315	455	32	9.3	265	15	705	970	68	8.5

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P c) = 0.10.

TABLE IV CONT'D: DOUBLE SAMPLING LOT INSPECTION TABLES—BASED ON STATED VALUES OF "AVERAGE OUTGOING QUALITY LIMIT"

TABLE DA-7

Average Outgoing Quality Limit = 7.0%

Process Average %	0–.14						.15–1.40					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–5	All	0	–	–	–	–	All	0	–	–	–	–
6–25	5	0	–	–	–	35.5	5	0	–	–	–	35.5
26–50	7	0	4	11	1	38.5	7	0	4	11	1	38.5
51–100	7	0	5	12	1	36.5	7	0	5	12	1	36.5
101–200	8	0	5	13	1	36.0	9	0	10	19	2	30.5
201–300	8	0	6	14	1	35.5	9	0	11	20	2	30.0
301–400	9	0	11	20	2	30.0	9	0	11	20	2	30.0
401–500	9	0	11	20	2	30.0	9	0	11	20	2	30.0
501–600	9	0	12	21	2	29.5	9	0	12	21	2	29.5
601–800	9	0	12	21	2	29.5	9	0	12	21	2	29.5
801–1000	9	0	12	21	2	29.5	11	0	18	29	3	25.5
1001–2000	10	0	11	21	2	29.0	11	0	18	29	3	25.5
2001–3000	10	0	11	21	2	29.0	11	0	19	30	3	25.5
3001–4000	10	0	11	21	2	29.0	11	0	19	30	3	25.5
4001–5000	10	0	11	21	2	29.0	11	0	19	30	3	25.5
5001–7000	10	0	12	22	2	28.5	11	0	19	30	3	26.0
7001–10,000	10	0	12	22	2	28.5	12	0	26	38	4	23.5
10,001–20,000	10	0	12	22	2	28.5	12	0	27	39	4	23.0
20,001–50,000	10	0	12	22	2	28.5	13	0	36	49	5	21.0
50,001–100,000	10	0	12	22	2	28.5	13	0	36	49	5	21.0

Process Average %	1.41–2.80						2.81–4.20					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–5	All	0	–	–	–	–	All	0	–	–	–	–
6–25	5	0	–	–	–	35.5	5	0	–	–	–	35.5
26–50	7	0	4	11	1	38.5	7	0	4	11	1	38.5
51–100	9	0	9	18	2	31.5	9	0	9	18	2	31.5
101–200	10	0	15	25	3	27.5	10	0	15	25	3	27.5
201–300	10	0	17	27	3	27.0	11	0	24	35	4	25.0
301–400	11	0	17	28	3	26.0	11	0	25	36	4	24.5
401–500	11	0	17	28	3	26.0	12	0	32	44	5	23.0
501–600	12	0	25	37	4	24.0	21	1	34	55	6	21.0
601–800	12	0	25	37	4	24.0	22	1	43	65	7	20.0
801–1000	12	0	25	37	4	24.0	23	1	42	65	7	19.8
1001–2000	12	0	25	37	4	24.0	23	1	47	70	8	19.2
2001–3000	13	0	34	47	5	21.5	24	1	56	80	9	18.7
3001–4000	13	0	34	47	5	21.5	24	1	66	90	10	18.2
4001–5000	13	0	35	48	5	21.5	26	1	74	100	11	17.7
5001–7000	14	0	46	60	6	20.0	37	2	78	115	12	16.5
7001–10,000	25	1	50	75	8	18.5	38	2	87	125	13	16.2
10,001–20,000	25	1	55	80	8	18.0	38	2	97	135	14	16.0
20,001–50,000	25	1	60	85	9	17.8	38	2	107	145	15	15.8
50,001–100,000	26	1	69	95	10	17.5	39	2	111	150	16	15.7

Process Average %	4.21–5.60						5.61–7.00					
	Trial 1		Trial 2				Trial 1		Trial 2			
Lot Size	n_1	c_1	n_2	n_1+n_2	c_2	p_t %	n_1	c_1	n_2	n_1+n_2	c_2	p_t %
1–5	All	0	–	–	–	–	All	0	–	–	–	–
6–25	5	0	–	–	–	35.5	5	0	–	–	–	35.5
26–50	8	0	8	16	2	34.0	8	0	8	16	2	34.0
51–100	10	0	14	24	3	28.5	10	0	14	24	3	28.5
101–200	11	0	22	33	4	25.5	18	1	28	46	6	24.0
201–300	19	1	28	47	6	23.5	20	1	40	60	8	22.0
301–400	20	1	35	55	7	22.0	29	2	41	70	9	21.0
401–500	22	1	43	65	8	20.5	31	2	59	90	11	19.5
501–600	23	1	52	75	9	19.5	32	2	63	95	12	19.0
601–800	32	2	53	85	10	19.0	41	3	74	115	14	18.3
801–1000	33	2	62	95	11	18.4	43	3	82	125	15	17.5
1001–2000	34	2	81	115	13	17.8	55	4	115	170	20	16.5
2001–3000	44	3	91	135	15	17.0	65	5	135	200	23	16.0
3001–4000	47	3	103	150	16	16.0	75	6	175	250	28	15.0
4001–5000	50	3	135	185	19	15.0	100	8	225	325	35	14.0
5001–7000	65	4	155	220	22	14.0	115	9	265	380	40	13.5
7001–10,000	75	5	180	255	25	13.6	130	10	310	440	45	13.0
10,001–20,000	85	6	210	295	29	13.2	150	12	370	520	52	12.5
20,001–50,000	100	7	235	335	33	12.9	180	14	460	640	63	12.0
50,001–100,000	110	8	300	410	40	12.6	205	16	555	760	75	11.7

TABLE DA-10

Average Outgoing Quality Limit = 10.0%

Process Average % → Lot Size	0–.20 Trial 1 n_1	c_1	Trial 2 n_2	n_1+n_2	c_2	p_t %	.21–2.00 Trial 1 n_1	c_1	Trial 2 n_2	n_1+n_2	c_2	p_t %
1–3	All	0	–	–	–	–	All	0	–	–	–	–
4–15	3	0	–	–	–	50.0	3	0	–	–	–	50.0
16–50	5	0	3	8	1	53.5	5	0	3	8	1	53.5
51–100	5	0	3	8	1	55.0	6	0	8	14	2	43.0
101–200	5	0	4	9	1	52.0	7	0	7	14	2	42.0
201–300	7	0	7	14	2	42.5	7	0	7	14	2	42.5
301–400	7	0	7	14	2	42.5	7	0	7	14	2	42.5
401–500	7	0	8	15	2	40.0	7	0	8	15	2	40.0
501–600	7	0	8	15	2	40.0	8	0	13	21	3	35.0
601–800	7	0	8	15	2	40.5	8	0	13	21	3	35.0
801–1000	7	0	8	15	2	40.5	8	0	13	21	3	35.0
1001–2000	7	0	8	15	2	40.5	8	0	14	22	3	34.0
2001–3000	7	0	8	15	2	41.0	8	0	14	22	3	34.0
3001–4000	7	0	8	15	2	41.0	8	0	14	22	3	34.5
4001–5000	7	0	8	15	2	41.0	8	0	14	22	3	35.0
5001–7000	7	0	8	15	2	41.0	9	0	18	27	4	32.5
7001–10,000	7	0	8	15	2	41.0	9	0	18	27	4	32.5
10,001–20,000	7	0	8	15	2	41.0	9	0	18	27	4	32.5
20,001–50,000	7	0	8	15	2	41.0	9	0	18	27	4	32.5
50,001–100,000	8	0	14	22	3	33.5	9	0	25	34	5	30.0

Process Average % → Lot Size	2.01–4.00 Trial 1 n_1	c_1	Trial 2 n_2	n_1+n_2	c_2	p_t %	4.01–6.00 Trial 1 n_1	c_1	Trial 2 n_2	n_1+n_2	c_2	p_t %
1–3	All	0	–	–	–	–	All	0	–	–	–	–
4–15	3	0	–	–	–	50.0	3	0	–	–	–	50.0
16–50	5	0	3	8	1	53.5	6	0	6	12	2	48.0
51–100	6	0	8	14	2	43.0	7	0	11	18	3	38.5
101–200	7	0	12	19	3	38.0	8	0	16	24	4	35.5
201–300	7	0	13	20	3	37.0	8	0	17	25	4	35.0
301–400	8	0	17	25	4	35.0	8	0	22	30	5	34.0
401–500	8	0	18	26	4	34.0	15	1	23	38	6	30.5
501–600	8	0	18	26	4	34.0	16	1	28	44	7	28.5
601–800	8	0	18	26	4	34.5	16	1	28	44	7	29.0
801–1000	9	0	18	27	4	33.0	16	1	34	50	8	28.0
1001–2000	9	0	23	32	5	31.0	17	1	38	55	9	27.5
2001–3000	9	0	24	33	5	30.0	17	1	48	65	10	26.0
3001–4000	9	0	24	33	5	30.5	24	2	46	70	11	25.0
4001–5000	10	0	29	39	6	29.5	26	2	54	80	12	23.5
5001–7000	16	1	29	45	7	28.5	27	2	63	90	13	22.5
7001–10,000	17	1	38	55	8	26.0	27	2	68	95	14	22.0
10,001–20,000	17	1	38	55	8	26.0	28	2	77	105	15	22.0
20,001–50,000	18	1	42	60	9	25.5	28	2	87	115	17	21.5
50,001–100,000	18	1	52	70	10	24.5	36	3	99	135	20	21.0

Process Average % → Lot Size	6.01–8.00 Trial 1 n_1	c_1	Trial 2 n_2	n_1+n_2	c_2	p_t %	8.01–10.00 Trial 1 n_1	c_1	Trial 2 n_2	n_1+n_2	c_2	p_t %
1–3	All	0	–	–	–	–	All	0	–	–	–	–
4–15	3	0	–	–	–	50.0	3	0	–	–	–	50.0
16–50	6	0	6	12	2	48.0	6	0	6	12	2	48.0
51–100	7	0	11	18	3	38.5	7	0	16	23	4	36.5
101–200	13	1	20	33	6	33.5	14	1	24	38	7	32.0
201–300	14	1	26	40	7	31.5	19	2	29	48	9	31.0
301–400	15	1	30	45	8	31.0	21	2	44	65	12	29.0
401–500	16	1	39	55	9	28.5	22	2	53	75	13	27.0
501–600	22	2	38	60	10	27.5	28	3	52	80	14	26.5
601–800	22	2	43	65	11	27.0	29	3	56	85	15	26.0
801–1000	24	2	56	80	13	25.5	36	4	69	105	18	24.5
1001–2000	24	2	61	85	14	25.0	45	5	95	140	23	23.0
2001–3000	33	3	72	105	16	23.0	50	6	115	165	27	22.0
3001–4000	41	4	99	140	21	21.5	70	8	150	220	34	20.5
4001–5000	44	4	111	155	22	20.0	80	9	195	275	41	19.0
5001–7000	50	5	120	170	24	19.5	90	10	240	330	47	18.0
7001–10,000	60	6	145	205	28	18.5	110	12	265	375	53	17.5
10,001–20,000	70	7	165	235	32	18.0	125	14	320	445	62	17.0
20,001–50,000	80	8	205	285	39	17.5	140	16	355	495	69	16.8
50,001–100,000	85	8	245	330	44	17.0	150	17	390	540	77	16 6

n_1 = Size of First Sample; n_2 = Size of Second Sample; entry of "All" indicates that each piece in lot is to be inspected.
c_1 = Allowable Defect Number for First Sample; c_2 = Allowable Defect Number for First and Second Samples Combined.
p_t = Lot Tolerance Per Cent Defective corresponding to a Consumer's Risk (P_C) = 0.10.

Television Transmission Over Wire Lines*

By M. E. STRIEBY and J. F. WENTZ

Intercity networks appear vital to the success of television broadcasting. Experiments with wire lines for this purpose and for local transmission of present-day television signals are reported herein. The design and construction of the equipment used are described and its performance characteristics given.

The intercity lines discussed involve carrier transmission over coaxial cable with repeaters which pass a net band of about $2\frac{3}{4}$ megacycles. For local intracity connections video transmission of about a 4 mc band is obtained over existing telephone plant or by means of special low attenuation cable. Various circuit arrangements including the facilities used in bringing scenes from the Republican Convention in Philadelphia to the N.B.C. in New York are shown together with their overall television transmission characteristics.

Introduction

IF THE development of television broadcasting follows in the footsteps of its predecessor in the sound broadcasting field, networks for interconnecting television stations will be very important. In fact many students[1] of the problem believe that such networks are a virtual necessity because of the expected high cost of programs.

Considerable progress in the development of a wire line technique for this purpose has been made in connection with the Bell System's study of coaxial conductor systems for use in wide band telephony. Data previously published[2, 3, 4] have been supplemented recently by certain tests and experiments in the transmission of 441-line television images, the results of which are presented in this paper. This will cover the transmission characteristics of facilities both for intercity and local distribution, including the wire lines which were used during the television broadcast in New York of the proceedings of the Republican Convention in Philadelphia during the last week of June, 1940. This broadcast was undertaken jointly by the National Broadcasting Company and Bell System Companies as an experiment in the furtherance of the television art. A large part of the experimental facilities used were manufactured by the Western Electric Company.

Long Haul Coaxial Systems

For long-distance broad-band transmission, coaxial systems have certain natural advantages which have been previously pointed out. In common

* Presented before the A.I.E.E., January 30, 1940.

with all long-distance systems for multiplex telephony, the carrier method of transmission is essential and has been found to be relatively straightforward. For long-distance television transmission, the carrier method is necessary with the present coaxial lines and coaxial repeaters, due to the fact that satisfactory long-distance transmission cannot be obtained at the very low frequencies involved in a video television signal. Hence for

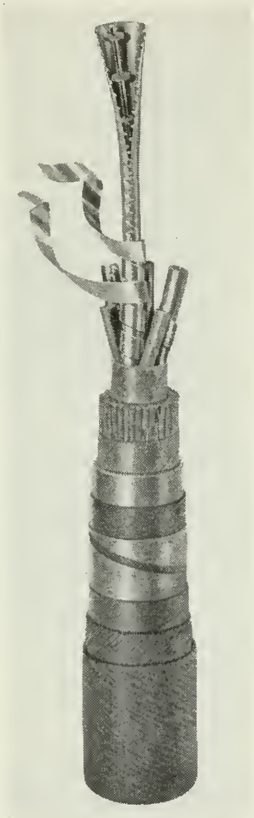


Fig. 1—Photograph of coaxial cable

television the entire signal must be raised bodily to a higher frequency. The modulating means developed for this purpose will be described in detail later. In this section we will confine discussion to the transmission of a broad band of frequencies independently of how this band is used.

Cable

The transmission characteristics of ideal coaxial cables have long been known. The properties of practical structures so far built including matters of cross-talk

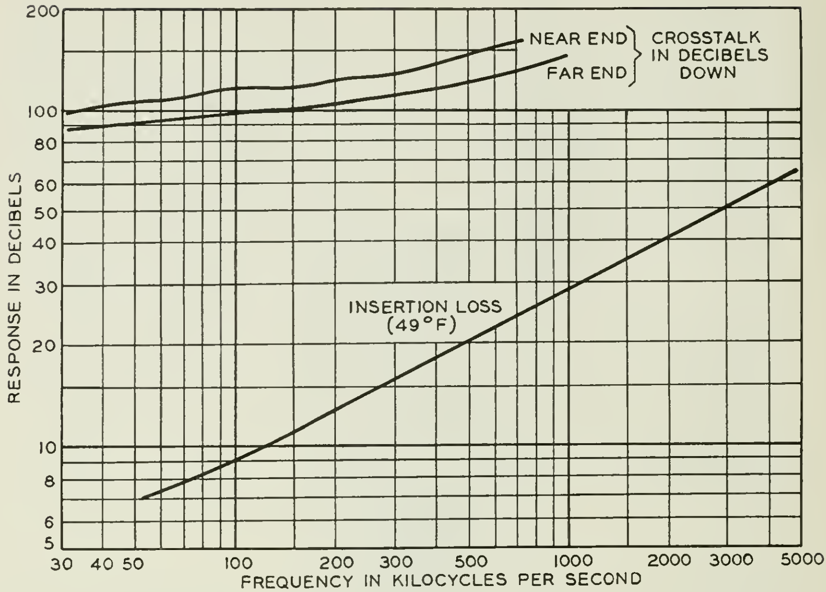


Fig. 2—Attenuation, crosstalk

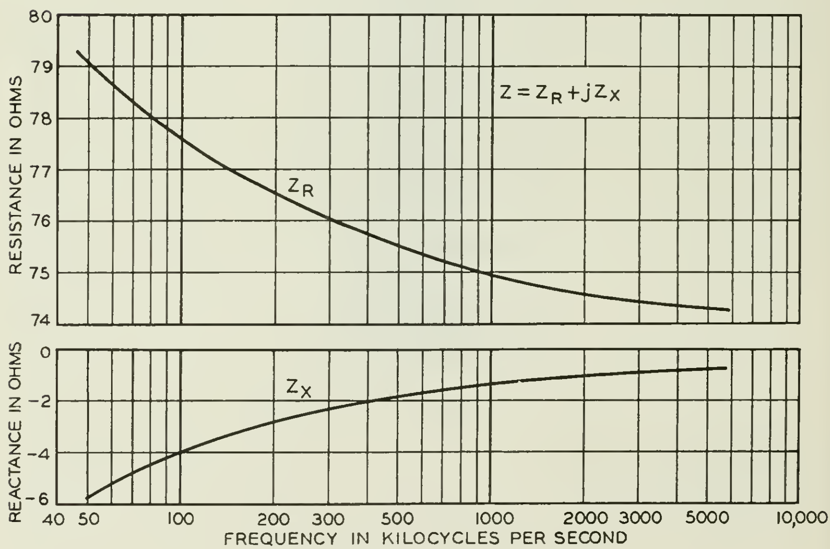


Fig. 3—Impedance of 5 miles

or shielding are also now well understood[5,6]. Certain mechanical improvements in construction have been made recently[7] and may be illustrated by a photograph,

Fig. 1, of the recently installed Baltimore-Washington coaxial cable. A similar construction was used in a cable completed last summer between Stevens Point, Wisconsin and Minneapolis, Minnesota.

These cables each contain 4 coaxial units. Two of these are used to provide a normal broad-band system having one pipe for each direction of transmission. The other two provide spare facilities for each direction. The construction of the coaxial unit itself can be seen from the photograph to use a single longitudinal copper tape for the outer conductor. This is formed into a tube which is held to a fixed diameter by the width of the tape and is prevented from collapsing by the interlocking of its saw-toothed edges. Two layers of steel tape provide the needed support against buckling and also give additional shielding. This construc-

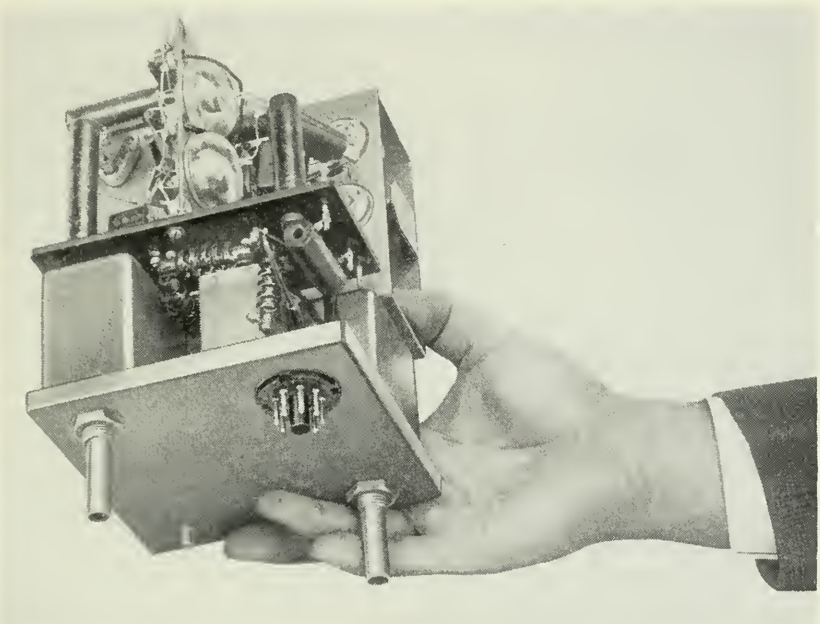


Fig. 4—3-Megacycle amplifier in hand

tion results in somewhat improved transmission characteristics and lower manufacturing costs as compared with other types of construction with which we have experimented. Improvements in transmission include lower attenuation, due to a reduction in the effective resistance of the outer conductor, and a smoother impedance frequency characteristic due to greater mechanical uniformity. In spite of the thinner outer conductor satisfactory crosstalk characteristics are obtained.[8] Typical attenuation, crosstalk and impedance characteristics of this cable as a function of frequency are shown in Figs. 2 and 3 for a 5-mile length of installed cable.

Repeaters

The band width of a coaxial system, at least over regions which we have studied, is limited only by the amplifiers with which it is provided. The amplifiers which have been built most recently for use in these systems are known as "3-megacycle amplifiers" and were intended to provide about a 2-megacycle band of suitable characteristics for telephone purposes or about a $2\frac{3}{4}$ megacycle band suitable for television transmission.

Figure 4 shows one of these amplifiers. It is a three-stage feedback device using two small pentodes in parallel in each stage. The mathematical design of the circuit is beyond the scope of this paper and has been treated elsewhere.[9] This type of pentode has an initial transconductance of from 2000 to 2500 mi-

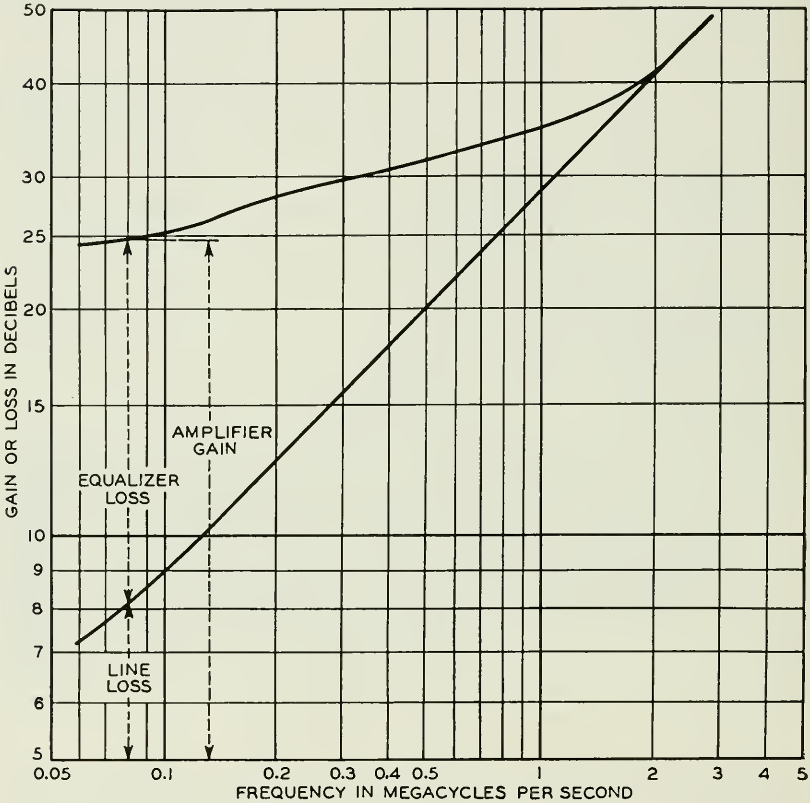


Fig. 5—Repeater gain line loss and equalization characteristics

cromhos and an output power of .1 to .2 watt at 130 volts as used in this system. These tubes are in parallel only to give added reliability. The gain of this amplifier is very roughly the complement of the line loss as a function of frequency. With this amplifier and the cable described above, these repeater sections are about $5\frac{1}{4}$ miles in length. As illustrated in Fig. 5, the difference between the gain and line loss is made up by a line equalizer so that to a first approximation, zero loss in transmission is obtained at all frequencies within the band over each repeater section. About 30 db of feedback is effective over the telephone frequency band (i.e. up to 2000 kc) around the entire amplifier with about 10 db additional around the final stage. From 2 mc up to 3 mc the feedback gradually falls off about 10 db. This arrangement gives the high degree of transmission stability and linearity required for long telephone systems with hundreds of amplifiers in tandem, and satisfactorily meets present requirments. Limited experience with television transmission so far indicates satisfactory performance. The linearity is illustrated in Fig. 6 which shows measurements of 2nd and 3rd order modulation products of a 1000-kc signal in a typical amplifier at various signal levels. As in previous coaxial systems, power for operating the amplifiers is transmitted at 60 cycles over the coaxial cable itself from main stations located at about 50-mile intervals.

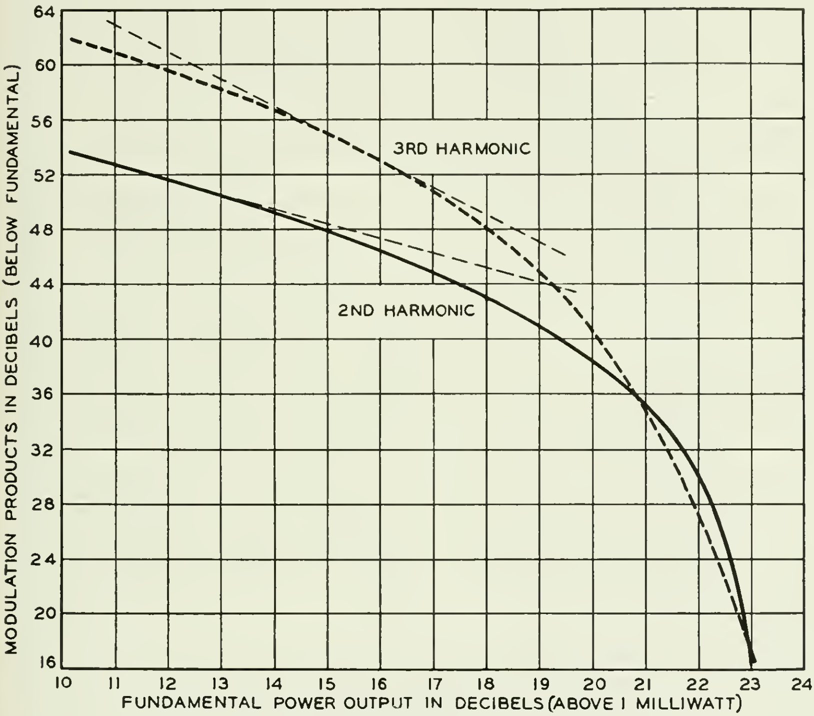


Fig. 6—Amplifier linearity

Regulation

In order to compensate for changes in attenuation due to temperature change of the copper conductors, the gain of the amplifier is regulated automatically by a device located at each amplifier point which is operated from a pilot channel. In this system a pilot frequency of 2064 kc is transmitted along the line with the signal. At the output of each amplifier, a high-impedance highly selective crystal filter is bridged on the circuit to select the pilot frequency. This is then amplified, rectified, and used to control the output of an oscillator. The oscillator output in turn is used to control the resistance of one element in the feedback circuit of the amplifier. This variable element is a very tiny thermistor[10] made up of certain oxides which have a very large negative temperature coefficient of resistance. The regulator is "back-acting" and maintains a substantially constant output voltage at the pilot frequency over a range of about 9 db in input voltage. The feedback circuit of the amplifier is so designed that the changes in the resistance of the thermistor produce changes in gain over the entire frequency band in such a way as to compensate for the changes in loss in the coaxial conductors, as illustrated in Fig. 7. Changes there shown are for ±70° F., which is about the maximum which is expected in a repeater section, even though the cable is of the aerial type.

In a long system, it has not been feasible to make the accuracy of equalization and regulation in each 5-mile section sufficient to give the desired overall uniformity of transmission. Hence, certain supplementary adjustment is required. Devices for such adjustment have been installed at 50-mile points on the Stevens Point-Minneapolis system with satisfactory results. Also, two additional pilot

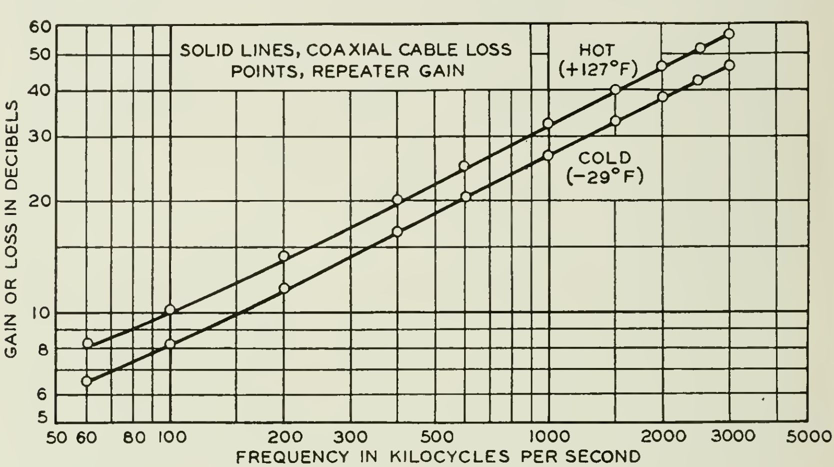


Fig. 7—Regulation hot and cold vs. frequency

channels have been provided, one at 64 kc and one at 3096 kc. These serve to indicate the circuit performance and the need for manual adjustment. These pilots could be used to actuate automatic regulators if desired. For longer systems, it is expected that additional, and necessarily more complicated, supplementary devices will be required at intervals of perhaps 200 to 500 miles.

Performance

A complete repeater containing amplifiers for each direction of transmission, automatic regulators, equalizers, power supply and various automatic alarm features is mounted in a box about 2 x 2 x 1 ft. as shown in a photograph (Fig. 8). Measurements on the overall performance of systems with many such repeaters in tandem indicate a high degree of transmission stability and freedom from noise. In the neighborhood of the pilot frequency the transmission variations are in the order of .1 db. At other frequencies there are slow drifts due to aging of tubes which, when they reach a few db, will require readjustment. These changes are now effected manually at the attended stations.

Interference from all sources, both external and internal, is very low in this system. The largest contributions of such interference are from tube noise and from thermal agitation in the conductors and circuit elements. The effect of interference from external sources so far encountered is lower than the above, although the presence of radio broadcasting stations can be detected. Intermodula-

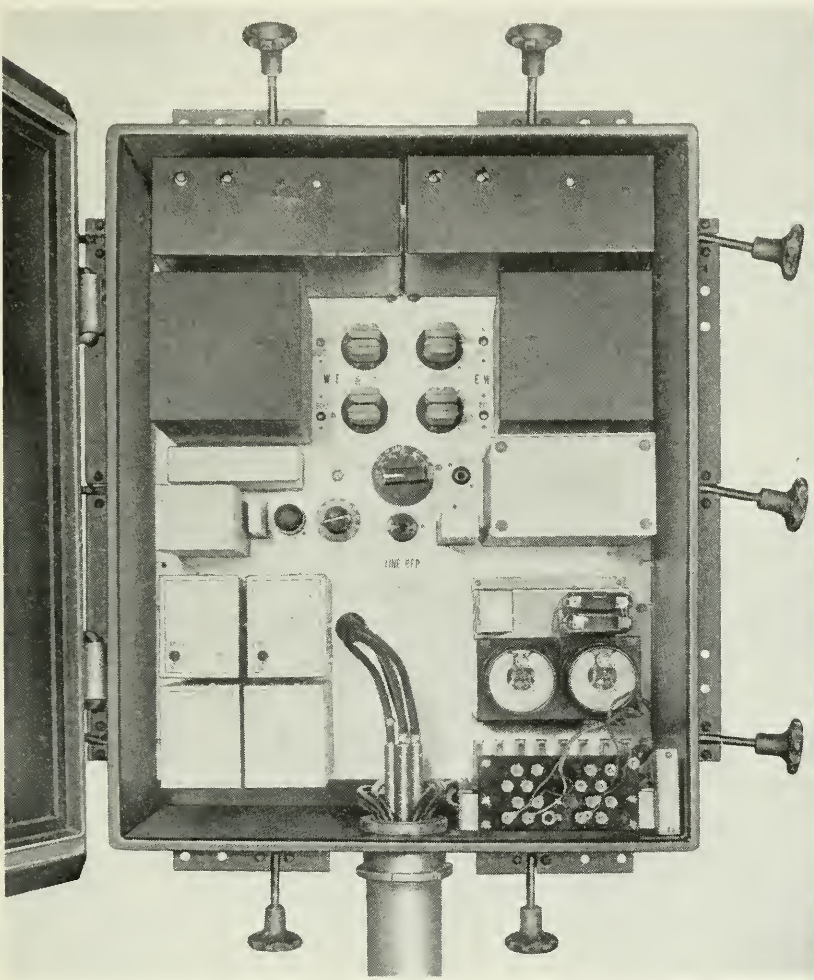


Fig. 8—Complete repeater

tion of signals traversing the system simultaneously has been very carefully measured because of its importance for multichannel telephony and television. In telephony, because of the large number of modulation products, principally 2nd and 3rd order, these appear as random interference.

The method of measurement of interference from all sources was to transmit over the system a wide-band signal having a continuous spectrum such as thermal noise. At the sending end a narrow-band elimination filter was inserted. At the far end the noise was measured within that same band.

The total noise so measured depends upon the signal energy levels at the input and the output of the repeaters, the former controlling the effect of line and resistance noise and the latter controlling the effect of modulation. These levels in turn are a function of repeater spacing. The tests that have been made indicate that it is practicable to keep this type of interference within desirable limits on long telephone or television circuits.

Due to the 60-cycle power supply used on the system, power frequency modulation products require special attention. Sixty-cycle sidebands are produced on all signals transmitted due to the traces of nonlinearity in the system. As these are very small in magnitude and result mostly in a 120-cycle component they are unimportant for telephony. However, in the television transmission system used, this component is larger because of the presence of a strong carrier and one or more pilot channels. Also, 120-cycle sidebands produce a very disturbing type of horizontal bar pattern across the picture. This type of interference will increase as the circuit length is increased, and may become more visible as receiving tubes are improved. On systems so far available for test, however, it has been possible to hold this type of interference within acceptable limits, on present day television broadcast images.

Distortion in Television Images

Departures from ideal transmission in the line, equipment or in a radio path produce distortion in the form of negative or positive fringes or "ghosts." These occur when there is a lack of proportionality between phase shift and frequency through the system. This trouble in television images is perhaps more easily understood if one thinks of it as an actual difference in time of transmission of various parts of the signal. In discussing this matter in this paper, we will use the term "delay" to mean the time of transmission of the envelope of a modulated wave. This quantity is often more accurately referred to as "envelope delay".[11] If this quantity varies too widely there is an actual difference in the time of transmission of various parts of the signal, producing distortion in the form of fringes or "ghosts" which are exhibited by many television images today.

Band Width

A band width of about 3 mc is required to give equal resolution in the vertical and horizontal directions in a 441-line, 30-frame interlaced image. Recent experiments[12] with out-of-focus moving pictures have shown not only that the eye is quite insensitive in its requirement for equal detail in the two directions but also that the loss of detail due to a narrowing of the frequency band from 4 mc to $2\frac{3}{4}$ mc will pass unnoticed by many careful observers at normal viewing distance.

Television on Coaxial Systems

As mentioned above, no practical method has been found for transmitting television over long-haul coaxial circuits in the video frequency range. By the carrier method, however, the video frequency band may be raised to a region suitable for transmission. To conserve frequency space, single-sideband transmission, of course, is desirable. The actual method chosen involves also a modest vestigial band since it appears impracticable to select a single sideband involving video frequencies as low as 45 cycles in any other way. The present coaxial amplifiers pass a band from about 64 kc to about 3100 kc. The region useful for television, however, appeared to be somewhat less than three megacycles on account of the difficulty of equalizing the delay distortion near the lower edge of this band. About 100 kc was allotted to obtain proper shaping of the vestigial sideband. The carrier was therefore placed at about 300 kc and a net television band of about $2\frac{3}{4}$ mc was obtained. If we attempt to move a 3 mc video band up 300 kc in a single step of modulation, the result is an overlapping of the sidebands which hopelessly distorts the signal. Two steps of modulation are therefore resorted to as shown in Fig. 9.

The energy of a television system is concentrated in the lower frequencies or, in a carrier system, near the carrier. To take most advantage of the coaxial system, the carrier should be at the low end where the full feedback in the amplifiers is available. The four lines in Fig. 9 illustrate the four stages of modulation, two at the transmitting terminal and two at the receiving terminal. As can be seen the signal is first modulated with a carrier of about 8 megacycles and the lower sideband, part of the carrier, and a portion of the upper sideband, are selected by a band filter. This signal is then modulated again with a carrier of about 8.3 megacycles and the lower sideband again selected. In this position of the signal, which is the position at which it will be transmitted over the coaxial line, the frequency which corresponds to d.c. in the video signal is at 311 kc, the main sideband extends from 311 to 3111 kc and the vestigial sideband from 311 kc down to 200 kc.

The receiving terminal is in general the inverse of the transmitting terminal and will not be discussed in detail. The sideband shaping[4] is accomplished by the four filters, two at the transmitting terminal and two similar ones at the receiving terminal, acting in conjunction. The result is that at the final stage of demodulation the contribution from the vestigial sideband when added to the contribution from the shaped portion of the main sideband gives back very nearly an undistorted video signal. This last stage of demodulation is accomplished in a linear detector. The carrier amplitude at the input terminals of this detector is about six db greater than the amplitude of the video envelope of the modulated signal, the amount of carrier which was mixed with the sidebands at the output of the first modulator having been adjusted to achieve this result. The reason for using this amount of transmitted carrier is the relatively narrow vestigial sideband—111 kc vs. a main sideband of about $2\frac{3}{4}$ mc. With such a narrow vestigial sideband the quadrature component of the carrier en-

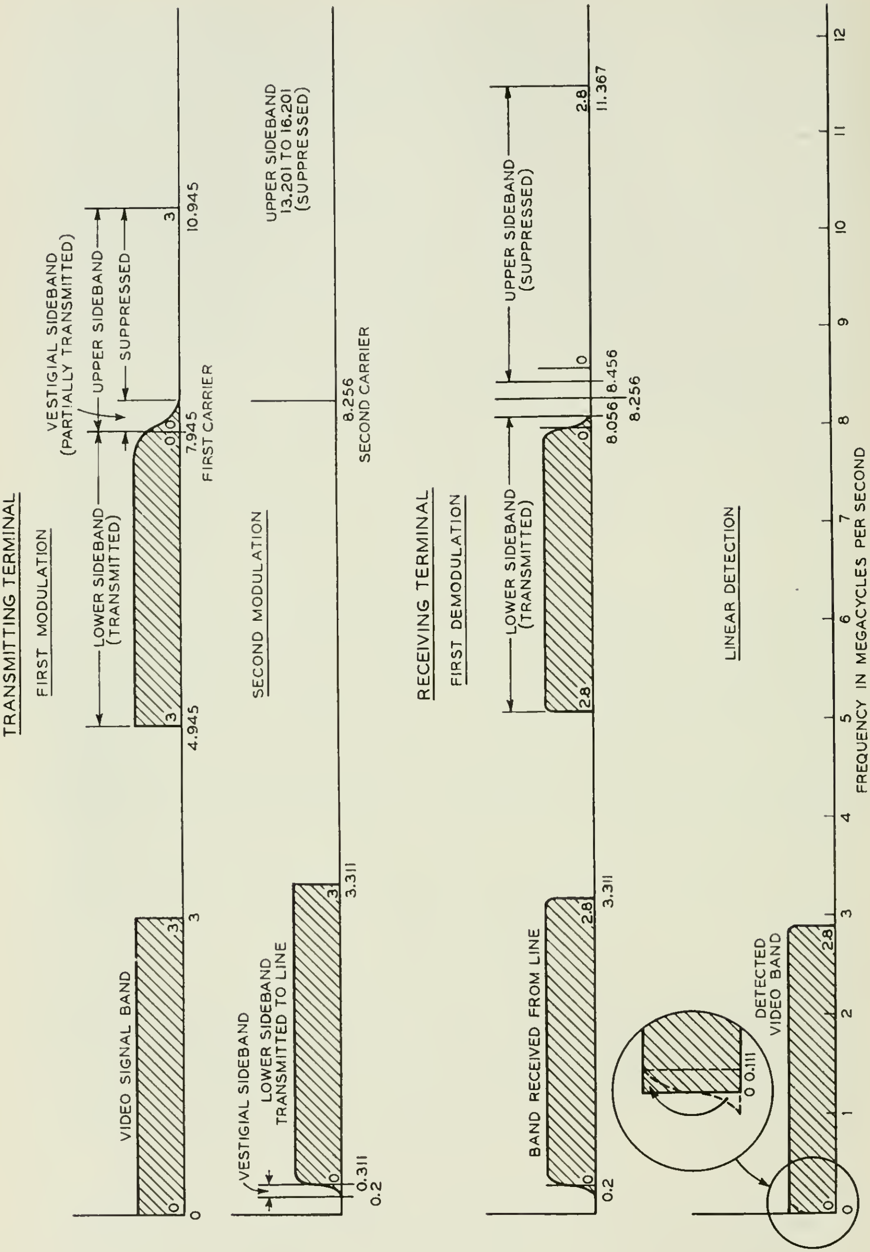


Fig. 9—Modulation steps diagram

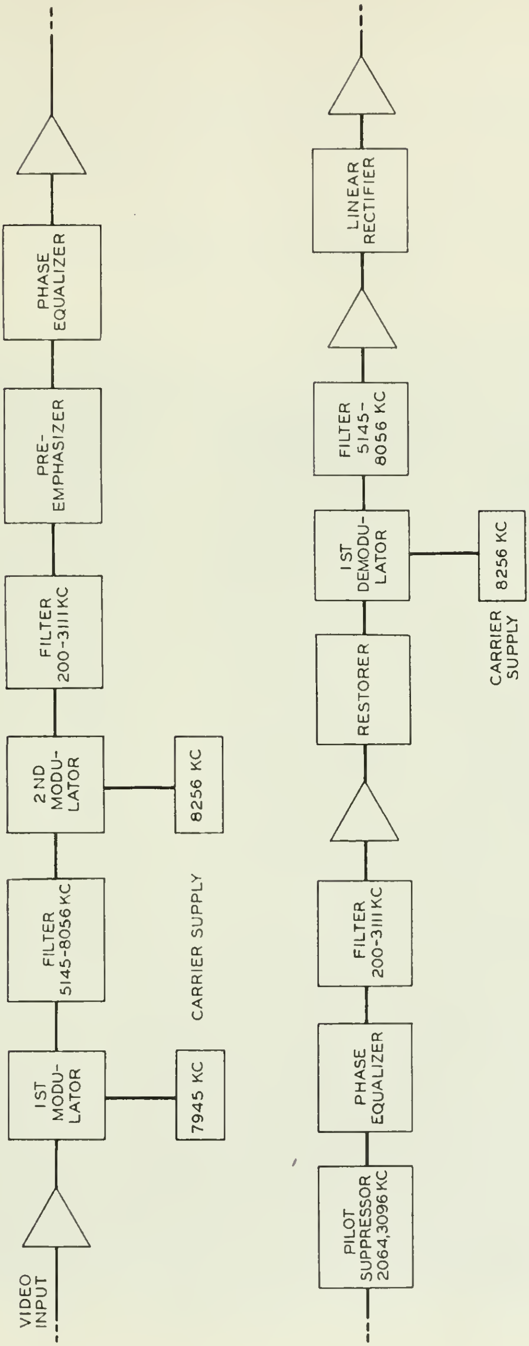


Fig. 10—Box diagram of carrier terminals

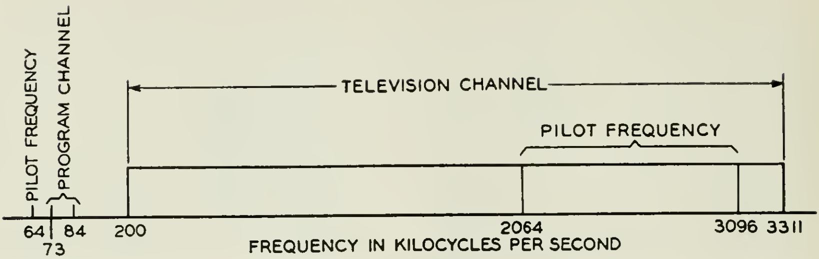


Fig. 11—Frequency allocation on line

Fig. 12—Carrier terminals

velope is relatively large, resulting in objectionable distortion at sharp changes in the picture signal if the greater ratio of carrier to sideband is not employed.[13]

Figure 10 shows a box diagram of the terminal arrangements. In addition to ordinary video amplifiers and modulators and filters mentioned above, a "pre-emphasizer" and a "restorer" are shown. These networks partially equalize the energy in the various components of the signal, and thus help to override the noise and spurious modulation products introduced by the line and amplifiers. A phase equalizer is also shown which, in conjunction with a similar equalizer at the receiving end, is designed to correct for the phase distortion in both the transmitting and the receiving terminals. Before transmission over the coaxial, pilot frequencies of 64 kc, 2064 kc and 3096 kc are added, as well as a program channel from 73–84 kc. Figure 11 shows the frequency allocation of the television signal and its associated channels on the coaxial line.

At the receiving end the pilot frequencies and the program channel must be removed. The 64-kc pilot and the program channel are eliminated by the 200-3111 kc filter which precedes the first demodulator. The 2064-kc and 3096-kc pilots, however, are within the transmitted television band. The frequency allocation was so chosen as to place them approximately in the center of the "empty energy regions"[14] of the television spectrum where they can be eliminated by sharp selective networks without appreciably distorting adjacent television signal components.

Three carrier television terminals are shown in the photograph, Fig. 12. The one on the right is a transmitting terminal, the two on the left receiving terminals. Each terminal occupies one six-foot relay rack bay and is complete with power supply and means for adjustment.

Short Haul Lines for Television

For the pickup or transmission of television within cities or metropolitan areas, it appears to be more economical, as would be expected, not to use the carrier method described above but to transmit "video" frequency signals over cable circuits. For this purpose existing telephone cables may be used or special cables may be provided. In either case amplifiers and special equalizers are required which will overcome the attenuation and delay distortion of the cable circuits. Because of high-frequency crosstalk usually only a small fraction of the circuits in any existing telephone cable can be used simultaneously.

Video Amplifiers and Equalizers

Television pickup and broadcasting equipment is quite naturally designed on an unbalanced (i.e. one side grounded) basis. Unbalanced amplifiers for the video band have been available for some time.[15] New amplifier designs have been worked out for use with balanced lines. In general, the problem is to provide approximately zero loss and constant delay between unbalanced terminals a mile or more apart. Thus, an unbalanced to balanced amplifier is required at the sending end, the converse at the receiving end. If the circuit is long, balanced amplifiers are most convenient for use at intermediate points. The equalization problem has been successfully met even if ordinary telephone cables are used. A series of variable equalizers have been experimented with which have several degrees of flexibility. A variety of circuits ranging in length up to 9 miles have

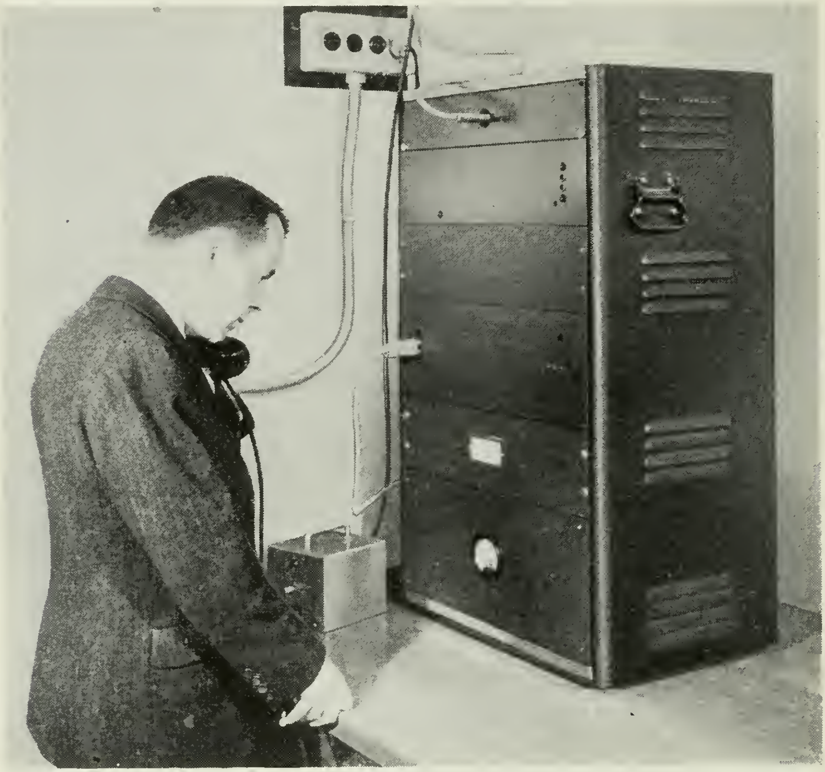


Fig. 13—Photograph of Video amplifier

been equalized with this arrangement with considerable success. A typical amplifier, equalizer, and power supply are shown in the photograph, Fig. 13.

Telephone Cables

Ordinary fine wire paper insulated cables have very high attenuation at the frequencies required. Typical values for loss and net loss after amplification and equalization are shown in Fig. 14. Experience has shown that the noise levels in such cables even at the higher video frequencies are rather high so that amplifiers are required at intervals of a mile or even less. Local telephone cables are usually laid out with many branches. At high frequencies these branches introduce irregularities similar to those produced by obstacles along a radio path which cause delay distortion. Plant changes are frequently required to obtain a clean circuit free from such bridged taps. When amplifiers and proper equalizers are added, however, substantially flat transmission is obtained as shown in the figure.

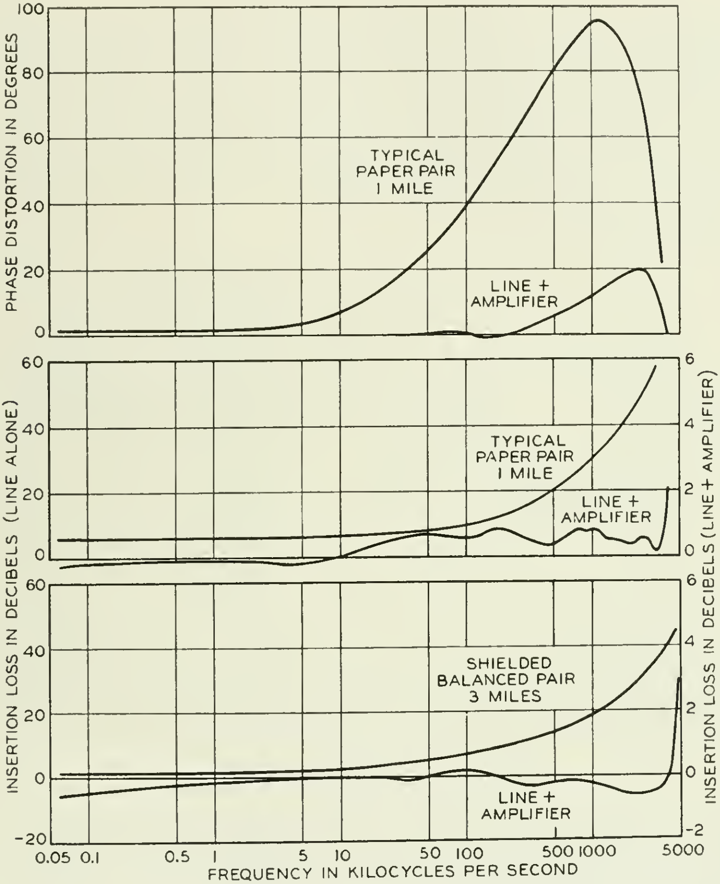


Fig. 14—Transmission characteristics of cables

Phase distortion characteristics of a typical cable circuit are also shown in Fig. 14. After the amplifiers and equalizers are added, the phase distortion is made substantially negligible.

Coaxial Cables

Coaxial cables may be used for video transmission for short distances but power or other low-frequency interference may introduce serious problems.

Coaxial units of the size discussed above have been used in a few cases a mile or so in length. Even for such distances, however, it has been found desirable to reduce the power interference by balancing it out. One method which we have used is shown in Fig. 15. This has given an improvement at power frequencies of the order of 50 db in certain cases.

Balanced Shielded Cables

The ideal type of transmission line for video signals combines the balance feature with low attenuation and a high-frequency shield. The distance over which such cables could be used appears to depend upon the perfection of balanced video amplifiers and the equalization, although power interference may also present difficulties. Such cables have been built using a pair of wires and a disc type of insulation analogous to the coaxial structure described above. Attenuation measurements on a 3-mile test length installed in New York City are shown in Fig. 14. This figure also shows the net result after amplifiers and equalizers were

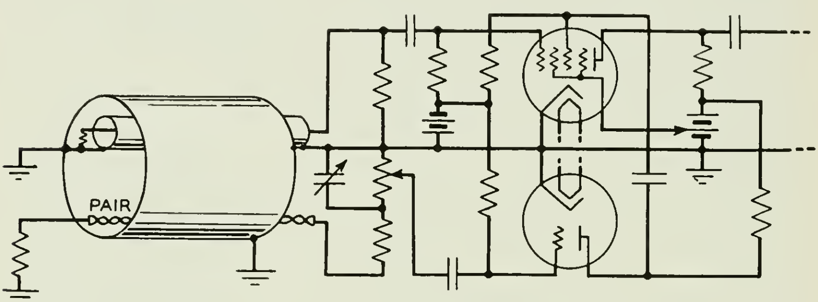


Fig. 15—60-Cycle balance on coaxial line

added. The attenuation of this special type of cable is such that amplifiers would be required at only about 5-mile intervals. The useful range of such a cable for video transmission has not been determined but in any case it should be considerably greater than that of the paper-insulated telephone cable circuit.

Experiment in Network Broadcasting

During the last week of June 1940, the proceedings of the Republican Convention in Philadelphia were broadcast in New York by television. The facilities used included the 3-megacycle coaxial system plus certain video connections at each end as shown in Fig. 16.

Because of the interest in this circuit and its good performance in transmitting 441-line television, the overall attenuation and delay characteristics are given on Fig. 17. It will be noticed that a net band of about $2\frac{3}{4}$ megacycles was transmitted and that over most of that band the delay distortion did not exceed ± 0.2 microsecond. The random noise, modulation and other distortions introduced by the wire line network appeared to be unimportant when viewed on a commercial television receiver.

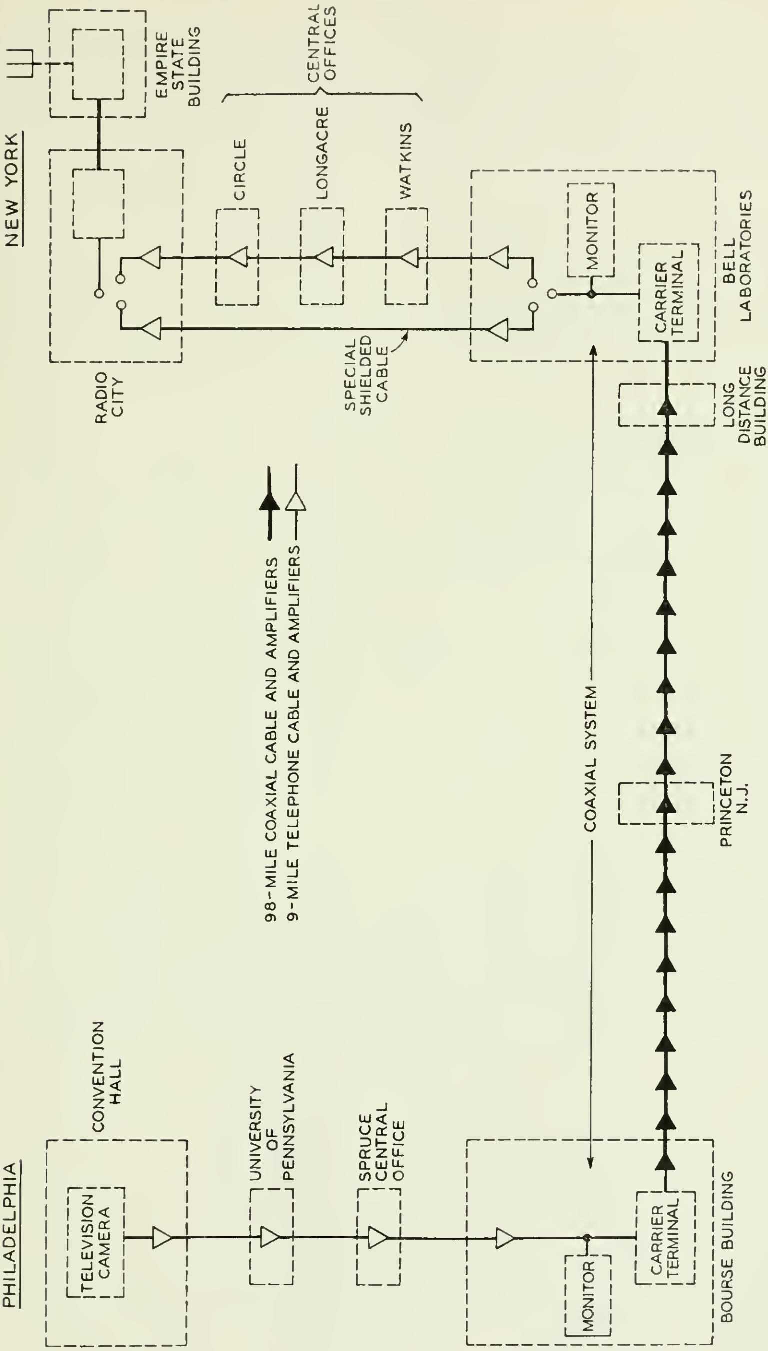


Fig. 16—Map of New York-Philadelphia network

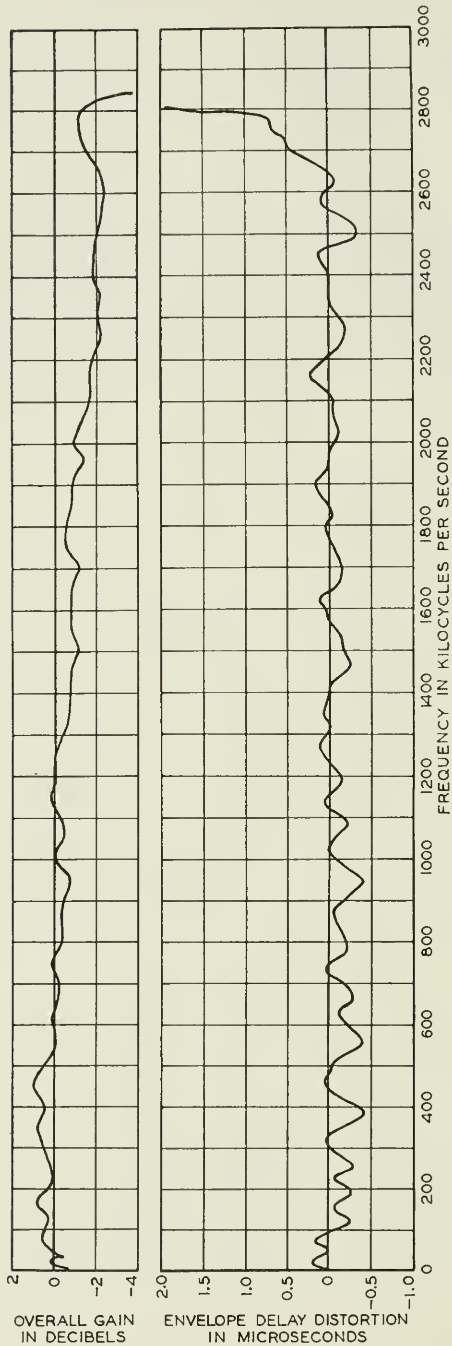


Fig. 17—Attenuation and delay—New York-Philadelphia network

Conclusion

The experiments so far made in the transmission of present-day television indicate that wire lines can be provided at least for moderate size intercity networks; also, that such lines if properly equalized for delay and attenuation do not materially alter or distort the transmission of present-day 441-line images, even though the frequency band is somewhat narrower than the nominal 4-mc band.

The use of ordinary telephone cables for local television connections also has been found to be feasible for all of the conditions so far tested. The $2\frac{3}{4}$ mc television transmission experiments over wire lines reported herein have proved very successful. Experiments with wider band coaxial systems are being undertaken.

Bibliography

1. "Television Broadcasting," A Book by L. R. Lohr, McGraw-Hill, 1940, Chapter VIII.
2. "Systems for Wide Band Transmission over Coaxial Lines," L. Espenschied and M. E. Strieby, *B. S. T. J.*, Oct. 1934.
3. "Million Cycle Telephone System," M. E. Strieby, *B. S. T. J.*, Jan. 1937.
4. "Coaxial Cable System for Television Transmission," M. E. Strieby, *B. S. T. J.*, July, 1938.
5. "Crosstalk between Coaxial Transmission Lines," S. A. Schelkunoff and T. M. Odarenko, *B. S. T. J.*, Apr. 1937.
6. "Crosstalk in Coaxial Cables—Analysis Based on Short Circuited and Open Tertiaries," K. E. Gould, *B. S. T. J.*, July 1940.
7. "Stevens Point-Minneapolis Coaxial Cable," O. S. Markuson, *Bell Labs. Record*, Jan. 1940.
8. "Crosstalk between Coaxial Conductors in Cable," R. P. Booth and T. M. Odarenko, *B. S. T. J.*, July 1940.
9. "Relations between Attenuation and Phase in Feedback Amplifier Design," H. W. Bode, *B. S. T. J.*, July 1940.
10. "Thermistors—Their Characteristics and Uses," G. L. Pearson, *Bell Labs. Record*, Dec. 1940.
11. "Delay Measurements," H. Nyquist and S. Brand, *B. S. T. J.*, July, 1930.
12. "The Subjective Sharpness of Simulated Television Images," M. W. Baldwin, *B. S. T. J.*, October, 1940.
13. "Effect of Quadrature Component in Single Sideband Transmission," H. Nyquist and K. W. Pfleger, *B. S. T. J.*, Jan. 1940.
14. "A Theory of Scanning and its Relation to the Characteristics of the Transmitted Signal in Telephotography and Television," P. Mertz and F. Gray, *B. S. T. J.*, July 1934.
15. "Analysis and Design of Video Amplifiers," S. W. Seeley and C. N. Kimball, *R. C. A. Review*, Oct. 1937.

Insulation of Telephone Wire with Paper Pulp*

By J. S. LITTLE

A method has been developed for economically manufacturing insulated wire for telephone exchange area cable by making the paper on the wire. Further, this method has made it possible to increase the number of wires in a full sized cable by 175% over the number in use in 1914. Developments now under way indicate that suitable insulation can be made to replace certain textiles in some classes of wire and that the use of this process may therefore be still further extended in the not so distant future.

Introduction

IN 1887 the leading telephone engineers attempted to standardize telephone cables and specifications, finally deciding upon # 18 B & S gauge wire covered with two wrappings of cotton and twisted into pairs. A maximum cable size of 52 pairs in a two-inch diameter cable sheath 97% lead, 3% tin, and $\frac{1}{8}''$ thick was permitted under the specifications. The grounded capacity of such cable was 0.20 mf. per mile. In 1891 the Western Electric Company had made successful application of manila rope paper as insulating material for dry core cable and by drying this paper immediately before covering with lead by the newly developed extrusion process the core could be kept dry without the old impregnation with hot paraffin. A great improvement in electrical properties resulted from this change, the electrostatic capacity dropping to approximately one-half its former value. The use of manila paper made from old rope from this time on grew in use for insulating purposes (Fig. 1). The telephone demand was increasing all the time, and since the supply of old rope depended in a large measure on maritime sources of supply the price began to increase. Improvements in telephone instruments, together with increased demand for telephones, permitted the use with economy of more and more pairs of finer and finer wires in a given diameter of cable. This trend can be readily seen if we follow the change in maximum number of pairs used at different dates. In 1888—50 pairs of 18-gauge wire were used, 1896—180 pairs 19-gauge, 1912—909 pairs 22-gauge, 1914—1212 pairs 24-gauge, 1928—1818 pairs 26-gauge, and in 1939—1515 pairs 24-gauge and 2121 pairs 26-gauge (Fig. 2). The increasing number of wires demanded thinner and thinner and better and better paper. As the cable demand increased, increased insulating speeds were necessary to aid in keeping down the cost

* Reprinted, with minor changes, from *Wire and Wire Products*, October 1939.

due to the higher priced papers. Increased flexibility of paper without sacrificing strength and greater uniformity were required in these new thinner papers. Considerable time and money were spent in attempting to reduce the amount of manila fibre due to its price and increased scarcity and to substitute cheaper fibres of wood and cotton. It was finally found that mixtures of 45% rope, 40% wood and 15% cotton could be used for all

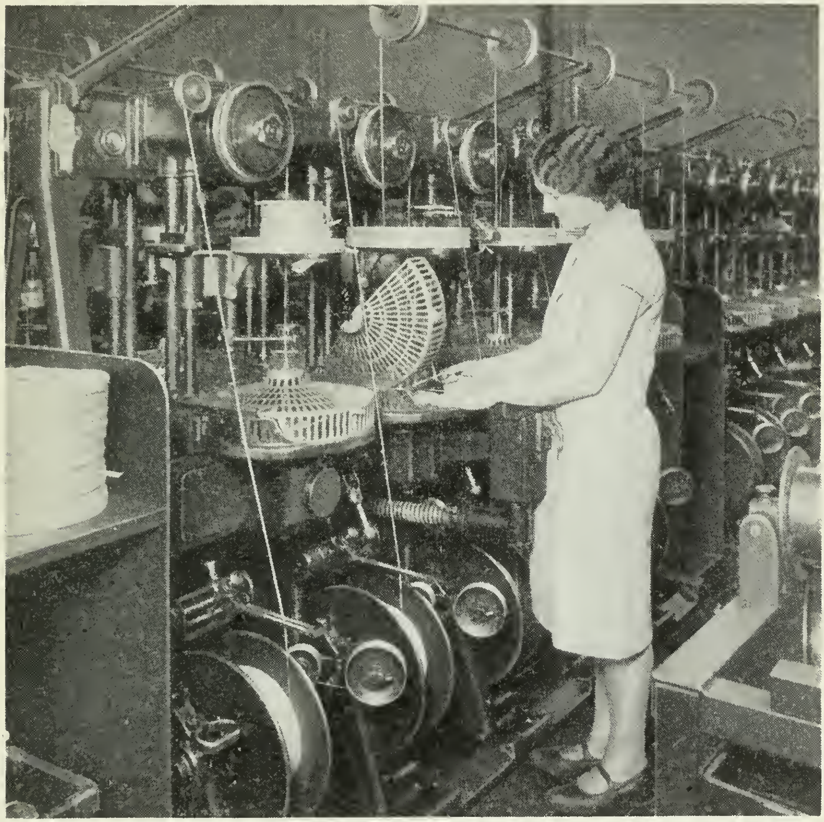


Fig. 1—Paper strip insulating machine

but the very finest insulating papers and that as high as 80% wood and 20% manila rope could be used for the coarser wrapping papers.

In spite of these changes and the improved paper making technique developed by the industry the use of paper $\frac{1}{4}''$ x .0025″ for insulating 26-gauge wire was not entirely satisfactory from a manufacturing point of view. About 1920 some of our engineers began developing the idea of manufacturing the paper right on the wire. If this were possible there seemed to be no reason electrically why wood pulp would not make a suitable insulating material, and from the mechanical standpoint many of the difficulties involved in wrapping the insulation would be eliminated.

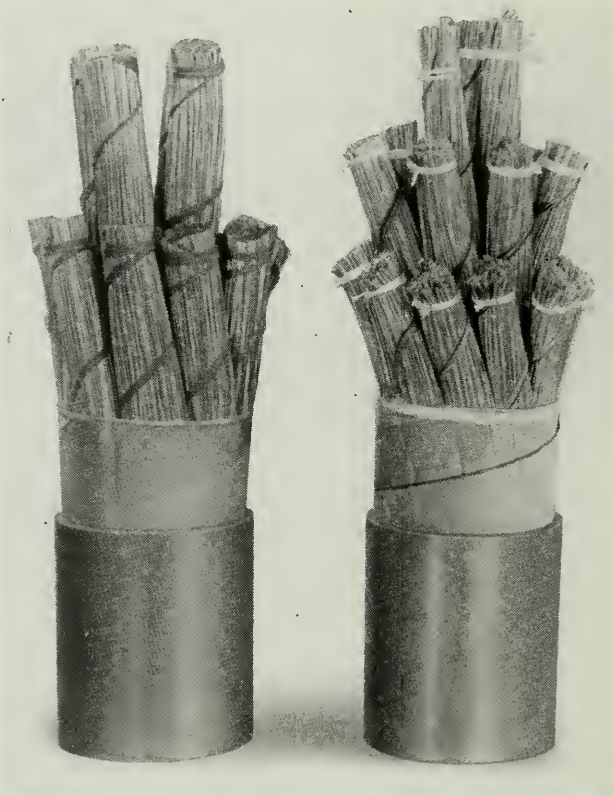


Fig. 2—Comparison of 1212 (left) and 2121 (right) pair cables

The Development of Pulp Insulation

Pulp Machine

The first crude experiments on insulating wire with pulp were done by pouring a suspension of pulp over a wire backed up by a fine mesh screen and after the water was drained away lifting the wire up together with whatever fibers clung to it and then rolling the wire on a flat surface. These samples gave an idea of the type of product to be expected and looked so interesting that a study of equipment and methods was authorized. It developed that the machine most adaptable for our purpose was the standard single cylinder paper machine in use in the paper making industry.

The essentials of this machine are a vat for holding a thin pulp suspension and a hollow cylinder covered with fine mesh screen immersed in the vat. Suitable dams at the ends prevent the pulp suspension passing into the interior of the cylinder. As this cylinder rotates on its axis the water flows through the screen and deposits pulp on its surface. This pulp mat is then picked up by an endless felt belt which is brought into contact with the surface of the cylinder by means of a soft rubber roll which presses it firmly against the pulp mat on the surface of the cylinder. The pulp mat adheres

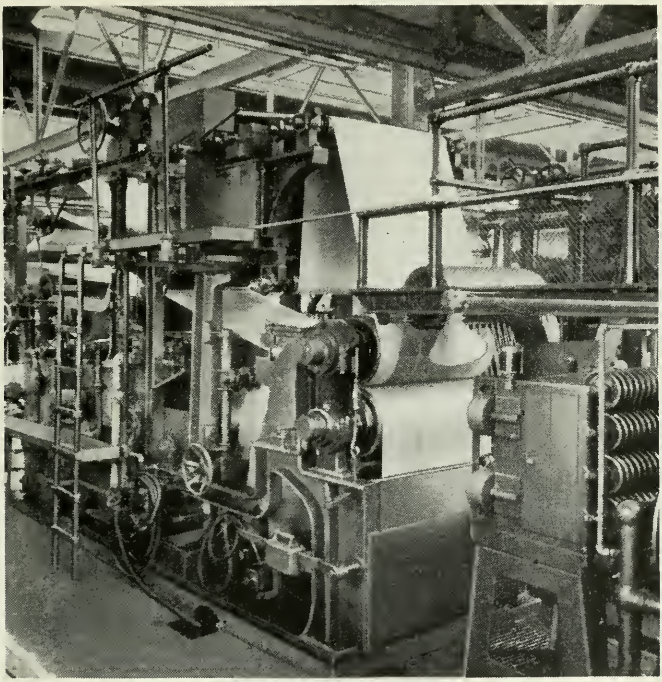


Fig. 3—Forming end of pulp insulating machine

to the felt and together they travel over suction rolls and through squeeze rolls where the excess water is removed. The fibers are thus firmly pressed together so that a sheet of wet paper is formed. After drying and calendering the paper appears in its usual form.

The idea of embedding a wire in the sheet as the pulp was deposited on the cylinder formed the basis of the present development. Usually the paper machines produce a continuous sheet eight or nine feet wide so that it became necessary to devise ways and means of producing sheets only about one quarter inch wide to supply the necessary material for insulating

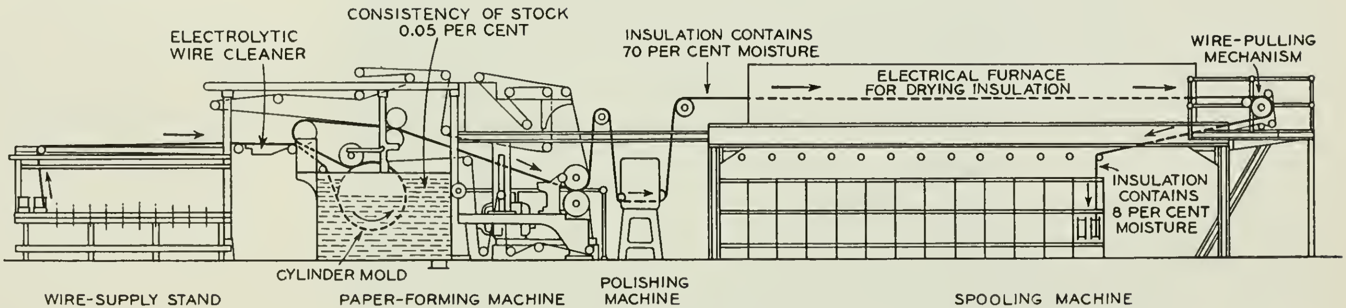


Fig. 4—Schematic of pulp process

wires of 22-gauge and smaller. The most practical size of commercial machine (Fig. 3) was no less than three feet wide but by painting annular rings on the cylinder screen surface the effect of a series of small cylinders all immersed in a single vat could be produced. This was the scheme finally adopted for preparing the paper making machine and we have standardized on a cylinder three feet long with enough rings to simultaneously produce sixty sheets of paper approximately $\frac{1}{4}''$ in width. The layout of the resultant machine is shown schematically in Fig. 4.

Pulp Supply

Kraft pulp is among the toughest of the wood fibers as well as one of the cheapest. It is prepared by an alkaline process and our experience indicated that this process produced pulp of a greater degree of permanence than the acid processes unless special treatments were used. The chief drawback to its use was its color, brown or tan, which necessitated a change in the color code in the cables. Fortunately, cable designs could be made using fewer colors than had previously been employed so that this obstacle was not serious. Standard paper making beating equipment was purchased and used for preparing the pulp to form the sheets although special beating technique for our purpose had to be developed. The older beating method consists of grinding the fibers in the presence of water under a heavy roll. By this continuous maceration the pulp is softened and fibrilated and made suitable for paper making. The longer the grinding the more parchmentlike the final paper becomes, and as we desire as porous a paper as possible it is necessary to control the beating to a point where good strong paper will be made but will still contain a high degree of porosity. Within the last few years a continuous beating system has been developed to replace the original batch system. In this method the pulp mixed with water is run through a preliminary hydrofiner grinder where the pulp is partially beaten before being stored in a large tank. From this tank it is then fed to the various machines and colored by adding the proper dye. A further refiner in the line to each machine finishes the beating for the particular insulation being made in that position. Study showed that fiber from different sources of wood supply handled differently so that standardization of sources of supply had to be made and methods of test developed to check on new fibers or new sources of pulp.

Due to the small thin sheets made on the machine, the amount of pulp required per unit of time is extremely small. No commercial means of measuring such quantities accurately had been developed and it was necessary to spend considerable time in this study. The suspension of pulp to be measured contains only 1.5% fiber and this is further diluted to .05% in the machine vat. The actual quantity of liquid measured is about 8 gallons per minute. The device most recently adopted is similar to the jaws of a pair of pliers held between two stationary guides. As the jaws are separated more liquid flows through them and as they are closed the flow is cut down. A vernier scale adjustment makes close and accurate settings possible when used with a constant head. In the older system the dye for coloring the pulp is added in the beater but the newer system more recently put to use in the Kearny, New Jersey Plant supplies the dye as needed so that only uncolored pulp need be stored in tanks and color changes can be rapidly made with little loss of stock.

Wire Supply

A machine of this size and difficulty of control necessitated a continuous supply of wire to avoid large losses in junk and lost time. It was necessary to devise methods of continuous feed, and to do this wire supplied on spools was utilized. On the earliest machines spools 8″ x 8″, containing sixty pounds of bare copper wire, were used, the wire being removed over the head of the spool by means of a flier. At each supply position two spools were placed side by side and a flier placed on one. When the first spool was emptied to the last few turns, an operator, by means of a special hook, pulled out one turn and brazed it to the outer end of the other spool. A flier was then placed on the second spool, and when the braze was reached the transfer to the new spool took place.

Using the new 400-pound spools from the new Kearny Wire Mill, a larger supply space is needed (Fig. 5), and the two spools per position are set opposite one another instead of side by side. The inner end of wire on these spools is brought out and coiled in the head of the spool so that the two spools can be brazed at any time the operator wishes and the wire will be completely used from each spool. Either a stationary flat ring or a rotating type flier can be used for removing the wire. The latter type has certain operating advantages which at present warrant its introduction and use, although the flat disc has so far been used. With the disc type take-off, a tensioning device consisting of a system of three small rollers is used. One roller can be varied in size and as the three come together the slip between them supplies the tension. With the flier type take-off, a tension device is not essential.

Wire Cleaning

In our early efforts to make insulated wire by this process very erratic results were obtained in the continuity of the insulation. It was finally found that small traces of drawing compound left on the wire made it difficult for the wet pulp to adhere to it during the subsequent polishing operations. Therefore, it became necessary to clean all the wire. Con-

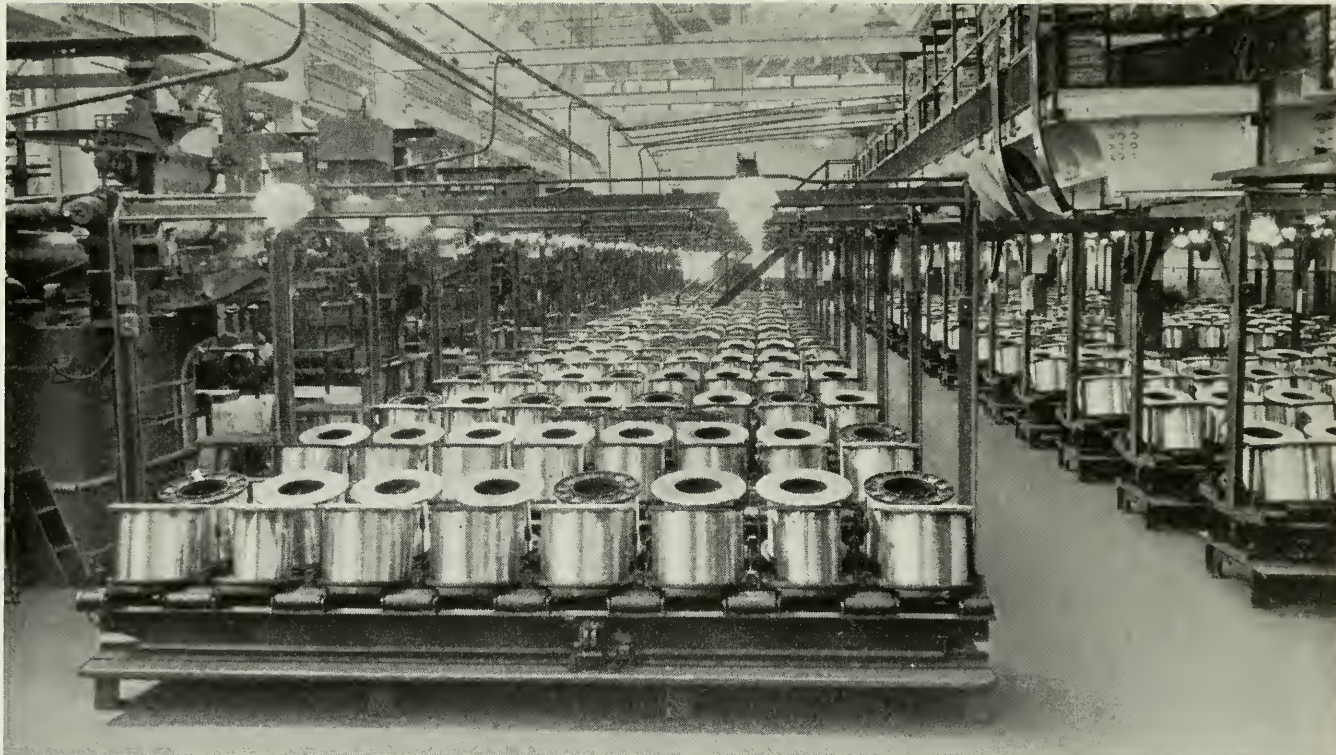


Fig. 5—General view of wire supply

siderable difficulties in designing a suitable cleaner were experienced, but ultimately the use of alternating current together with an alkaline cleaning bath, was found most suitable. The wire passing to the machine comes in contact with the surface of a cleaning solution for a short distance. Electrical contact with the wire is made through guide rolls and the current flows from the wire through the solution to the container. Originally a mixture of cyanides was used as the cleansing agent and the current flow was held to about 8 amperes per square inch surface at 12 volts. Recently a more effective non-poisonous cleansing agent has been developed by using sodium ortho silicate and ivory soap. The passage of the current in either case heats the solution and liberates a rather violent evolution of gas at the surface of the wire. With the soap solution a foam is built up which is continually floated off, carrying the grease, copper, dust, etc. to the sewer. This method keeps the cleaner from concentrating the dirt and consequently eliminates frequent cleaning both of the cleaner and the screen on the cylinder which formerly used to get plugged up with particles of grease carried over from the cleaner by the wire.

Embedding the Wire in the Pulp

From the cleaner the wire is guided into the cylinder machine. It is extremely important at this point that the wire be guided into the center of the small sheets and at such a point on the periphery of the drum that some pulp is deposited below and some over the wire. After passing around the cylinder the wire travels along with the felt and pulp through the presses and finally emerges at the last press embedded in a small sheet of wet paper (Fig. 6). It was found that poor pick-up of the fibers often occurred unless the surface tension of the water was lowered by some means. Ordinary soap is used for this purpose. Approximately ten pounds per thousand pounds of pulp are dissolved in the storage tanks to give effective results and to smooth out the pick-up to give a high degree of uniformity to the weight of pulp per unit length of wire.

Polishing

Polishing of the insulation on the wire is brought about by passing the wire and pulp sheet over polishing blocks which are rotated rapidly around the wire as an axis. Three blocks are used and are so placed that the wire is slightly deflected from its course as it passes first over one block, then the second and finally the last (Fig. 7). The rapid rotation of the polishing head produces a light rubbing action on the sheet which is rolled down without tearing and results in a good round smooth wrapping of wet paper about the wire. With the wire running at a linear speed of 130 feet per minute the polishers are rotated at 5000 r.p.m. to give satisfactory insulation.

Drying of the Insulation

The method of drying the insulation is very important. In the early experiments low-temperature air drying, high-temperature air drying and finally moderate-temperature-controlled humidity drying were studied.

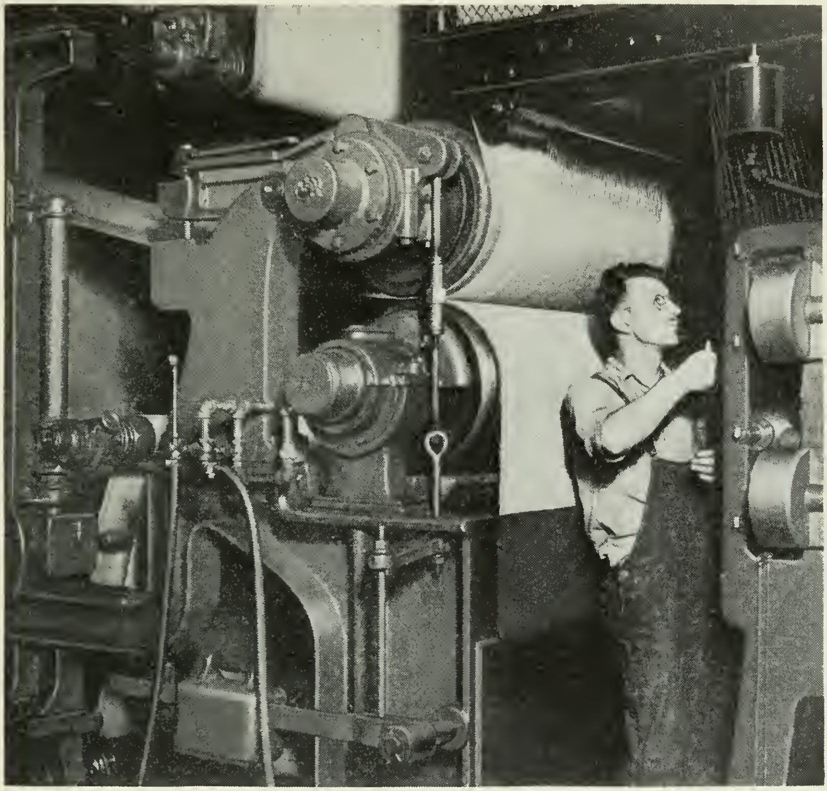


Fig. 6—Wire and pulp from presses

A temperature of about 180°F. and 20% humidity was finally adopted. For a number of years this method was used for experimental cables but it was impossible to get electrostatic capacities below .095 mf. per mi. With such values it appeared that the use of pulp would be strictly limited to certain sizes of wire and certain cables. Study indicated that lack of porosity and close adhesion of the pulp to the wire were large factors in this difficulty and steps were taken to determine what could be done to improve these values. It was found that by drying the wire at very high speeds by passing it rapidly through high temperatures, the natural shrinkage of the pulp could be greatly reduced and that increased porosity could be obtained. Results on capacitance from such wire were markedly better, and so high-temperature radiant-heat drying was introduced into the process. In this method a box type electric furnace with a heating chamber approximately 26 feet long, 3 feet wide and 8 inches high is used. The wire passes through

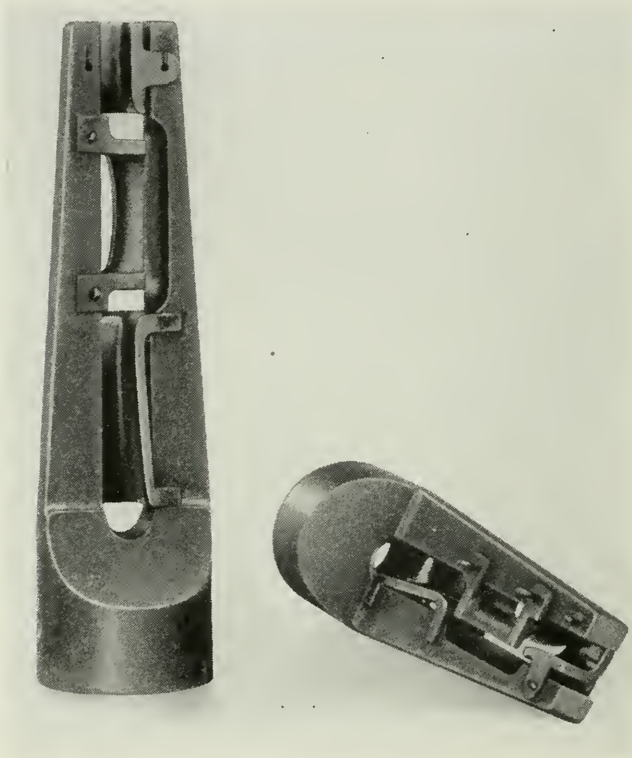


Fig. 7—Individual polisher

this furnace horizontally. In the first third of the furnace 1500°F. is maintained, 1200°F. in the second third, and 800°F. in the last third. The water is literally exploded out of the pulp in this process and leaves a soft porous insulation which is easily stripped from the wire. Electrostatic capacitance values of about .072 mf. per mi. on 24-gauge cables are obtained with this method of drying and improved centering of the wire and roundness of insulation. These values are practically the same as those obtained with wrapped paper. At wire insulating speeds of 140 feet per minute the insulation is dried in approximately 11 seconds. Since in case of a shutdown the wire is immediately burned off, a band of nichrome tape is kept in the furnace at all times so that the wires can be tied to it and carried through for restringing. Broken wires are strung in by tying them to adjacent wires to be carried through.

Reeling up the Finished Wire

Because it is necessary to shift from full to empty spools without shutting down the machine, dual take-up positions are supplied just as dual feed

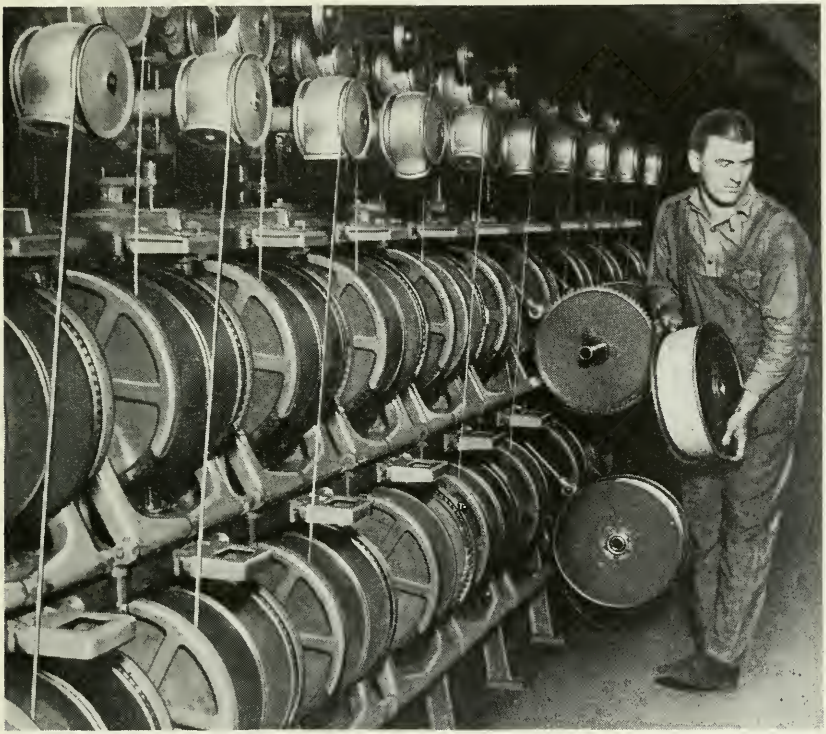


Fig. 8—Insulated wire take-up

positions are used. The take-up spools are rotated through a slipping disc clutch, the pressure on the clutch being controlled by the tension in the wire. Before reaching the take-up spool the wire is passed over a tension drum made up of two capstan pulleys separated by a movable housing enclosing and fastened to a coiled clock spring. By running the wire first around one pulley then reversing it by passing it around a pulley on the spring housing and then around the second capstan pulley any tension variation in the wire causes the spring housing to rotate. The rotation of this housing is communicated by a system of rods to the clutch so that it tends to speed up or reduce the speed of the take-up spool. With the long coiled spring wound to a definite tension a predetermined pull on the wire can be maintained. Two spools are driven simultaneously side by side through suitable gears. Each spool, however, is held on a separate arbor which can be pulled out of mesh with the driving gear so that the take-up spool can be stopped and removed. When it is desirable to do this the wire being taken up is simply switched over to the other spool and when a few turns have been taken up the wire between the spools is cut so that the first spool can be removed from the machine (Fig. 8). Sixty spools are run at one time at an average speed of 140 feet per minute or a total wire footage of 8400 feet per minute of running time. Improved beating, better pulp, better cleaning and improved drying and polishing as well as better trained employees in the last few years have greatly improved the product over the original and simplified the control of the process.

Types of Wire Insulated

As mentioned in the first few paragraphs the trend in telephone cable construction has been toward finer and finer wire. The insulating equipment and process described are particularly well adapted to apply coatings of pulp from six to ten mils in thickness to gauges of wire between 19 and 30-gauge. Changes in the mechanical equipment would be necessary for handling wire finer than 30-gauge or wire heavier than 19-gauge. As little demand for these gauges exists in exchange area telephone circuits, no attempt has been made to adapt the machine to these sizes. However, use of the process can be extended quite widely both in the type of materials used for insulating and kind of wire covered, if demand for such extension exists. So far the development of this insulation process has made it possible to produce wires with insulations so thin that 1515-pair cables of 24-gauge wire and 2121-pair cables of 26-gauge wire are now commercially available to the telephone companies with no increase in external diameter of the lead sheath now used. More effective use of existing underground ducts can therefore be made, eliminating possible large expenditures by the telephone companies for such facilities.

Design and Operation of New Copper Wire Drawing Plant*

A new wire mill for the drawing of copper wire is described. The speeds attained are close to the theoretical limit set by the breaking strength of the wire under the centrifugal stress of winding. The No. 1 machine which draws from rod down to No. 16 A.W.G. and has 10 dies operates at 6000 ft. a minute. The No. 2 machine redraws to finished sizes of No. 19 A.W.G. down to No. 30, possesses 12 dies, and operates at 10,000 to 12,000 ft. a minute. With the single installation at the Western Electric Company at Kearny, N. J., over 2,500,000 pounds of annealed wire are now delivered monthly to the insulating machines for processing into lead covered cable. Part I deals with the design of the machines; Part II with the wire mill installation and operation.

PART I—DESIGN AND OPERATION OF HIGH SPEED COPPER WIRE DRAWING MACHINES

By H. BLOUNT

Introduction

COPPER wire is used extensively in the making of facilities for communication purposes, the Bell Telephone System alone now using over 40 billion conductor feet per year. It is essential that this wire be of high quality with deviations in diameter kept to the minimum so that the apparatus with which it is to be used will function properly.

A study made some years ago showed it would be economical for Western Electric to manufacture its wire, with the possibility of greater production by increasing the speed of drawing. The equipment provided at that time operated at speeds much higher than were then in general use.

A few years later it became evident that the speeds selected were far from the ultimate at which wire could be drawn, and another development was started to determine a practical and economical speed, resulting in the design, construction, and placing into operation of two sizes of wire drawing machines. One, which will draw rod to sizes as small as No. 16 A.W.G., is called the No. 1 and is of 10 die capacity, designed to operate at 6000 ft. per minute. Figures 1, 2, and 3 show the front and rear views of this machine. A second machine for redrawing to finished sizes No. 19 A.W.G., and smaller, is called the No. 2, and is of 12 die capacity, designed to operate at 10,000 and 12,000 ft. per minute. Figure 4 shows the front

* Reprinted, with minor changes, from *Wire and Wire Products*, October 1940. This paper was presented at the Wire Association Convention in Cleveland, Ohio, October 24, 1940, receiving Honorable Mention in Recognition of its Contribution to the Research Literature of the Wire Industry during the Year 1940.

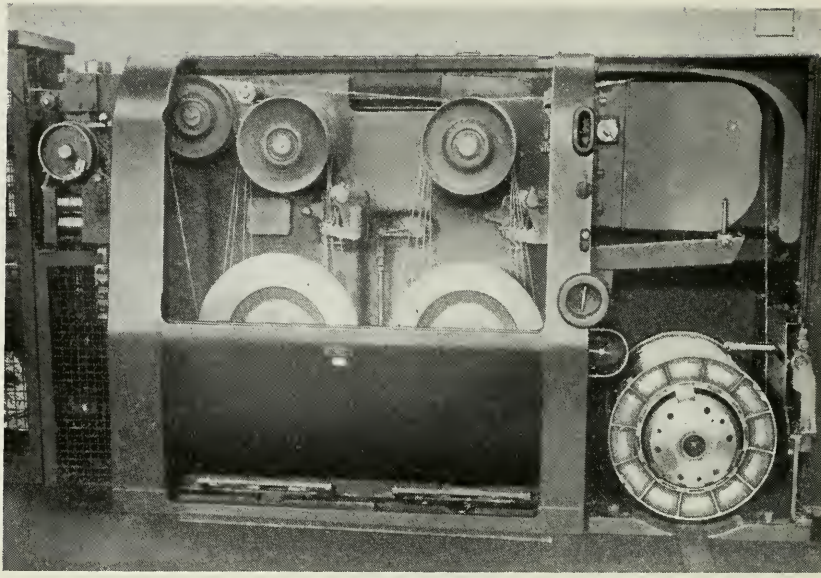


Fig. 1—No. 1 wire drawing machine—front view

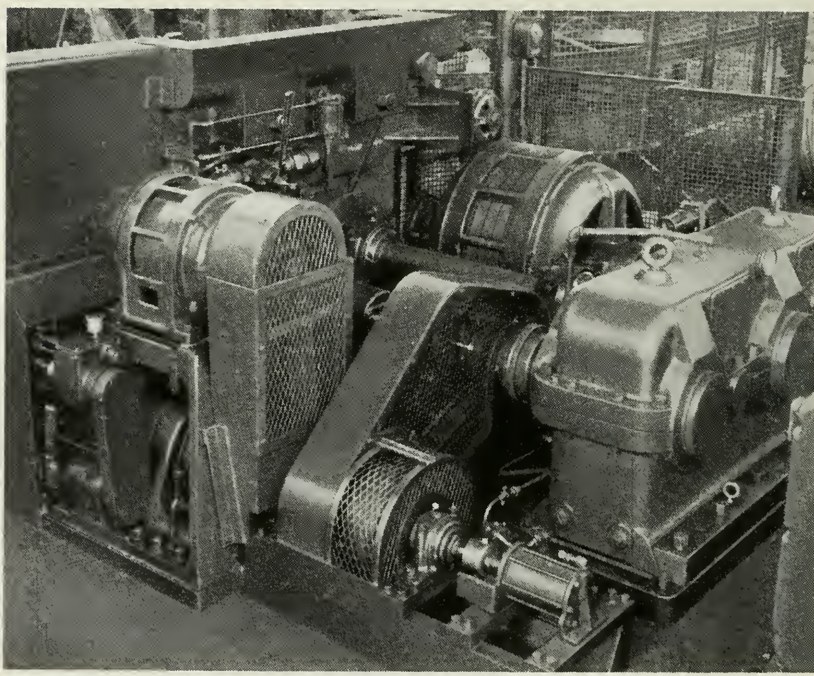


Fig. 2—No. 1 wire drawing machine—rear view

view of this machine. The design features outlined in the following text deal largely with the No. 2 Machine. Similar features are incorporated in the No. 1 Machines and reference is made to changes in design applicable only to that machine.

The Problem

Continuity of operation is essential to higher drawing speeds; therefore, wire should be delivered in as large a unit package as practicable to secure

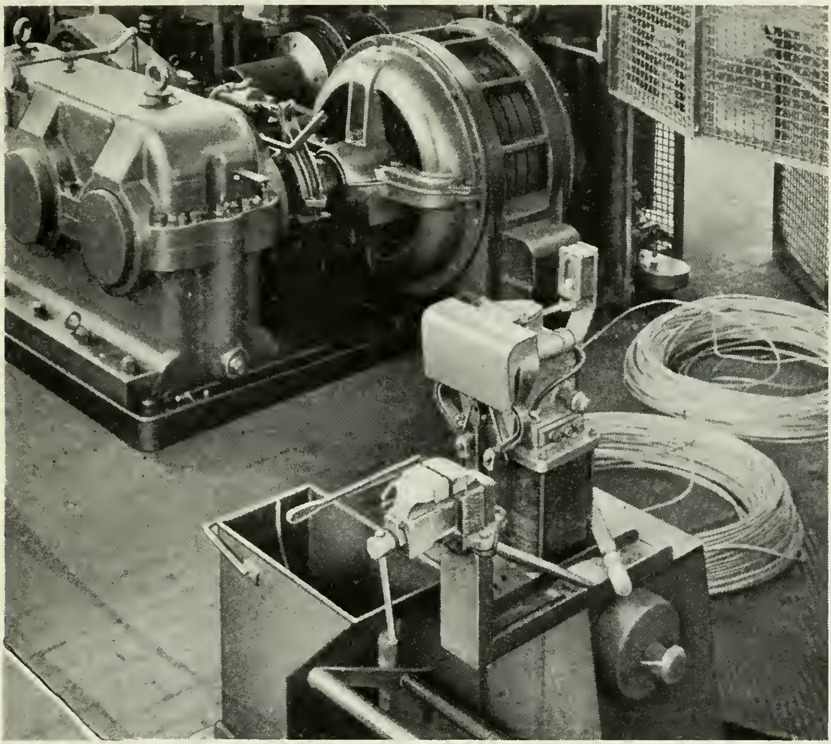


Fig. 3—No. 1 wire drawing machine—rear view

this continuity with a minimum of scrap at the subsequent operations. A survey of the wire using equipment showed that reels as large as 18″ diameter could be used with a capacity of 400 lbs. of wire. This requires that a suitable drive be introduced on the reel takeups of the wire drawing machines to allow for gradual deceleration of speed as the reels fill up with wire.

The drive for the takeup reel should be capable of producing a uniform tension in the wire as taken up for its entire length on the reel. This tension must be controllable for the different sizes of wire, in order that after being annealed it can be easily removed at the subsequent operation.

The application of torque motors on several installations at Western Electric to secure uniform tension in the product being taken upon a reel had demonstrated this to be a very satisfactory form of takeup drive, as the motor will slow down with build up of wire on the reel without changing

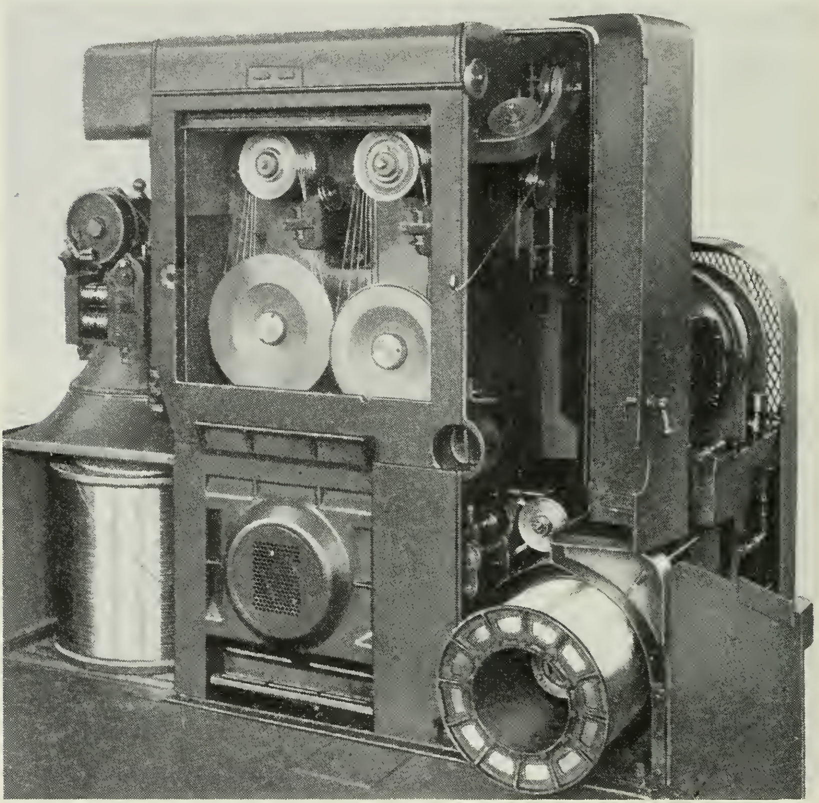


Fig. 4—No. 2 wire drawing machine—front view

the wire tension to any appreciable extent. By changing the stator voltage of this type of motor, the torque can be regulated to give tension suitable for drawing the various sizes of wire within the capacity of the machine. By a proper selection of motors, variations in speed, without undue heating, can be obtained with a ratio of 1–1.8, which is of sufficient range to permit the use of reels of 400 lb. capacity.

The next step was to determine the maximum speed at which wire might be drawn and the speed at which it could be wound onto a reel. The maximum drawing speed was considered to be that speed where the stress set up by centrifugal force would equal the safe stress for copper of 25,000 lbs. per sq. in. A maximum drawing speed of 27,400 ft. per minute was determined by the following calculations:

Let W = Weight of drawn copper wire per cu. in. in lbs. = .3212 (A.I.E.E.)

V = Speed of wire in feet per second.

G = Acceleration due to gravity.

1. Stress in Wire due to Centrifugal Forces = S

$$S = \frac{12 \times W \times V^2}{G} = \frac{12 \times .3212 \times V^2}{32.2} = .1197\ V^2$$

2. Maximum Wire Speed Considering Only Stress Due to Centrifugal Force. The speed at which S would produce a stress of 25,000 lbs. per sq. in.

$$V = \sqrt{\frac{25000}{.1197}} = 456 \text{ f.p.s. or } 27400 \text{ f.p.m.}$$

With the possibility of a range of speed of 1 to 1.8, the stress set up in the reel rim at a wire speed of 27,400 f.p.m. would equal 62,000 lbs. per square inch when the wire is being taken up on the core of the reel, and the rim running 80% faster. Since this speed and resulting stress are above the safe limit for low carbon steel, a speed of 12,000 f.p.m. was selected which provided a factor of safety of approximately five to one. The stresses set up in the wire and reel rims for the various speeds are shown on diagram, Fig. 5.

The horsepower requirement of the torque motor for the takeup is made up of three components:

1. Tension in Wire
2. Bearing Friction for Takeup
3. Reel Windage.

Wire should be taken up on the reel under sufficient tension to offset that created by centrifugal force. The tension in the wire resulting from centrifugal force is shown on diagram, Fig. 6, and is determined by taking the stress in copper wire at 12,000 ft. per minute, Fig. 5, and multiplying this by the area of each size of wire. The tension in the wire changes for each size of wire and remains practically constant throughout the entire reel; therefore, the horsepower required to take up wire on the reel remains constant from the core to the outside of the reels, the speed of the reel slowing down with the build-up of wire on the reel. The lower curves on diagram, Fig. 7, show the horsepower requirements for taking up wire of different sizes.

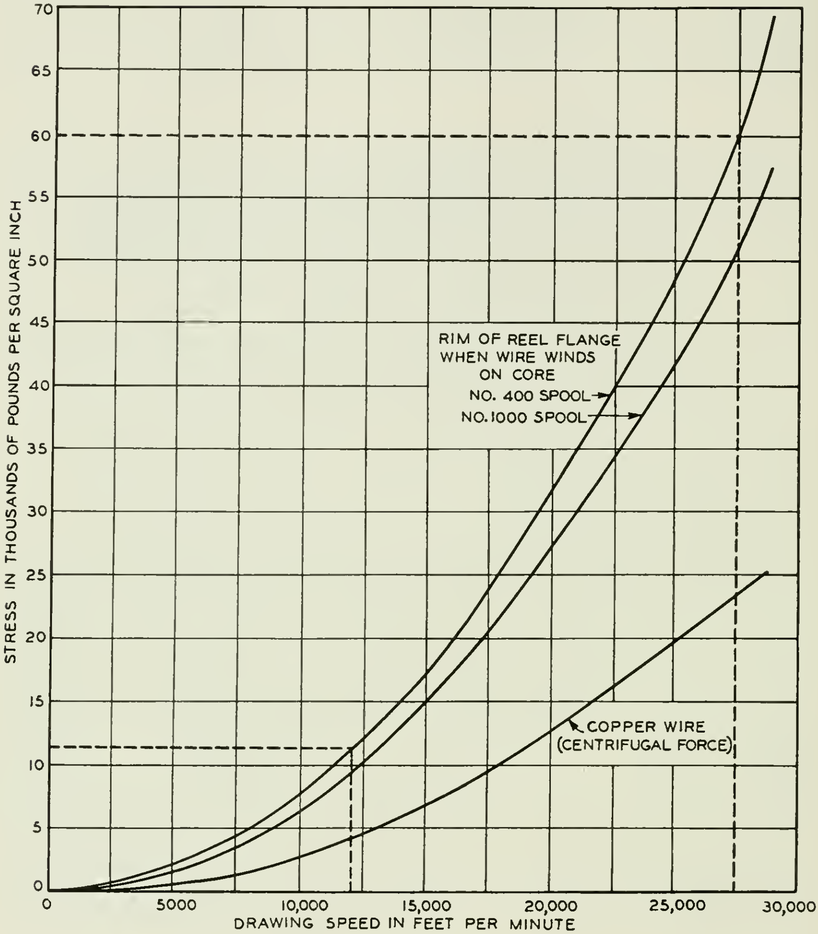


Fig. 5—Stresses produced in wire and takeup reel at various wire drawing speeds

Power Requirement

The horsepower required to overcome bearing friction was calculated and is constant for all sizes of wire. The windage is governed by the design of the reel and the horsepower was determined by test for the minimum and maximum reel speeds. The data for these components are shown by the upper curves of diagram, Fig. 7.

The constant horsepower requirement for uniform tension when converted into torque shows that the torque increases as the wire on the reel builds up due to the lengthening of the radius arm.

The decreasing horsepower requirement to overcome windage and friction when converted into torque shows that the torque decreases with this build-up of wire due to the slowing down of the reel when using a uniform

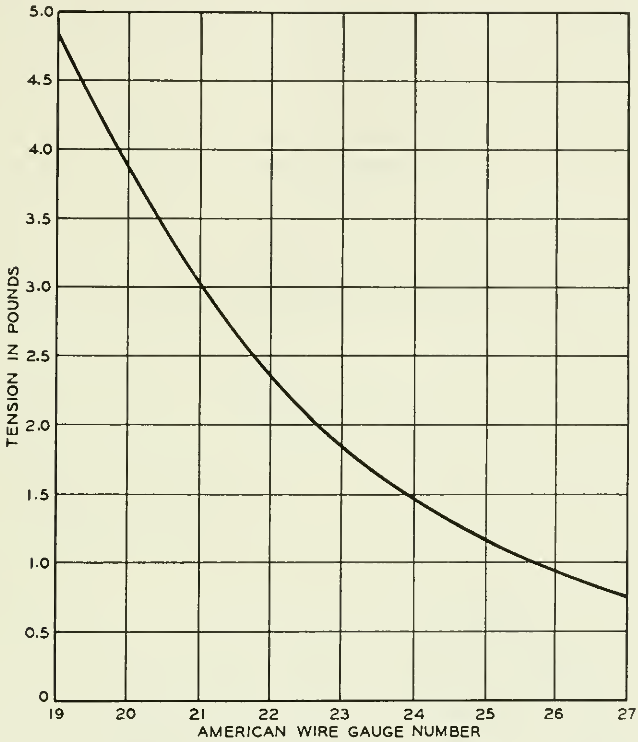


Fig. 6—Tension in wire due to stress set up at wire speed of 12,000 F.P.M.

speed of drawing. This decrease will be at a faster rate than the increase resulting from tension.

The calculated torques when plotted for the different sizes of wire within the scope of the machine show curves gradually diverging as the reel decreases in speed. To simplify the electrical control it was found these curves could be made parallel and still be within the allowable variation of wire tension and the required tolerances of the supplier of the electrical equipment. It was decided to select as the base curve that condition which would be most favorable to the making of the smallest size of wire and to use the average torque value from an empty to a full reel for the different tension requirements for the other curves. Therefore the composite curves as shown by diagram, Fig. 8, show the result of this compromise.

The curves showing the results of the test run of the takeup motors are shown by Fig. 9, which demonstrates how closely the motor manufacturer met the requirements of Fig. 8, which are superimposed for reference.

The minimum of slip between the wire and capstans has been incorporated

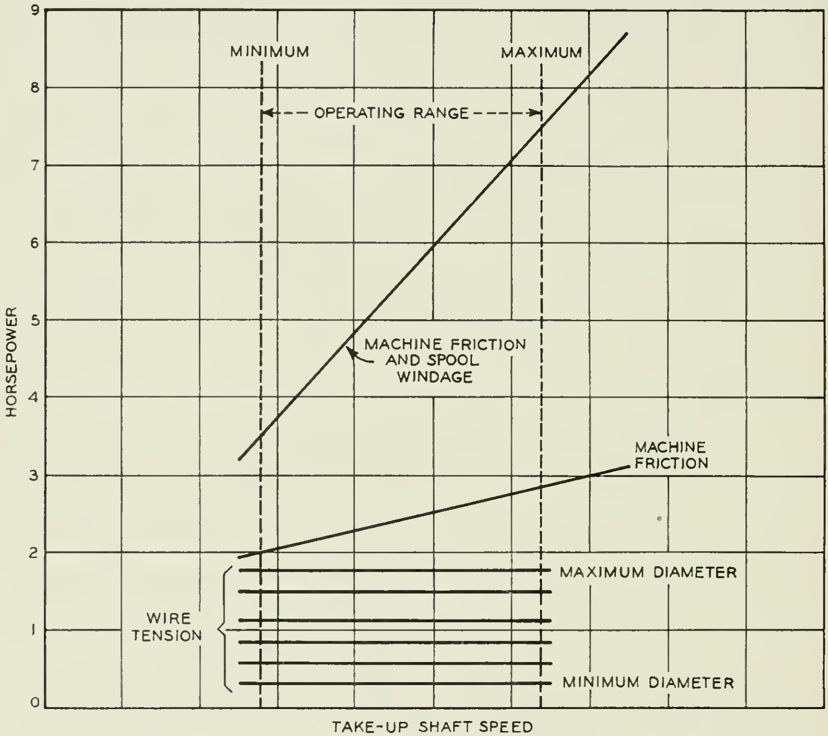


Fig. 7—Power requirements for torque motor for takeup of wire sizes No. 19 A.W.G. and smaller to secure uniform tension

into the design to secure the greatest economy of power. Each reduction of one size A.W.G. increases the length by 26%; and a ratio of 23% between each capstan step has been found most economical. To further reduce power required, ratios of 25% have been used, but because of the uneven wear of diamonds this ratio is disturbed and excessive breaks occur at the location where the die has worn the fastest. With the ratio of 25%, dies must be kept more evenly matched for reduction in area, and the expense of rematching dies, and the loss of production during the period of rematching, make this expense greater than the power charge with the 2% greater slip.

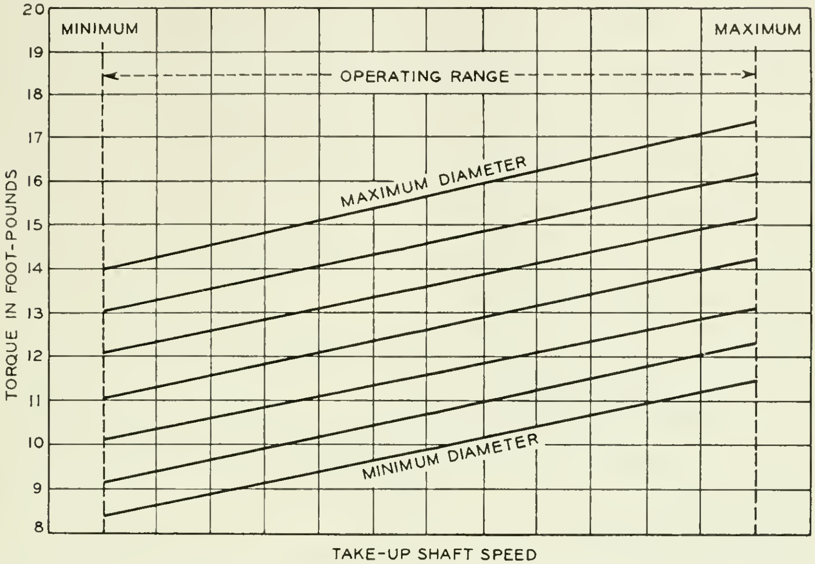


Fig. 8—Torque requirements of takeup motor with tension, friction and windage combined for wire sizes No. 19 A.W.G. and smaller

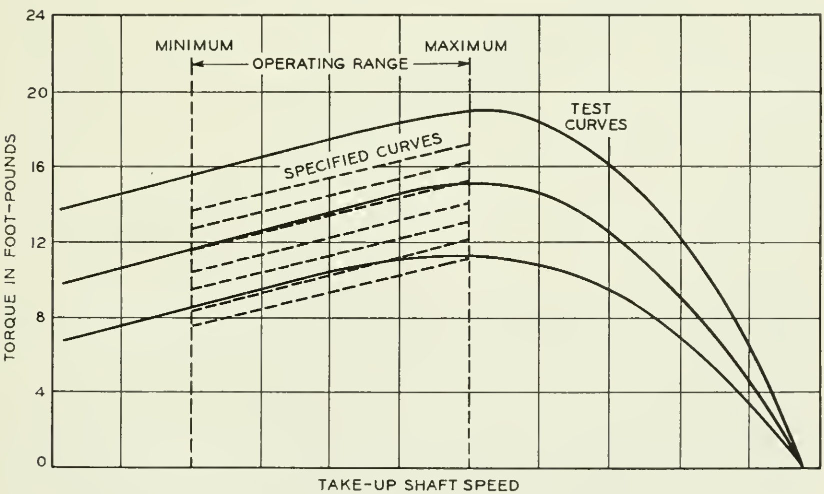


Fig. 9—Speed torque curves for takeup motor

Inasmuch as it is economical to maintain dies within definite ratios of reduction of area, by the same token it is also necessary to keep the diameter of the capstan steps within like proportion.

The die pulls and power required to draw copper wire of any size are determined by referring to the chart, Figs. 10 and 11, applicable for Tungsten Carbide and Diamond Dies respectively.

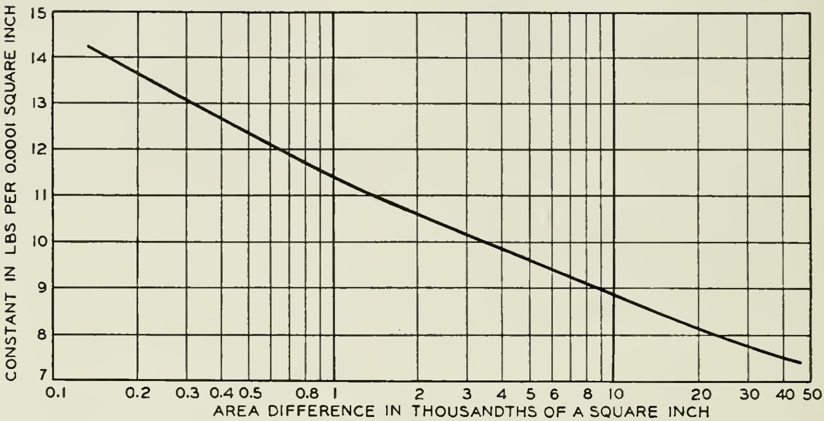


Fig. 10—Constants for determination of die pulls for tungsten carbide dies drawing copper

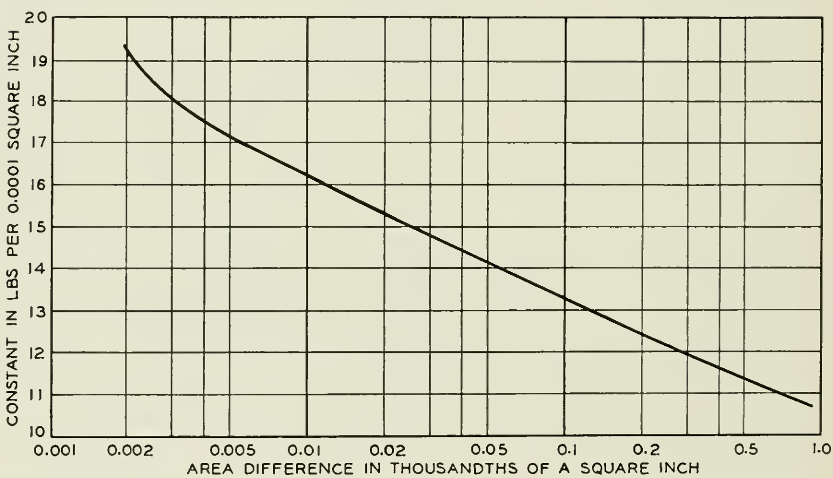


Fig. 11—Constants for determination of die pulls for diamond dies drawing copper

In using these charts the diameters of wire being drawn from and to are selected. The difference in areas represented by these diameters is determined, the curve is then chosen in the range of this difference, and by reading up from this difference to the proper range curve, the constant can be determined per .0001 sq. in. area reduction, which, when multiplied by the difference in area, will give the total pounds pull required to draw to the size selected.

Capstan diameters are determined to secure minimum slip, and from the die diameters selected the pull through the dies is calculated. The horsepower for the main motor is determined from these die pulls and capstan speeds, to which is added machine and motor losses.

As the drawing machine is required to start up and accelerate under full load, a high torque squirrel cage induction motor was selected. This type

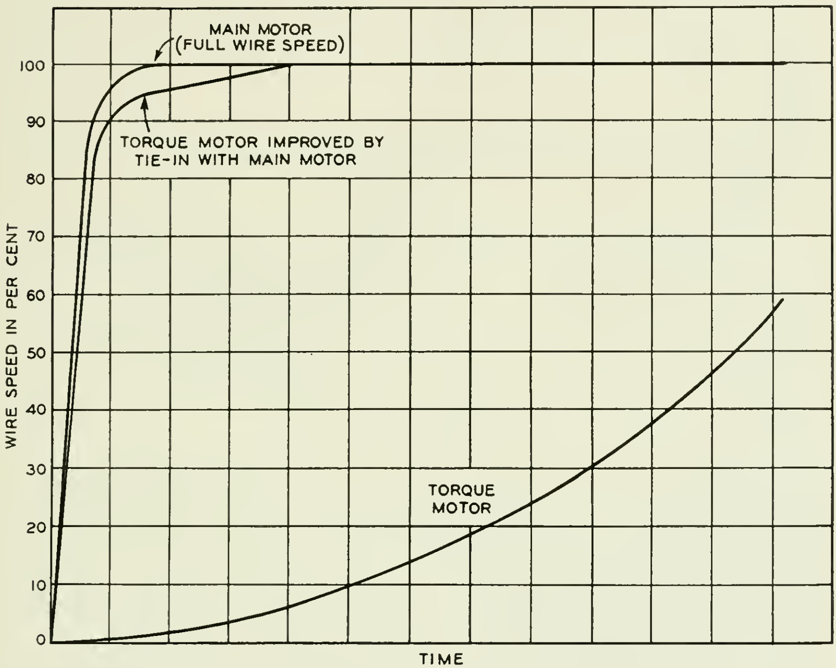


Fig. 12—Relative acceleration curves of main and takeup motors

of motor accelerates to full speed very rapidly, whereas the acceleration of the torque motor for the takeup is very much extended and would therefore result in very high slip between the wire and capstan. It was therefore necessary to introduce some auxiliary means to assist the torque motor to come up to speed. This has been effectively accomplished through the introduction of a magnetic clutch for coupling together the main and takeup motors during the starting period. This clutch is energized as soon as the starting button is operated and before the contactors for the main motor make contact. A time relay releases the magnetic clutch as soon as the main motor is up to speed. During this acceleration period on the No. 2 Machines a slip of 5% between wire and capstan occurs when the capstans are new, and no slip when the capstans are reduced to the minimum diameter. Curves showing the relative acceleration between main and takeup motors are shown on diagram, Fig. 12, which also represents the improvement of acceleration by the tie-in. An electric time clock is connected into the motor circuit for stopping the machine when the 400 lb. reel is full, a time setting being made for each size of wire. On the No. 1 Machine the takeup is accelerated by the magnetic clutch to full reel speed of the 1000 lb. reel and the contact made by the time clock re-energizes the clutch so that the takeup will slow down in synchronism with the main motor.

An under current relay is also interposed in the motor circuit to stop the machine should a break occur while drawing.

Lubrication

Introduction of oil lubrication introduced difficulties in securing effective sealing against oil leakage. It has been our experience that commercial seals are effective when used on shafts revolving at surface speeds below 1200 f.p.m., but above this speed they were inadequate. For the capstan bearings the seals have to be effective in both directions to prevent the leakage of mineral oil from the bearings into the wire drawing compound, and also to prevent the wire drawing compound from mixing with the lubricating oil. This has been accomplished very effectively by the use of multiple slingers, the design of which is shown by Fig. 13. The two front slingers throw off the compound which drains back into the compound system and the two rear slingers do the same with the oil. There is no friction and corresponding wear between surfaces and only occasionally can small drops of compound be seen in the drain reservoir, which shows the effectiveness of this type of seal. As an extra precaution against contamination of oil with wire drawing compound, only a small amount of oil is permitted to flow to the capstan bearings, sufficient for adequate lubrication. This is drained to a reservoir and clarified before re-use.

Another form of seal is shown by Fig. 14. This is used at the takeup arbor where oil was driven through any commercial gasket material by centrifugal force. The oil was thrown out into the inside of the reels and caused discoloration during the annealing. This has been effectively sealed by making a ring of dead soft copper wire carefully joined. The end cap is bevelled to force the ring to the inside of the arbor and against the edge of the bearing, making a tight three-point contact. These rings are never used more than once.

Reels:—The reels are provided with a magazine on the outside of the

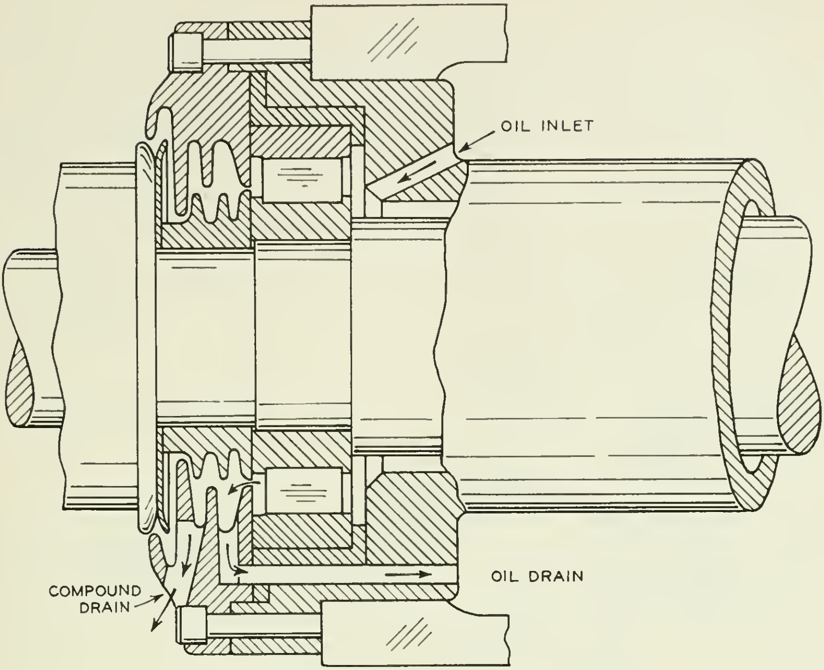


Fig. 13—Labyrinth oil seal

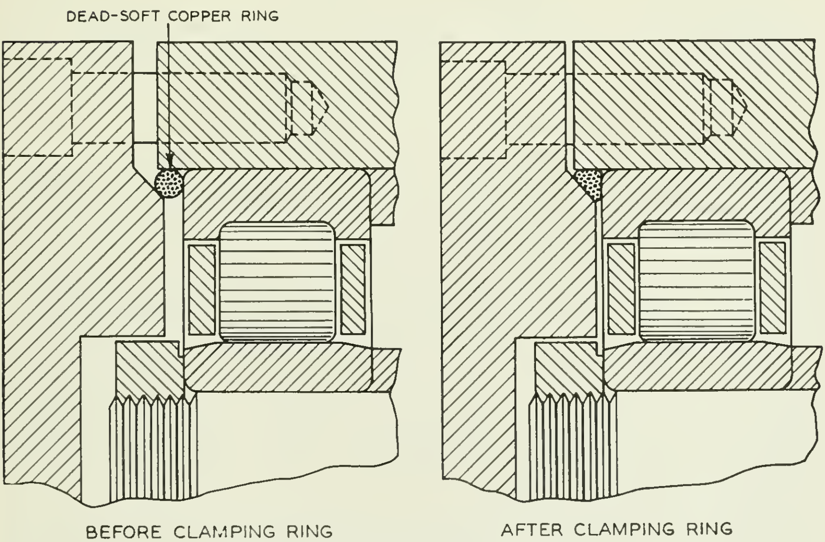


Fig. 14—Assembly of oil retaining ring

flange, in which magazine approximately 10 to 12 feet of the inside end of the wire is stored. At the subsequent operation this inside end is removed from the storage and joined to the outside end on the next reel, by which means continuous production with a minimum of scrap is secured.

Consideration was given to a reel machined all over to get it in correct balance. However, it was realized that distortion would occur as a result of annealing the wire on the reel, and from the handling. Therefore it was decided to provide a very substantial construction of the takeup unit, with bearings of ample capacity to provide for any eccentric loading which might result from spools which had become irregular by use.

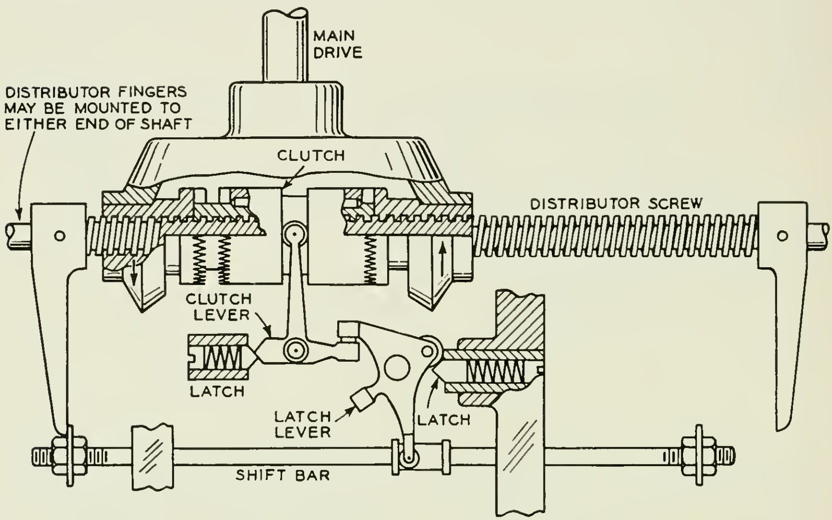


Fig. 15—Diagrammatic view of distributor

Welding facilities are provided to join the ends of wire on successive supply reels at the No. 2 Machines. A special hood permits the transfer of supply from the emptied reel to the succeeding reels. The contour of this hood had to be developed to reduce the noise from the wire which is whipping around. The noise is further minimized by suitable ribbing of the hood, irregularly spaced to break up the frequency of vibration.

A roller conveyor is installed beneath the hood with a capacity of three 1000-lb. reels. These reels are up-ended with magazine down before welding the ends of two reels. After emptying the front reel, the three reels are pushed forward, the empty one being discharged at the front, leaving space for another full reel at the rear.

A turntable is furnished at the supply end of the No. 1 Machine, which can be seen in Fig. 3, with capacity for four coils sufficient for one full takeup reel. In operation, the bottom of the coil of rod is welded to the top of the succeeding coil, thus securing continuity of drawing. With the emptying of one position, the table is revolved through one quarter of a revolution, leaving the empty space for reloading the next coil which is then welded to the one ahead, the welding and loading being performed during the period of drawing.

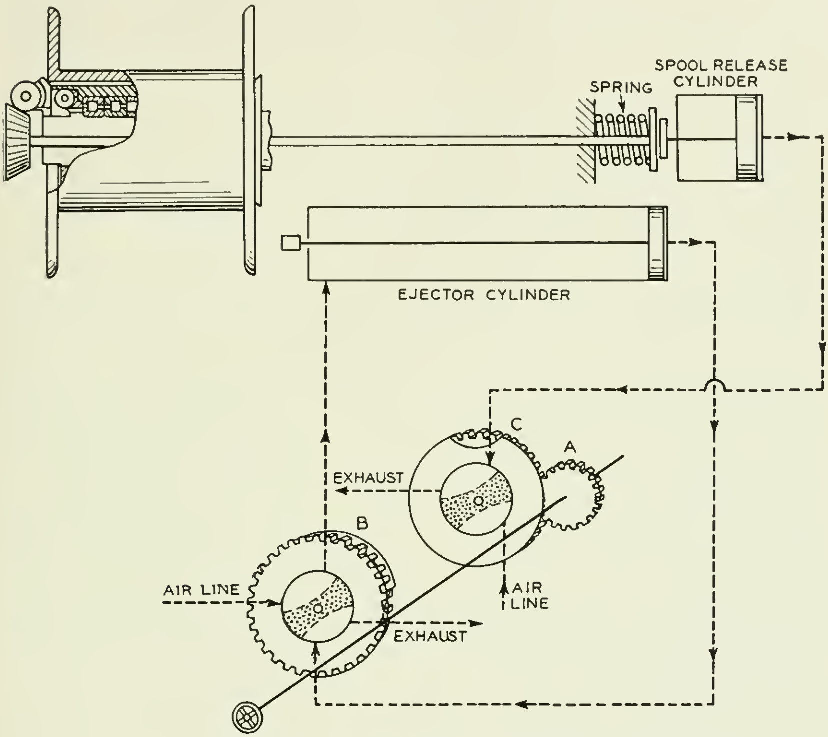


Fig. 16—Diagram of clamping and releasing mechanism

A slow speed stringing block and die support are provided for stringing the larger dies onto the wire and by depression of a foot switch the machine can be operated at slow speed for stringing the dies and wire into the machine.

Good distribution of wire on the reels is essential to permit of easy removal of wire from the reel. The distributor designed for this machine is of the reversing screw type with a reversing clutch as shown by diagram, Fig. 15. At the speed with which wire is being delivered to the reel, any pause at either end of the traverse would result in considerable build up of wire against the flanges. This type of distributor is practically instant-reversing.

Breaking of wire during drawing is not frequent, the average weight of wire on the takeup reel being well over 1000 lbs. for the No. 1 Machine and 300 lbs. for the No. 2 Machine.

Replacements and Safety Features

The design provides for readily replaceable unit assemblies so that the machines are out of production for the minimum of time when any repairs are necessary.

To reduce the effect of vibration to a minimum, the main frame of the machine was constructed to keep as much weight as possible close to the floor and thus secure a low center of gravity. All parts revolving at high speed are given a dynamic balance. Welded construction was not as readily adaptable as castings, and would have been noisier.

Safety features have been incorporated into the design for the protection of the operators. Doors are provided so that the wire is fully enclosed during the drawing process and all revolving parts are amply protected. The clamping of the reel on the arbor is effected through spring pressure, air being used for releasing and ejecting the reel. An interlock is provided between the reel release and ejector as shown by Fig. 16, Air Valve "C" effecting the reel release and air valve "B" controlling the ejector. Only one valve can be operated at a time, and they must be operated in proper sequence. The master control "A" is left in contact with the gear segment of Valve "C" until it is fully opened and the reel released. When "A" registers with the segmental opening, it can be withdrawn and moved over into mesh with segmental gear on Valve "B"; the master Control "A" cannot be disengaged from "B" until the ejector plunger is back in correct position. Additional safety was introduced into the reels by making the flanges of an alloy casting, changing the factor of safety from 4-1 to 8-1.

The use of high speed machinery with large capacities of the takeup unit and introducing the minimum of slip between the wire and capstan has resulted in meeting the performances anticipated from this development.

PART II—EQUIPPING AND OPERATING THE NEW WIRE MILL

By J. E. WILTRAKIS

Alloy and Diamond Dies

The experience gained in operating the older Hawthorne and Point Breeze wire mills demonstrated the importance of providing and maintaining dies of high quality. The hardest materials, alloys such as tungsten carbide and flawless diamonds, are used in these dies.

The alloy dies are used in the No. 1 drawing machine where the wire surface and resulting die wear are relatively small per pound of wire produced.

Diamond dies are used exclusively in the No. 2 machine. Definite problems were solved in maintaining dies to rigid specifications which include correct die contours, a finely polished surface, and definite die pull values.

The cross section of a diamond die, Fig. 17, illustrates the general contour found to be most satisfactory for high speed wire drawing. The approach blends smoothly into the reduction angle where the wire is reduced in diameter one AW gage. The bearing is approximately 40% of the wire diameter. With the use of a contour projector, 100X enlargements of die impressions are periodically made to control the process.

Well graded diamond dust is used to enlarge the hole in the die and for polishing operations. Dust graded by flotation methods, closely checked, offers the best results.

For final polishing 6 micron diameter dust is used. A 30X wide angle binocular microscope is used to check the various stages of die making operations and of inspection as shown in Fig. 18.

The following die pull requirements have been set up for each gage when reducing wire one AW gage size:

AWG Size	Pounds Pull	AWG Size	Pounds Pull
15	75	21	21
16	60	22	17
17	49	23	13.5
18	40	24	11
19	32	25	9
20	25	26	7

After grouping dies of a certain diameter according to the pounds pull required, they are matched into sets for use in the No. 2 drawing machines. Records are kept of the characteristics and output of each die.

The increase in speeds up to 12,000 f.p.m. does not appear to have an appreciable effect on die wear. In other words, the same quality and quantity of wire can be obtained from high speeds as from low speeds if (1) the dies are made to definite specifications, (2) the dies are matched into sets, and (3) the drawing machine factors are the same.

The drawing machines have been designed and are maintained with the view of overcoming some of the serious causes of short die life. Long die life is not only obtained by good die shop practice but also control of the following machine factors; (1) smooth drawing capstans and minimum slip, (2) minimum whip of wire entering dies, (3) adequate lubrication of capstans and cooling of dies, and (4) elimination of foreign particles from the drawing compound.

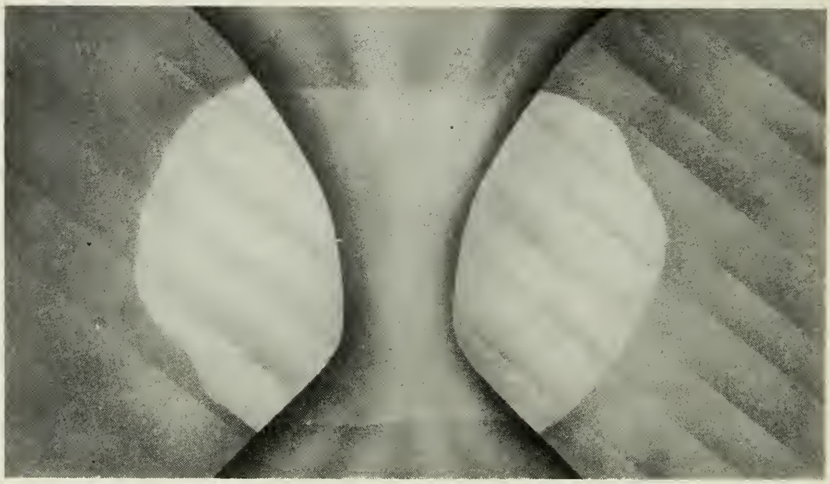


Fig. 17—Schematic showing cross-section of diamond die used in high speed wire drawing

Drawing Compound and Equipment

A one-story building is used for manufacturing wire. In the basement the drawing compound tanks, piping, heat exchangers, pumps, power services and controls are installed. The compound solution used to lubricate and cool the capstans and dies in the drawing machines consists of a homogenized solution of soap, fat and oil mixed with water. This compound returns to a self-cleaning distributing launder in an enclosed steel tank. The launder consists of a pipe with slots evenly depositing the compound over the entire width of the tank. The copper sludge settles to the bottom and the lighter impurities rise to the surface to be held back by a skimmer plate. The clarified super-natant solution rises over a dam into the pump suction chamber to be pumped at the rate of 200 gallons per minute to each No. 1 machine and 100 gallons per minute to each No. 2 machine. The heat from the clean compound is removed in heat exchangers as the compound is delivered to the machines. The compound is maintained at approximately 130°F by a closed recirculating water system thermostatically controlled, Fig. 19.

Layout of Principal Wire Mill Equipment

The building used for wire drawing is ideally situated adjacent to the cable manufacturing unit and has facilities for water, rail and motor truck

Fig. 18—Microscopic examination of diamond die polish by die maker

deliveries. The area is easily ventilated and has excellent illumination provided by mercury vapor lamps, close to a high ceiling yet providing an average of more than 20-foot candle illumination. Stroboscopic effect is practically eliminated by staggering the lamps over separate phases of the three-phase circuit.

The No. 1 machines are located adjacent to the copper rod receival area. The No. 2 drawing machines, nine of them in a row, are placed in the center of the building. Along the wall, five annealing bases for the electric bell type furnace are located. A bridge type crane handles all the material between the No. 2 machines, annealing and inspection. This layout, Fig. 20, of the equipment makes possible quick and easy transfer of material

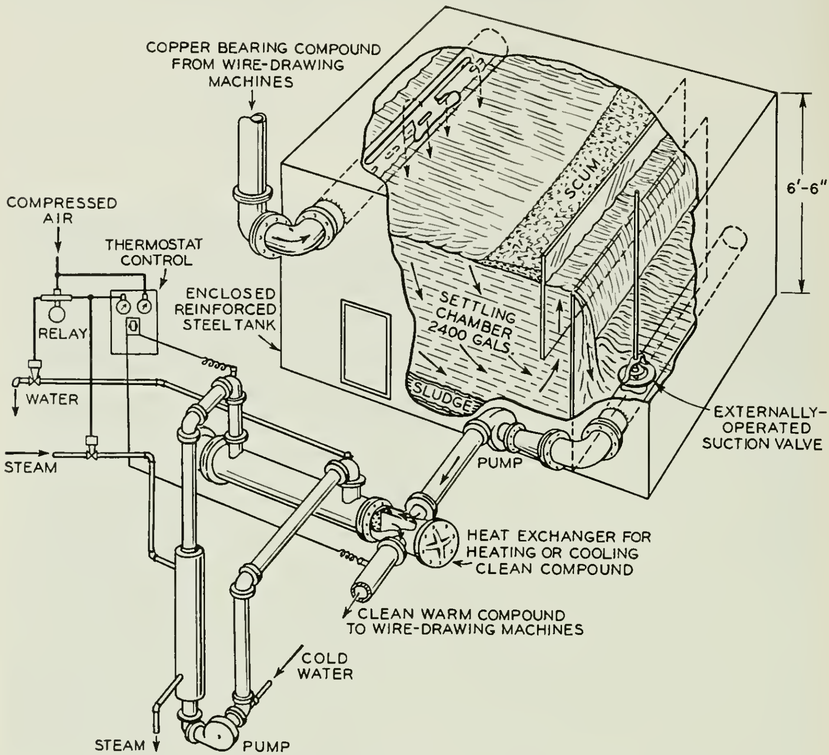


Fig. 19—Sectional sketch of compound system

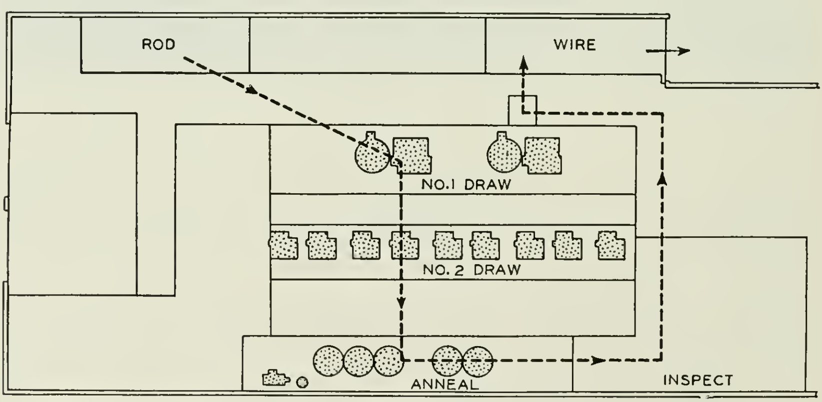


Fig. 20—First floor layout of wire mill

between operations, low inventories and multiple operation of machines by operators. Capacity can be increased without rearrangement. Adequate space has been provided to facilitate maintenance. The entire distance from the receival of rod to the wire shipping area is 100 feet. One electric truck and the crane just mentioned, suffice to handle and transport all materials in the building.

On either side of the main flow of material, space is provided for storage of rod, shop maintenance machines and racks.

Processes in the Wire Mill and Flow of Material

Copper rod is delivered on double prongs of an electric truck, approximately 4,000 pounds at a time, and is placed adjacent to each of the No. 1 drawing machines, Fig. 21. Here each 250-pound coil is placed on the floor of the eight-foot diameter supply table. A maximum of four coils is maintained on the table at a time. The rod ends are electro-welded to form a continuous supply. As rod from one coil is converted to wire, the operator pushes a button and rotates the table 90° to locate the next coil. This process of supplying coils, welding rod ends and rotating the supply table is repeated while the machine continues to fill the 1000-pound reel with 14 gage (.064″) wire at 5000 f.p.m.

When the machine automatically stops, the operator opens the spooler compartment and actuates an air operated mechanism which releases and pushes the two-foot diameter 1000-pound reel off the take-up arbor. An empty reel is placed on the arbor and locked. The guard is closed and the push button starts the machine with no additional attention on the part of the operator, who returns to the welding operation after placing the filled reel in the storage area.

The 1000-pound reel must be up-ended before it can be placed under the supply compartment of the No. 2 machine. The up-ending device, Fig. 22, consists of two floor castings, a pneumatic hoist and cables. The operator first rolls the large reel on the first floor casting and then actuates the pneumatic hoist. The cables hinge upward two castings like covers of a partly closed book, forming 45° angles with the floor. At this position the weight of the reel settles onto the second casting. The operator releases the air and the reel is gently lowered upon floor rollers. The axis of the reel is now vertical.

One end of the copper wire is electro-welded to the wire end of one of the two reels in the supply compartment. As the machine empties the first reel, the operator pushes the second and third reels into the supply position within a compartment. A continuous supply is thereby provided with safety and ease of handling.

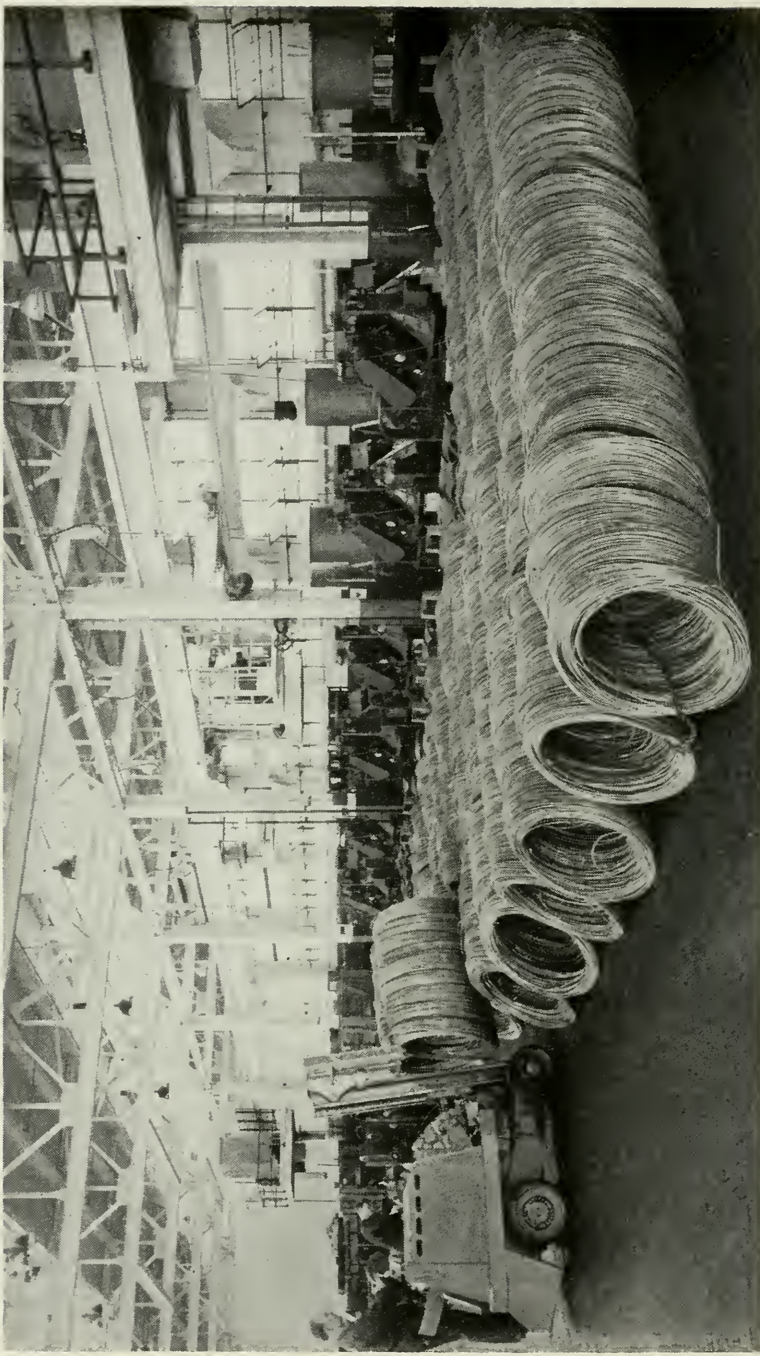


Fig. 21 — Electric truck delivering copper rod

The duties of the No. 2 machine operator, Fig. 23, principally consist of furnishing several machines with supply wire, removing filled reels of drawn cable wire, gaging wire, starting the machines and periodically adjusting for tension. Breaks are infrequent as evidenced by the fact that the average weight of reels shipped was over 340 pounds. When these

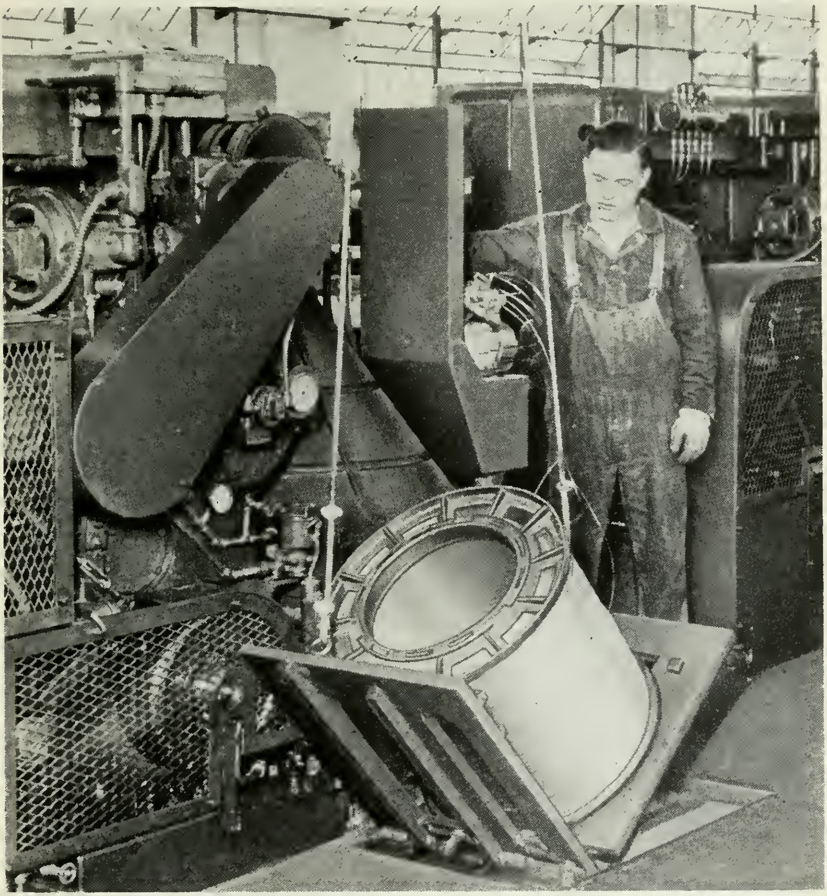


Fig. 22—Up-ending device for 1000-pound supply reel at back of No. 2 machine

breaks occur or when a change is made in the die sets, this operator also strings up the machine.

On these machines wire is drawn at 10,000 f.p.m. The importance of the various mechanical and electrical details mentioned in the first section of this paper can therefore be visualized. One of the No. 2 machine has, for the past year, operated with certain refinements at the finishing speed

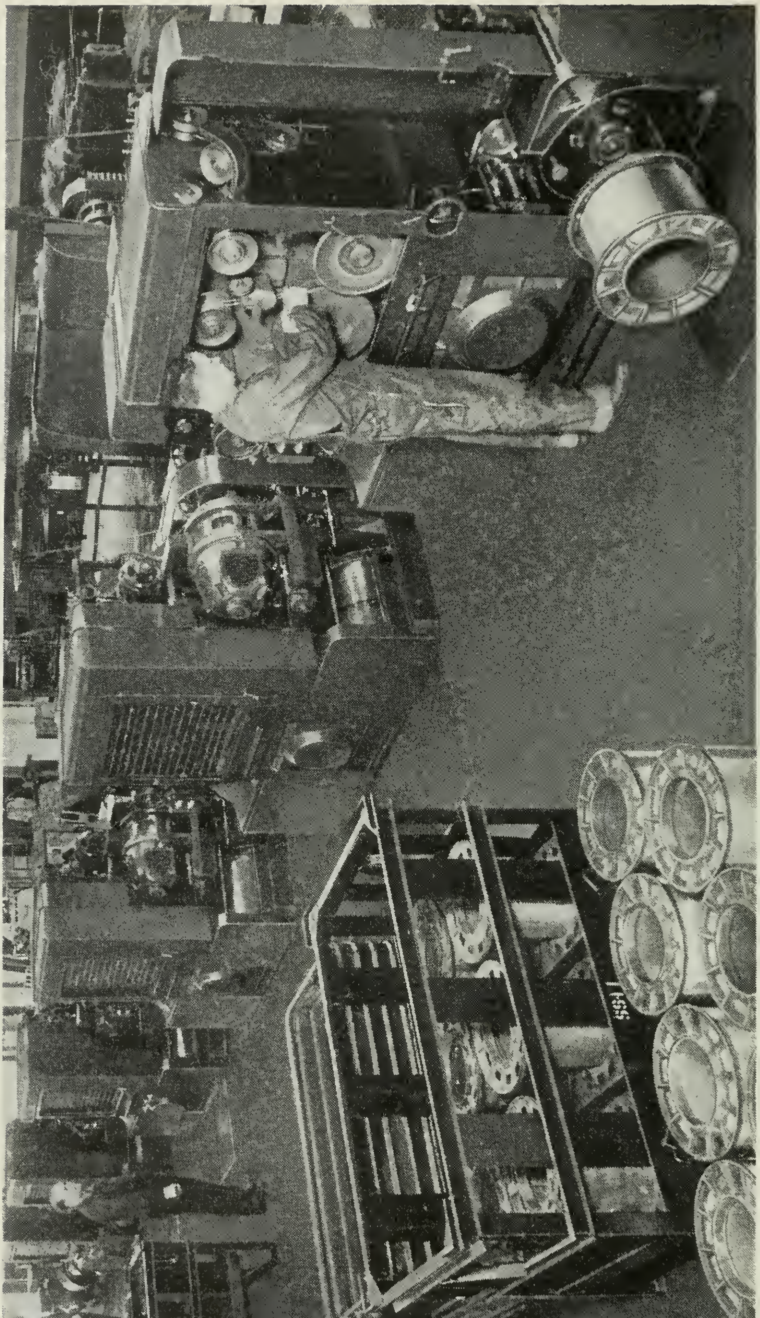


Fig. 23—Line-up of No. 2 machines

of 12,000 feet. The data being collected so far are favorable and it is expected that this study will justify the conversion of additional machines to the higher speed.

After the take-up reel is released and pushed off the arbor by the air operated mechanism, it is rolled to the area below the bridge crane and up-ended by hand.

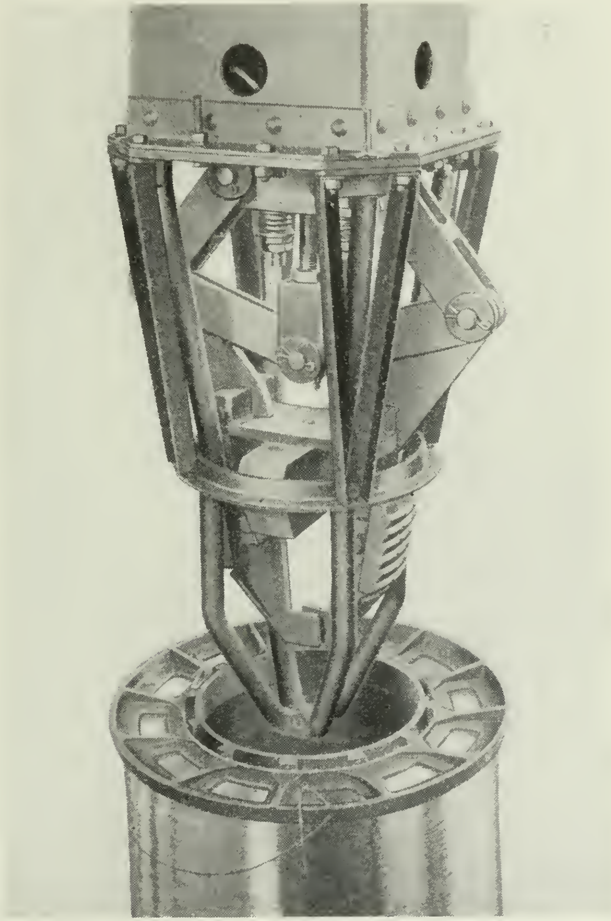


Fig. 24—Solenoid operated chuck grapple being located to lift reel of wire

Multi-Purpose Crane

The movements of the bridge crane and grapples are controlled from the crane cab by the operator. A six-ton grapple handles the baskets of wire, the electric furnace bell and the furnace details. The crane is also equipped with an auxiliary hoist to lift the wire reels into the annealing basket. To this hoist, a locating device has been attached together with solenoid operated, internal expanding jaws which engage in the wire reel core, Fig.

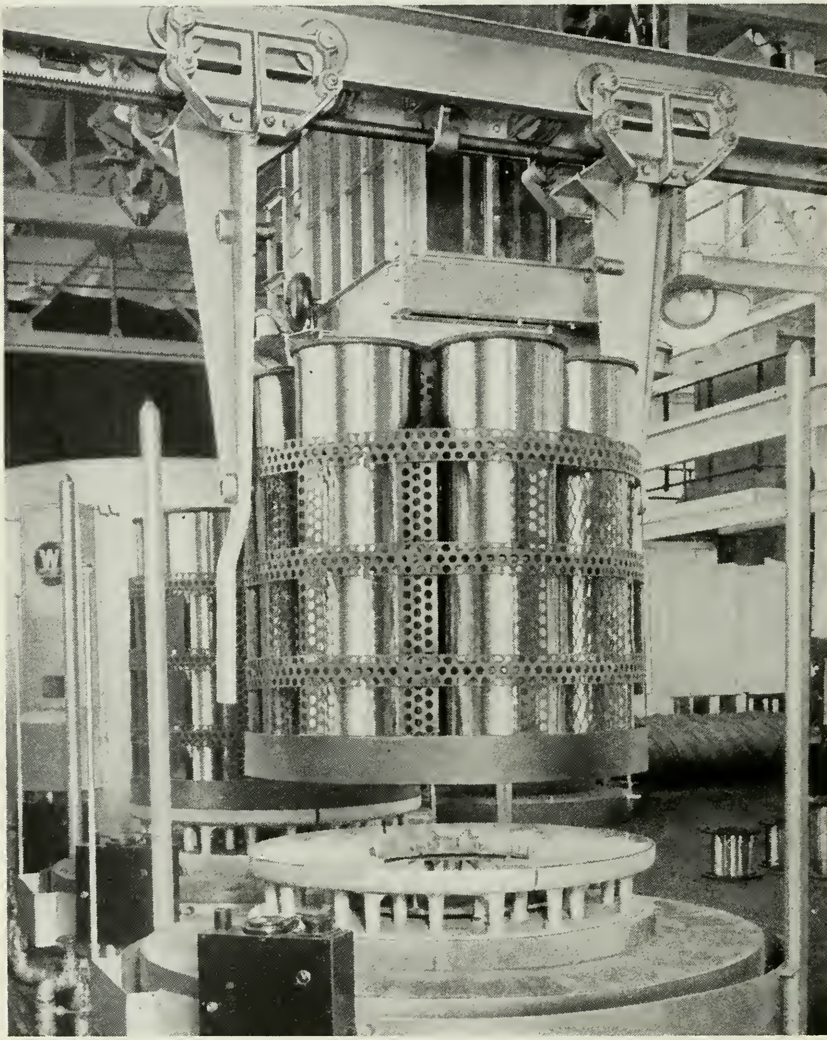


Fig. 25—Six-ton crane grapple placing basket of wire on annealing base

24. With safety and facility of operation, twenty-eight reels, a total of 10,000 pounds of wire, are loaded into a 56″ diameter light-weight perforated steel basket, Fig. 25.

Batch Type Electric Annealing Furnace

The operation of electric batch type annealing furnaces and the use of reducing gas atmospheres, with an average composition of about $1\frac{1}{2}\%$ CO, 2% H_2 and 14.5% CO_2, produced by combusting city gas, are generally known to the wire industry. Certain provisions in the Kearny installation may be of interest.

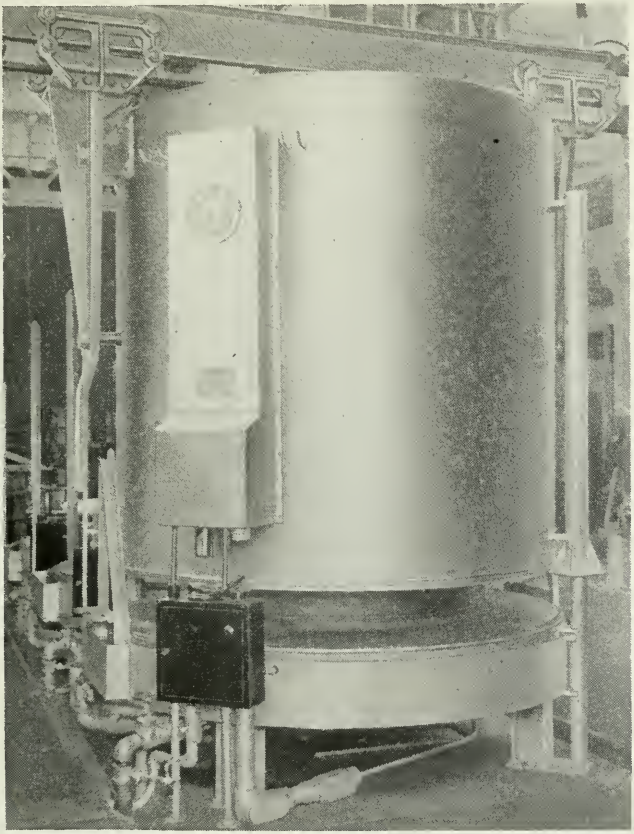


Fig. 26—Electric furnace bell, with automatic plugging equipment, being placed over covered charge

Details of the annealing baskets and bases, the steel alloy retorts used to cover and water-seal the charges, and the electric furnace moved from base to base have been designed so that the arms of the six-ton crane grapple can engage, handle and move all these items.

The electrical connections to the furnace bell are made automatically.

This design consists of control and power plugs located on the exterior of the furnace and a floor stand with positions for electrical receptacles. Two pins align these units, one of which opens the receptacle covers as the furnace is lowered over the retort, Fig. 26.

Features such as these make it possible to perform all the furnace and crane operations, to deliver wire to the inspection area, to load skids for shipment of wire with a minimum of effort on the part of the operator. He attends these operations from a crane cab and as required, operates and adjusts the gas, water and drain valves from floor positions.

In the event of power, gas or generating equipment failure, automatic indicating equipment summons the operator who then connects an 8%

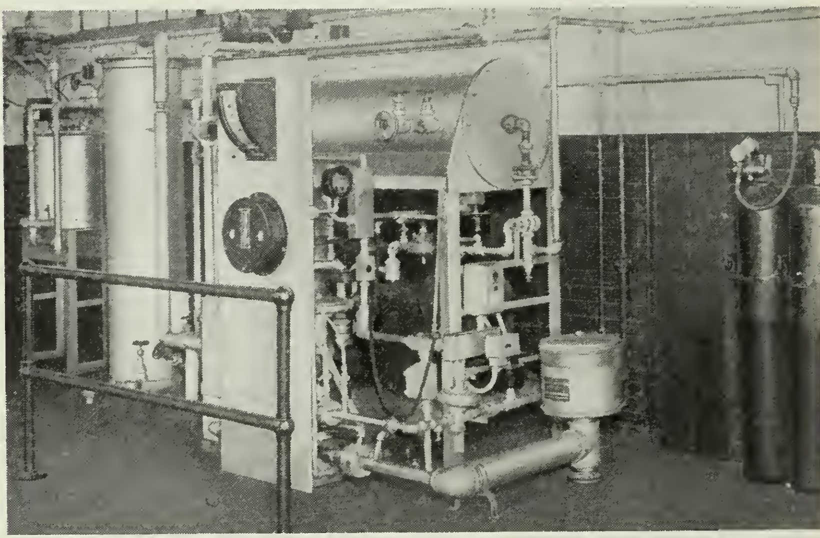


Fig. 27—Gas generator and N_2 tanks used as standby equipment

H_2-92% N_2 mixture into the annealing gas lines. This has provided inexpensive stand-by equipment and constant production of bright annealed copper wire, Fig. 27.

Capacity and Results

This mill is set up to produce wire on a three-shift basis. The equipment of the type described, including space for rod and wire storage, occupies an area of approximately 14,000 square feet. Training time was not excessive for the average operator—efficiencies of 80% to 90% being attained in a few months. Rotation of operators to the next shift every two weeks has worked out satisfactorily.

Periodic checks and maintenance of electrical circuits and apparatus, with adjustments of mechanical assemblies before major repairs arise, have kept repair costs low. Additional training and experience should further reduce maintenance and repair time. The use of diamond dies in diameters

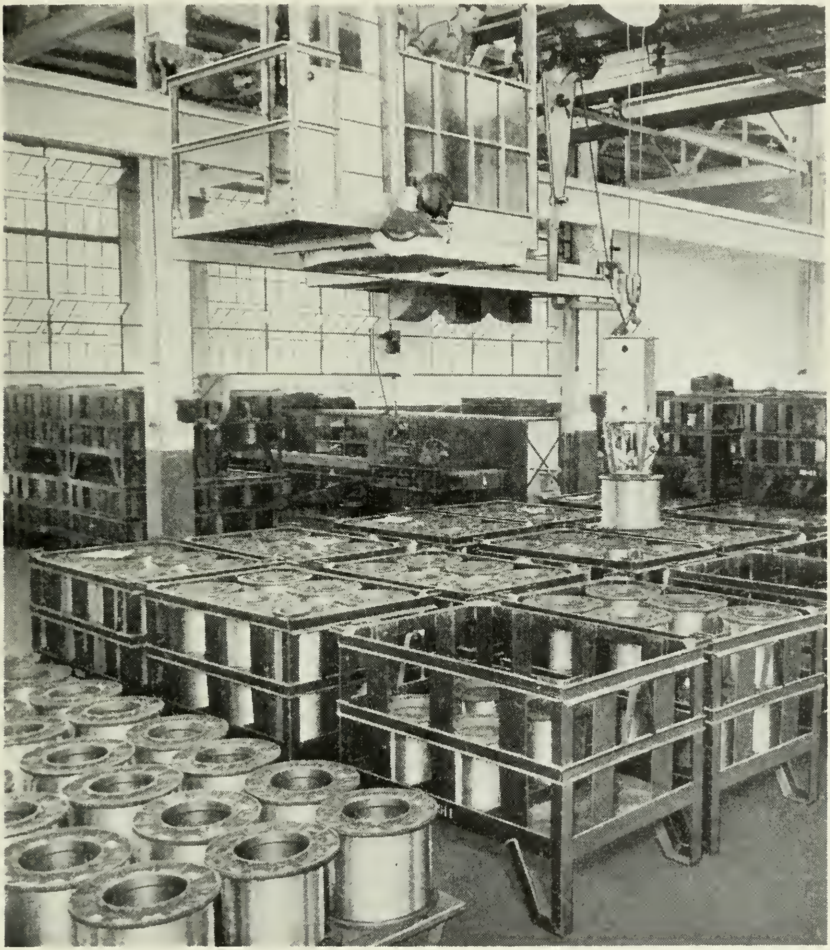


Fig. 28—Inspection and shipping area

up to and including 15 AWG has been found economical for high-speed machines. Cracked dies are negligible when properly mounted clear stones are used. After the first year of operation, the Wire Mill has bettered the anticipated performance objectives.

Acknowledgements

In setting up and operating this project at Kearny, the engineering group responsible was greatly assisted by the Wire Mill experience and developments at the Hawthorne and Point Breeze Works and by the recommendations and designs of the material handling and factory planning engineers at Kearny. This cooperation, together with that obtained from the men on the machines and the maintenance groups, has been reflected in the results.

Abstracts of Technical Articles by Bell System Authors

Two papers by Reverend Thomas Bayes—A facsimile publication from the Philosophical Transactions, Vol. LIII, for the year 1763. This facsimile has been prepared under the direction of W. Edwards Deming, Senior Mathematician of the Bureau of the Census, Washington, from a copy of the Philosophical Transactions in the possession of the Naval Observatory in Washington. An interesting foreword to the volume has been supplied by Edward C. Molina of the Bell Telephone Laboratories. The volume is available at the Department of Agriculture, Washington, D. C., price $1.00.

The Subjective Sharpness of Simulated Television Images.[1] M. W. BALDWIN, JR. Small-sized motion pictures, projected out of focus in simulation of the images reproduced by home television receivers, are used in a statistical study of the appreciation of sharpness. Sharpness, in the subjective sense, is found to increase more and more slowly as the physical resolution of the image is increased. Images of present television grade are shown to be within a region of diminishing return with respect to resolution. Equality of horizontal and vertical resolutions is found to be a very uncritical requirement on the sharpness of an image, especially of a fairly sharp one.

Synchronized Frequency Modulation.[2] W. H. DOHERTY. Probably the foremost practical problem in FM transmitter design is that of stabilization of the mean or carrier frequency. Crystal stability is required, but the direct use of a crystal would necessarily give rise to a conflict between the factors which stabilize the frequency and those which are to produce the desired variation.

In Synchronized Frequency Modulation, which makes its first appearance in the 1000-Watt Western Electric 503A-1 Radio Transmitting Equipment, this problem is solved by associating the crystal indirectly with the system in a monitoring role which ignores the rapid frequency variations due to modulation and responds only to variations in the mean frequency. This is done by taking a sample of the output of the frequency-modulated electric oscillator and shrinking the spectrum down through a succession of frequency dividers to about 1/8,000th of the transmitted carrier frequency. It then consists of a strong central carrier (about 5,000 cycles) with a few degrees of phase modulation. This is then compared with a

[1] *Proc. I.R.E.*, October 1940.
[2] *Pick-Ups*, August 1940.

crystal standard (likewise about 5,000 cycles) in a device which produces a rotating magnetic field at the difference frequency. An armature which follows this field controls the tuning condenser of the original electric oscillator, coming to rest when exact synchronism is attained. The small phase vibrations accompanying modulation are not followed because of the inertia of the system.

The stability thus obtained for the mean frequency is identically that of a crystal oscillator. Since the actual control is mechanical, no sustaining voltage is required, so that failures in the control system do not result in sudden departures in frequency. Mechanical control, moreover, completely relieves the modulating elements of any connection with the stabilization of the mean frequency, so that the modulation range is not restricted. This and other refinements in design permit frequency excursions of hundreds of kilocycles with extremely low distortion.

Ultra-Short-Wave Transmission Over a 39-Mile "Optical" Path.[3] C. R. Englund, A. B. Crawford, and W. W. Mumford. Continuous records of ultra-short-wave transmission on wave-lengths of 2 and 4 meters, over a good "optical" path, have shown variations in the received signal strength. These variations can be explained as being due to wave interference; an interference which varies with the changes in the composition of the troposphere.

Some of the variations are due to changes in the dielectric-constant gradient of the atmosphere near the earth. Other variations are explicable in terms of reflections from the discontinuities at the boundaries of different air masses. The diurnal and annual meteorological factors which affect the transmission are discussed.

A Decade of Progress in the Use of Electronic Tubes. Part I—In the Field of Communication.[4] S. B. Ingram. The dependency of the art of communication on the science of electronics is so great as to make a review of progress in electronics almost of necessity a review of the field of communications itself. While it is true that the early forms of telephone and radio communication advanced to a degree without the use of electronic devices as we know them today, the recognition of the vacuum tube as an amplifier and generator of high-frequency alternating currents in the years just preceding the first World War marked a turning point in the development of the communication art. From that day to this the progress of electronics and communications has gone hand in hand. The need of the communications engineer for new electronic tools has kept him continually

[3] *Proc. I.R.E.*, August 1940.
[4] *Electrical Engineering*, Transactions section, December 1940.

urging the electronics engineer to improve old devices and to originate new ones, and each time the efforts of the latter have been rewarded with success the fruits of his work have been immediately applied to produce new and more startling miracles of long-distance communication.

Because of the close relationship of electronics and communications it is necessary in reviewing the progress of the last decade to keep in mind that it is progress in electronics and not in communications which is our theme. It will be necessary to survey the trends in communications during the period under review, but then it will be necessary to ask to what extent the progress which has been made is due to advances in the electronic field and what advances in the electronic devices themselves have laid the foundation of this progress. There has been no attempt made to make this review comprehensive in the sense that it include all items of progress which are of individual interest. To do so would make it merely a catalog of these many advances and an index to the periodical literature of the subject. Rather the object has been to trace the most significant trends of development in the various fields and to emphasize those lines of advance which appear to be most closely related to the general direction of progress in the several fields of electrical communication.

The Location of Hysteresis Phenomena in Rochelle Salt Crystals.[5] W. P. Mason. Measurements of the elastic properties of an unplated crystal, the piezoelectric constant f_{14}, and the clamped dielectric constant of a Rochelle salt crystal show that practically all hysteresis and dissipation effects are associated with the clamped dielectric properties of the crystal. A theoretical formulation of the equations of a piezoelectric crystal has been made which takes account of the dissipation effects. The formulation is given for the polarization theory. The frequency variation of the clamped dielectric constant when interpreted by Debye's theory of dielectrics, modified to take account of hysteresis losses, indicates that there are two components, one of which has associated with it a high viscous resistance, whereas the other one does not. The non-viscous component has a dielectric constant of about 100 at 0°C and is probably due to the displacement of the ions in the lattice structure. The viscous component has a dielectric constant of about 140 at 0°C and is probably due to the dipoles of the Rochelle salt. Both components have higher dielectric constants and hysteresis between the Curie points indicating a cooperative action of the molecules for both components in this temperature region.

A New Broadcast-Transmitter Circuit Design for Frequency Modulation.[6] J. F. Morrison. The problem of generating wide-band frequency-modu-

[5] *Phys. Rev.*, October 15, 1940.
[6] *Proc. I.R.E.*, October 1940.

lated waves is first reviewed in order to ascertain specifically the desired performance capabilities for a commercial transmitter circuit. The factors which influence or limit these performance capabilities in the two methods available for the generation of frequency-modulated waves, compensated phase modulation, and direct frequency modulation, are then explored. It is found that each method possesses desirable fundamental characteristics not present in the other, but with the circuits now generally employed with either method the modulation characteristics and carrier frequency stability are interrelated so that one has a limiting effect upon the other.

A new circuit is described in which these two important characteristics are independent of each other. Owing to this independence and to other circuit refinements the modulation capabilities are unrestricted with low distortion over an exceedingly wide range.

A balanced electric oscillator operating at one-eighth the radiated frequency is modulated by balanced reactance-control tubes and negative feedback is used to minimize amplitude modulation and harmonic distortion. A system of *frequency division* is employed together with a crystal-controlled oscillator and synchronous motor in such a manner as to control mechanically the mean frequency of the modulated wave with the same stability as that of the crystal-controlled oscillator. The carrier, or mean, frequency stability is that of a single crystal-controlled oscillator and is independent of any other circuit variations. A carrier frequency stability of 0.0025 per cent is possible without the use of temperature-controlled crystals or apparatus.

Neutron Studies of Order in Fe-Ni Alloys.[7] F. C. Nix, H. G. Beyer and J. R. Dunning. Neutron transmission measurements are used to study order in Fe-Ni alloys. The difference in neutron transmission between fully annealed and quenched alloys when plotted against the nickel content displays a broad peak around Ni_3Fe and falls to vanishingly small values near 35 atomic per cent Ni and pure Ni. The higher the degree of order the greater the neutron transmission. The substitution of 2.3 atomic per cent Mo or 4.1 atomic per cent Cr for Fe in the annealed 78 atomic per cent Fe-Ni alloy caused a decrease in the neutron transmission, relative to the annealed 78 atomic per cent Fe-Ni alloy, of 15.6 and 21.2 per cent, respectively. The cold working of an annealed binary 75 atomic per cent Ni alloy, a treatment known to produce disorder, gave rise to a decrease of 20.6 per cent in neutron transmission. These results demonstrate that neutron techniques serve as a useful tool to study order in Fe-Ni alloys, and suggest that they can be extended to study other solid state phenomena.

[7] *Phys. Rev.*, December 15, 1940.

Contributors to this Issue

Harry Blount, educated in England. Western Electric Company, 1903–. Mr. Blount has been engaged in the rolling of copper rods, drawing of wire, and application of protective coatings on wire.

H. F. Dodge, S.B., Massachusetts Institute of Technology, 1916; Instructor, Electrical Engineering, 1916–17; A.M., Columbia University, 1922. Engineering Department, Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Earlier associated with the development of telephone instruments and allied devices, Mr. Dodge as Quality Results Engineer is now engaged in quality assurance work, particularly the application of statistical methods to inspection and quality engineering.

John S. Little, S.B., Chemical Engineering, Massachusetts Institute of Technology, 1915; Operating Engineer, DuPont Company, 1915–1916; Instructor, Chemical Engineering Practice, Massachusetts Institute of Technology, 1916–1917; Lieutenant and Captain, Chemical Warfare Service, 1917–1919; Research Engineer, Pulp and Paper, Brown Company, Berlin, N. H., 1919–1922. Western Electric Company, 1922–. Mr. Little has been engaged continuously in development work in cable since joining the Western Electric Company. Recently, wire drawing and vacuum tube manufacturing development have been added to his activities.

H. G. Romig, A.B., Pacific University, 1921; University of Washington, 1922; A.M., University of California, 1923; Ph.D., Columbia University, 1939; Teaching Fellow in Physics, University of California, 1922–24; Instructor in Mathematics and Physics, San Jose State Teachers College, 1924–26. Bell Telephone Laboratories, 1926–. Mr. Romig's work has been in the field of application of statistical methods to inspection and quality engineering.

M. E. Strieby, A.B., Colorado College, 1914; B.S., Harvard, 1916; B.S. in E.E., Massachusetts Institute of Technology, 1916. New York Telephone Company, Engineering Department, 1916–17; Captain, Signal Corps, U. S. Army, A.E.F., 1917–19. American Telephone and Telegraph Company, Department of Development and Research, 1919–29; Bell Telephone Laboratories; 1929–1940; American Telephone and Telegraph Company, Long Lines Department, 1940–. Mr. Strieby has been associated with various phases of transmission work, more particularly during the last ten years with the development of coaxial systems for telephone and television. At the present time he is Engineer of Transmission of the Long Lines Department.

J. F. Wentz, E.E., Lehigh University, 1917; A.M., Columbia, 1923; First Lieutenant, Infantry, U. S. Army, 1917–19. Western Electric Company, Engineering Department, 1919–24; Bell Telephone Laboratories, 1924–. Mr. Wentz' early work in the Western Electric Company was on high tension fuses and protectors. Upon his transfer to the Research Department he was engaged in the development of permalloy loaded submarine cables. In 1931 he was assigned to the study of transmission properties of coaxial conductors and the development of suitable measuring methods. In 1940 he was made High Frequency Transmission Engineer in charge of the development of coaxial systems for telephone and television.

J. E. Wiltrakis, B.S., Loyola University, 1929. Western Electric Company, Engineering Department, 1920–. Mr. Wiltrakis, formerly engaged on manufacturing capacity, cost reduction and plant location studies, was in charge of engineering and operating the new Kearny Wire Mill.

VOLUME XX APRIL, 1941 NUMBER 2

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

50c per copy *$1.50 per Year*

THE BELL SYSTEM TECHNICAL JOURNAL

Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.

SUBSCRIPTIONS

Subscriptions are accepted at $1.50 per year. Single copies are 50 cents each
The foreign postage is 35 cents per year or 9 cents per copy.

The Bell System Technical Journal

Vol. XX *April, 1941* *No. 2*

STEADY STATE SOLUTIONS OF TRANSMISSION LINE EQUATIONS

S. O. Rice

Methods of obtaining the steady state voltages and currents in a uniform transmission line consisting of several parallel wires are described in Part I. This line may or may not be acted upon by an externally impressed field distributed along its length. A square matrix Γ, which is a generalization of the propagation constant γ for a single circuit, is introduced. Matrix expressions obtained for the voltages and currents involve Γ in much the same way as the corresponding single circuit expressions involve γ. In Part II similar methods are described for obtaining the voltages and currents in a transmission line composed of a number of multi-terminal symmetrical sections connected in tandem. Expressions for the voltages and currents in a line composed of unsymmetrical sections are also given. These sections may or may not contain generators.

THE transmission lines considered here are of two kinds, namely the uniform transmission line, and the transmission line consisting of a number of identical sections connected in tandem. The problem discussed is that of determining the steady state electrical behavior of these lines when the terminal conditions are given. Often there arises the problem of determining the currents induced in a uniform transmission line by an arbitrary impressed field of some fixed frequency or of determining the currents produced by generators placed in the branches of the sections if the line is of the second kind. This is the type of problem with which we shall be particularly concerned.

In dealing with the uniform transmission line it is found convenient to introduce a matrix Γ, which is a generalization of the propagation constant γ for a single wire with ground return, or for a single circuit. This enables us to obtain matrix expressions for the currents and voltages which are similar in form to the single circuit expressions.

A similar situation exists for the transmission line composed of a number of symmetrical sections. However, when the sections are unsymmetrical the corresponding procedure does not appear to yield a corresponding simplification and the formulas are considerably more complicated than in the symmetrical case.

This paper is divided into two parts corresponding to the two kinds of transmission lines. The first part discusses the uniform line. After a statement of the transmission equations in matrix form, expressions for the voltages and currents are given. Two methods of evaluating these expressions are described. The first is based upon a property possessed by many transmission systems, namely that the various modes of propagation have nearly the same speed. The second method is based upon equations which may be obtained by the formal application of a theorem due to Sylvester. The first part concludes with the proof that these two methods lead to the correct results.

After a short introduction the second part discusses the difference equations which govern the transmission in a line composed of multi-terminal sections. The sections may contain generators. Expressions for the voltages and currents in a symmetrical section line, i.e. a line whose sections are symmetrical, are stated and proved in much the same order as the corresponding expressions for the uniform line. A discussion of the unsymmetrical section line concludes the second part.

A sketch of the solution of the uniform transmission line equations by the classical method is given in Appendix I. In Appendices II and III methods are described for solving the symmetrical section line difference equations. These methods are similar to the one of Appendix I. The method of Appendix III uses section constants which may be obtained from measurements made at one end of a typical section.

Part I

Uniform Transmission Lines

1.1 *Differential Equations*

For the sake of convenience in writing down equations we shall assume that the particular line under consideration consists of three parallel wires with ground return, or of three parallel circuits, denoted by the subscripts a, b, and c respectively. The differential equations for this line in an arbitrary impressed field are[1]

$$
\begin{aligned}
\frac{dv_a}{dx} &= -Z_{aa}i_a - Z_{ab}i_b - Z_{ac}i_c + l_a(x) \\
\frac{dv_b}{dx} &= -Z_{ba}i_a - Z_{bb}i_b - Z_{bc}i_c + l_b(x) \\
\frac{dv_c}{dx} &= -Z_{ca}i_a - Z_{cb}i_b - Z_{cc}i_c + l_c(x)
\end{aligned}
\tag{1.1}
$$

[1] These equations are given in substance by J. R. Carson and R. S. Hoyt, *B.S.T.J.*, Vol. 6, pp. 495–545 (1927). Equations (1.2) are equivalent to their equation (90) and equations (1.1) may be obtained by combining their equations (83), (84), and (94). We shall use the term "impressed field" to mean a field distributed along the line. According to our convention there is no impressed field when the line is energized only at the terminals.

and

$$\frac{di_a}{dx} = -Y_{aa}v_a - Y_{ab}v_b - Y_{ac}v_c + t_a(x)$$

$$\frac{di_b}{dx} = -Y_{ba}v_a - Y_{bb}v_b - Y_{tc}v_c + t_b(x) \tag{1.2}$$

$$\frac{di_c}{dx} = -Y_{ca}v_a - Y_{cb}v_b - Y_{cc}v_c + t_c(x)$$

where $Z_{ab} = Z_{ba}$, $Y_{ab} = Y_{ba}$, etc. If we are dealing with three parallel wires $l_a(x)$, $l_b(x)$, $l_c(x)$ are the longitudinal components of the electric force of the impressed field at the wire surfaces; $t_a(x)$, $t_b(x)$, $t_c(x)$ are specified by the admittance of the direct leakage paths and the values of the impressed potentials at the wires. If there are no direct leakage paths the t's are zero.

In order to put these equations in matrix form[2] we introduce the column matrices

$$v = \begin{bmatrix} v_a \\ v_b \\ v_c \end{bmatrix}, \quad i = \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}, \quad l(x) = \begin{bmatrix} l_a(x) \\ l_b(x) \\ l_c(x) \end{bmatrix}, \quad t(x) = \begin{bmatrix} t_a(x) \\ t_b(x) \\ t_c(x) \end{bmatrix}, \tag{1.3}$$

and the symmetrical square matrices

$$Z = \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} \qquad Y = \begin{bmatrix} Y_{aa} & Y_{ab} & Y_{ac} \\ Y_{ba} & Y_{bb} & Y_{bc} \\ Y_{ca} & Y_{cb} & Y_{cc} \end{bmatrix} \tag{1.4}$$

The equations (1.1) and (1.2) may now be written as

$$\frac{dv}{dx} = -Zi + l(x)$$

$$\frac{di}{dx} = -Yv + t(x) \tag{1.5}$$

and these are the equations to be solved.

When there is no impressed field equations (1.5) give

$$\frac{d^2v}{dx^2} = ZYv$$

$$\frac{d^2i}{dx^2} = YZi \tag{1.6}$$

[2] Cf. L. A. Pipes, *Phil. Mag.*, Vol. 24 (1937), p. 97.

and the analogy with the one circuit case leads us to put

$$\Gamma^2 = ZY, \qquad \Gamma = \sqrt{ZY} \tag{1.7}$$

where Γ is a square matrix representing a generalization of the propagation constant. Putting aside for the moment the question of interpreting the square root, we note that interchanging the rows and columns in $\Gamma^2 = ZY$ gives

$$\Gamma'^2 = Y'Z' = YZ, \qquad \Gamma' = \sqrt{YZ} \tag{1.8}$$

where the primes denote transposition. Y' and Z' are equal to Y and Z respectively because of their symmetry. We thus expect Γ' to be associated with the propagation of i in the same way that Γ is associated with the propagation of v.

1.2 Statement of Results for an Infinite Line—No Impressed Field

It is shown that when there is no impressed field the voltages and currents at any point x in a transmission line extending from $x = 0$ to $x = \infty$ are given by

$$\begin{aligned} v(x) &= e^{-x\Gamma}v(o) = e^{-x\Gamma}Z_o i(o) \\ i(x) &= e^{-x\Gamma'}i(o) \\ v(x) &= Z_o i(x) \end{aligned} \tag{1.9}$$

where $e^{-x\Gamma}$ is the square matrix defined by the convergent series of matrices[3]

$$e^{-x\Gamma} = I - \frac{x\Gamma}{1!} + \frac{x^2\Gamma^2}{2!} - \frac{x^3\Gamma^3}{3!} + \cdots \tag{1.10}$$

and $e^{-x\Gamma'}$ is the transposed of $e^{-x\Gamma}$. I denotes the unit matrix. Z_o is a square matrix and is called the characteristic impedance matrix:

$$Z_o = \Gamma^{-1}Z = \Gamma Y^{-1} \tag{1.11}$$

Additional expressions of the same type for Z_o are given by equations (1.45). The matrix $e^{-x\Gamma}Z_o$, being of the nature of a transfer impedance, is symmetrical.

The matrices $e^{-x\Gamma}$ and Z_o may be computed in several ways, the choice depending upon the circumstances. The first method to be described is useful when x is not too large and when the propagation constants of the various modes of propagation are nearly equal to each other. In the case of open-wire lines these propagation constants are grouped around the value $j\omega/v$ where v is of the order of 180,000 miles per second. The second method may be used for all cases, including those for which the series in

[3] Frazer, Duncan and Collar, "Elementary Matrices," Cambridge University Press, §2.5. In the work which follows, this text will be referred to as "F.D.C."

the first method converge too slowly to be of value. However, it requires the solution of an mth degree equation and the determination of the m modes of propagation where m is the number of circuits. For $m = 2$ this is no handicap and the method is quite convenient. In this case the method is closely related to one described by John Riordan in an unpublished memorandum.

First Method: Multiply the matrices Z and Y together to obtain ZY. Choose the number γ^2 in

$$ZY = I\gamma^2 + R, \tag{1.12}$$

where I is the unit matrix, so that the elements of R are small in comparison with γ^2. For many transmission lines it is possible to do this. Γ may be obtained by using the binomial theorem to expand the square root in the formula

$$\Gamma = \sqrt{ZY} = \gamma(I + \gamma^{-2}R)^{\frac{1}{2}}, \tag{1.13}$$

where γ is that square root of γ^2 whose real and imaginary parts are non-negative. In carrying out the work it is convenient to introduce the matrix S whose elements are small in comparison with unity.

$$\Gamma = \gamma(I + S) \tag{1.14}$$

To compute S, first compute the matrix $R/2\gamma^2$ and then use the power series

$$\begin{aligned} S = \left(\frac{R}{2\gamma^2}\right) - \frac{1}{2}\left(\frac{R}{2\gamma^2}\right)^2 + \frac{1}{2}\left(\frac{R}{2\gamma^2}\right)^3 - \frac{5}{8}\left(\frac{R}{2\gamma^2}\right)^4 \\ + \frac{7}{8}\left(\frac{R}{2\gamma^2}\right)^5 - \frac{21}{16}\left(\frac{R}{2\gamma^2}\right)^6 + \cdots \end{aligned} \tag{1.15}$$

This series will usually converge rapidly. The matrix $e^{-x\Gamma}$ is given by

$$e^{-x\Gamma} = e^{-z} \cdot e^{-zS} \tag{1.16}$$

where z is a number, $z = \gamma x$, and e^{-zS} is to be computed from

$$e^{-zS} = I - \frac{zS}{1!} + \frac{(zS)^2}{2!} - \frac{(zS)^3}{3!} + \cdots \tag{1.17}$$

$e^{-x\Gamma'}$ is obtained from $e^{-x\Gamma}$ by interchanging the rows and columns. The characteristic impedance matrix may be obtained from (1.11),

$$Z_o = \Gamma Y^{-1},$$

after computing Γ from S as in (1.14).

If only $e^{-x\Gamma}$ is required the following series may be used.

$$e^{-x\Gamma} = \sum_{p=0}^{\infty} \left(\frac{Rx}{2\gamma}\right)^p \frac{b_p(z)}{p!} \tag{1.18}$$

where R, γ, and z have the same meaning as above and the coefficients are computed from

$$b_o = e^{-z}, \qquad b_1(z) = -e^{-z}, \qquad b_2(z) = e^{-z}\left(1 + \frac{1}{z}\right)$$

$$b_{p+2}(z) = b_p(z) - \frac{2p+1}{z} b_{p+1}(z)$$

In the first term of the series $\left(\frac{Rx}{2\gamma}\right)^0$ denotes I.

Second Method: Γ, $e^{-x\Gamma}$ and Z_o may be regarded as functions of the square matrix ZY. In order to express these functions in a form suitable for calculation we apply Sylvester's theorem[4]. The characteristic matrix of ZY is

$$f(\gamma^2) = \gamma^2 I - ZY \tag{1.19}$$

where now γ^2 is regarded as a variable instead of a fixed number as in the first method. We shall suppose that ZY is a square matrix of order m and that the roots $\gamma_1^2, \gamma_2^2, \cdots \gamma_m^2$ of the characteristic function, i.e. of the determinantal equation

$$|f(\gamma^2)| = 0, \tag{1.20}$$

are distinct. Let the matrix $F(\gamma^2)$ be the adjoint of $f(\gamma^2)$ and denote the derivative of the characteristic function by

$$|f(\gamma^2)|^{(1)} = \frac{d}{d(\gamma^2)} |f(\gamma^2)| \tag{1.21}$$

Since $\gamma_1^2, \gamma_2^2, \cdots \gamma_m^2$ are all different $|f(\gamma_r^2)|^{(1)}$ is unequal to zero for $r = 1, 2, \cdots m$. Sylvester's theorem says that if $P(ZY)$ is any polynomial in ZY then

$$P(ZY) = \sum_{r=1}^{m} N(\gamma_r^2) P(\gamma_r^2) \tag{1.22}$$

where $P(\gamma_r^2)$ is a scalar (and thus deviates from our convention that capital letters denote square matrices). $N(\gamma_r^2)$ is a square matrix:

$$N(\gamma_r^2) = \frac{F(\gamma_r^2)}{|f(\gamma_r^2)|^{(1)}} \tag{1.23}$$

When $m = 2$, $N(\gamma_2^2)$ is equal to $I - N(\gamma_1^2)$.

[4] F.D.C. §3.9. The u and λ of the reference are the ZY and γ^2 of the present section.

Applying (1.22) to Γ, $e^{-x\Gamma}$ and Z_o even though they are not polynomials in ZY gives results which may be verified to be true.

$$
\begin{aligned}
\Gamma &= \sqrt{ZY} = \sum N(\gamma_r^2)\gamma_r \\
e^{-x\Gamma} &= e^{-x\sqrt{ZY}} = \sum N(\gamma_r^2)e^{-x\gamma_r} \\
Z_o &= (ZY)^{\frac{1}{2}}Y^{-1} = \sum N(\gamma_r^2)\gamma_r Y^{-1} \\
e^{-x\Gamma}Z_o &= \sum N(\gamma_r^2)\gamma_r e^{-x\gamma_r}Y^{-1}
\end{aligned}
\tag{1.24}
$$

where the summations extend from $r = 1$ to $r = m$ and $\gamma_1, \gamma_2, \cdots \gamma_m$ are the square roots of $\gamma_1^2, \gamma_2^2, \cdots \gamma_m^2$ respectively whose real parts are non-negative. $\gamma_1, \gamma_2, \cdots \gamma_m$ are also the propagation constants of the "normal modes" of propagation. Some light is thrown on the physical significance of the matrix $N(\gamma_r^2)$ by supposing that only the rth normal mode is being propagated on the transmission line. $N(\gamma_r^2)$ is such that it can be expressed as a column matrix times a row matrix. The voltages in circuits 1, 2, $\cdots$ m are proportional to the first, second, $\cdots$ mth elements, respectively of the column matrix. The currents in circuits 1, 2, $\cdots$ m are proportional to the corresponding elements in the row matrix.

1.3 Results for Any Uniform Line—No Impressed Field

When the length of the line is finite the voltages and currents may be expressed as

$$
\begin{aligned}
v(x) &= \cosh x\Gamma\, v(o) - \sinh x\Gamma\, Z_o i(o) \\
i(x) &= -\sinh x\Gamma'\, Z_o^{-1} v(o) + \cosh x\Gamma'\, i(o)
\end{aligned}
\tag{1.25}
$$

where Z_o and Γ have the same meaning as before. The matrices $\sinh x\Gamma Z_o$ and $\sinh x\Gamma' Z_o^{-1}$ are symmetrical. The square matrices $\cosh x\Gamma$ and $\sinh x\Gamma$ are defined by the series

$$
\begin{aligned}
\cosh x\Gamma &= I + \frac{x^2\Gamma^2}{2!} + \frac{x^4\Gamma^4}{4!} + \cdots \\
\sinh x\Gamma &= \frac{x\Gamma}{1!} + \frac{x^3\Gamma^3}{3!} + \cdots
\end{aligned}
\tag{1.26}
$$

$\cosh x\Gamma'$ is obtained by interchanging the rows and columns of $\cosh x\Gamma$ and $\sinh x\Gamma'$ is obtained similarly from $\sinh x\Gamma$. Solving (1.25) for $v(o)$ and $i(o)$ gives

$$
\begin{aligned}
v(o) &= \cosh x\Gamma\, v(x) + \sinh x\Gamma\, Z_o i(x) \\
i(o) &= \sinh x\Gamma'\, Z_o^{-1} v(x) + \cosh x\Gamma'\, i(x)
\end{aligned}
$$

As in the case of the infinite line, we have two ways of computing the coefficients of $v(o)$ and $i(o)$ in the expressions (1.25) for $v(x)$ and $i(x)$.

First Method: Choose a number γ^2 and compute the matrices R, S, Γ, Z_o as described in the first method for the infinite line. The matrix $e^{x\Gamma}$ is given by

$$e^{x\Gamma} = e^{z} \cdot e^{zS}$$

where $z = \gamma x$ and e^{zS} is computed from the series

$$e^{zS} = I + \frac{zS}{1!} + \frac{z^2 S^2}{2!} + \cdots$$

If the elements of zS are so large that the series converges slowly it may be worthwhile to divide zS by 16, say, compute $\exp\left(\frac{zS}{16}\right)$ from the series, and then obtain e^{zS} by four matrix multiplications. When e^{zS} is known its inverse e^{-zS} can be computed and $e^{-x\Gamma}$ obtained from (1.16). The hyperbolic functions are given by

$$\begin{aligned} \cosh x\Gamma &= \tfrac{1}{2}\,(e^{x\Gamma} + e^{-x\Gamma}) \\ \sinh x\Gamma &= \tfrac{1}{2}\,(e^{x\Gamma} - e^{-x\Gamma}) \end{aligned} \tag{1.27}$$

which follow from the series definitions of the various matrices.

If only the coefficients in (1.25) are required we may choose γ^2 and compute R and powers of the matrix $Rx/2\gamma$. Then the coefficients in (1.25) are given by

$$\begin{aligned} \cosh x\Gamma &= \sum_{p=0}^{\infty} \left(\frac{Rx}{2\gamma}\right)^p \frac{a_p(z)}{p!} \\ \sinh x\Gamma\, Z_o &= \sum_{p=0}^{\infty} \left(\frac{Rx}{2\gamma}\right)^p \frac{a_{p+1}(z)}{p!\gamma}\, Z \\ \sinh x\Gamma' Z_o^{-1} &= \sum_{p=0}^{\infty} \left(\frac{R'x}{2\gamma}\right)^p \frac{a_{p+1}(z)}{p!\gamma}\, Y \end{aligned} \tag{1.28}$$

where R' is the transposed of R, and the scalar coefficient $a_p(z)$ is a function of $z = \gamma x$ given by

$$\begin{aligned} a_o(z) &= \cosh z \qquad a_1(z) = \sinh z \\ a_2(z) &= \cosh z - \frac{\sinh z}{z} \\ a_{p+2}(z) &= a_p(z) - \frac{2p+1}{z}\, a_{p+1}(z), \end{aligned} \tag{1.29}$$

and it is understood that $(Rx/2\gamma)^0 = I$.

Second Method: Compute the propagation constants $\gamma_1, \gamma_2, \cdots \gamma_m$ and the square matrices $N(\gamma_r^2)$ given by (1.23) as in the second method for the infinite line. Then

$$
\begin{aligned}
\cosh x\Gamma &= \Sigma\, N(\gamma_r^2) \cosh x\gamma_r \\
\sinh x\Gamma\, Z_o &= \sinh x\Gamma\, \Gamma Y^{-1} \\
&= \Sigma\, N(\gamma_r^2) \sinh x\gamma_r\, \gamma_r Y^{-1} \\
\sinh x\Gamma'\, Z_o^{-1} &= \sinh x\Gamma'\, \Gamma'^{-1} Y \\
&= \sum N'(\gamma_r^2) \frac{\sinh x\gamma_r}{\gamma_r} Y
\end{aligned}
\tag{1.30}
$$

where $N'(\gamma_r^2)$ is the transposed of $N(\gamma_r^2)$, $N(\gamma_r^2)$ being defined by (1.23), and the summations extend from $r = 1$ to $r = m$.

When the transmission line consists of perfectly conducting wires strung on perfect insulators over a perfectly conducting earth the magnetic and electrostatic fields are related so as to make Z equal to $\gamma_o^2 Y^{-1}$ where

$$\gamma_o = j\omega/c,$$

ω being 2π times the frequency and c the speed of light.

It is interesting to apply the first method of solution to this line. Even though the proof of the first method, which is given in §1.10, does not cover this case there seems to be little doubt that the correct answer is obtained. We have

$$ZY = \gamma_o^2 I$$

Choosing $\gamma = \gamma_o$ gives $R = 0$ and therefore $S = 0$. It follows that

$$
\begin{aligned}
&\Gamma = \gamma_o I,\ Z_o = \Gamma^{-1} Z = \gamma_o^{-1} Z \\
&\cosh x\Gamma = \cosh (x\gamma_o I) = \cosh x\gamma_o\, I \\
&\sinh x\Gamma\, Z_o = \sinh x\gamma_o\, \gamma_o^{-1} Z \\
&\sinh x\Gamma'\, Z_o^{-1} = \sinh x\gamma_o\, \gamma_o Z^{-1}
\end{aligned}
$$

When these are put into equations (1.25) the expressions for $v(x)$ and $i(x)$ in a perfect transmission line are obtained:

$$
\begin{aligned}
v(x) &= \cosh x\gamma_o v(o) - \frac{\sinh x\gamma_o}{\gamma_o} Zi(o) \\
i(x) &= -\gamma_o \sinh x\gamma_o Z^{-1} v(o) + \cosh x\gamma_o i(o)
\end{aligned}
\tag{1.31}
$$

1.4 Results for Any Uniform Line—Impressed Field

The differential equations to be satisfied in this case are given by (1.5). A solution which reduces to $v(o)$ and $i(o)$ at $x = 0$ is

$$\begin{aligned} v(x) &= \cosh x\Gamma\, v(o) - \sinh x\Gamma\, Z_o i(o) \\ &\quad + \int_0^x \cosh (x - \xi)\Gamma\, l(\xi)\, d\xi - \int_0^x \sinh (x - \xi)\Gamma\, Z_o\, t(\xi)\, d\xi \\ i(x) &= -\sinh x\Gamma'\, Z_o^{-1} v(o) + \cosh x\Gamma'\, i(o) \\ &\quad - \int_0^x \sinh (x - \xi)\Gamma' Z_o^{-1} l(\xi)\, d\xi + \int_0^x \cosh (x - \xi)\Gamma' t(\xi)\, d\xi \end{aligned} \tag{1.32}$$

The matrices $\cosh x\Gamma$, $\sinh x\Gamma$ and Z_o are the same as the ones discussed in §1.2 and §1.3. The elements of the integral[5] of a matrix U (U is not necessarily a square matrix) are given by the integrals of the corresponding elements of U.

In many cases of practical interest the impressed field varies exponentially with respect to x. The column matrices $l(x)$ and $t(x)$ may then be expressed as

$$l(x) = e^{-x\theta} \begin{bmatrix} \lambda_a \\ \lambda_b \\ \lambda_c \end{bmatrix} \qquad t(x) = e^{-x\theta} \begin{bmatrix} \tau_a \\ \tau_b \\ \tau_c \end{bmatrix} \tag{1.33}$$

where the λ's and τ's are constants and θ is the propagation constant of the impressed field in the direction of the line. The integrations in the expressions (1.32) may be performed with the result

$$\begin{aligned} v(x) &= \cosh x\Gamma\, v(o) - \sinh x\Gamma\, Z_o i(o) \\ &\quad + \tfrac{1}{2} (e^{x\Gamma} - e^{-x\theta} I)\, (\Gamma + \theta I)^{-1} (\lambda - Z_o \tau) \\ &\quad - \tfrac{1}{2} (e^{-x\Gamma} - e^{-x\theta} I)\, (\Gamma - \theta I)^{-1} (\lambda + Z_o \tau) \\ i(x) &= -\sinh x\Gamma'\, Z_o^{-1} v(o) + \cosh x\Gamma'\, i(o) \\ &\quad + \tfrac{1}{2} (e^{x\Gamma'} - e^{-x\theta} I)\, (\Gamma' + \theta I)^{-1} (\tau - Z_o^{-1} \lambda) \\ &\quad - \tfrac{1}{2} (e^{-x\Gamma'} - e^{-n\theta} I)\, (\Gamma' - \theta I)^{-1} (\tau + Z_o^{-1} \lambda) \end{aligned} \tag{1.34}$$

provided that the inverse matrices exist. The matrix $(e^{x\Gamma'} - e^{-x\theta} I)(\Gamma' + \theta I)^{-1}$ is the transposed of $(e^{x\Gamma} - e^{-x\theta} I)(\Gamma + \theta I)^{-1}$, etc. If one of these matrices, say $\Gamma - \theta I$, has no inverse then it is necessary to evaluate the

[5] F.D.C. §2.10.

corresponding integral in some other way. Thus it may be advantageous to use the formula

$$-(e^{-x\Gamma} - e^{-x\theta}I)(\Gamma - \theta I)^{-1} = e^{-x\Gamma}\int_0^x e^{\xi\Gamma - \xi\theta I}\,d\xi$$
$$= e^{-x\Gamma}\left[xI + \frac{x^2}{2!}(\Gamma - \theta I) + \frac{x^3}{3!}(\Gamma - \theta I)^2 + \cdots\right] \tag{1.35}$$

Two special cases of (1.34) are of interest. When the line is shorted at both ends, $v(o) = v(x) = 0$, where x is the line length, and

$$i(o) = \tfrac{1}{2} Z_o^{-1} (\sinh x\Gamma)^{-1}[(e^{x\Gamma} - e^{-x\theta}I)(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau)$$
$$- (e^{-x\Gamma} - e^{-x\theta}I)(\Gamma - \theta I)^{-1}(\lambda + Z_o\tau)]$$

$$i(x) = \frac{e^{-x\theta}}{2} Z_o^{-1} (\sinh x\Gamma)^{-1}[(e^{x\Gamma} - e^{x\theta}I)(\Gamma - \theta I)^{-1}(\lambda + Z_o\tau)$$
$$- (e^{-x\Gamma} - e^{x\theta}I)(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau)]$$

When the line is terminated in its characteristic impedance at both ends, $v(o) = -Z_oi(o)$, $v(x) = Z_oi(x)$, and

$$i(o) = \tfrac{1}{2}(I - e^{-x\theta}e^{-x\Gamma'})(\Gamma' + \theta I)^{-1}(Z_o^{-1}\lambda - \tau)$$
$$i(x) = -\tfrac{1}{2}(e^{-x\Gamma'} - e^{-x\theta}I)(\Gamma' - \theta I)^{-1}(Z_o^{-1}\lambda + \tau)$$

The matrices occurring in the expressions (1.34) for $v(x)$ and $i(x)$ may be computed by the first or second method described for the uniform line in the absence of an impressed field. The second method involves the use of expansions similar to

$$(e^{x\Gamma} - e^{-x\theta}I)(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau)$$
$$= \sum N(\gamma_r^2)\left(\frac{e^{x\gamma_r} - e^{-x\theta}}{\gamma_r + \theta}\right)\left(\lambda - \frac{Z}{\gamma_r}\tau\right)$$
$$(e^{x\Gamma'} - e^{-x\theta}I)(\Gamma' + \theta I)^{-1}(Z_o^{-1}\lambda - \tau) \tag{1.36}$$
$$= \sum N'(\gamma_r^2)\left(\frac{e^{x\gamma_r} - e^{-x\theta}}{\gamma_r + \theta}\right)\left(\frac{Y}{\gamma_r}\lambda - \tau\right)$$

where the summations run from $r = 1$ to $r = m$ and $N'(\gamma_r^2)$ is the transposed of the square matrix $N(\gamma_r^2)$ given by (1.23). In obtaining these expansions by Sylvester's theorem, Z_o in the first is replaced by $\Gamma^{-1}Z$ and Z_o^{-1} in the second by $\Gamma'^{-1}Y$.

If we assume that an impressed field acts upon the perfect transmission line of equations (1.31), we see that $t(x) = 0$ because there are no direct leakage paths. We may also write

$$(e^{x\Gamma} - e^{-x\theta}I)(\Gamma + \theta I)^{-1} = (e^{x\gamma_o}I - e^{-x\theta}I)(\gamma_o I + \theta I)^{-1}$$
$$= \frac{e^{x\gamma_o} - e^{-x\theta}}{\gamma_o + \theta} I$$

From this and similar equations it follows that

$$\begin{aligned} v(x) &= \cosh x\gamma_o v(o) - \frac{\sinh x\gamma_o}{\gamma_o} Zi(o) \\ &\qquad + \frac{1}{2}\left[\frac{e^{x\gamma_o} - e^{-x\theta}}{\gamma_o + \theta} - \frac{e^{-x\gamma_o} - e^{-x\theta}}{\gamma_o - \theta}\right]\lambda \\ i(x) &= -\gamma_o \sinh x\gamma_o Z^{-1} v(o) + \cosh x\gamma_o i(o) \\ &\qquad - \frac{\gamma_o}{2}\left[\frac{e^{x\gamma_o} - e^{-x\theta}}{\gamma_o + \theta} + \frac{e^{-x\gamma_o} - e^{-x\theta}}{\gamma_o - \theta}\right] Z^{-1}\lambda \end{aligned} \tag{1.37}$$

1.5 *Results for Infinite Uniform Line—Impressed Field*

When the line extends from $x = 0$ to $x = \infty$ and the impressed field is such that the voltages and currents remain finite at $x = \infty$, the appropriate solutions may be obtained from the results of §1.4 by a limiting process. The condition that $v(x)$ remain finite suggests that the coefficient of $e^{x\Gamma}$ be zero in the expression (1.32) for $v(x)$. This gives a relation between $v(o)$ and $i(o)$ which must be satisfied:

$$v(o) = Z_o i(o) - \int_0^\infty e^{-\xi\Gamma} [l(\xi) - Z_o t(\xi)] \, d\xi \tag{1.38}$$

If the impressed field varies exponentially with x expression (1.34) gives

$$v(o) = Z_o i(o) - (\Gamma + \theta I)^{-1}(\lambda - Z_o \tau) \tag{1.39}$$

Expressions for $v(x)$ and $i(x)$ may be obtained by using relations (1.38) and (1.39) in (1.32) and (1.34) respectively. As these are somewhat lengthy we shall state only two which follow from (1.39).

$$\begin{aligned} v(x) &= e^{-x\Gamma} v(o) \\ &\quad + \tfrac{1}{2}(e^{-x\Gamma} - e^{-x\theta}I)[(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau) \\ &\qquad\qquad - (\Gamma - \theta I)^{-1}(\lambda + Z_o\tau)] \\ i(x) &= Z_o^{-1} e^{-x\Gamma} v(o) \\ &\quad + \tfrac{1}{2}Z_o^{-1}(e^{-x\Gamma} + e^{-x\theta}I)(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau) \\ &\qquad - \tfrac{1}{2}Z_o^{-1}(e^{-x\Gamma} - e^{-x\theta}I)(\Gamma - \theta I)^{-1}(\lambda + Z_o\tau) \end{aligned} \tag{1.40}$$

Two similar expressions may be obtained in which the initial current $i(o)$ instead of $v(o)$ appears on the right. If the line is terminated in its characteristic impedance at $x = 0, v(o) = -Z_o i(o)$, and the voltages and currents produced by the impressed field are

$$\begin{aligned} v(o) &= -\tfrac{1}{2}(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau) \\ i(o) &= \tfrac{1}{2} Z_o^{-1}(\Gamma + \theta I)^{-1}(\lambda - Z_o\tau) \end{aligned} \tag{1.41}$$

As in §1.4 these expressions may be computed by the first and second methods described in §1.3. For example, the application of the second method to the relation (1.39) which must exist between $v(o)$ and $i(o)$ in an infinite line gives

$$v(o) = \sum_{r=1}^{m} N(\gamma_r^2)\left[\frac{Z}{\gamma_r}\, i(o) - \frac{1}{\gamma_r + \theta}\left(\lambda - \frac{Z}{\gamma_r}\,\tau\right)\right] \tag{1.42}$$

where $N(\gamma_r^2)$ is the square matrix (1.23).

1.6 *Outline of Proofs*

The proof of the results which have been stated is divided into three parts. In the first part it is shown that if Γ is a matrix such that (a) its square is ZY and (b) every element in the matrix $e^{-x\Gamma}$ approaches zero as $x \to \infty$, then the expressions for $v(x)$ and $i(x)$ involving Γ and Z_o satisfy the transmission line equations. In the second part of the proof it is shown that if certain requirements are met Γ as obtained by the first method satisfies the conditions (a) and (b) and hence the expressions for $v(x)$ and $i(x)$ given by the first method are correct. The third part of the proof discusses a general procedure which may be used to prove the equations which constitute the second method.

Both the second and the third parts of the proof are based upon the solution of the transmission line equations which is sketched in Appendix I. This solution assumes that the propagation constants of the normal modes of propagation are unequal, and our proofs are limited accordingly. However, considerations of continuity seem to show that the first method is valid even when two or more propagation constants are equal. Under the same circumstances the second method suggests the use of the confluent form of Sylvester's theorem.[6]

1.7 *Relations Obtained by Considering An Infinite Line*

We suppose that we are going to deal with transmission lines possessing the non-singular, symmetrical impedance and admittance matrices Z and Y. We further suppose that, by some means or other, we have determined a matrix Γ which satisfies the two conditions; (a) the square of Γ is

$$\Gamma^2 = ZY, \tag{1.43}$$

and (b) every element in the matrix $e^{-x\Gamma}$ approaches zero as $x \to \infty$.

[6] F.D.C. §3.10.

Consider a line extending from $x = 0$ to $x = \infty$, there being no impressed field. Viewing the line at $x = 0$ as an n terminal network shows that there is a symmetrical matrix Z_o such that $v(o) = Z_o i(o)$. Let this be taken as the definition of the characteristic impedance matrix Z_o. We shall show from the differential equations of the line that

1. The voltages and currents in the infinite line are given by

$$\begin{aligned} v(x) &= e^{-x\Gamma} v(o) \\ i(x) &= e^{-x\Gamma'} i(o) \end{aligned} \tag{1.44}$$

2. The matrix Z_o satisfies the relations

$$\begin{aligned} Z_o &= \Gamma^{-1} Z = Z\Gamma'^{-1} = \Gamma Y^{-1} = Y^{-1}\Gamma' \\ Z_o^{-1} &= Z^{-1}\Gamma = \Gamma' Z^{-1} = Y\Gamma^{-1} = \Gamma'^{-1} Y \end{aligned} \tag{1.45}$$

$$v(x) = Z_o i(x) \tag{1.46}$$

3. The matrices Z_o, Z, and Y obey the commutation rules

$$\begin{aligned} \Phi(\Gamma) Z_o &= Z_o \Phi(\Gamma') \\ \Phi(\Gamma) Z &= Z \Phi(\Gamma') \\ Y\Phi(\Gamma) &= \Phi(\Gamma') Y \end{aligned} \tag{1.47}$$

where $\Phi(\Gamma)$ is any square matrix, such as $e^{-x\Gamma}$, representable as a convergent power series in Γ with scalar coefficients. Furthermore, the matrices $\Phi(\Gamma) Z_o$, $\Phi(\Gamma) Z$, and $Y\Phi(\Gamma)$ are symmetrical.

The differential equations of the transmission line are

$$\frac{dv}{dx} = -Zi, \qquad \frac{di}{dx} = -Yv, \qquad \frac{d^2 v}{dx^2} = ZYv \tag{1.48}$$

the third following from the first two when i is eliminated. That $v(x) = e^{-x\Gamma} v(o)$ is a solution of the third equation may be verified by direct substitution and differentiation[7]. Since this expression for $v(x)$ approaches zero as $x \rightarrow \infty$ and reduces to $v(o)$ at $x = 0$, it represents the voltages in an infinite transmission line. Hence the first equation in (1.44) is true. Setting it in the first differential equation of (1.48), putting $x = 0$, replacing $v(o)$ by $Z_o i(o)$, and noting that $i(o)$ may be regarded as an arbitrary column gives

$$\Gamma Z_o = Z \tag{1.49}$$

Since Γ was assumed to be non-singular, Z_o is equal to $\Gamma^{-1} Z$. Z is symmetrical and the reciprocity theorem for electrical networks requires that Z_o

[7] The differentiation of the exponential function is discussed in F.D.C. §2.7.

be symmetrical, hence

$$Z_o = \Gamma^{-1}Z = Z\Gamma'^{-1}$$

The first group of equations in (1.45) follow from this together with the expression $\Gamma^2\Gamma'^{-1}$ for Z obtained from (1.43). The second group in (1.45) is obtained from the first group.

The commutation rule for Z_o is obtained from (1.49) together with the equation obtained from (1.49) by transposition. Since Z is symmetrical

$$\Gamma Z_o = Z_o\Gamma', \qquad \Gamma^2 Z_o = \Gamma Z_o \Gamma' = Z_o\Gamma'^2,$$

$$\Gamma^n Z_o = Z_o \Gamma'^n$$

and the first of equations (1.47) follow from this. The second and third of equations (1.47) may be obtained similarly from the relations (1.45). The matrix $\Phi(\Gamma)Z_o$ is symmetrical since its transposed is $Z_o\,[\Phi(\Gamma)]'$ and this is equal to $Z_o\Phi(\Gamma') = \Phi(\Gamma)Z_o$. A similar argument applies to the other matrices in (1.47).

The expression for $i(x)$ in (1.44) may be obtained by Maclaurin's expansion. Setting $x = 0$ in the second differential equation of (1.48),

$$\left(\frac{di}{dx}\right)_o = -Yv(o) = -YZ_o i(o) = -\Gamma' i(o)$$

where we have used the equality between the first and last members of the first equation of (1.45) and where the subscript 0 denotes the value of the derivative at $x = 0$. Repeated differentiation gives

$$\frac{d^2 i}{dx^2} = -Y\frac{dv}{dx} = YZi = \Gamma'^2 i$$

$$\left(\frac{d^3 i}{dx^3}\right)_o = \Gamma'^2\left(\frac{di}{dx}\right)_o = -\Gamma'^3 i(o)$$

and so on. Hence

$$i(x) = \left[I - \frac{x\Gamma'}{1!} + \frac{x^2\Gamma'^2}{2!} - \cdots\right] i(o)$$

$$= e^{-x\Gamma'} i(o)$$

Equation (1.46) may now be obtained by using the commutation rule for Z_o:

$$v(x) = e^{-x\Gamma} v(o) = e^{-x\Gamma} Z_o i(o)$$

$$= Z_o e^{-x\Gamma'} i(o) = Z_o i(x)$$

This completes the proof of equations (1.44) to (1.47).

1.8 Proof of Relations for Any Uniform Line—Impressed Field

Here it is shown that if a matrix Γ satisfies the two conditions of §1.7 and if Z_o is the characteristic impedance matrix defined there, then the voltages and currents in any uniform line are given by the expressions (1.32). If suitable conditions are fulfilled the relation (1.38) between $v(o)$ and $i(o)$ for an infinite line may be obtained from (1.32).

First of all, $v(x)$ and $i(x)$ reduce to the required values of $v(o)$ and $i(o)$ at $x = 0$. All that remains to be shown is that $v(x)$ and $i(x)$ as given by (1.32) are solutions of the transmission line equations (1.5). By substituting (1.32) in (1.5) and using the formulas

$$\frac{d}{dx} \cosh x\Gamma = \Gamma \sinh x\Gamma = \sinh x\Gamma \, \Gamma$$

$$\frac{d}{dx} \sinh x\Gamma = \Gamma \cosh x\Gamma = \cosh x\Gamma \, \Gamma$$

which follow immediately from the series definitions (1.26) of the hyperbolic functions, we obtain two matrix equations corresponding to the two differential equations. The terms in these equations involving $v(o)$ may be canceled out provided

$$\begin{aligned} \Gamma \sinh x\Gamma &= Z \sinh x\Gamma' \, Z_o^{-1} \\ \Gamma' \cosh x\Gamma' \, Z_o^{-1} &= Y \cosh x\Gamma \end{aligned} \tag{1.50}$$

and these are seen to be true from (1.45) and (1.47). The terms involving $i(o)$ may be canceled by a similar argument. The terms involving $l(x)$ may be canceled provided

$$\int_0^x \sinh (x - \xi)\Gamma \, \Gamma l(\xi) \, d\xi = \int_0^x Z \sinh (x - \xi)\Gamma' Z_o^{-1} l(\xi) \, d\xi$$

$$\int_0^x \Gamma' \cosh (x - \xi)\Gamma' Z_o^{-1} l(\xi) \, d\xi = \int_0^x Y \cosh (x - \xi)\Gamma \, l(\xi) \, d\xi$$

and these are seen to be true when x in (1.50) is replaced by $(x - \xi)$. The terms involving $t(x)$ may be similarly canceled. Thus we have verified that $v(x)$ and $i(x)$ as given by (1.32) are solutions of the transmission line equation provided that the commutation rules (1.47) and the relations (1.45) involving Z_o of §1.7 are satisfied. This is the case when Γ is such that (a) Γ^2 is equal to ZY and also (b) every element in $e^{-\Gamma x}$ approaches zero as $x \longrightarrow \infty$.

In order to establish equation (1.38) for the Γ of §1.7 several assumptions regarding the impressed field are required. Writing the hyperbolic functions in the first of equations (1.32) in exponential form and premultiplying both sides by $2e^{-\Gamma x}$ gives

$$2e^{-x\Gamma}v(x) = \left[v(o) - Z_o i(o) + \int_0^x e^{-\xi\Gamma}[l(\xi) - Z_o t(\xi)]\,d\xi\right] + e^{-2x\Gamma}[v(o) + Z_o i(o)] + e^{-x\Gamma}\int_0^x e^{-(x-\xi)\Gamma}[l(\xi) + Z_o t(\xi)]\,d\xi$$

When $x \longrightarrow \infty$ equation (1.38) is obtained provided that the impressed field and the terminal conditions at the far end are such that (a) $v(x)$ remains finite, (b) the integral in (1.38) converges, and (c), the last expression on the right in the equation above approaches zero as $x \longrightarrow \infty$.

1.9 Derivation of Equations (1.25)

Although equations (1.25) may be obtained by setting $l(x) = t(x) = 0$ in §1.8, it is of some interest to derive them directly. By repeated differentiation of the equations

$$\frac{dv}{dx} = -Zi, \qquad \frac{di}{dx} = -Yv \tag{1.48}$$

the second, third and higher order derivatives may be obtained. Using these in Maclaurin's expansion about $x = 0$ gives

$$\begin{aligned} v(x) &= \left[I + \frac{x^2}{2!}ZY + \frac{x^4}{4!}(ZY)^2 + \cdots\right]v(o) - \left[\frac{x}{1!}I + \frac{x^3}{3!}ZY + \frac{x^5}{5!}(ZY)^2 + \cdots\right]Zi(o) \\ i(x) &= -\left[\frac{x}{1!}I + \frac{x^3}{3!}YZ + \frac{x^5}{5!}(YZ)^2 + \cdots\right]Yv(o) + \left[I + \frac{x^2}{2!}YZ + \frac{x^4}{4!}(YZ)^2 + \cdots\right]i(o) \end{aligned} \tag{1.51}$$

These series converge for all values of x and could be used for computation were it not for the unfortunate fact that in most problems a great many terms would be required for a satisfactory answer. For the time being, let Γ be any matrix whose square is ZY. The definitions (1.26) of the hyperbolic functions enable us to write (1.51) as

$$\begin{aligned} v(x) &= \cosh x\Gamma\, v(o) - \sinh x\Gamma\, \Gamma^{-1}Zi(o) \\ i(x) &= -\sinh x\Gamma'\, \Gamma'^{-1}Yv(o) + \cosh x\Gamma'\, i(o) \end{aligned} \tag{1.52}$$

If in addition to being a matrix whose square is ZY, Γ is also such that every element in $e^{-x\Gamma}$ approaches zero as $x \longrightarrow \infty$, then we may use the relations (1.45) for Z_o and obtain (1.25).

Incidentally, when we put $ZY = I\gamma^2 + R$ in (1.51) and rearrange the terms so as to get a power series in R we get the series (1.28).

1.10 Proof of the First Method

The first method consists essentially of determining Γ from the series expansion of (1.13):

$$\Gamma = \gamma \sum_{n=0}^{\infty} \frac{(-\frac{1}{2})_n}{n!} \frac{(-R)^n}{\gamma^{2n}}, \tag{1.53}$$

where $(-\frac{1}{2})_n = (-\frac{1}{2})(\frac{1}{2})(\frac{3}{2}) \cdots (n - \frac{3}{2})$ when $n > 0$ and $(-\frac{1}{2})_o = 1$, and then computing Z_o and the required exponential and hyperbolic functions of $x\Gamma$. From §1.7 and §1.8 it follows that the first method gives the correct result provided that Γ as determined by (1.53) satisfies the conditions: (a) its square is equal to ZY and (b) every element of $e^{-x\Gamma}$ approaches zero as $x \rightarrow \infty$.

These two conditions are satisfied by the matrix

$$\Gamma = PGP^{-1} \tag{1.54}$$

where P and G are matrices defined by equations (A1.1) and (A1.3) of Appendix I, G being a diagonal matrix whose rth element is γ_r. For from (A1.9) the square of Γ is

$$\Gamma^2 = PG^2P^{-1} = ZY$$

Furthermore,

$$\begin{aligned} e^{-x\Gamma} &= \sum_0^{\infty} \frac{(-x)^n}{n!} (PGP^{-1})^n \\ &= P \sum_0^{\infty} \frac{(-x)^n}{n!} G^n P^{-1} \\ &= PM(x)P^{-1} \end{aligned} \tag{1.55}$$

where $M(x)$ is diagonal matrix (A1.5) whose rth element is $e^{-\gamma_r x}$. Since the real part of γ_r is positive and the elements of P are independent of x it follows that the second condition is satisfied.

It will now be shown that PGP^{-1} may be expanded in the series (1.53) provided that γ may be chosen so as to make all of the points $\zeta_r = \dfrac{\gamma_r}{\gamma}$, $r = 1, 2, \cdots m$, in the complex ζ plane lie within that loop of the lemniscate $|\zeta^2 - 1| = 1$ which contains the point $\zeta = 1$. For then we may write the rth element in G as a convergent series:

$$\begin{aligned} \gamma_r &= \gamma\left(1 + \frac{\gamma_r^2 - \gamma^2}{\gamma^2}\right)^{\frac{1}{2}} \\ &= \gamma \sum_{n=0}^{\infty} \frac{(-\frac{1}{2})_n}{n!} \frac{(\gamma^2 - \gamma_r^2)^n}{\gamma^{2n}} \end{aligned} \tag{1.56}$$

and PGP^{-1} may be written as a convergent infinite series, the nth term of which contains the matrix (assuming only three circuits for the sake of simplicity)

$$P\begin{bmatrix} \gamma_1^2-\gamma^2 & 0 & 0 \\ 0 & \gamma_2^2-\gamma^2 & 0 \\ 0 & 0 & \gamma_3^2-\gamma^2 \end{bmatrix}^n P^{-1} = R^n, \tag{1.57}$$

where the equality follows from the definition (1.12) of R and equation (A1.9) of Appendix I. This series for PGP^{-1} is exactly the same as the series (1.53), and this completes the proof of the first method.

The equations (1.18) and (1.28) which are incidental to the first method, will now be established for the case in which the matrix Γ occurring in them is equal to PGP^{-1}. For then we have equation (1.55) and the equations

$$\cosh x\Gamma = P\begin{bmatrix} \cosh x\gamma_1 & 0 & 0 \\ 0 & \cosh x\gamma_2 & 0 \\ 0 & 0 & \cosh x\gamma_3 \end{bmatrix} P^{-1} \tag{1.58}$$

$$\sinh x\Gamma\, Z_o = \sinh x\Gamma\, \Gamma^{-1} Z$$

where $\sinh x\Gamma\, \Gamma^{-1}$ may be expressed in the same fashion as $\cosh x\Gamma$, the rth element of the diagonal matrix being $\dfrac{\sinh x\gamma_r}{\gamma_r}$. The elements in the diagonal matrices occurring in these expressions may be expanded in series by replacing γ_r by its representation (1.56), assuming $\left|\dfrac{\gamma_r^2}{\gamma^2} - 1\right| < 1$, and using[8]

$$e^{-z\sqrt{1+r}} = \sum_0^\infty \left(\frac{rz}{2}\right)^p \frac{(-)^p}{p!} \sqrt{\frac{2z}{\pi}}\, K_{p-\frac{1}{2}}(z)$$

$$\cosh z\sqrt{1+r} = \sum_0^\infty \left(\frac{rz}{2}\right)^p \frac{1}{p!} \sqrt{\frac{\pi z}{2}}\, I_{p-\frac{1}{2}}(z)$$

$$\frac{\sinh z\sqrt{1+r}}{\sqrt{1+r}} = \sum_0^\infty \left(\frac{rz}{2}\right)^p \frac{1}{p!} \sqrt{\frac{\pi z}{2}}\, I_{p+\frac{1}{2}}(z)$$

where $I_{p-\frac{1}{2}}(z)$ and $K_{p-\frac{1}{2}}(z)$ are Bessel functions of the first and second kinds, respectively, for imaginary argument. Equations (1.18) and (1.28) are obtained when equation (1.57) and the Bessel function recurrence relations are used.

[8] These are special cases of formulas given in "Theory of Bessel Functions," by G. N. Watson, page 141.

1.11 *Proof of the Second Method*

To establish the second method we must prove the various formulas which are used. These formulas all involve the square matrix $N(\gamma_r^2)$ defined by (1.23).

Since $N(\gamma_r^2)$ is proportional to $F(\gamma_r^2)$ it follows that $N(\gamma_r^2)$ may be expressed as

$$N(\gamma_r^2) = p_r\rho_r \cdots \tag{1.59}$$

where p_r is the column matrix defined in Appendix I and ρ_r is a row matrix specified by p_r and $N(\gamma_r^2)$. Applying Sylvester's theorem to the unit matrix gives

$$I = \sum N(\gamma_r^2) = \sum p_r\rho_r = [p_1, p_2, p_3]\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix}$$

where the two matrices on the extreme right are partitioned square matrices. From the definition of P in Appendix I it follows that

$$[p_1, p_2, p_3] = P, \qquad \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} = P^{-1} \tag{1.60}$$

These relations enable us to verify the equations (1.24) when Γ is equal to PGP^{-1}. Thus for the first of equations (1.24)

$$\Gamma = PGP^{-1} = [p_1, p_2, p_3]G\begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix}$$

$$= [p_1, p_2, p_3]\begin{bmatrix} \gamma_1\rho_1 \\ \gamma_2\rho_2 \\ \gamma_3\rho_3 \end{bmatrix} = \sum p_r\gamma_r\rho_r = \sum N(\gamma_r^2)\gamma_r$$

The second equation of (1.24) follows likewise from the expression (1.55) for $e^{-x\Gamma}$.

The third equation of (1.24) follows at once from the first when we use (1.45), $Z_o = \Gamma Y^{-1}$. The fourth equation is obtained by writing

$$e^{-x\Gamma}Z_o = PM(x)P^{-1}PGP^{-1}Y^{-1}$$

$$= PM(x)GP^{-1}Y^{-1}$$

and proceeding as in the case of the first equation.

All of the other equations connected with the second method may be proved in a similar way. Incidentally, the formulas obtained by the second method are closely related to the "special form of solution" described in §6.5 of F.D.C.

PART II

TRANSMISSION LINES COMPOSED OF MULTI-TERMINAL SECTIONS

2.1 Introductory

Some transmission systems may be regarded as consisting of a number of identical sections connected in tandem. The question of determining the steady state electrical behavior of such a system from a knowledge of the properties of a single section will be considered here.

Each section will have a certain number, say $m + 1$, terminals on its left end and an equal number on the right. The case in which there are only two terminals ($m = 1$) has been completely worked out, and some studies of more general cases have been made. The ones which most nearly approach the point of view of the present paper are those due to S. Koizumi[9].

In the present work difference equations are used to solve the general case in much the same manner as they have been used in studying the two-terminal case. This approach differs from that used by Koizumi and throws additional light on the problem.

In several lists of formulas, particularly in Appendix IV, I have included a number of results due to Koizumi for the sake of completeness.

2.2 Transmission Equations for a Typical Section

We consider the equations connecting the input and output currents and voltages for the nth section which is shown in Fig. 1. The directions which are assumed for positive current flow are indicated by arrows. The leads marked 0 play a special role in that all the voltages are

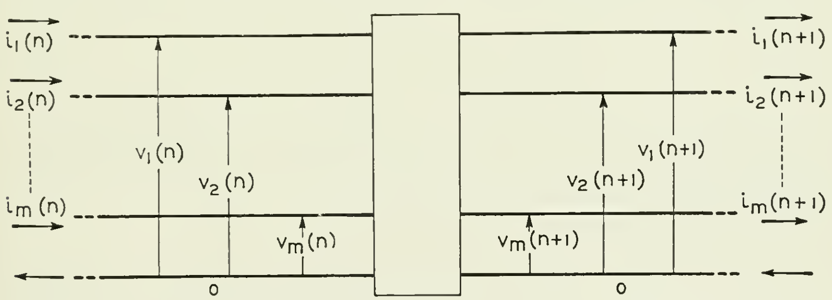


FIG. 1

measured with respect to them, and the currents which they carry are the sum of the currents flowing into or out of the remaining terminals at the end under consideration. In applications to transmission lines the terminals 0 would correspond to the ground or the cable sheath.

The currents and voltages shown in Fig. 1 are related by a number of

[9] Archiv für Electrotechnik, Vol. 33, pp. 171–188, 609–622 (1939). See also a paper by M. G. Malti and S. E. Warschawski, Trans. A.I.E.E., Vol. 56, pp. 153–158 (1937).

sets of $2m$ linear equations which may be conveniently written in matrix form. One such set is

$$\begin{aligned} v(n) &= Z_{11}i(n) - Z_{12}i(n+1) + v^\circ(n) \\ v(n+1) &= Z_{21}i(n) - Z_{22}i(n+1) + u^\circ(n) \end{aligned} \tag{2.1}$$

Z_{11}, Z_{12}, Z_{21}, Z_{22} are square matrices of order m whose elements are impedances. $v(n)$ and $i(n)$ are the column matrices

$$v(n) = \begin{bmatrix} v_1(n) \\ v_2(n) \\ \cdot \\ \cdot \\ v_m(n) \end{bmatrix} \qquad i(n) = \begin{bmatrix} i_1(n) \\ i_2(n) \\ \cdot \\ \cdot \\ i_m(n) \end{bmatrix}$$

The column matrices $v^\circ(n)$ and $u^\circ(n)$ arise from generators which may be acting within the nth section. If both ends of the section are open circuited so that $i(n) = i(n+1) = 0$ the equations show that $v(n) = v^\circ(n)$, $v(n+1) = u^\circ(n)$. Consequently, $v^\circ(n)$ and $u^\circ(n)$ give the open circuit voltages produced on the left and right ends of the nth section by the internal generators. If the section is a passive network then $v^\circ(n) = u^\circ(n) = 0$ and they do not appear in the equations. The subscripts on the square matrices, the Z's, are chosen so as to preserve the analogy for the simple case $m = 1$, where the left and right ends of the section are denoted by the subscripts 1 and 2, respectively.

Solving the equation (1.1) for $i(n)$ and $i(n+1)$ gives

$$\begin{aligned} i(n) &= Y_{11}v(n) + Y_{12}v(n+1) + i^\circ(n) \\ -i(n+1) &= Y_{21}v(n) + Y_{22}v(n+1) - j^\circ(n) \end{aligned} \tag{2.2}$$

where the elements of the Y's are admittances and $i^\circ(n)$, $j^\circ(n)$ are the currents produced by the internal sources when the terminals on the right and left are short-circuited so that $v(n) = v(n+1) = 0$.

A third set of equations is

$$\begin{aligned} v(n) &= Av(n+1) + Bi(n+1) - Bj^\circ(n) \\ i(n) &= Cv(n+1) + Di(n+1) - Cu^\circ(n) \end{aligned} \tag{2.3}$$

Solving these equations for $v(n+1)$ and $i(n+1)$ gives

$$\begin{aligned} v(n+1) &= D'v(n) - B'i(n) + B'i^\circ(n) \\ i(n+1) &= -C'v(n) + A'i(n) + C'v^\circ(n) \end{aligned} \tag{2.4}$$

There are a great many relations between the square matrices appearing in the equations (2.1) to (2.4). These are discussed in Appendix IV.

For symmetrical sections $Y_{21} = Y_{12}$, $Y_{22} = Y_{11}$, $Z_{21} = Z_{12}$, $Z_{22} = Z_{11}$ and equations (2.1) and (2.2) become

$$\begin{aligned} v(n) &= Z_{11}i(n) - Z_{12}i(n+1) + v^\circ(n) \\ v(n+1) &= Z_{12}i(n) - Z_{11}i(n+1) + u^\circ(n) \end{aligned} \tag{2.5}$$

$$\begin{aligned} i(n) &= Y_{11}v(n) + Y_{12}v(n+1) + i^\circ(n) \\ -i(n+1) &= Y_{12}v(n) + Y_{11}v(n+1) - j^\circ(n) \end{aligned} \tag{2.6}$$

Eliminating $i(n)$ from (2.5) and $v(n)$ from (2.6) and using, from (A4.4), $A = Z_{11}Z_{12}^{-1} = -Y_{12}^{-1}Y_{11}$ leads to the difference equations

$$v(n+1) + v(n-1) - 2Av(n) = B[i^\circ(n) - j^\circ(n-1)] \tag{2.7}$$

$$i(n+1) + i(n-1) - 2A'i(n) = C[v^\circ(n) - u^\circ(n-1)] \tag{2.8}$$

Since we also have $B' = B$, $C' = C$, $D' = A$ for symmetrical sections equations (2.4) become

$$\begin{aligned} v(n+1) &= Av(n) - Bi(n) + Bi^\circ(n) \\ i(n+1) &= -Cv(n) + A'i(n) + Cv^\circ(n) \end{aligned} \tag{2.9}$$

We assume that the distribution of the sources in the branches of a symmetrical network need not be symmetrical with respect to the two ends, even though the impedances of the branches are.

2.3 *Statement of Results for Infinite Symmetrical Section Line—Passive*

When the sections are passive the equations to be solved are, from (2.9),

$$\begin{aligned} v(n+1) &= Av(n) - Bi(n) \\ i(n+1) &= -Cv(n) + A'i(n) \end{aligned} \tag{2.10}$$

If the line extends from $n = 0$ to $n = \infty$ the solution is

$$\begin{aligned} v(n) &= e^{-n\Gamma}v(o) \\ i(n) &= e^{-n\Gamma'}i(o) \\ v(n) &= Z_o i(n) \end{aligned} \tag{2.11}$$

where the matrix $e^{-\Gamma}$ is such that (a) the equation

$$e^{-\Gamma} + e^{\Gamma} = 2A \tag{2.12}$$

is satisfied, e^{Γ} being the inverse of $e^{-\Gamma}$, and (b) all the elements of the matrix $e^{-n\Gamma}$ approach zero as $n \to \infty$. In dealing with sections we shall never have occasion to consider Γ itself but only its exponential and associated functions. The characteristic impedance matrix Z_o is defined by the relation between the initial currents and voltages in an infinite line

$$v(o) = Z_o i(o) \tag{2.13}$$

A formal solution of (2.12) may be obtained by writing it as

$$\cosh \Gamma = A \tag{2.14}$$

Then

$$\begin{aligned} e^{-\Gamma} &= \cosh \Gamma - \sinh \Gamma \\ &= A - (A^2 - I)^{\frac{1}{2}} = A - (BC)^{\frac{1}{2}} \end{aligned}$$

where the square root is to be chosen so that condition (b) for $e^{-\Gamma}$ is satisfied. The characteristic impedance matrix Z_o is given by equations (2.34) of which the following two are representative.

$$Z_o = (\sinh \Gamma)^{-1} B = \sinh \Gamma \, C^{-1} \tag{2.15}$$

where $\sinh \Gamma$ is given by $2 \sinh \Gamma = e^{\Gamma} - e^{-\Gamma}$.

The wide variety of sections makes it appear unlikely that there is a general method of determining $e^{-\Gamma}$ analogous to the first method discussed for the uniform line. However, in some cases rapidly convergent series for $e^{-\Gamma}$ and e^{Γ} may be obtained. For example, suppose that the elements of $(2A)^{-1}$ are small compared to those of 2A. Then, from (2.12),

$$\begin{aligned} e^{\Gamma} &= 2A - (2A)^{-1} - (2A)^{-3} - 2(2A)^{-5} - \cdots \\ e^{-\Gamma} &= (2A)^{-1} + (2A)^{-3} + 2(2A)^{-5} + \cdots \end{aligned}$$

Again, if $A^2 - I = BC$ is expressible as $I\gamma^2 + R$ where the elements of R are small in comparison with γ^2, we have (cf. equations (1.14), (1.15))

$$e^{\Gamma} = A + \gamma \left[I + \frac{R}{2\gamma^2} - \frac{1}{2}\left(\frac{R}{2\gamma^2}\right)^2 + \cdots \right]$$

$$e^{-\Gamma} = A - \gamma \left[I + \frac{R}{2\gamma^2} - \frac{1}{2}\left(\frac{R}{2\gamma^2}\right)^2 + \cdots \right]$$

Finally, it follows from a comparison of equations (2.11) and (A2.12) that a suitable $e^{-\Gamma}$ is given by

$$e^{-\Gamma} = P\Lambda P^{-1}, \qquad e^{-\Gamma'} = Q\Lambda Q^{-1} \tag{2.16}$$

where P, Q and Λ are the matrices designated by the same symbols in Appendices II and III.

The formal application of Sylvester's theorem leads to a method of solving the symmetrical section line which is analogous to the second method discussed for the uniform line. Thus, if $P(A)$ is any polynomial in A, then

$$P(A) = \sum_{r=1}^{m} N(\zeta_r) P(\zeta_r) \tag{2.17}$$

where $P(\zeta_r)$ is not a square matrix but a scalar and $N(\zeta_r)$ is the square matrix

$$N(\zeta_r) = \frac{F(\zeta_r)}{|f(\zeta_r)|^{(1)}}. \tag{2.18}$$

$F(\zeta)$ is the adjoint of the characteristic matrix

$$f(\zeta) = I\zeta - A \tag{2.19}$$

and $\zeta_1, \zeta_2, \cdots \zeta_m$ are the roots, assumed to be unequal, of the characteristic equation

$$|f(\zeta)| = 0.$$

The denominator in the expression for $N(\zeta_r)$ is the derivative of the characteristic function:

$$|f(\zeta_r)|^{(1)} = \left[\frac{d}{d\zeta}|f(\zeta)|\right]_{\zeta=\zeta_r}$$

The formal application of Sylvester's theorem then gives

$$\begin{aligned} \cosh\Gamma &= A = \Sigma N(\zeta_r)\zeta_r \\ e^{-\Gamma} &= A - (A^2 - I)^{\frac{1}{2}} = \Sigma N(\zeta_r)\lambda_r \\ e^{-n\Gamma} &= \Sigma N(\zeta_r)\lambda_r^n \\ Z_o &= (\sinh\Gamma)^{-1}B = \Sigma \frac{N(\zeta_r)\,2}{(\lambda_r^{-1} - \lambda_r)} B \end{aligned} \tag{2.20}$$

where $N(\zeta_r)$ is given by (2.18), the summations run from $r = 1$ to $r = m$, and λ_r is related to ζ_r through

$$2\zeta_r = \lambda_r + \lambda_r^{-1}, \qquad \lambda_r = \zeta_r - \sqrt{\zeta_r^2 - 1} \tag{2.21}$$

where the sign of the square root is chosen so that $|\lambda_r| < 1$. λ_r is related to $e^{-\Gamma}$ in the same way that ζ_r is related to $\cosh\Gamma$.

2.4 *Results for Any Symmetrical Section Line—Passive*

The solutions of equations (2.10) which reduce to the given values $v(o)$, $i(o)$ at $n = 0$ are

$$\begin{aligned} v(n) &= \cosh n\Gamma\, v(o) - \sinh n\Gamma\, Z_o i(o) \\ i(n) &= -\sinh n\Gamma'\, Z_o^{-1}\, v(o) + \cosh n\Gamma'\, i(o) \end{aligned} \tag{2.22}$$

where $e^{-\Gamma}$ and Z_o are the matrices of §2.3. These may be put in slightly different form by using the relations

$$\sinh n\Gamma\, Z_o = Z_o \sinh n\Gamma'$$

$$\sinh n\Gamma'\, Z_o^{-1} = Z_o^{-1} \sinh n\Gamma$$

When (2.22) are interpreted by Sylvester's theorem we obtain

$$\begin{aligned} v(n) &= \Sigma N(\zeta_r)\left[\tfrac{1}{2}(\lambda_r^{-n} + \lambda_r^n)v(o) - \frac{\lambda_r^{-n} - \lambda_r^n}{\lambda_r^{-1} - \lambda_r} Bi(o)\right] \\ i(n) &= \Sigma N'(\zeta_r)\left[-\frac{\lambda_r^{-n} - \lambda_r^n}{\lambda_r^{-1} - \lambda_r} Cv(o) + \tfrac{1}{2}(\lambda_r^{-n} + \lambda_r^n)i(o)\right] \end{aligned} \tag{2.23}$$

where $N'(\zeta_r)$ is the transposed of $N(\zeta_r)$ and $N(\zeta_r)$ is given by (2.18) and the summations run from $r = 1$ to $r = m$.

2.5 Results for Any Symmetrical Section Line—Active

When the sections contain generators the equations to be solved are those of (2.9). The solutions corresponding to the initial values $v(o)$ and $i(o)$ are, for $n \geqq 1$,

$$\begin{aligned} v(n) &= \cosh n\Gamma\, v(o) - \sinh n\Gamma\, Z_o i(o) \\ &\quad + \sum_{p=1}^{n} \{\cosh (n - p)\Gamma\, Bi^\circ(p - 1) - \sinh (n - p)\Gamma\, Z_o C v^\circ(p - 1)\} \\ i(n) &= -\sinh n\Gamma'\, Z_o^{-1} v(o) + \cosh n\Gamma'\, i(o) \\ &\quad + \sum_{p=1}^{n} \{\cosh (n - p)\Gamma'\, Cv^\circ(p - 1) - \sinh (n - p)\Gamma'\, Z_o^{-1} Bi^\circ(p - 1)\} \end{aligned} \tag{2.24}$$

These may be simplified somewhat by replacing $Z_o C$ and $Z_o^{-1} B$ by $\sinh \Gamma$ and $\sinh \Gamma'$, respectively.

The series in the above expressions may be summed when the generators are such that

$$v^\circ(n) = e^{-n\theta} i^\circ, \qquad i^\circ(n) = e^{-n\theta} v^\circ \tag{2.25}$$

where θ is a scalar and i° and v° are column matrices whose elements are independent of n. Thus

$$\begin{aligned} v(n) &= \cosh n\Gamma\, v(o) - \sinh n\Gamma\, Z_o i(o) \\ &\quad + \tfrac{1}{2}(e^{n\Gamma} - e^{-n\theta} I)(e^{\Gamma} - e^{-\theta} I)^{-1}(Bi^\circ - Z_o C v^\circ) \\ &\quad + \tfrac{1}{2}(e^{-n\Gamma} - e^{-n\theta} I)(e^{-\Gamma} - e^{-\theta} I)^{-1}(Bi^\circ + Z_o C v^\circ) \\ i(n) &= -\sinh n\Gamma'\, Z_o^{-1} v(o) + \cosh n\Gamma'\, i(o) \\ &\quad + \tfrac{1}{2}(e^{n\Gamma'} - e^{-n\theta} I)(e^{\Gamma'} - e^{-\theta} I)^{-1}(Cv^\circ - Z_o^{-1} Bi^\circ) \\ &\quad + \tfrac{1}{2}(e^{-n\Gamma'} - e^{-n\theta} I)(e^{-\Gamma'} - e^{-\theta} I)^{-1}(Cv^\circ + Z_o^{-1} Bi^\circ) \end{aligned} \tag{2.26}$$

provided that the inverse matrices exist.

We may interpret these expressions by Sylvester's theorem. For example,

$$v(n) = \sum_{r=1}^{m} N(\zeta_r)\left[\tfrac{1}{2}(\lambda_r^{-n} + \lambda_r^{n})v(o) - \frac{\lambda_r^{-n} - \lambda_r^{n}}{\lambda_r^{-1} - \lambda_r} Bi(o) \right.$$
$$+ \frac{1}{2}\frac{\lambda_r^{-n} - e^{-n\theta}}{\lambda_r^{-1} - e^{-\theta}}\left(Bi^\circ - \frac{2BCv_o}{\lambda_r^{-1} - \lambda_r}\right) \tag{2.27}$$
$$\left. + \frac{1}{2}\frac{\lambda_r^{n} - e^{-n\theta}}{\lambda_r - e^{-\theta}}\left(Bi^\circ + \frac{2BCv_o}{\lambda_r^{-1} - \lambda_r}\right)\right]$$

where $N(\zeta_r)$ is given by (2.18).

When the line extends to $n = \infty$ and the sources and end conditions satisfy suitable conditions we have the relation

$$v(o) = Z_o i(o) - \sum_{p=1}^{\infty} e^{-p\Gamma}[Bi^\circ(p-1) - Z_o Cv^\circ(p-1)] \tag{2.28}$$

When the impressed field is of the form (2.25) this becomes

$$v(o) = Z_o i(o) - (e^{\Gamma} - e^{-\theta}I)^{-1}(Bi^\circ - Z_o Cv^\circ) \tag{2.29}$$

provided that the inverse matrix exists. Expressions for $v(n)$ and $i(n)$ in such an infinite line may be obtained by using (2.28) or (2.29) in (2.24) or (2.26).

Applying Sylvester's theorem to (2.29) gives

$$v(o) = \sum_{r=1}^{m} N(\zeta_r)\left(\frac{2Bi(o)}{\lambda_r^{-1} - \lambda_r} - \frac{Bi^\circ}{\lambda_r^{-1} - e^{-\theta}} + \frac{\lambda_r^{-1} - \lambda_r}{2(\lambda_r^{-1} - e^{-\theta})}v^\circ\right) \tag{2.30}$$

The last term within the braces may be replaced by

$$\frac{2BCv^\circ}{(\lambda_r^{-1} - \lambda_r)(\lambda_r^{-1} - e^{-\theta})}$$

2.6 *Derivation of the Properties of an Infinite Line*

We shall consider a symmetrical section line which is specified by the equations

$$\begin{aligned} v(n+1) &= Av(n) - Bi(n) \\ i(n+1) &= -Cv(n) + A'i(n) \end{aligned} \tag{2.10}$$

From these equations and the relations $A^2 - BC = I$, $AB = BA'$, $A'C = CA$ of (A4.6) it follows that

$$\begin{aligned} v(n+1) + v(n-1) &= 2Av(n) \\ i(n+1) + i(n-1) &= 2A'i(n) \end{aligned} \tag{2.31}$$

If $e^{-\Gamma}$ is a matrix satisfying the conditions of §2.3, namely, (a) $e^{-\Gamma}$ satisfies the equation

$$2 \cosh \Gamma = e^{\Gamma} + e^{-\Gamma} = 2A \tag{2.14}$$

and (b), every element in $e^{-n\Gamma}$ approaches zero as $n \longrightarrow \infty$, and if Z_o is defined by $v(o)$ and $i(o)$ for an infinite line as in (2.13), then

1. In an infinite line

$$\begin{aligned} v(n) &= e^{-n\Gamma} v(o), \\ i(n) &= e^{-n\Gamma'} i(o), \end{aligned} \tag{2.32}$$

$$v(n) = Z_o i(n) \tag{2.33}$$

2. The characteristic impedance matrix Z_o is given by

$$\begin{aligned} Z_o &= (\sinh \Gamma)^{-1} B = B(\sinh \Gamma')^{-1} = C^{-1} \sinh \Gamma' = \sinh \Gamma\, C^{-1} \\ Z_o^{-1} &= B^{-1} \sinh \Gamma = \sinh \Gamma'\, B^{-1} = (\sinh \Gamma')^{-1} C = C(\sinh \Gamma)^{-1} \end{aligned} \tag{2.34}$$

3. The matrices Z_o, B and C obey the commutation rules

$$\begin{aligned} \Phi(e^{\Gamma}) Z_o &= Z_o \Phi(e^{\Gamma'}) \\ \Phi(e^{\Gamma}) B &= B \Phi(e^{\Gamma'}) \\ C\Phi(e^{\Gamma}) &= \Phi(e^{\Gamma'}) C \end{aligned} \tag{2.35}$$

where $\Phi(e^{\Gamma})$ is a square matrix representable as a sum of powers of $e^{\pm\Gamma}$. The matrices $\Phi(e^{\Gamma})\, Z_0$, $\Phi(e^{\Gamma})\, B$, and $C\Phi(e^{\Gamma})$ are symmetrical.

To prove these statements we proceed as follows: By direct substitution into (2.31) it is seen that $v(n) = e^{-n\Gamma} v(o)$ is a solution by virtue of condition (a) satisfied by $e^{-\Gamma}$. Since, by condition (b), $v(n) \longrightarrow o$ as $n \longrightarrow \infty$ it follows that $v(n)$ is the voltage in an infinite line. Similarly, $i(n) = e^{-n\Gamma'} i(o)$ is the current in such a line. Substituting the expressions (2.32) for $v(n)$ and $i(n)$ into the difference equations (2.10), setting $n = 0$, using the definition of Z_o, and regarding $v(o)$ and $i(o)$ as arbitrary columns gives

$$\begin{aligned} e^{-\Gamma} &= A - BZ_o^{-1} \\ e^{-\Gamma'} &= -CZ_o + A' \end{aligned} \tag{2.36}$$

Applying condition (a) in the form of (2.14) to these equations gives

$$BZ_o^{-1} = \sinh \Gamma, \qquad CZ_o = \sinh \Gamma' \tag{2.37}$$

Since the sections are symmetrical, B and C are symmetrical matrices, and from the reciprocal theorem for networks it follows that Z_o is also symmetrical. These remarks and (2.37) lead to (2.34). Setting the expressions (2.32) for $v(n)$ and $i(n)$ in the second of the difference equations (2.10) using the definition of Z_o, and regarding $i(o)$ as an arbitrary column gives

$$e^{-(n+1)\Gamma'} = -Ce^{-n\Gamma}Z_o + A'e^{-n\Gamma'}$$

$$(A' - e^{-\Gamma'})e^{-n\Gamma'} = Ce^{-n\Gamma}Z_o$$

Replacing $A' - e^{-\Gamma'}$ by CZ_o, as follows from the case $n = 0$, and premultiplying by C^{-1} gives

$$Z_oe^{-n\Gamma'} = e^{-n\Gamma}Z_o$$

and this leads to the first of equations (2.35). From (2.34) and the relations $AB = BA'$, $CA = A'C$ we have

$$\sinh \Gamma\, B = B \sinh \Gamma' \qquad \cosh \Gamma\, B = B \cosh \Gamma'$$

$$C \sinh \Gamma = \sinh \Gamma'\, C \qquad C \cosh \Gamma = \cosh \Gamma'\, C$$

Addition and subtraction leads to

$$e^{\pm\Gamma}B = Be^{\pm\Gamma'} \qquad Ce^{\pm\Gamma} = e^{\pm\Gamma'}C$$

from which the last two of equations (2.35) follow. Since each of equations (2.35) expresses the equality of a matrix and its transposed, it follows that the matrices are symmetrical.

Equation (2.33), which is almost self-evident on physical grounds, follows from

$$v(n) = e^{-n\Gamma}v(o) = e^{-n\Gamma}Z_oi(o)$$

$$= Z_oe^{-n\Gamma'}i(o) = Z_oi(n).$$

2.7 *Proof of Relations for Any Symmetrical Section Line*

The expressions (2.24) for $v(n)$ and $i(n)$ in a line whose sections contain generators may be verified to satisfy the difference equations (2.9). The expressions (2.34) for Z_o and the commutation rules (2.35) for B and C are used in the verification. Setting $n = 1$ in the expressions for $v(n)$ and $i(n)$ gives the difference equations (2.9) and hence $v(n)$ and $i(n)$ are the solutions which correspond to the initial values $v(o)$ and $i(o)$.

In order to derive the relation (2.28) between $v(o)$ and $i(o)$ for an infinite line we put the hyperbolic functions in the expression (2.24) for $v(n)$ in exponential form and multiply through by $2e^{-n\Gamma}$

$$\begin{aligned} 2e^{-n\Gamma}v(n) = v(o) - Z_oi(o) &+ \sum_{p=1}^{n} e^{-p\Gamma}[Bi^\circ(p-1) - Z_oCv^\circ(p-1)] \\ &+ e^{-2n\Gamma}[v(o) + Z_oi(o)] \\ &+ e^{-n\Gamma}\sum_{p=1}^{n} e^{-(n-p)\Gamma}[Bi^\circ(p-1) + Z_oCv^\circ(p-1)] \end{aligned}$$

Hence, letting $n \to \infty$ and using condition (b) satisfied by $e^{-\Gamma}$, equation (2.28) is obtained provided that (i) the terminal conditions at the far end are such that $v(n)$ remains finite, (ii) the sum in (2.28) converges, and (iii) the expression in the last line in the equation just above approaches zero as $n \to \infty$.

The results obtained by the formal application of Sylvester's theorem may be verified by using the results of Appendix II and writing $N(\zeta_r)$ as the product of a column matrix and a row matrix. They may also be verified more directly. For example, setting $n = 0$ in the expressions (2.23) for $v(n)$ and $i(n)$ in any passive symmetrical section line and using

$$\sum_{r=1}^{m} N(\zeta_r) = I, \tag{2.38}$$

which follows from Sylvester's theorem, we see that $v(n)$ and $i(n)$ reduce to the appropriate values $v(o)$ and $i(o)$ at $n = 0$. Substituting $v(n)$ and $i(n)$ into the difference equations (2.10) and using

$$\begin{aligned} BC &= A^2 - I \\ (I\zeta_r - A)N(\zeta_r) &= N(\zeta_r)(I\zeta_r - A) = 0 \\ BN'(\zeta) &= N(\zeta)B \\ CN(\zeta) &= N'(\zeta)C, \end{aligned} \tag{2.39}$$

shows that they are solutions. The second of the relations (2.39) follows from the fact that $N(\zeta_r)$ is proportional to the adjoint $F(\zeta_r)$ of $f(\zeta_r)$. In the third and fourth relations

$$N(\zeta) = \frac{F(\zeta)}{|f(\zeta)|^{(1)}}$$

which is in agreement with the definition (2.18) of $N(\zeta_r)$. To establish the third relation we start from,[10]

$$\begin{aligned} (\zeta I - A)\, F(\zeta) &= I\,|f(\zeta)| \\ (\zeta I - A)\, N(\zeta) &= I\,|f(\zeta)|/|f(\zeta)|^{(1)} \end{aligned}$$

Postmultiplication by B gives

$$(\zeta I - A)\, N(\zeta)\, B = B\,|f(\zeta)|/|f(\zeta)|^{(1)}$$

We also have

$$\begin{aligned} (\zeta I - A')F'(\zeta) &= I\,|f(\zeta)| \\ (\zeta I - A')N'(\zeta) &= I\,|f(\zeta)|/|f(\zeta)|^{(1)} \end{aligned}$$

[10] F.D.C. §3.5.

Premultiplication by B and use of $BA' = AB$ gives

$$(\zeta I - A)BN'(\zeta) = B\,|\,f(\zeta)\,|/|\,f(\zeta)\,|^{(1)}$$

Hence, the third equation in (2.39) holds except possibly for $\zeta = \zeta_r$, and from the concept of continuity it holds there also. The fourth equation in (2.39) may be proved in the same manner.

The expression (2.20) for Z_o may be obtained by letting n become very large in the expression (2.23) for $v(n)$. $v(o)$ and $i(o)$ must be related so that $v(n)$ remains finite. Since $|\lambda_r| < 1$ and the λ_r's are unequal the coefficients of λ_r^{-n} must vanish. This requires

$$N(\zeta_r)v(o) = \frac{2N(\zeta_r)Bi(o)}{\lambda_r^{-1} - \lambda r}$$

Summing r from 1 to m and using (2.38) gives the required expression for Z_o.

2.8 The Unsymmetrical Section Line

The method used here is analogous to those described in Appendices I and II for the uniform line and the symmetrical section line. The other methods apparently do not lead to the simplification which occurs in the symmetrical case.

Equations (2.2) and (2.1) lead to the difference equations

$$Y_{12}v(n+2) + [Y_{11} + Y_{22}]v(n+1) + Y_{21}v(n) = -i^\circ(n+1) + j^\circ(n) \quad (2.40)$$

$$Z_{12}i(n+2) - [Z_{11} + Z_{22}]i(n+1) + Z_{21}i(n) = v^\circ(n+1) - u^\circ(n) \quad (2.41)$$

Both of these equations are of the form

$$Gx(n+2) + Hx(n+1) + G'x(n) = g(n) \quad (2.42)$$

in which G and H are square matrices of order m, H being symmetrical and G' being the transposed of G. When the sections are passive equations (2.40) and (2.41) become

$$Y_{12}v(n+2) + [Y_{11} + Y_{22}]v(n+1) + Y_{21}v(n) = 0 \quad (2.43)$$

$$Z_{12}i(n+2) - [Z_{11} + Z_{22}]i(n+1) + Z_{21}i(n) = 0 \quad (2.44)$$

In the passive, unsymmetrical case the expressions for $v(n)$ and $i(n)$ are of the form

$$\begin{aligned} v(n) &= P\Lambda^n a + \bar{P}\Lambda^{-n}\bar{a} \\ i(n) &= Q\Lambda^n a - \bar{Q}\Lambda^{-n}\bar{a} \end{aligned} \quad (2.45)$$

Comparison with (A2.8) shows that in the symmetrical case $\bar{P} = P$ and $\bar{Q} = Q$. The minus signs over $\bar{P}$, $\bar{Q}$, and $\bar{a}$ indicate that they are associated with propagation in the negative direction. The propagation constants of the m modes of propagation are the same in the positive as in the negative direction, as indicated by the appearance of Λ^n and Λ^{-n} in (2.45). Corresponding to any given propagation constant say λ_r, there are two modes of propagation, one in a positive direction and the other in the negative direction. The distribution of the voltages corresponding to these two modes are given by the rth columns in P and $\bar{P}$, respectively. The fact that P and $\bar{P}$ differ shows that the distributions differ according to the direction of propagation even though the propagation constant is the same. Λ is still the diagonal matrix defined in (A2.3) but now the computation of the elements λ_r is more difficult than when the section is symmetrical. They are defined as the roots of the equation

$$| G\lambda^2 + H\lambda + G' | = 0 \tag{2.46}$$

which are less than unity in absolute value. The second of the equations (A4.5) shows that the roots of (2.46) are the same whether the Z's or the Y's are used in place of G and H. Of course, this is to be expected on physical grounds. The third of the equations (A4.5) may be used to show that the roots of (2.46) are also the roots of

$$\begin{vmatrix} \lambda A - I & \lambda B \\ \lambda C & \lambda D - I \end{vmatrix} = 0 \tag{2.47}$$

From the form of (2.46) it follows that if λ_r is a root so is λ_r^{-1}. This fact may be used to simplify the determination of λ_r. When the substitution

$$2\zeta = \lambda + \lambda^{-1}, \qquad \lambda = \zeta - \sqrt{\zeta^2 - 1} \tag{A2.4}$$

is made equation (2.46) may be written as

$$0 = | (G + G')\zeta + H + (G' - G)\sqrt{\zeta^2 - 1} |$$

$0 = | (G + G')\zeta + H |$

$+ (\zeta^2 - 1)$ times the sum of $\frac{m(m-1)}{2!}$ determinants each obtained by replacing two columns of $|(G + G')\zeta + H|$ by the corresponding columns of $(G - G')$

$+ (\zeta^2 - 1)^2$ times the sum of $\frac{m(m-1)(m-2)(m-3)}{4!}$ determinants each obtained by replacing four columns of $|(G + G')\zeta + H|$ by the corresponding columns of $(G - G')$

$+ \cdots$

The last equation is a polynomial of degree m in ζ which is to be solved for its roots $\zeta_1, \zeta_2, \cdots \zeta_m$. For simplicity we assume that these roots are distinct. λ_r is then determined from ζ_r by the relations (A2.4), the sign of the radical being chosen so that $|\lambda_r| < 1$ as in the symmetrical case. In his second paper Koizumi has given a procedure which amounts to an alternative method of determining Λ.

We shall first assume that the Y's are known and that our equations are

$$\begin{aligned} i(n) &= Y_{11}v(n) + Y_{12}v(n+1) \\ -i(n+1) &= Y_{21}v(n) + Y_{22}v(n+1) \end{aligned} \tag{2.48}$$

As described above Λ may be computed from the determinantal equation

$$|f(\lambda)| = 0$$

where $f(\lambda)$ represents the matrix

$$f(\lambda) = Y_{12}\lambda^2 + (Y_{11} + Y_{22})\lambda + Y_{21} \tag{2.49}$$

Let p_r be proportional to any non-zero column in $F(\lambda_r)$ where $F(\lambda)$ is the adjoint of $f(\lambda)$ and let $\bar{p}_r'$ be proportional to any non-zero row of $F(\lambda_r)$. Then the matrices P and $\bar{P}$ in the expressions (2.45) for $v(n)$ and $i(n)$ are given by

$$\begin{aligned} P &= [p_1, p_2, \cdots p_m] \\ \bar{P} &= [\bar{p}_1, \bar{p}_2, \cdots \bar{p}_m] \end{aligned} \tag{2.50}$$

where $\bar{p}_r$ is the column obtained by transposing the row $\bar{p}_r'$. The matrices Q and $\bar{Q}$ are obtained from P and $\bar{P}$ by means of the equations

$$\begin{aligned} Q &= Y_{11}P + Y_{12}P\Lambda = -Y_{22}P - Y_{21}P\Lambda^{-1} \\ \bar{Q} &= -Y_{11}\bar{P} - Y_{12}\bar{P}\Lambda^{-1} = Y_{22}\bar{P} + Y_{21}\bar{P}\Lambda \end{aligned} \tag{2.51}$$

which are derived from (2.45) and (2.48).

The properties of the individual columns of P and $\bar{P}$ lead to the relations

$$\begin{aligned} Y_{12}P\Lambda^2 + (Y_{11} + Y_{22})P\Lambda + Y_{21}P &= 0 \\ Y_{12}\bar{P}\Lambda^{-2} + (Y_{11} + Y_{22})\bar{P}\Lambda^{-1} + Y_{21}\bar{P} &= 0 \end{aligned} \tag{2.52}$$

and these guarantee that the difference equations (2.48) will be satisfied when the expressions (2.45) for $v(n)$ and $i(n)$ are used.

When the Z's are known instead of the Y's the procedure is much the same. The difference equations are

$$\begin{aligned} v(n) &= Z_{11}i(n) - Z_{12}i(n+1) \\ v(n+1) &= Z_{21}i(n) - Z_{22}i(n+1) \end{aligned} \tag{2.53}$$

and the equation to determine the λ_r's is

$$|f(\lambda)| = 0$$

where now $f(\lambda)$ represents the matrix

$$f(\lambda) = Z_{12}\lambda^2 - (Z_{11} + Z_{22})\lambda + Z_{21} \tag{2.54}$$

Let q_r be proportional to any non-zero column in $F(\lambda_r)$ where $F(\lambda)$ is the adjoint of $f(\lambda)$ and let $\bar{q}_r'$ be proportional to any non-zero row of $F(\lambda_r)$. The matrices Q and $\bar{Q}$ in the expressions (2.45) for $v(n)$ and $i(n)$ are given by

$$\begin{aligned} Q &= [q_1, q_2, \cdots q_m] \\ \bar{Q} &= [\bar{q}_1, \bar{q}_2, \cdots \bar{q}_m] \end{aligned} \tag{2.55}$$

where $\bar{q}_r$ is the column obtained by transposing the row $\bar{q}_r'$. From (2.45) and (2.53) equations for P and $\bar{P}$ in terms of Q and $\bar{Q}$ are obtained:

$$\begin{aligned} P &= Z_{11}Q - Z_{12}Q\Lambda \phantom{^{-1}} = -Z_{22}Q + Z_{21}Q\Lambda^{-1} \\ \bar{P} &= -Z_{11}\bar{Q} + Z_{12}\bar{Q}\Lambda^{-1} = Z_{22}\bar{Q} - Z_{21}\bar{Q}\Lambda \end{aligned} \tag{2.56}$$

The difference equations (2.53) are satisfied by our expressions for $v(n)$ and $i(n)$ by virtue of the relations

$$\begin{aligned} Z_{12}Q\Lambda^2 - (Z_{11} + Z_{22})Q\Lambda + Z_{21}Q &= 0 \\ Z_{12}\bar{Q}\Lambda^{-2} - (Z_{11} + Z_{22})\bar{Q}\Lambda^{-1} + Z_{21}\bar{Q} &= 0 \end{aligned} \tag{2.57}$$

which are a consequence of the properties of the individual columns of Q and $\bar{Q}$.

If the system extends to $n = +\infty$ and if the voltages and currents are to remain finite at $n = \infty$ the elements of $\bar{a}$ must be zero and the expressions (2.45) for $v(n)$ and $i(n)$ become

$$\begin{gathered} v(n) = P\Lambda^n a = P\Lambda^n P^{-1} v(o) \\ i(n) = Q\Lambda^n a = Q\Lambda^n Q^{-1} i(o) \\ v(n) = PQ^{-1} i(n), \qquad i(n) = QP^{-1} v(n) \end{gathered} \tag{2.58}$$

where we have assumed that P^{-1} and Q^{-1} exist. We accordingly introduce the characteristic impedance and admittance matrices Z_o and Y_o associated with propagation in the positive direction, i.e., in the direction of n increasing.

$$\begin{gathered} v(n) = Z_o i(n), \qquad i(n) = Y_o v(n), \qquad Z_o = Y_o^{-1} \\ Z_o = PQ^{-1} = Z_{11} - Z_{12}Q\Lambda Q^{-1} = -Z_{22} + Z_{21}Q\Lambda^{-1}Q^{-1} \\ Y_o = QP^{-1} = Y_{11} + Y_{12}P\Lambda P^{-1} = -Y_{22} - Y_{21}P\Lambda^{-1}P^{-1} \end{gathered} \tag{2.59}$$

Incidentally, since Z_o must be a symmetrical matrix the above equations show that $Z_{12}Q\Lambda Q^{-1}$ and $Z_{21}Q\Lambda^{-1}Q^{-1}$ are symmetrical. Z_o and Y_o satisfy the relations

$$Z_oCZ_o + Z_oD - AZ_o - B = 0 \qquad Y_oBY_o + Y_oA - DY_o - C = 0$$

$$(Z_{22} + Z_o)Z_{12}^{-1}(Z_{11} - Z_o) = Z_{21}, \qquad (Y_{22} + Y_o)Y_{12}^{-1}(Y_{11} - Y_o) = Y_{21} \quad (2.60)$$

$$Z_oQ\Lambda Q^{-1} = P\Lambda P^{-1}Z_o \qquad Y_oP\Lambda P^{-1} = Q\Lambda Q^{-1}Y_o$$

The characteristic and admittance matrices $\bar{Z}_o$ and $\bar{Y}_o$ associated with propagation in the negative direction are introduced in a similar way. Suppose the system extends to $n = -\infty$. Then $a = o$ and

$$\begin{aligned} v(n) &= \bar{P}\Lambda^{-n}\bar{a} = \bar{P}\Lambda^{-n}\bar{P}^{-1}v(o) \\ i(n) &= \bar{Q}\Lambda^{-n}\bar{a} = -\bar{Q}\Lambda^{-n}\bar{Q}^{-1}i(o) \end{aligned} \quad (2.61)$$

$$v(n) = -\bar{P}\bar{Q}^{-1}i(n), \qquad i(n) = -\bar{Q}\bar{P}^{-1}v(n)$$

Hence we write

$$v(n) = -\bar{Z}_oi(n), \qquad i(n) = -\bar{Y}_ov(n)$$

$$\bar{Z}_o = \bar{P}\bar{Q}^{-1} = -Z_{11} + Z_{12}\bar{Q}\Lambda^{-1}\bar{Q}^{-1} = Z_{22} - Z_{21}\bar{Q}\Lambda\bar{Q}^{-1} \quad (2.62)$$

$$\bar{Y}_o = \bar{Q}\bar{P}^{-1} = -Y_{11} - Y_{12}\bar{P}\Lambda^{-1}\bar{P}^{-1} = Y_{22} + Y_{21}\bar{P}\Lambda\bar{P}^{-1}$$

$\bar{Z}_o$ and $\bar{Y}_o$ satisfy the relations

$$\begin{aligned} &\bar{Z}_oC\bar{Z}_o - \bar{Z}_oD + A\bar{Z}_o - B = 0 \qquad \bar{Y}_oB\bar{Y}_c - \bar{Y}_oA + D\bar{Y}_o - C = 0 \\ &(Z_{11} + \bar{Z}_o)Z_{21}^{-1}(Z_{22} - \bar{Z}_o) = Z_{12} \qquad (Y_{11} + \bar{Y}_o)Y_{21}^{-1}(Y_{22} - \bar{Y}_o) = Y_{12} \end{aligned} \quad (2.63)$$

The fact that $Q'(Z_o + \bar{Z}_o)\bar{Q} = P'(Y_o \dotplus \bar{Y}_o)\bar{P}$ is a diagonal matrix may be used as a check on computations.

When the expressions (2.45) for $v(n)$ and $i(n)$ are placed in (2.3), $j^\circ(n)$ and $u^\circ(n)$ being zero, we obtain the relations

$$\begin{aligned} P\Lambda^{-1} &= AP + BQ \qquad & P\bar{\Lambda} &= A\bar{P} - B\bar{Q} \\ Q\Lambda^{-1} &= CP + DQ \qquad & \bar{Q}\Lambda &= -C\bar{P} + D\bar{Q} \end{aligned} \quad (2.64)$$

When the typical section contains generators the difference equation to be solved is of the form (2.42)

$$Gx(n + 2) + Hx(n + 1) + G'x(n) = g(n) \quad (2.42)$$

This is true for symmetrical as well as unsymmetrical sections, G being a symmetrical matrix in the former case so that $G' = G$. The expressions for $v(n)$ and $i(n)$ are those of (2.45) with the particular solutions added:

$$\begin{aligned} v(n) &= P\Lambda^na + \bar{P}\Lambda^{-n}\bar{a} + u(n) \\ i(n) &= Q\Lambda^na - \bar{Q}\Lambda^{-n}\bar{a} + j(n) \end{aligned} \quad (2.65)$$

where P, $\bar{P}$, Q, $\bar{Q}$ are determined as before and $u(n)$ and $j(n)$ depend upon the generators.

Here we shall consider only the physically important case in which the voltages of the generators in the nth section are proportional to $e^{-n\theta}$ where θ is a constant. In this case $g(n)$ may be expressed as

$$g(n) = ge^{-n\theta} \tag{2.66}$$

where g is a column matrix whose elements are independent of n. A particular solution is obtained by assuming

$$x(n) = ye^{-n\theta}$$

Setting this in (2.42) gives

$$(Ge^{-2\theta} + He^{-\theta} + G')y = g$$

Hence a particular solution is

$$x(n) = (Ge^{-2\theta} + He^{-\theta} + G')^{-1}ge^{-n\theta} \tag{2.67}$$

This method fails when θ is equal to one of the roots $\lambda_1, \cdots \lambda_m, \lambda_1^{-1}, \cdots \lambda_m^{-1}$. In this case, a particular integral may be obtained by a method similar to one described in §5.11 of F.D.C.

APPENDIX I

Classical Solution of Uniform Transmission Line Equations

For the sake of convenience we again assume that there are three circuits in the transmission line. The equations to be solved are:

$$\frac{dv}{dx} = -Zi, \qquad \frac{di}{dx} = -Yi \tag{1.48}$$

We adopt here the notation associated with equations (1.19) and (1.20), $f(\gamma^2)$ being the characteristic matrix of ZY, $F(\gamma^2)$ the adjoint of $f(\gamma^2)$, and $\gamma_1^2, \gamma_2^2, \gamma_3^2 (m = 3)$ being the roots, supposed distinct, of $|f(\gamma^2)| = 0$. The propagation constants $\gamma_1, \gamma_2, \gamma_3$ are those square roots of $\gamma_1^2, \gamma_2^2, \gamma_3^2$ which in physical systems have a positive real part.

The solution may be constructed[11] as follows: Let the column p_r be proportional (with any convenient constant of proportionality) to any non-zero column of $F(\gamma_r^2)$. The non-zero columns of $F(\gamma_r^2)$ are proportional to each other according to a theorem in matrix algebra.[12] Construct the square matrix P from the columns p_1, p_2, p_3:

$$P = [p_1, p_2, p_3] \tag{A1.1}$$

[11] The method is that described in F.D.C. §5.7(a) and §5.10
[12] F.D.C. §3.5 Theorem D.

and obtain the square matrix Q from P:

$$Q = Z^{-1}PG = YPG^{-1} \tag{A1.2}$$

where G is the diagonal matrix

$$G = \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{bmatrix} \tag{A1.3}$$

The voltages and currents at any point x are

$$\begin{aligned} v(x) &= PM(x)a + PM(-x)\bar{a} \\ i(x) &= QM(x)a - QM(-x)\bar{a} \end{aligned} \tag{A1.4}$$

where a and $\bar{a}$ are arbitrary column matrices associated with propagation in the positive and negative directions of x and $M(x)$ is the diagonal matrix

$$M(x) = \begin{bmatrix} e^{-\gamma_1 x} & 0 & 0 \\ 0 & e^{-\gamma_2 x} & 0 \\ 0 & 0 & e^{-\gamma_3 x} \end{bmatrix} \tag{A1.5}$$

The values of a and $\bar{a}$ are to be determined from the boundary conditions. When the line extends to $x = \infty$

$$\begin{aligned} v(x) &= PM(x)P^{-1}v(o) = Z_0 i(x) \\ i(x) &= QM(x)Q^{-1}i(o) \end{aligned} \tag{A1.6}$$

where the characteristic impedance matrix Z_0 is given by

$$\begin{aligned} Z_0 = PQ^{-1} &= PG^{-1}P^{-1}Z = PGP^{-1}Y^{-1} \\ &= ZQG^{-1}Q^{-1} = Y^{-1}QGQ^{-1} \end{aligned} \tag{A1.7}$$

Since $v = p_r e^{\gamma_r x}$ and $i = q_r e^{\gamma_r x}$, where q_r is the rth column of Q, are solutions the differential equations give

$$(I\gamma_r^2 - ZY)p_r = 0, \qquad (I\gamma_r^2 - YZ)q_r = 0 \tag{A1.8}$$

and from these it follows that

$$P^{-1}ZYP = Q^{-1}YZQ = G^2 \tag{A1.9}$$

The relations (A1.8) may be used to prove the following:

1. The elements in the rth column of Q are proportional to those in the non-zero rows of $F(\gamma_r^2)$.
2. The matrix $P'Q$ is a diagonal matrix and from this it follows that if ψ is any diagonal matrix

$$(P\psi P^{-1})' = Q\psi Q^{-1} \tag{A1.10}$$

3. The characteristic impedance matrix Z_0 satisfies the relation

$$Z = Z_0 Y Z_0 \tag{A1.11}$$

4. The inverse matrices P^{-1} and Q^{-1} always exist if $\gamma_1, \gamma_2, \gamma_3$ are distinct.

APPENDIX II

Classical Solution of Symmetrical Section Line Equations—I

The method of this section is very similar to that of Appendix I. The equations to be solved are (2.10) or one of the sets

$$\begin{aligned} v(n) &= Z_{11}i(n) - Z_{12}i(n+1) \\ v(n+1) &= Z_{12}i(n) - Z_{11}i(n+1) \end{aligned} \tag{A2.1}$$

$$\begin{aligned} i(n) &= Y_{11}v(n) + Y_{12}v(n+1) \\ -i(n+1) &= Y_{12}v(n) + Y_{11}v(n+1) \end{aligned} \tag{A2.2}$$

which are obtained from (2.5) and (2.6). We shall use the notation associated with equation (2.19), $f(\zeta)$ being the characteristic matrix of A, $F(\zeta)$ the adjoint of $f(\zeta)$, and $\zeta_1, \zeta_2, \cdots \zeta_m$ the roots, assumed unequal, of the characteristic equation $|f(\zeta)| = 0$. The diagonal matrices Λ and Σ are defined by

$$\begin{aligned} \Lambda &= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ 0 & & & \lambda_m \end{bmatrix}, \\ \Sigma &= \begin{bmatrix} \sqrt{\zeta_1^2 - 1} & 0 & \cdots\cdots & 0 \\ 0 & \sqrt{\zeta_2^2 - 1} & & \\ 0 & & & \sqrt{\zeta_m^2 - 1} \end{bmatrix} \end{aligned} \tag{A2.3}$$

where

$$2\zeta_r = \lambda_r + \lambda_r^{-1}, \qquad \lambda_r = \zeta_r - \sqrt{\zeta_r^2 - 1} = \frac{1}{\zeta_r + \sqrt{\zeta_r^2 - 1}} \tag{A2.4}$$

In general, electrical energy will be dissipated in the typical section and from the physical significance of λ_r, as seen from equations (A2.8) below, it follows that the sign of the radical in (A2.4) may be chosen so that $|\lambda_r| < 1$. Since $\sqrt{\zeta_r^2 - 1} = \zeta_r - \lambda_r = \frac{1}{2}(\lambda_r^{-1} - \lambda_r)$ it follows that

$$\Sigma = \tfrac{1}{2}(\Lambda^{-1} - \Lambda) \tag{A2.5}$$

Let the column matrix s_r be proportional to any non-zero column in $F(\zeta_r)$ where $F(\zeta)$ is the adjoint of $f(\zeta)$. (It follows from the theory of matrices that the non-zero columns of $F(\zeta_r)$ differ from each other only by a multiplying factor.) The matrix S is then formed by taking s_1 to be the first column, s_2 the second and so on.

$$S = [s_1, s_2, \cdots s_m] \tag{A2.6}$$

Similarly let the row matrix t_r' be proportional to any nonvanishing row of $F(\zeta_r)$ and form the matrix T where

$$T = [t_1, t_2, \cdots t_m] \tag{A2.7}$$

in which t_r is the column matrix obtained by transposing t_r'.[13]

Solving our difference equations for the passive case by the customary method gives the expressions

$$\begin{aligned} v(n) &= P\Lambda^n a + P\Lambda^{-n}\bar{a} \\ i(n) &= Q\Lambda^n a - Q\Lambda^{-n}\bar{a} \end{aligned} \tag{A2.8}$$

for the voltages and the currents. P and Q are square matrices and a and $\bar{a}$ are column matrices whose elements are determined by the boundary conditions. a and $\bar{a}$ are of the same nature as constants of integration. The minus sign over $\bar{a}$ indicates that it is associated with propagation in the negative direction, i.e., in the direction of n decreasing.

P and Q may be chosen in a number of ways, each choice requiring different values of a and $\bar{a}$ to represent the same system. In all cases, however, the rth column of P may be expressed as $\alpha_r s_r$ where α_r is a scalar multiplier which may depend upon r. Similarly the rth column of Q may be expressed as $\beta_r t_r$. When either P or Q has been chosen the other one is fixed since equations (A2.2) and (A2.1) require

$$\begin{aligned} Q &= Y_{11}P + Y_{12}P\Lambda = -Y_{11}P - Y_{12}P\Lambda^{-1} \\ P &= Z_{11}Q - Z_{12}Q\Lambda = -Z_{11}Q + Z_{12}Q\Lambda^{-1} \end{aligned} \tag{A2.9}$$

Some useful choices are,

$$\begin{aligned} &1.\quad P = S, && Q = -Y_{12}S\Sigma = B^{-1}S\Sigma \\ &2.\quad P = S\Sigma && Q = Z_{12}^{-1}S = CS \\ &3.\quad Q = T, && P = Z_{12}T\Sigma = C^{-1}T\Sigma \\ &4.\quad Q = T\Sigma && P = -Y_{12}^{-1}T = BT \end{aligned} \tag{A2.10}$$

The particular choice to be made depends upon the system of difference equations which is being used. In choices 1 and 2, T is not required and in 3 and 4, S is not required. However, if both S and T are known some of

[13] Methods of determining s_r and t_r' are available. A description will be found in F.D.C. §4.12.

the matrix multiplication may be avoided. Taking choice 1 as an example, we may determine the rth row of Q from the expression $\beta_r t_r$. To determine β_r only one element in the rth column of $-Y_{12}S\Sigma$ need be known, for β_r is the quotient obtained by dividing this element by the corresponding element in t_r. The product $P'Q$ must be a diagonal matrix, and the same is true of $S'T$. This may serve to check computations.

That the expressions for $v(n)$ and $i(n)$ given by (A2.8) and (A2.10) satisfy the transmission equations (A2.1), (A2.6) and (2.10) may be verified by direct substitution and use of

$$S(\Lambda + \Lambda^{-1}) = 2AS \qquad T(\Lambda + \Lambda^{-1}) = 2A'T \tag{A2.11}$$

These relations follow from the properties of the individual columns of S and T.

When the system extends to $n = \infty$ $\bar{a}$ must be zero in order that the voltages and currents may remain finite. This is true because λ_r is chosen so that $|\lambda_r| < 1$. From equations (A2.8) it follows that

$$\begin{aligned} v(n) &= P\Lambda^n a = P\Lambda^n P^{-1} v(o) \\ i(n) &= Q\Lambda^n a = Q\Lambda^n Q^{-1} i(o) \\ v(n) &= PQ^{-1} i(n) \qquad i(n) = QP^{-1} v(n) \end{aligned} \tag{A2.12}$$

the reciprocal matrices Q^{-1} and P^{-1} always exist when the sections are symmetrical and the roots $\zeta_1, \zeta_2, \cdots \zeta_m$ distinct. The last equations in (A2.12) suggest the introduction of the characteristic impedance and admittance matrices Z_o and Y_o:

$$\begin{aligned} v(n) &= Z_o i(n), \qquad i(n) = Y_o v(n), \qquad Z_o = Y_o^{-1}. \\ Z_o &= PQ^{-1} = Z_{11} - Z_{12}Q\Lambda Q^{-1} = -Z_{11} + Z_{12}Q\Lambda^{-1}Q^{-1} \\ &= S\Sigma^{-1}S^{-1}B = S\Sigma S^{-1}Z_{12} \\ &= Z_{12}T\Sigma T^{-1} = BT\Sigma^{-1}T^{-1} \\ Y_o &= QP^{-1} = Y_{11} + Y_{12}P\Lambda P^{-1} = -Y_{11} - Y_{12}P\Lambda^{-1}P^{-1} \\ &= -Y_{12}S\Sigma S^{-1} = CS\Sigma^{-1}S^{-1} \\ &= T\Sigma^{-1}T^{-1}C = -T\Sigma T^{-1}Y_{12} \end{aligned} \tag{A2.13}$$

Not all of the expressions for Z_o and Y_o obtainable from (A2.10) have been included in (A2.13). Z_o and Y_o are symmetrical matrices. Although P and Q are arbitrary to some extent the same is not true of Z_o and Y_o. Computed values of Z_o and Y_o may be checked by use of the relations

$$\begin{aligned} A^2 - I &= (Z_o Z_{12}^{-1})^2 = (Y_{12}^{-1} Y_o)^2 \\ Z_o C Z_o &= B, \qquad Y_o B Y_o = C \\ Y_o Z_{12} &= -Y_{12} Z_o \end{aligned} \tag{A2.14}$$

Sometimes it is desirable to terminate a line consisting of a finite number of sections by a network which simulates an infinite line. As is known, the elements in one such network may be obtained from Y_o. Every terminal is joined to every other terminal, including the return terminal (denoted by the subscript o), by the branches of this network. The admittance of the branch connecting terminal i to terminal j, $i \neq 0$, $j \neq 0$, is $-y_{ij}$ where y_{ij} is the element in the ith row and jth column of Y_o. The admittance of the branch connecting terminal i to terminal o is $y_{i1} + y_{i2} + \cdots + y_{ii} + \cdots + y_{im}$, i.e., it is the sum of all the elements whose first subscript is i.

APPENDIX III

Classical Solution of Symmetrical Section Line Equations—II

When the electrical properties of a typical symmetrical section are to be determined by measurement, equations (A2.1) and (A2.2) show that Z_{11} and Y_{11} may be obtained by measurements at one end. In order to obtain Y_{12} and Z_{12} measurements have to be made at both ends. Expressions for $v(n)$ and $i(n)$ will now be given which depend only upon Z_{11} and Y_{11} and hence are useful in case the measurements are restricted to one end.

The method is based upon the equations

$$\begin{aligned} v(n+2) + v(n) &= Z_{11}[i(n) - i(n+2)] \\ i(n+2) + i(n) &= Y_{11}[v(n) - v(n+2)] \end{aligned} \tag{A3.1}$$

which may be derived from (A2.1) and (A2.2). Combining these equations leads to

$$[I - Z_{11}Y_{11}][v(n+2) + v(n-2)] + 2\,[I + Z_{11}Y_{11}]v(n) = 0$$

$$[I - Y_{11}Z_{11}][i(n+2) + i(n-2)] + 2\,[I + Y_{11}Z_{11}]i(n) = 0$$

The first step in the solution is to solve the equation

$$|\,\mu I - Z_{11}Y_{11}\,| = 0 \tag{A3.2}$$

for its roots $\mu_1, \mu_2, \cdots \mu_m$ which we shall suppose are distinct. The diagonal matrices M and $M^{\frac{1}{2}}$ are defined by

$$M = \begin{bmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & & \\ \vdots & & & \\ 0 & \cdots & \cdots & \mu_m \end{bmatrix}, \qquad M^{\frac{1}{2}} = \begin{bmatrix} \mu_1^{\frac{1}{2}} & 0 & \cdots & 0 \\ 0 & \mu_2^{\frac{1}{2}} & & \\ \vdots & & & \\ 0 & \cdots & \cdots & \mu_m^{\frac{1}{2}} \end{bmatrix} \tag{A3.3}$$

and Λ is defined as in (A2.3) where λ_r is given by

$$\lambda_r = \sqrt{\frac{\mu_r^{\frac{1}{2}} - 1}{\mu_r^{\frac{1}{2}} + 1}}, \qquad \mu_r = \left[\frac{1 + \lambda_r^2}{1 - \lambda_r^2}\right]^2 \tag{A3.4}$$

The sign of $\mu_r^{\frac{1}{2}}$ is chosen so that $|\lambda_r| < 1$, and this is the value to be used in $M^{\frac{1}{2}}$. However, there is an ambiguity in the sign of λ_r which is inherent in this method. A relation between Λ and $M^{\frac{1}{2}}$ is

$$M^{\frac{1}{2}} = (I + \Lambda^2)(I - \Lambda^2)^{-1} \tag{A3.5}$$

Let u_r be proportional to any non-zero column and w_r' be proportional to any non-zero row of the matrix adjoint to $[\mu_r I - Z_{11} Y_{11}]$ and form the matrices

$$U = [u_1, u_2, \cdots u_m]$$

$$W = [w_1, w_2, \cdots w_m]$$

(cf. equations (A2.6) and (A2.7) for S and T) where w_r is the column obtained from w_r'.

The voltages and currents are given, as before, by

$$\begin{aligned} v(n) &= P\Lambda^n a + P\Lambda^{-n}\bar{a} \\ i(n) &= Q\Lambda^n a - Q\Lambda^{-n}\bar{a} \end{aligned} \tag{A2.8}$$

and there is again a number of ways in which P and Q may be chosen. In all cases the rth column of P may be expressed as $\alpha_r u_r$ and the rth column of Q as $\beta_r w_r$. The equations fixing Q when P is chosen and vice versa are, from equations (A3.1)

$$\begin{aligned} Q &= Y_{11} P M^{-\frac{1}{2}} \\ P &= Z_{11} Q M^{-\frac{1}{2}} \end{aligned} \tag{A3.6}$$

where $M^{-\frac{1}{2}}$ is the inverse of $M^{\frac{1}{2}}$. Equations (A3.6) may also be obtained from (A2.10).

Suitable choices for P and Q are

$$\begin{aligned} &1.\ P = U, && Q = Y_{11} U M^{-\frac{1}{2}} = Z_{11}^{-1} U M^{\frac{1}{2}} \\ &2.\ P = U M^{\frac{1}{2}}, && Q = Y_{11} U = Z_{11}^{-1} U M \\ &3.\ Q = W, && P = Z_{11} W M^{-\frac{1}{2}} = Y_{11}^{-1} W M^{\frac{1}{2}} \\ &4.\ Q = W M^{\frac{1}{2}}, && P = Z_{11} W = Y_{11}^{-1} W M \end{aligned} \tag{A3.7}$$

$P'Q$ and $U'W$ must be diagonal matrices. That the expressions for $v(n)$ and $i(n)$ just derived satisfy the difference equations (A3.1) may be verified by making use of

$$UM = Z_{11} Y_{11} U, \qquad WM = Y_{11} Z_{11} W \tag{A3.8}$$

Equations (A3.8) follow from the properties of the individual columns of U and W. The characteristic impedance and admittance matrices are given by

$$
\begin{aligned}
Z_o &= PQ^{-1} = Z_{11}QM^{-\frac{1}{2}}Q^{-1} = PM^{\frac{1}{2}}P^{-1}Y_{11}^{-1} \\
&= UM^{-\frac{1}{2}}U^{-1}Z_{11} = UM^{\frac{1}{2}}U^{-1}Y_{11}^{-1} \\
&= Z_{11}WM^{-\frac{1}{2}}W^{-1} = Y_{11}^{-1}WM^{\frac{1}{2}}W^{-1} \\
Y_o &= QP^{-1} = Y_{11}PM^{-\frac{1}{2}}P^{-1} = QM^{\frac{1}{2}}Q^{-1}Z_{11}^{-1} \\
&= Y_{11}UM^{-\frac{1}{2}}U^{-1} = Z_{11}^{-1}UM^{\frac{1}{2}}U^{-1} \\
&= WM^{-\frac{1}{2}}W^{-1}Y_{11} = WM^{\frac{1}{2}}W^{-1}Z_{11}^{-1}
\end{aligned} \tag{A3.9}
$$

The matrices Z_o and Y_o may be checked by means of the relations

$$Z_oY_{11} = Z_{11}Y_o\,, \qquad Z_oY_{11}Z_o = Z_{11}\,, \qquad Y_oZ_{11}Y_o = Y_{11} \tag{A3.10}$$

Another set of solutions may be based upon the equations

$$
\begin{aligned}
2v(n) &= -Z_{12}[i(n+1) - i(n-1)] \\
2i(n) &= Y_{12}[v(n+1) - v(n-1)]
\end{aligned} \tag{A3.11}
$$

which are derivable from (A2.1) and (A2.2). Combining these equations gives, upon using $Y_{12}^{-1}Z_{12}^{-1} = -BC$,

$$
\begin{aligned}
v(n+2) - 2v(n) + v(n-2) &= 4BCv(n) \\
i(n+2) - 2i(n) + i(n-2) &= 4CBi(n)
\end{aligned} \tag{A3.12}
$$

However, we shall not consider these equations here beyond pointing out that they lead to

$$
\begin{aligned}
P &= Z_{12}Q\Sigma, \qquad & Q &= -Y_{12}P\Sigma \\
P\Sigma^2 &= BCP, \qquad & Q\Sigma^2 &= CBQ
\end{aligned}
$$

which may also be derived from (A2.10).

APPENDIX IV

Relations Between the Square Matrices of a Multi-Terminal Section

When the reciprocal theorems of network theory are applied to equations (2.1) and (2.2) it is found that Z_{11}, Z_{22}, Y_{11}, Y_{22} are symmetrical and

$$Z_{21} = Z_{12}'\,, \qquad Y_{21} = Y_{12}' \tag{A4.1}$$

i.e., Z_{21} and Y_{21} are the transposed matrices of Z_{12} and Y_{12}, respectively.

Solving equations (2.1) for the currents and comparing the result with (2.2) shows that

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}^{-1} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}, \qquad \begin{bmatrix} i^\circ(n) \\ -j^\circ(n) \end{bmatrix} = -\begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}\begin{bmatrix} v^\circ(n) \\ u^\circ(n) \end{bmatrix}$$

These are partitioned matrices. The square matrices have $2m$ rows and columns and the column matrices have $2m$ elements. The first of these relations may be written as

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{A4.2}$$

where I denotes the unit diagonal matrix of order m. Multiplying the two matrices on the left together and equating the elements of the product to the elements on the right gives

$$\begin{aligned} Z_{11}Y_{11} + Z_{12}Y_{21} &= I \\ Z_{11}Y_{12} + Z_{12}Y_{22} &= 0 \\ Z_{21}Y_{11} + Z_{22}Y_{21} &= 0 \\ Z_{21}Y_{12} + Z_{22}Y_{22} &= I \end{aligned} \tag{A4.3}$$

Transposing the matrices in these equations leads to other relations. Thus, from the first we obtain $Y_{11}Z_{11} + Y_{12}Z_{21} = I$. These equations also yield expressions for the Y's in terms of the Z's and vice versa.

A somewhat similar treatment involving equations (2.1) and (2.3) leads to expressions for the Z's in terms of A, B, C and D. The Y's may be likewise expressed. These relations are given in the following table.

$$\begin{aligned} Y_{11} &= DB^{-1} & Y_{11}^{-1} &= Z_{11} - Z_{12}Z_{22}^{-1}Z_{21} \\ Y_{12} &= C - DB^{-1}A = -B'^{-1} & Y_{12}^{-1} &= Z_{21} - Z_{22}Z_{12}^{-1}Z_{11} \\ Y_{21} &= -B^{-1} & Y_{21}^{-1} &= Z_{12} - Z_{11}Z_{21}^{-1}Z_{22} \\ Y_{22} &= B^{-1}A & Y_{22}^{-1} &= Z_{22} - Z_{21}Z_{11}^{-1}Z_{12} \\ Z_{11} &= AC^{-1} & Z_{11}^{-1} &= Y_{11} - Y_{12}Y_{22}^{-1}Y_{21} \\ Z_{12} &= AC^{-1}D - B = C'^{-1} & Z_{12}^{-1} &= Y_{21} - Y_{22}Y_{12}^{-1}Y_{11} \\ Z_{21} &= C^{-1} & Z_{21}^{-1} &= Y_{12} - Y_{11}Y_{21}^{-1}Y_{22} \\ Z_{22} &= C^{-1}D & Z_{22}^{-1} &= Y_{22} - Y_{21}Y_{11}^{-1}Y_{12} \end{aligned}$$

$$\begin{aligned} A &= Z_{11}Z_{21}^{-1} = -Y_{21}^{-1}Y_{22} \\ B &= Z_{11}Z_{21}^{-1}Z_{22} - Z_{12} = -Y_{21}^{-1} \\ C &= Z_{21}^{-1} = Y_{12} - Y_{11}Y_{21}^{-1}Y_{22} \\ D &= Z_{21}^{-1}Z_{22} = -Y_{11}Y_{21}^{-1} \end{aligned} \tag{A4.4}$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} D' & -B' \\ -C' & A' \end{bmatrix} \qquad \begin{matrix} AD' - BC' = I \\ CD' = DC' \end{matrix} \qquad \begin{matrix} AB' = BA' \\ DA' - CB' = I \end{matrix}$$

$$\begin{bmatrix} v^\circ(n) \\ -i^\circ(n) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} u^\circ(n) \\ -j^\circ(n) \end{bmatrix}$$

The following equations in which λ is an arbitrary scalar multiplier may be verified by equating coefficients of powers of λ and using the relations just given.

$$(Z_{21} - \lambda Z_{11})(Y_{11} + \lambda Y_{12}) = (\lambda Z_{12} - Z_{22})(\lambda Y_{22} + Y_{21})$$

$$\begin{aligned}(\lambda^2 Z_{12} - \lambda Z_{11} - \lambda Z_{22} + Z_{21})(Y_{11} + \lambda Y_{12})\\ = (\lambda Z_{12} - Z_{22})(\lambda^2 Y_{12} + \lambda Y_{11} + \lambda Y_{22} + Y_{21})\end{aligned} \tag{A4.5}$$

$$\begin{bmatrix} -Y_{21} & 0 \\ 0 & Z_{21} \end{bmatrix} \begin{bmatrix} \lambda A - I & \lambda B \\ \lambda C & \lambda D - I \end{bmatrix} = \begin{bmatrix} \lambda Y_{22} + Y_{21} & \lambda I \\ \lambda I & \lambda Z_{22} - Z_{21} \end{bmatrix}$$

Sometimes it is of interest to obtain the elements of Y_{12}, say, when Z_{11}, Z_{22}, Y_{11}, Y_{22} are known. Relations helpful in studying this problem are

$$\begin{aligned} Y_{11}Z_{11}Y_{12} &= Y_{12}Z_{22}Y_{22}, & Y_{11}Z_{11}Y_{11} - Y_{11} &= Y_{12}Z_{22}Y_{21} \\ Y_{12}Y_{22}^{-1}Y_{21} &= Y_{11} - Z_{11}^{-1} & Z_{12} &= -Z_{11}Y_{12}Y_{22}^{-1} \\ Y_{21}Y_{11}^{-1}Y_{12} &= Y_{22} - Z_{22}^{-1} & Z_{21} &= -Y_{12}^{-1}(Y_{11}Z_{11} - I) \end{aligned}$$

When the typical section is symmetrical some simplification takes place and we have

$$\begin{aligned} Y_{11} &= Y_{22} & Z_{11} &= Z_{22} & A &= D' & AB &= BA' \\ Y_{12} &= Y_{21} & Z_{12} &= Z_{21} & B &= B' & A'C &= CA \\ & & & & C &= C' & A^2 - BC &= I \end{aligned} \tag{A4.6}$$

$$Z_{11}Y_{11} + Z_{12}Y_{12} = I \qquad A'B^{-1}A - C = B^{-1}$$

$$Z_{11}Y_{12} + Z_{12}Y_{11} = 0$$

APPENDIX V

Properties of the Matrix $G\lambda^2 + H\lambda + G'$

The matrix

$$f(\lambda) = G\lambda^2 + H\lambda + G' \tag{A5.1}$$

which entered the discussion of the case of unsymmetrical sections has a number of interesting properties which are given below. G and H are square matrices with m rows each, and H is required to be symmetrical. As before, we shall denote by $\lambda_1, \cdots \lambda_m, \lambda_1^{-1}, \cdots \lambda_m^{-1}$ the $2m$ roots of the determinantal equation

$$|f(\lambda)| = 0$$

and we shall suppose these roots to be distinct. Let the column k_r and the row l_r be such that

$$k_r l_r = F(\lambda_r) \tag{A5.2}$$

where $F(\lambda)$ is the matrix adjoint to $f(\lambda)$, and let the square matrices K and L be defined by

$$K = [k_1, k_2, \cdots k_m], \qquad L = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{bmatrix} \tag{A5.3}$$

Comparison of (A5.3) and (2.50) suggests that when G and H are expressed in terms of the Y's we have the relations

$$K = P, \qquad L = \bar{P}' \tag{A5.4}$$

The method of choosing the column p_r and the row $\bar{p}_r'$ shows that they are related by

$$p_r \bar{p}_r' = \gamma_r F(\lambda_r)$$

instead of (A5.2) where γ_r may turn out to be any non-zero constant, and consequently equations (A5.4) are not satisfied in general. Nevertheless K and L may be regarded as particular choices for P and $\bar{P}'$. In the same way K and L may be regarded as particular choices for Q and $\bar{Q}'$ when G and H are expressed in terms of the Z's. There is still some arbitrariness connected with K and L since the product $k_r l_r$ is unchanged when the k_r is multiplied by some number and l_r is divided by the same number.

The relations which correspond to (2.52) and (2.57) are

$$\begin{aligned} GK\Lambda^2 + HK\Lambda + G'K &= 0 \\ GL'\Lambda^{-2} + HL'\Lambda^{-1} + G'L' &= 0 \end{aligned} \tag{A5.5}$$

where Λ is the diagonal matrix whose elements are $\lambda_1, \lambda_2, \cdots \lambda_m$. These relations are consequences of the properties of k_r and l_r. Two more relations may be obtained by transposition. From the first of (A5.5) and the transposed of the second it follows that

$$\begin{aligned} GK\Lambda K^{-1} + H + G'K\Lambda^{-1}K^{-1} &= 0 \\ L^{-1}\Lambda LG + H + L^{-1}\Lambda^{-1}LG' &= 0 \end{aligned} \tag{A5.6}$$

where it is assumed that the reciprocal matrices K^{-1} and L^{-1} exist. Combinations similar to $K\Lambda K^{-1}$, $K\Lambda^{-1}K^{-1}$, etc. enter the expressions (2.59) for Z_o and Y_o.

By differentiating the equation

$$f(\lambda)F(\lambda) = \Delta(\lambda)I,$$

where $\Delta(\lambda)$ is the determinant

$$\Delta(\lambda) = |f(\lambda)| = |G\lambda^2 + H\lambda + G'|,$$

it may be proved that

$$\begin{aligned} GK\Lambda K^{-1} + H + L^{-1}\Lambda LG &= L^{-1}EK^{-1} \\ G'K\Lambda^{-1}K^{-1} + H + L^{-1}\Lambda^{-1}LG' &= -L^{-1}EK^{-1} \end{aligned} \tag{A5.7}$$

in which E is the diagonal matrix

$$E = \begin{bmatrix} \overset{(1)}{\Delta}(\lambda_1) & 0 & \cdots & 0 \\ 0 & \overset{(1)}{\Delta}(\lambda_2) & & \\ \vdots & & & \\ 0 & & & \overset{(1)}{\Delta}(\lambda_m) \end{bmatrix}$$

and

$$\overset{(1)}{\Delta}(\lambda_r) = \left[\frac{d}{d\lambda}\Delta(\lambda)\right]_{\lambda=\lambda_r}$$

Since the roots λ_r are assumed to be distinct, $\overset{(1)}{\Delta}(\lambda_r) \neq 0$.

We also have the equations

$$\begin{aligned} KE^{-1}L &= L'E^{-1}K' \\ GK\Lambda E^{-1}L - GL'\Lambda^{-1}E^{-1}K' &= I \end{aligned} \tag{A5.8}$$

The first equation of (A5.8) shows that $KE^{-1}L$ is a symmetrical matrix. From this and the second equation it follows that

$$GK\Lambda K^{-1} - GL'\Lambda^{-1}L'^{-1} = L^{-1}EK^{-1} \tag{A5.9}$$

From the first of equations (A5.7) and the second of (A5.6)

$$GK\Lambda K^{-1} - L^{-1}\Lambda^{-1}LG' = L^{-1}EK^{-1} \tag{A5.10}$$

and the comparison with (A5.9) shows that the matrix $GL'\Lambda^{-1}L'^{-1}$ is symmetrical. The other matrices of this type are also symmetrical as may now be seen from equations (A5.6) and (A5.7). Results of this sort may be obtained from physical principles by noting that Z_o and Y_o must be symmetrical matrices.

As an example of the application of these formulas we assume that G and H are expressed in terms of the Y's. Then we may take K and L' to be particular choices of P and $\bar{P}$ and equation (A5.9) becomes

$$Y_{12}P\Lambda P^{-1} - Y_{12}\bar{P}\Lambda^{-1}\bar{P}^{-1} = (\bar{P}')^{-1}EP^{-1}.$$

From equations (2.59) and (2.62)

$$Y_o + \bar{Y}_o = Y_{12}P\Lambda P^{-1} - Y_{12}\bar{P}\Lambda^{-1}\bar{P}^{-1}$$

and hence

$$\bar{P}'(Y_o + \bar{Y}_o)P = E.$$

For the more general choice of P and $\bar{P}$ allowed in §2.8 the diagonal matrix E is replaced by a general diagonal matrix. Similarly it follows that

$$\bar{Q}'(Z_o + \bar{Z}_o)Q$$

is a diagonal matrix.

Engineering Problems in Dimensions and Tolerances

By W. W. WERRING

Dimensional Units

The basic unit in most considerations of dimensions in the United States is the inch. The value of the inch is so important that many companies including the Bell System maintain in their measurement laboratories a standard yard bar calibrated against the standard at the National Bureau of Standards. In spite of this it is an interesting and curious fact that though all have been much concerned over the legal value of the dollar there has been little interest even among engineers in the exact legal value of the inch. Actually there is no single answer to so simple a question as "What is an inch?" In fact, we have changed from a British inch and our own legal meter, to our inch and the International meter and now through action of the American Standards Association we are actually using an inch based on conversion from the International meter which is neither our own legal inch or the British legal inch—and the British are using it too. Table I shows this history of the legal inch in the United States.

It will be seen that under the present status there exists a difference of two parts in a million between the legal inch and the inch used in the dimensional work of industry. This difference is more theoretical than real in small dimensions and industrial use. The bill before Congress, sponsored by the Bureau of Standards is intended to eliminate this as well as any possible ambiguity in the U. S. inch.

Decimal Dimensioning

In subdividing the inch the modern trend in industry is toward the use of decimals instead of the older common fractions although fractions continue to be used, especially for dimensions of certain materials such as iron pipe, lumber, phenol fiber. In fact even a special decimal system based on using only the tenths and fiftieths of an inch is being considerably discussed by general industry. This system would use a scale on which the smallest division is $\frac{1}{50}''$ or $.020''$ instead of $\frac{1}{64}'' = .0156''$. It is in use by the Ford Motor Company and the values shown in Table II are some of those used in place of common fractions. The decimal equivalents of these common fractions are also shown rounded to 3 decimal places in accordance with American Standard Rules for Rounding off Numerical Values Z25.1-1940.

In the Ford system one and two-digit decimals carry the general tolerance of $\pm$ .010″. When greater accuracy is required three-place decimals are used to express a minimum and a maximum value.

The adoption of decimal dimensioning for all drawings prepared at Bell Telephone Laboratories is being actively considered. However, adoption

TABLE I
History of United States Dimensional Standards

Year	Action	Resulting Dimensional Relationships
1830–36	Adoption for Customs Service and for distribution to individual states of standards intended to be the English yard based on a certain portion of an 82 inch bar imported in 1813. The portion selected was supposed to be identical with the English yard.	
1856	Official copy of new British Imperial Yard accepted as standard	International Meter = 39.370147 British Inch
1866	Congress declared metric units lawful and established legal equivalents	Legal Meter in U. S. = 39.37 British Inch
1893	Mendenhall Order set up International meter as the fundamental standard	International Meter = 39.37 U. S. Inch
1933	American Standards Association (Representing Industry) adopts 1 inch = 2.54 centimeters	International Meter = 39.370078 U. S. Inch
1937–41	Bill before Congress but held in committee for amendments	International Meter = 39.370078 U. S. Inch

TABLE II
Examples of Ford Decimals Compared to Common Fractions

Ford Decimal	Common Fraction	Decimal of Existing Common Fraction	American Standard Decimal Equivalents (3 Place)
.02	1/64	.015625	.016
.03	1/32	.03125	.031
.05	3/64	.046875	.047
.06	1/16	.0625	.062
.08	5/64	.078125	.078
.3	7/32	.21875	.219
.46	15/32	.46875	.469

of decimal dimensioning would not of itself result in any changes in our system for establishing tolerance values.

Raw Material Sizes

In contrast to this continued trend toward simplification and rationalization of our systems of dimensional units raw material supply is still complicated by a multitude of obsolete systems of gauge sizes in every day use.

Many in industry have probably grown used to the standard gauges in particular fields but though gauge numbers were undoubtedly initiated as a simplified identification the variety of gauges and the variety of names for the same gauge now merely increases confusion. Sheet metals are handled in terms of a number of gauges such as B&S gauge, U. S. standard gauge and BWG gauge; and sheet soft rubber is even designated in decimals of $\frac{1}{64}$ such as $\frac{4.3}{64}''$. It has become good practice to specify sizes by decimal dimension values and not by gauge numbers and holes by actual decimal size rather than by drill numbers. The actual sizes used, however, are determined in many cases by the values corresponding to old gauge numbers long used commercially, though in large running items mills will and do manufacture to any specified decimal size. For some time it has been the practice of material manufacturers and other large industries thus to discontinue the use of gauge numbers though still using the decimal values of gauge sizes.

There is now under way an effort, organized under committee B32 of the American Standards Association, to eliminate the old wire and sheet metal gauge systems entirely and set up a rational series of American standard thicknesses for all metal sheets and preferred diameters for wire, and insure availability in these sizes. The basic conception of a rational series of sizes is that a uniform degree of choice should be presented between successive sizes. Therefore each size should differ from the next by a fixed percentage. The series should therefore be geometric. A variety of geometric series could be used but in order to permit extending the series indefinitely by shifting the decimal point, the particular series based on the root of 10 has been established internationally as the Preferred Numbers Series for standard sizes. The 5 series is one having 5 numbers between 1 and 10 (or between 10 and 100) and is produced by using as the multiplier the fifth root of 10; the 10 series is produced by multiplying by the 10th root of 10; the 20 series by multiplying by the 20th root of 10 etc. The complete Preferred Numbers Series is explained and listed in various forms in American Standard Z17.1-1936.

The subcommittee working on the sheet metal sizes has recently issued a proposed American Standard of preferred thicknesses for all uncoated flat metals thinner than .250″. These thicknesses are all decimals based on the 20 series of preferred numbers rounded in the standard manner to 3 decimal places. The Preferred Numbers and the proposed thicknesses are shown by Table III. It happens that this series closely approximates the Brown and Sharp gauge used in the nonferrous metals which simplifies that portion of the changeover. If this proposed American Standard is generally approved, as now appears most promising, we will be able to choose thicknesses of any metal interchangeably without the restrictions of ancient gauge sizes es-

TABLE III

Decimal Series of Preferred Numbers 10–100			Proposed Preferred American Standard Thicknesses		
5 Series $\sqrt[5]{10} = 1.6$	10 Series $\sqrt[10]{10} = 1.25$	20 Series $\sqrt[20]{10} = 1.12$	Under .010	.010 to .100	.1120 to .250
10	10	10		.010	
		11.2		.011	.112
	12.5	12.5		.012	.125
		14		.014	.140
16	16	16		.016	.160
		18		.018	.180
	20	20		.020	.200
		22.4		.022	.224
25	25	25		.025	
		28		.028	
	31.5	31.5		.032	
		35.5		.036	
40	40	40	.004	.040	
		45		.045	
	50	50	.005	.050	
		56		.056	
63	63	63	.006	.063	
		71	.007	.071	
	80	80	.008	.080	
		90	.009	.090	
100	100	100		.100	

tablished for reasons which were possibly good and sufficient but which certainly have long been forgotten. Meanwhile, another subcommittee is investigating the possibility of applying a similar series to the diameters of wire. Probably diameters to 4 decimal places will be required.

Dimensional Tolerances

Part Tolerances

Regardless of the dimension decided upon in a design it is obvious that it cannot be regularly manufactured to the exact size. Certain manufacturing variations or tolerances must be expected and these introduce a large share of our dimensional problems.

The usual statement on tolerances is that the larger the tolerance allowed the cheaper the part is to manufacture and, therefore, the tolerance specified should be the widest that will permit functioning. However, this is generally true only of overall tolerances which define the manufacturing methods that may be used. It is true in the sense that apparatus is inexpensive to manufacture if it can be so designed that its functioning is largely independent of variations in dimensions. However, such design is not usually achieved and in much apparatus fairly good overall accuracy of dimensions and fit is necessary for uniform functioning. The problem of setting tolerances then becomes one of distributing certain tolerances over various dimensions and different parts. This is a very difficult problem and in the case of any individual tolerance a larger value does not necessarily mean lower apparatus cost and may even mean the reverse.

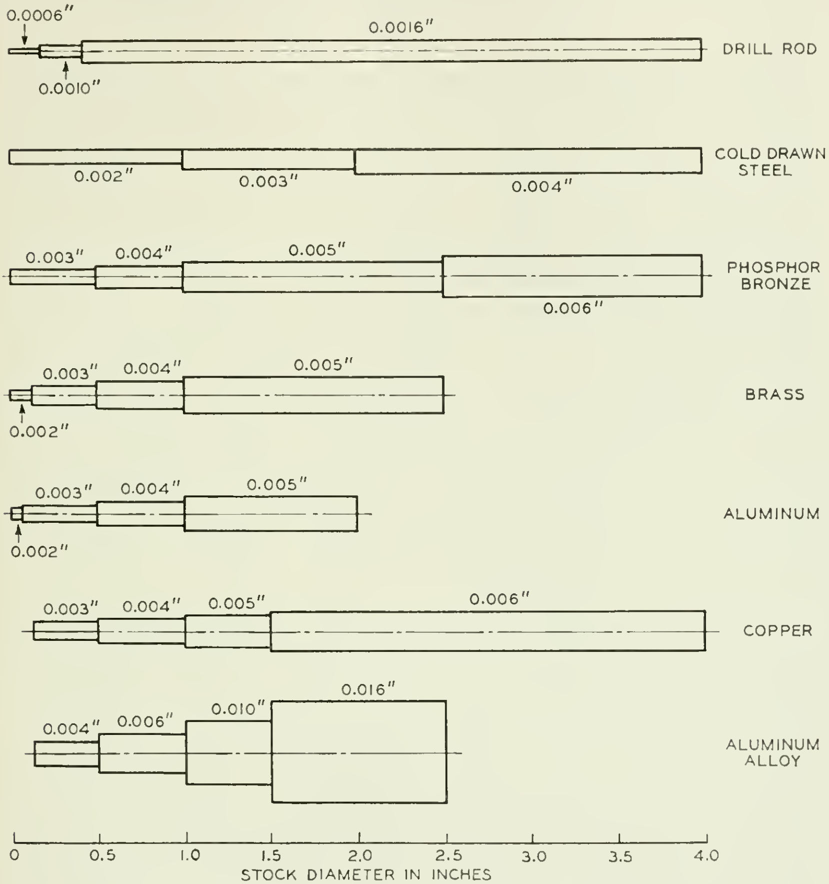


Fig. 1—Total diameter tolerances of commercial round stock

This is easily demonstrated in the case of part tolerances on dimensions which correspond to the dimensions of raw materials. Figure 1 shows the tolerances of commercial grades of round stock. If, for example, engineering requirements dictate the use of a particular material there is no gain in specifying larger tolerances than those to which it is regularly furnished and doing so may require greater accuracy in the mating part. There may even be economy in the use of higher priced material produced to closer tolerances, as for example, drill rod instead of cold drawn steel through economy in the manufacture of associated parts. Similarly manufacture of cantilever springs from sheet stock produced to closer tolerances may reduce the cost of subsequent adjustments. Therefore, when individual part tolerances are involved consideration must always be given to the size tolerances of raw materials.

The same situation exists in the case of tolerances on dimensions produced by a manufacturer's own tools. While close overall limits will require greater overall accuracy of the tools provided and greater frequency of set-ups the most economical distribution of tolerances will be that based upon the normal tolerances that can be expected from various manufacturing operations. Certain degrees of accuracy are inherent in certain types of machines and tools and allowing variations not in proportion to these values serves little if any purpose. Also there are types of combination tools and automatic machines, familiar in mass production practice with which wide tolerances are not an economy because accuracy is required for locating or nesting the part for subsequent operations. Since the distribution of tolerances involves such complex factors of manufacturing method and cost as these, it is desirable for the designing engineer to determine and to indicate unmistakably the effect of tolerances upon functioning and, where interchangeability of individual parts in service is not involved, to allow manufacturing considerations to determine the distribution of tolerances in an assembly.

It is apparent that considerable study of the requirements for functioning of the design, of available materials and the limitations of manufacturing process are required to establish the most economic balance between performance of the apparatus and the required tolerances. Consideration should be given to these tolerance factors in cooperation with manufacturing engineers in an early stage of a design problem so that they may influence the trend of design. This step may avoid the necessity for slow and costly manufacturing developments and delays in starting production. However, completely rigid adherence to the status quo of tolerances is not necessary in long range planning of major design projects. In such cases the trend of progress in materials and manufacture should be determined and anticipated. For example, some cantilever spring design requiring narrow control has been based on sheet material produced to tolerances not commercially available at the time but made so by the time it was needed for production. The extent of progress in this direction is shown by Fig. 2.

Similar progress in manufacturing technique can also be expected. For example, the development of broaching from a comparatively crude operation to the precision method it is today is recent and outstanding.

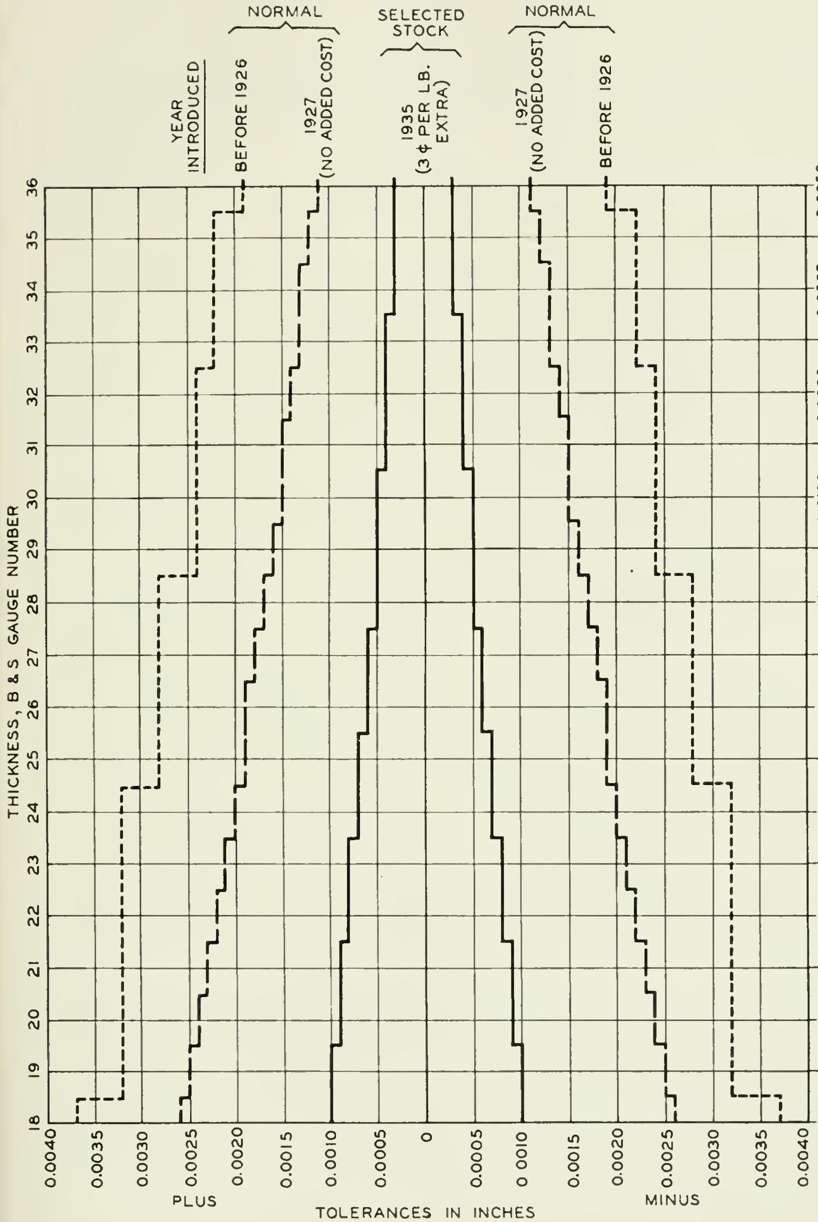


Fig. 2—Improvement in thickness tolerances for brass sheet 1926–1939

Cumulative Assembly Tolerances

Another problem in choice of tolerances is in those cases where a considerable number of parts are additively assembled into a unit as in the case of "spring pileups" used on electrical contacting apparatus such as relays and switches. These consist of considerable numbers of sheet metal springs and insulators alternating and clamped by screws. If the overall tolerance on such an assembly must be taken as the sum of the tolerances of the individual parts various courses of action are presented, the extremes of which are:

1. Very small tolerances must be maintained on the individual parts or
2. Adequate space must be provided in the apparatus for extremely large variations in the assembly.

Small tolerances on the individual parts may be extremely expensive and large space allowances and provisions in associated parts for variations in the assembly may be a serious design handicap.

However, it is recognized that there is obviously small probability that all minimum or all maximum parts will appear in any one assembly. It has been found satisfactory in certain types of such pileups to assume that the maximum dimensional variation that will actually be encountered in an assembly will not be greater than 70% of the sum of the part tolerances. A similar situation exists in many kinds of assemblies or associations of tolerances.

The statistical relationships involved in this problem are indicated by Fig. 3. The curves show the percentage of the cumulative part tolerances within which 99.7% of the assemblies may be expected to be found with different numbers of similar units in the assembly. The solid line is deduced from theoretical relationships. It assumes that the parts are all of one kind, that the parts going to assembly are controlled, of normal distribution and the limits are rationally set to represent the actual conditions. The dotted curves have been deduced from relationships which have been proposed as representing rectangular and triangular distributions of individual part tolerances. The curves may not be truly representative of specific cases because of inconsistent selection of limits or erratic distributions. However, they indicate that the 70% rule on pileups is probably on the safe side in most cases and that closer design of assembly or less restrictive tolerances and cheaper manufacture of piece-parts might be readily possible either (1) by better control, (2) by actual mixing of lots of piece parts or (3) even merely by knowledge of the actual statistical distribution of part dimensions.

The three points indicated in Fig. 3 show the results of a limited experiment in which pileups were assembled from 2083 individual insulators of $\frac{1}{32}''$ phenol fiber taken from factory stock. The establishment of curves by this type of experiment using a sufficiently large and representative sample would be practicable and would permit considerable condensation

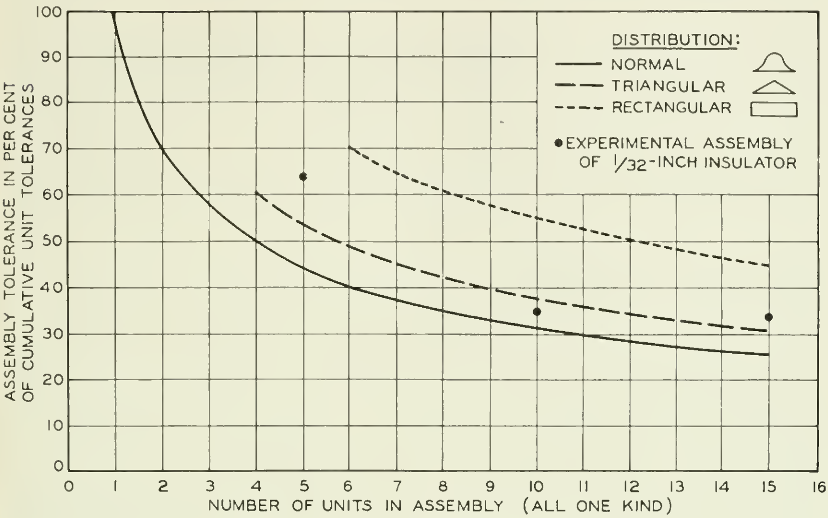


Fig. 3—Statistical relationship of overall tolerance on an assembly and the sum of the individual part tolerances

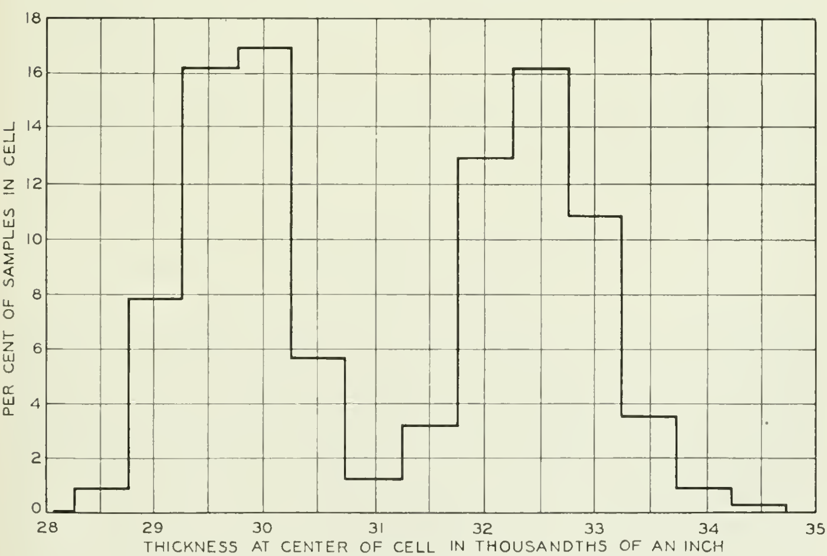


Fig. 4—Distribution of thickness in 2083 pileup insulators

of design on a sound basis. In this particular case the parts used apparently came from only two different sheets of fiber as indicated by the distribution of thickness of the individual parts shown by Fig. 4.

Further statistical analysis of this type of situation is needed together with experimental determination of the distribution of dimensions actually encountered in specific cases.

The distribution of dimensions in a product arises from a variety of causes. One type of cause is the variations such as those in the dimensions and physical properties of raw material which may produce different product dimensions even from a particular tool. A different and more systematic type of cause is the change in the dimensions of tools as a result of wear. The practice followed in establishing tool wear allowances will therefore affect the limits and statistical distribution of part dimensions during the life of the tool. Some designers and some tool makers consider that the specification of a nominal value with plus and minus variations requires a different handling of initial tool dimensions and wear allowances than does the specification only of minimum and maximum limits for a part dimension. Equally good authority maintains that a manufacturer recognizes no difference. Establishment of standard practices in such matters is a needed step in determining the distribution of dimensions to be expected in machined parts. In the present absence of standards or of any consistent attitude on the subject it is necessary for designing and manufacturing engineers to reach an agreement in specific cases where this factor is important.

Such are the factors which determine the tolerances which can be obtained economically or which perhaps will be unavoidably encountered. It is necessary for a designer to keep informed of the interaction of these factors as his design crystallizes and he must also determine the effect of such tolerances upon functioning in order to complete a design which will function properly when assembled in quantity production.

Functional Dimensioning

Effect of Tolerances

If apparatus parts are minute or have complicated relative motions it is recognized that manufacturing drawings to the usual scale have serious limitations to their usefulness in the analysis of the effects of combinations of tolerances. In such cases designers frequently make layouts to larger scales or large scale adjustable models to investigate the effect of variations on functioning. Illustrations of this practice are numerous in the experience of most designers of small apparatus.

Even in large parts which are stationary in use the application of tolerances, in effect, establishes several possible positions for each element and may present problems similar to those involving motion. These are not easily recognized because of a curious limitation inherent in small scale drawings. This limitation is probably well known to most engineers but it is worthwhile to analyze it because it is important to be always aware of it.

This limitation is the fact that in drawings the shape of the part and the effect of all nominal dimensions are actually shown graphically whereas, it is possible to indicate tolerances numerically but not graphically. We are therefore apt to visualize the part as it is graphically shown, that is, without tolerances and to think of the numerical tolerances one at a time rather than in combinations as they affect each other and the shape of the part.

If any dimension, significantly affecting the design of a part, is changed the drawing is immediately corrected so that its meaning will be clear and the functioning of the part can be checked. This obviously facilitates design and manufacture. Yet because they cannot be shown directly by regular drawing methods, we have grown accustomed to not being shown the effect of tolerances or changes in tolerances upon the shape of the part. Nevertheless it is obvious that these effects are critical in the functioning of the part or tolerances would not be set. The fact that these critical features of the design are not actually graphically shown and therefore are not easily seen and understood on the drafting board is a serious detriment in working out a design and in all later analysis of it. The full effect of interrelated variations particularly if in three dimensional space may appear only after tools are in process or the first parts produced and this may be rather late for economy.

Originally this difficult analysis of the effect of tolerances upon functioning probably involved only the designer. The manufacturer tried to make the part as nearly as possible to the nominal values shown and variations from them were accidental. Tolerances were looked upon as an indication of the care required and as a means of inspection for acceptance or rejection. With increasingly complex manufacturing tools the permitted tolerances are utilized more and more in the design of tools to allow the greatest possible wear before defective parts are produced and the tools must be replaced. For mass production parts progressive step type tools are used in which a continuous strip of stock advances by various stages from blank sheet to finished part. Tools of this type are extremely expensive and in order to obtain maximum life full use of allowed variations is made in their design. Design of such tools and the gauges required to maintain quality in mass production therefore also requires analysis of the effect of combinations of variables upon the desired part. As the designer has presumably already made this analysis, and incidentally is best qualified to do it, economy and accuracy dictate that his analysis be transmitted to the manufacturing engineer. The problem is to find means by which he can indicate unmistakably on the drawing his analysis of the required functioning of the part and the manner in which he intends the tolerances to apply, in the event that there is any possibility of misunderstanding.

The essence of this problem and some of the possibilities of solution can best be seen by reference to drawings which illustrate the major points.

Figure 5 shows the drawing of a flat plate dimensioned from center lines but without any tolerances whatever. Some minor dimensions not involved in this discussion are omitted in the interest of simplification but the part shown is in every way a normal one. The meaning of the drawing is completely clear and can be interpreted in but one way no matter from

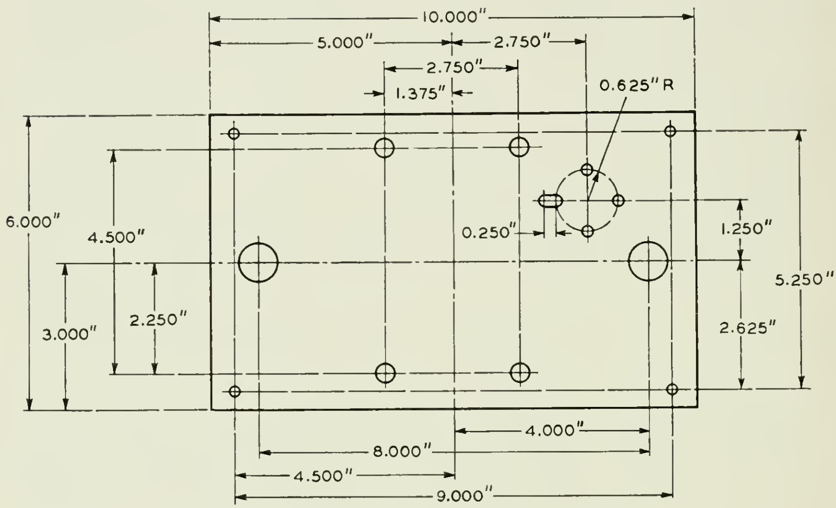


Fig. 5—Flat plate dimensioned without tolerances

what standpoint the analysis is made. The reason for this is obviously that but one value is shown for every dimension.

Figure 6 shows this same drawing dimensioned in exactly the same way with the exception that tolerances are shown for most of the dimensions. To the uninitiated it might appear to present no more problem than the previous drawing without tolerances because of the tendency to visualize the drawing in terms of the nominal dimensions only.

When the engineer analyzes the effect of the combinations of the various tolerances shown, interesting questions immediately arise. In the first place the combination of holes dimensioned 1.25″ ± .002″ from the center line appears to be definitely located because on the drawing the center line is shown in a definite position. Yet when the tolerances are considered the center line of this drawing could actually be shown in several different places as, for example:

1. It may be a line through the centers of the two large holes.
2. It may be a line anywhere from 2.992″ to 3.008″ from the outside edges.
3. It may be 2.247″ to 2.253″ from the small holes in the center of the plate.
4. It may be 2.615″ to 2.635″ from the holes numbered 2 and 4.

In brief, the center line which appears so definitely located on the drawing may actually be rather an indefinite location on the part when the various

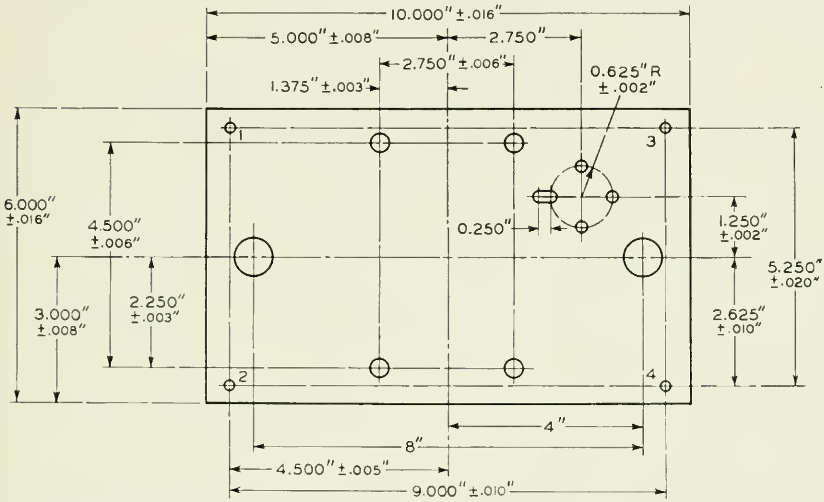


Fig. 6—Flat plate of Fig. 5 with the addition of tolerances

tolerances are considered. While the differences in the possible interpretations are in the order of thousandths of an inch nevertheless this order of magnitude is critical in this part or the indicated tolerances would not have been used. The interpretation of the center line which should be adopted will depend entirely upon the manner in which the part is intended to function and therefore should be indicated by the designing engineer. Obviously, not all designs or all dimensioning will present this difficulty but all should be studied from this viewpoint to determine whether or not they do.

Functional Datum Positions

When the type of uncertainty illustrated exists, it is necessary to indicate clearly the effect of tolerances on functioning by establishing the functional positions to which dimensions should refer. It may be difficult to do this graphically, in which case it is necessary to indicate by notes the particular interpretation which the designer intends. As an example, if the part of Figs. 5 and 6 functions by being located in position by means of the four holes numbered 1, 2, 3 and 4, the intentions of the designer are readily indicated by the following notes:

1. Functional datum line I is midway between the centers of holes 1 and 2 and the centers of holes 3 and 4.
2. Functional datum line II is perpendicular to datum line I at a point midway between the centers of holes 2 and 4.

These notes establish both horizontal and vertical center lines specifically in terms of the center of the one set of dimensions between the holes marked

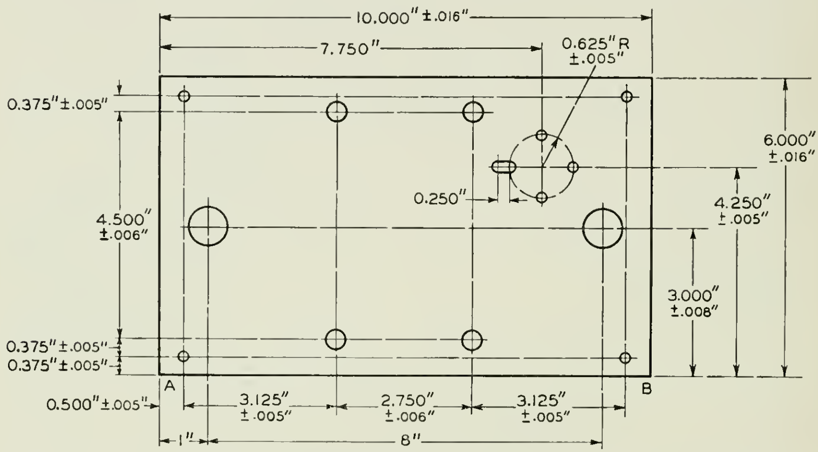


Fig. 7—Flat plate of Fig. 5 functionally dimensioned from outside edges with tolerances

1, 2, 3 and 4. The term functional datum line is suggested as completely descriptive but other equivalent terms might be used. This information could be indicated on the drawing without the use of notes by the adoption and use of some standard convention or symbol to indicate the particular dimension bisected by the center line.

If the functioning of this part were determined by location against the outside edges, this could be readily indicated by dimensioning the part as shown by Fig. 7 and using notes establishing the line A–B as one datum line and the perpendicular to it through A as the other.

In either of these cases the drawing becomes completely definite and subject to only one interpretation. In drawings of this type no change in the method of dimensioning may be required and the problem is solved simply by the addition of suitable notes or symbols indicating the intention of the designer as to functional datum lines.

It is sufficient to establish datum lines in the case of parts which are practically flat pieces with little depth but when a part has substantial depth it will be noted that center lines or other datum lines on a drawing really represent planes in space. In such parts it becomes necessary to establish datum planes rather than lines and three planes at right angles to each other are required.

Figure 8 illustrates such a part which might be an armature such as is used in many pieces of electrical contacting apparatus. In the typical operation of such a part its functioning is determined by the relation of its

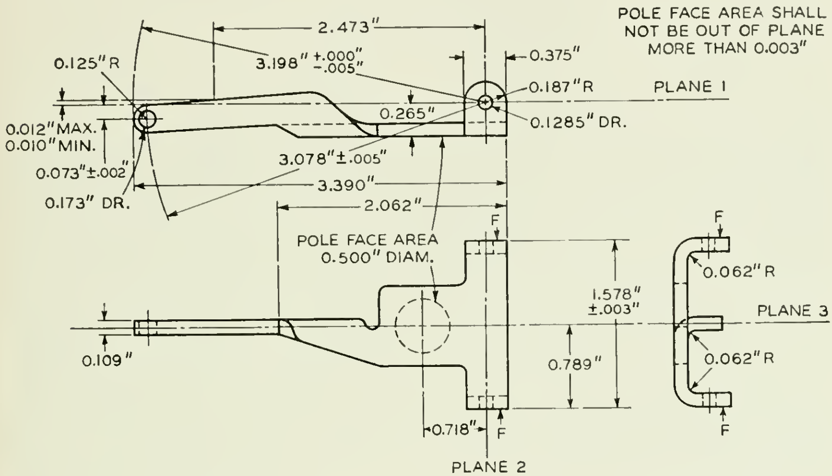


Fig. 8—Functional datum plane dimensioning of magnetic armature type of part

Functional datum plane I passes through the common axis of the two .1285 in. diameter holes and .265 in. above the pole face gauge position.

Functional datum plane II is perpendicular to plane I and passes through the common axis of the two .1285 in. diameter holes.

Functional datum plane III is perpendicular to planes I and II and passes midway between the finished surfaces which are 1.578 ± .003 in. apart.

various dimensions to the position of the pole face and the axis support. In order to indicate this on the drawing it is necessary to establish dimensioning as shown and add to the drawing the notes shown.

These notes establish three functional datum planes, the first through the axis at the point of support and a distance .265″ from the pole face area; the second at right angles to the first through the axis of support and the third at right angles to both the first and second and halfway between the finished surfaces 1.578″ apart. With these planes established the application of all the limits and tolerances shown is based on the operating position and analysis of the design is simplified. The drawing and the intentions of the design engineer cannot be misunderstood.

The clear expression of the designer's intentions by datum plane dimensioning will be appreciated by all concerned with the drawing or the resulting part. Inspection of the part is expedited no less than production. The inspector can usually by means of gauge blocks or simple fixtures set the part up on a surface plate as indicated by the drawings datum planes and positions. He can then establish the conformance of the part with the drawing by simple measurements to the indicated horizontal and vertical planes. When production quantities justify special gauges the required design of the gauge is established clearly by the datum planes.

Invariable or Gauge Dimensions

The drawing of Fig. 8 just described illustrates the use of gauge dimensions. The dimensions .265″ and .718″ and the indicated half-inch diameter for the pole face are all gauge dimensions without tolerances and some statement must be made or understanding reached that they are considered invariable and tolerances not permitted. They represent, it might be said, theoretical dimensions, on the drawing, or in practice they represent tools or gauging apparatus made to the highest standards of accuracy. These invariable dimensions are necessary in order to establish a starting point for the dimensioning of the part. It may appear at first that stating that a dimension has no manufacturing tolerance or variation is a hardship upon the manufacturer but this is not really so because the dimensions are not ones which are actually manufactured in the part. They represent usually dimensions built into tools or gauging equipment which are made to a precision greatly superior to that represented by part tolerances.

Invariable dimensions, or better, gauging dimensions or whatever it is proposed to call them are really not a new invention and it is possible to cite easily recognized examples. For instance, the dimension 2.473″ on Fig. 8 is an invariable gauging dimension not associated with the setting up of datum planes but typical of long standing use of invariable dimensions. We all can recall also the use of the term "theoretically correct position" and it is present practice in the case of vacuum tube bases and similar apparatus to designate the location of the contact studs in terms of a gauge having holes located on "true centers." Last but not least a minimum or maximum limit in its application is itself an invariable dimension.

In effect, datum lines or planes established when necessary by use of invariable or gauging dimensions remove the uncertainty as to the designer's intentions and prevent misunderstandings between design, production and inspecting engineers. Admittedly they do not completely solve all problems of dimensions as probably nothing will. They do, however, transfer whatever problems remain from the field of tolerances on finished product to the realm of tool making tolerances and gauging tolerances. The problem of how invariable is "invariable" remains but we are obviously then considering differences of an order of magnitude not usually vitally significant in the functioning of product parts. Theoretically, all "invariable" dimensions should be taken to the best accuracy of good gauging methods which means that any differences of opinion will be reduced at least to one-fifth and probably to one-tenth of the order of magnitude of those where tolerances themselves are involved.

It will be necessary to specially identify gauging dimensions on drawings to distinguish them from ordinary unlimited dimensions and to indicate that they are dimensions for gauges to which only gauge tolerances apply.

Practical Use of Datum Lines and Planes

It is not usual to establish datum lines on all drawings but if their use is necessary in the layout and design of the part they need to be permanently identified. This use of datum lines and planes on drawings, where necessary, may require somewhat greater drafting effort in the actual production of the drawing but their use results in a simplification of design and of the work of those subsequently using the drawings. It reduces the effort expended in analysis of drawings preparatory to the construction of tools and minimizes the possibility of misunderstandings or errors in tools. In products manufactured only intermittently it is particularly valuable as it minimizes the need for understandings and instructions supplementary to the drawings which may be forgotten between production periods or lost through shifts in personnel.

The overall economy in engineering effort and the reduction of the numerous possibilities of error more than compensate for the increase in the actual work of indicating datum positions, lines or planes upon drawings. In addition the choice of design of punches and dies and similar tools by production engineers is better guided by the designer's requirements if functional datum lines are clearly identified. An obvious example is the use of either the inside or outside of a punched and formed part as the starting point. In brief datum plane dimensioning is a more explicit expression on the drawing, of the designers "end point requirements".

When establishing datum planes, it is important to consider them in terms of the actual physical part rather than in terms of the drawing. Lines which appear as definite points on a drawing may not be actually part of the product when it is completed or may be on surfaces shown as a line on the drawing but rough or unfinished in the part. It is difficult to establish any set of rules covering what shall or shall not be done because each drawing and each part must be considered practically as an individual case. That this is so will be amply demonstrated by a serious study of even one part. However, there are obvious generalities which can be established and Fig. 9 shows some of them.

An example of functional datum plane analysis and dimensioning in three dimensions of a complicated part is shown by Fig. 10. This is the die cast frame for a special selector switch. It is the base upon which many interrelated parts and subassemblies are mounted. The proper functioning of the completely assembled switch depends in large measure on proper manufacture of this casting. In effect, the switch is designed around a vertical shaft passing through points P and Q and planes 1 and 2 are, therefore, established through the axis of this shaft. The production planning engineers intend to design the die and withdraw die plugs from such directions that the mounting surfaces will be smooth, flat and without

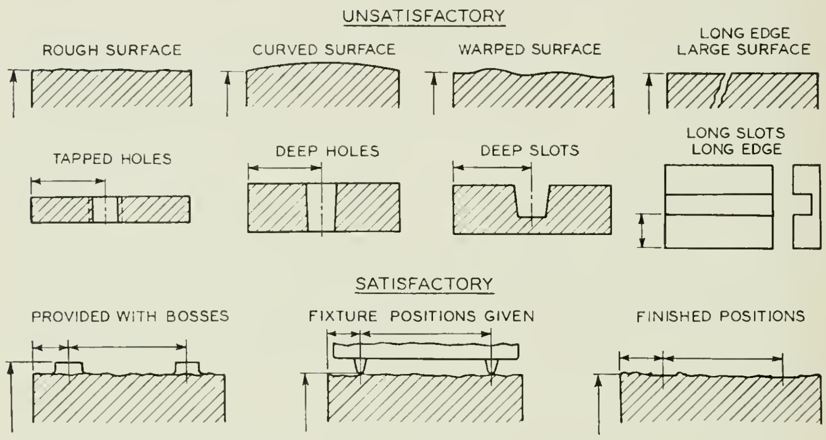


Fig. 9—Types of datum positions

any taper and they intend to use these surfaces as guiding points for their jigs and fixtures. It is for this reason that Plane 1 is established parallel to these mounting surfaces and an indicated distance from them. The other planes are established as shown on this drawing and described by the notes. With this arrangement of planes the designer's analysis in terms of Plane 1 is easily worked out and the reference of Plane 1 to the mounting surfaces permits the production or tool engineer to translate the design of the part into the design of his tools without necessity for further analysis and without the possibility of different interpretations. It will be noted that invariable or gauge dimensions are again used. The complete drawing of this part is very complicated and occupies a drawing practically 4 ft. x 6 ft. The perspective sketch shown and the accompanying notes are incorporated in the drawing as a separate view.

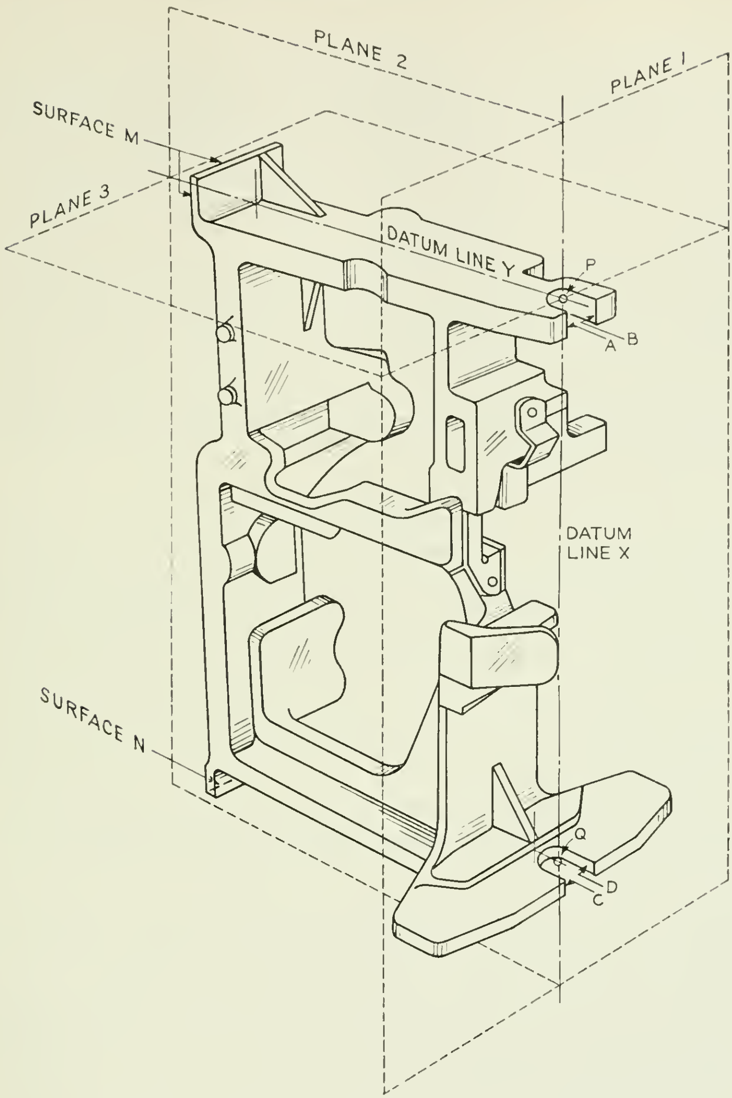


Fig. 10—Functional datum planes of complicated switch frame

Dimensions to datum line "X" or "Y" of the drawing of the frame refer to functional datum planes 1, 2 or 3 described below. Points "P" and "Q" are gauge points used in establishing these datum lines and planes. Points "P" and "Q" shall be half-way between the surfaces "A" and "B" and "C" and "D" respectively and 4.358 in. from the plane of surfaces "M" and "N" on the mounting lugs.

Datum line "X" shall pass through the points "P" and "Q".

Plane 1 shall be parallel to surfaces "M" and "N" and shall include datum line "X".

Plane 2 shall be perpendicular to plane 1 and shall also include datum line "X".

Plane 3 shall be perpendicular to plane 1 and to plane 2 at the point "P".

Datum line "Y" passes through point "P" and is the intersection of planes 2 and 3.

Required Standardization

It is not suggested that the drawings shown and the notes referred to represent a final practice on datum planes. A standard practice in designation of planes and standard terminology and understanding on gauge points and gauge dimensions is required. It will probably be desirable to adopt some symbol or designation for use on drawings to distinguish gauge dimensions which are invariable from ordinary unlimited dimensions to which manufacturing engineers for their own purposes usually add shop tolerances. One thing is certain and that is that datum planes, dimensions and tolerances when established should be primarily in terms of the required functioning of the apparatus. When that is done no one using the drawing in any capacity will have any doubts as to the designer's intention and this results in a great reduction in the discussions and analysis which might otherwise be necessary.

Summary

In summary it may be said that the whole approach to these problems in dimensions and tolerances should be on the basis of functioning. However, good engineering of dimensions and tolerances requires knowledge of what can reasonably be produced and the sources of reasonable tolerance values are:

1. Raw material limits including some knowledge of future trends and developments.
2. The normal accuracy of manufacture, also including anticipation of future improvement.
3. Discussion of trend of design with manufacturing engineers.

Solution of tolerance problems in the final design may involve all of the following steps:

1. Study of the effect of combinations of tolerances on functioning, allowing for statistical effects in accumulations of tolerances.
2. Discussion of this analysis with the production planning engineer because the analysis of tolerance combinations is important in the design of long life tools.
3. Indication of the results of such an analysis by the method of dimensioning drawings.
4. Indication on drawings of functional datum positions, lines or planes established on geometrically correct principles to permanently and unmistakably record the intentions of the designer regarding combinations of variations wherever this is necessary.

Time Division Multiplex Systems

By W. R. BENNETT

Introduction

THE idea of transmitting and receiving independent signals over a common line by means of synchronized switches at the terminals is quite old and has been used in multiplex telegraphy for many years. In general if N signal channels are to be provided over one line, the switching cycle includes N equal time intervals, one of which is allotted to each channel. Each channel is connected to the line throughout a part of its particular time interval and is disconnected throughout the remainder of the cycle. Absence of interference between the channels depends upon the fact that the channels are connected to the line throughout mutually exclusive time intervals. It is thus possible to avoid the use of channel band filters such as are necessary in carrier systems employing frequency as the basis of separation.

Application of time division multiplex methods to telephone channels has been proposed from time to time and some experiments have been made.[1,2,3,4,5,6] It is fairly evident that the concept of simple on-and-off switching giving alternately transmission and complete suppression for the signal from a particular channel on the line is inadequate for speech waves in actual telephone circuits. Imperfections in the transmission properties of the line tend to distort the wave form of the successive signal components and prolong the contribution of one signal into the time allotted for a different channel. It is the object of this paper to present a general quantitative discussion of the factors which enter into the transmission of any type of signal by a system of this kind. It has been found possible to arrive at definite criteria for such matters as the required switching frequency, the conditions to be imposed on contact time for good crosstalk suppression with economy of frequency band, and the transmission requirements which must be met by the intervening circuit to hold the interference between channels to tolerable values. The analysis leads directly to a physical viewpoint of the whole process which, to those familiar with the carrier and

[1] Patten and Minor, *U. S. Patent* 745,734, 1903.
[2] *Electrical World*, Dec. 5, 1903.
[3] Goldschmidt, *U. S. Patent* 1,087,113, Feb. 17, 1914.
[4] Poirson, *Soc. Fr. El.*, Apr. 1920.
[5] Marro, *L'Onde Electrique*, Oct. 1938.
[6] M. Cornilleau, *Revue de Telephones, Telegraphes et T. S. F.*, 13 (1935), pp. 625–643.

sideband philosophy of signal transmission, illuminates the manner in which departures from ideal amplitude and phase characteristics cause crosstalk between the several message channels. It further leads directly to other physical methods for producing and detecting a transmitted signal identical with the essential components derived in time division or switching processes.

A first step in the theoretical solution of the problem was taken by Dr. J. R. Carson, who, in an unpublished memorandum of May 25, 1920, derived quantitative relations between band width and interchannel interference in time division multiplex transmission. Applying Fourier series analysis to on-and-off switching, he showed that if the transmission medium had constant attenuation and linear phase shift for all frequencies below cutoff and no transmission of frequencies above the cutoff, the band width required for satisfactory multichannel telephony would be much wider than needed in conventional carrier methods. A further step was taken by Dr. H. Nyquist, who, in unpublished memoranda of August 24, 1936 and November 12, 1936,[7] showed that the width of band necessary may be reduced by providing a specially devised type of non-uniform transmission characteristic. In the following discussion, we shall see that a similar result can be obtained by control of the switching, and specific switching processes will be described which allow a flat transmission band of minimum width to be used.

In order to arrive at requirements which must be imposed on the various components of the system, we shall first give a theory of time division multiplex transmission in which both the switching processes and the transmission characteristic are completely general. Specific forms which give crosstalk suppression will then be discussed and effects of small departures estimated.

General Theory

We shall assume an N-channel system with a sinusoidal signal impressed on the j^{th} channel. An illustrative arrangement is shown in Fig. 1. Since the system is linear, we may represent currents and voltages by complex quantities with the understanding that the actual currents and voltages are the real components of the expressions used. Accordingly, let the signal voltage impressed on the j^{th} channel be

$$E_j(t) = E_j e^{i\omega_j t} \tag{1}$$

[7] Basic concepts used in Nyquist's analysis were included in his paper, "Certain Topics in Telegraph Transmission Theory," *A. I. E. E., Trans.*, April, 1928, pp. 617–644. Mention is also there made of the equivalence of signal shaping and equalizing in effect on reception of telegraph signals.

and let the switching between the j^{th} channel and the line at the sending end be represented by:

$$I_{sj}(t) = F_j(t)E_j(t), \tag{2}$$

where $I_{sj}(t)$ is the current flowing into the line from the j^{th} channel. The function $F_j(t)$ has the dimensions of an admittance and, in the arrangement shown in Fig. 1, is periodic in time with fundamental frequency $q = 2\pi/T$ radians per second, where T is the time occupied by one cycle of the switching operation. In the interests of economy of analysis, it is preferable for our purposes to assume for $F_j(t)$ a somewhat more general function of time

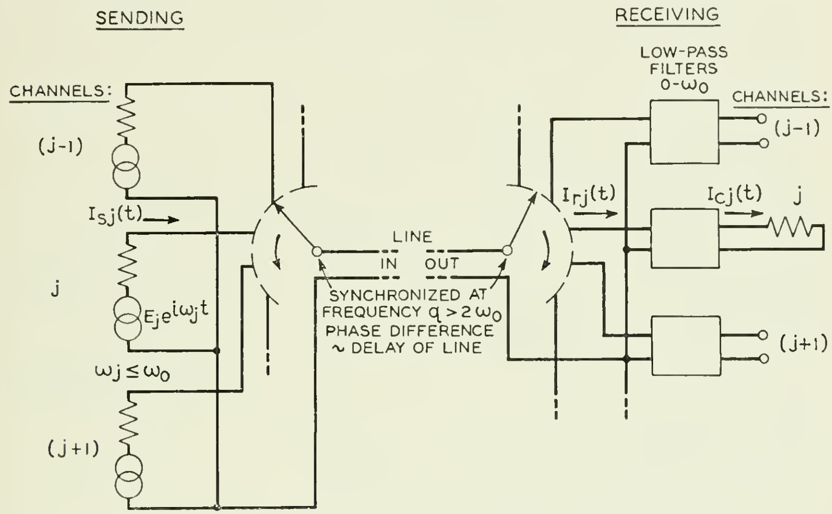


Fig. 1—Elementary arrangement for time division multiplex system

than is directly obtainable with the elementary arrangement of Fig. 1. We shall let

$$F_j(t) = \sum_{m=0}^{\infty} A_{mj} \cos\,[(\nu + mq)t - \theta_{mj}]. \tag{3}$$

To make the results applicable to Fig. 1, we merely let $\nu = 0$; then by the usual Fourier series analysis,

$$\left.\begin{aligned} &A_{0j} = a_0/2, \qquad \left.\begin{aligned} A_{mj}^2 &= a_{mj}^2 + b_{mj}^2 \\ \tan\theta_{mj} &= b_{mj}/a_{mj} \end{aligned}\right\} m > 0 \\ &\theta_{0j} = 0, \\ &a_{mj} = \frac{2}{T}\int_{t_1}^{T+t_1} F_j(t)\cos mqt\,dt \\ &b_{mj} = \frac{2}{T}\int_{t_1}^{T+t_1} F_j(t)\sin mqt\,dt,\ t_1 \text{ arbitrary} \end{aligned}\right\} \tag{4}$$

The wave (3) consists of the output of the circuit of Fig. 1 with all frequencies shifted by a constant amount ν radians per second; various means of accomplishing this result in the switching process will be discussed later. It is sufficient to point out here that such a shift in frequency is often desirable for optimum utilization of the transmission medium. Combining (2) and (3), we then have:

$$I_{sj}(t) = \frac{E_j}{2} \sum_{m=0}^{\infty} A_{mj}[e^{i(\nu+mq+\omega_j)t-i\theta_{mj}} + e^{-i(\nu+mq-\omega_j)t+i\theta_{mj}}] \tag{5}$$

It is clear from (5) that the result of the switching process is the production of upper and lower side frequencies from the signal on each harmonic of the switching frequency. It is also evident that if more than one signal component is superimposed, the resulting side frequencies constitute sidebands of the same nature as used in amplitude modulation systems. A significant difference between time division and amplitude modulation appears in that in the latter only one sideband or at most one pair of sidebands is transmitted, while the essential character of time division depends on the transmission of a plurality of sidebands. Thus if one pair of sidebands were selected from the output (5) by filtering, the time division process would merely be a particular way of generating the sidebands required in an amplitude modulation system.

The next step in a time division system is the transmission of the wave (5) over a line. The properties of the line in general may be specified by a complex transfer impedance, which we may express here by the ratio of open-circuit output voltage to input current:

$$E_r/I_s = Z(i\omega) \tag{6}$$

The result of applying the wave (5) to the line is then the open-circuit voltage:

$$\begin{aligned} E_{rj}(t) = {} & \frac{E_j}{2} \sum_{m=0}^{\infty} A_{mj} Z[i(\nu + mq + \omega_j)]\, e^{i(\nu+mq+\omega_j)t-i\theta_{mj}} \\ & + \frac{E_j}{2} \sum_{m=0}^{\infty} A_{mj} Z^*[i(\nu + mq - \omega_j)] e^{i(\nu+mq-\omega_j)t+i\theta_{mj}} \end{aligned} \tag{7}$$

In the above we have adopted the notation $Z^*(i\omega)$ to represent the conjugate of $Z(i\omega)$ and have made use of the fact that the response of a network to the applied wave $e^{-i\omega t}$ is the conjugate of the response to $e^{i\omega t}$.

At the receiving end another switching process takes place synchronously with that at the transmitting end. We shall assume that the switching process between the k^{th} channel and the line is represented by the relation

$$I_{rk}(t) = G_k(t) E_{rj}(t), \tag{8}$$

where $I_{rk}(t)$ is the current received in the k^{th} channel and $G_k(t)$ is a periodic function of time with fundamental frequency q. It is understood that j and k may be any two of the N channels. We shall express $G_k(t)$ in a manner analogous to the corresponding switching function at the transmitter, i.e.,

$$G_k(t) = \sum_{n=0}^{\infty} B_{nk} \cos\left[(\nu + nq)t - \Phi_{nk}\right] \tag{9}$$

Combining (7), (8), and (9), we find

$$I_{rk}(t) = \frac{E_j}{4} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} A_{mj} B_{nk} Z[i(\nu + mq + \omega_j)]$$

$$\left(e^{i[2\nu+(m+n)q+\omega_j]t-i(\theta_{mj}+\Phi_{nk})} + e^{i[(m-n)q+\omega_j]t-i(\theta_{mj}-\Phi_{nk})}\right)$$

$$+ \frac{E_j}{4} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} A_{mj} B_{nk} Z^*[i(\nu + mq - \omega_j)]$$

$$\left(e^{-i[2\nu+(m+n)q-\omega_j]t+i(\theta_{mj}+\Phi_{nk})} + e^{i[(n-m)q+\omega_j]t+i(\theta_{mj}-\Phi_{nk})}\right) \tag{10}$$

The received wave thus consists of a doubly infinite set of side frequencies involving harmonics of q. It is, however, possible to set up conditions under which the original signal may be selected and the frequencies involving the switching rate may be suppressed by filtering. If $\nu = 0$, such separation is possible provided

$$\omega_j < q/2, \tag{11}$$

for it then follows that a low-pass filter with cutoff frequency at $q/2$ will not pass any of the components with frequencies dependent on q. The condition (11) follows from the fact that the lowest frequency of (10) dependent on q is $q - \omega_j$, and hence we must make $q - \omega_j > \omega_j$ in order to separate ω_j from $q - \omega_j$. In other words the sidebands on adjacent harmonics must not overlap. If $\nu > 0$, the condition (11) also suffices as far as suppression of terms dependent on q are concerned, but an additional condition is required to suppress frequencies dependent on ν in the special case in which $\nu < q/2$, i.e., the case of ν less than the maximum allowable value of ω_j. For in the latter case the frequency $2\nu + (m + n)q - \omega_j$ is less than $q - \omega_j$ in the special case of $m = n = 0$. The additional condition needed is evidently either $A_{0j} = 0$ or $B_{0k} = 0$. If $\nu = 0$ or if $\nu > q/2$, this condition is unnecessary.

Assuming then that (11) is fulfilled, and that a low-pass filter with cutoff at $q/2$ is inserted in the output of each channel, we calculate for the typical channel output current:

$$I_{ck}(t) = Y_{jk} E_j e^{i\omega_j t}, \tag{12}$$

where the value of Y_{jk} is as follows:

Case 1, $\nu = 0$

$$Y_{jk} = A_{0j}B_{0k}Z(i\omega_j) + \frac{1}{4}\sum_{m=1}^{\infty} A_{mj}B_{mk}Z[i(mq + \omega_j)]e^{-i(\theta_{mj}-\Phi_{nk})} + \frac{1}{4}\sum_{m=1}^{\infty} A_{mj}B_{mk}Z^*[i(mq - \omega_j)]e^{i(\theta_{mj}-\Phi_{nk})} \quad (13)$$

Case 2, $\nu > 0$; A_{0j} or $B_{0k} = 0$ if $\nu < q/2$

$$Y_{jk} = \frac{1}{4}\sum_{m=0}^{\infty} A_{mj}B_{mk}Z[i(\nu + mq + \omega_j)]e^{-i(\theta_{mj}-\Phi_{mk})} + \frac{1}{4}\sum_{m=0}^{\infty} A_{mj}B_{mk}Z^*[i(\nu + mq - \omega_j)]e^{i(\theta_{mj}-\Phi_{mk})} \quad (14)$$

The combination of an N-channel time division multiplex system with low-pass filters in the receiving branches is thus found to be equivalent to a linear network having N pairs of input and output terminals with the transfer admittance from the j^{th} pair of input terminals to the k^{th} pair of output terminals given by Y_{jk} in (13) or (14). The transfer admittance is calculated by summing the contributions of upper and lower sidebands on harmonics of the switching frequency and is affected directly by the transmitting properties of the medium at the side band frequencies. The result we have obtained is of sufficient generality to include all cases we shall treat in this paper. We shall now proceed to specific examples.

On-and-Off Switching with Commutator

When an ideal commutator is used as a switching means, the switching functions for the N channels are identical except for a time displacement which is the same between all pairs of consecutive channels. This condition is expressed by:

$$F_j(t) = F_1[t - (j - 1)T/N] \quad (15)$$

Thus $F_1(t)$, the switching function for the first channel becomes a reference function, $F_2(t)$ is the same except for a time delay of T/N, $F_3(t)$ is delayed by $2T/N$, etc. Substitution of (15) in (3) gives the relations:

$$\left.\begin{aligned} A_{mj} &= A_{m1} \\ \theta_{mj} &= \theta_{m1} + (j - 1)2m\pi/N \end{aligned}\right\} \quad (16)$$

If we further suppose that the commutator makes contact between the typical channel and the line throughout a fraction x of the time interval

T/N allotted to that channel and breaks contact throughout the remainder of the switching cycle, we may write the reference switching function as:

$$F_1(t) = \begin{pmatrix} A, & -xT/2n < t < xT/2N \\ 0, & xT/2N < t < (2N - x)T/2N \end{pmatrix} \tag{17}$$

Hence from (4)

$$\left.\begin{aligned} A_{01} &= Ax/N, \\ A_{m1} &= \frac{2A}{m\pi} \sin \frac{mx\pi}{N}, \ m > 0 \\ \theta_{m1} &= 0 \end{aligned}\right\} \tag{18}$$

In the receiving device, the corresponding switching process should be delayed with respect to the transmitter by a time interval t_0 equal to the time of transmission of the line. Hence we write

$$\left.\begin{aligned} B_{01} &= Bx/N \\ B_{m1} &= \frac{2B}{m\pi} \sin \frac{mx\pi}{N}, \ m > 0 \\ \Phi_{m1} &= mqt_0 \end{aligned}\right\} \tag{19}$$

B_{mj} and θ_{mj} are related to B_{m1} and Φ_{m1} in a manner analogous to (16).

The time of transmission of a distorting line is not precisely definable, but may be represented for our purpose by a linear phase component of $Z(i\omega)$. That is, we write

$$Z(i\omega) = Z_0(i\omega)e^{-it_0\omega}, \tag{20}$$

where t_0 is the slope of a straight line giving the best linear approximation to the phase vs. frequency curve, and $Z_0(i\omega)$ is the impedance function remaining after the subtraction of $t_0\omega$ from the actual phase shift ordinates. Substituting (15)–(19) in (12), we find

$$\begin{aligned} Y_{jk} = ABe^{-it_0\omega_j} \Bigg(& \frac{x^2}{N^2} Z_0(i\omega_j) \\ & + \sum_{m=1}^{\infty} \frac{\sin^2 mx\pi/N}{m^2\pi^2} Z_0[i(mq + \omega_j)]e^{-i(j-k)2m\pi/N} \\ & + \sum_{m=1}^{\infty} \frac{\sin^2 mx\pi/N}{m^2\pi^2} Z_0^*[i(mq - \omega_j)]e^{i(j-k)2m\pi/N} \Bigg) \end{aligned} \tag{21}$$

If the attenuation of the line is constant throughout the range ω_j to $Mq + \omega_j$ and all frequencies above the latter value are suppressed, (21) becomes

$$Y_{jk} = \frac{ABx^2 Z_0 e^{-it_0\omega_j}}{N^2}\left[1 + 2\sum_{m=1}^{M}\left(\frac{\sin mx\pi/N}{mx\pi/N}\right)^2 \cos(j-k)2m\pi/N\right] \quad (22)$$

The crosstalk ratio or ratio of amplitude of signal received in the k^{th} channel to that received in the j^{th} channel when signal is transmitted in the j^{th} channel is, therefore,

$$\frac{Y_{jk}}{Y_{jj}} = \frac{1 + 2\sum_{m=1}^{M}\left(\frac{\sin m\pi\, x/N}{m\pi\, x/N}\right)^2 \cos 2m\pi(k-j)/N}{1 + 2\sum_{m=1}^{M}\left(\frac{\sin m\pi\, x/N}{m\pi\, x/N}\right)^2} \quad (23)$$

Results of calculations made for a 10-channel system from (23) for $x = 1$ and $x = .5$, corresponding to no lost time and half lost time respectively in switching are shown in Fig. 2. It may be noted that adjacent channel crosstalk with half lost time is equivalent to alternate channel crosstalk with no lost time. Examination of the curves reveals a number of significant facts, among which are:

1. Crosstalk is quite imperfectly suppressed when the band width of the line is smaller than the theoretical minimum—the width of one sideband multiplied by the number of channels.

2. As the band width of the line is increased above the theoretical minimum, improvement in crosstalk suppression increases slowly, so that in general the use of frequency range on the line is uneconomical compared with other systems. For example, with no lost time in switching, the band width of the line must be increased tenfold to suppress adjacent channel crosstalk by 40 *db*. This conclusion is, however, to be qualified as follows:

3. When the duration of contact is decreased (less of the available channel time used) definite optimum transmission band widths appear which give better crosstalk suppression than bands somewhat wider or narrower. This suggests the possibility of critical phase relations existing between the contributions from the various sidebands such that if the right number having proper amplitudes and phases can be combined, complete suppression of crosstalk may occur even when the transmitted band width is finite.

When x, the fraction of contact time used, is made to approach zero, the limit of the amplitude factor (18) for the typical harmonic of the switching function is $A_{m1} = 2Ax/N$, which is independent of m. This is consistent with the known fact that a wave consisting of periodically repeated sharp pulses is composed of a large number of harmonics of nearly equal amplitude. If we use very short contact durations in time division, we should accordingly expect a large number of sidebands of nearly equal amplitude. The combination of proper numbers and phases of these sidebands offers a key to the realization of a time division multiplex system giving good crosstalk suppression with economy of frequency band.

Suppose that the duration of contact time is made sufficiently small to realize approximately the limiting values $A_{m1} = 2Ax/N$, $B_{m1} = 2Bx/N$ in transmitting and receiving respectively for the first $2M + 1$ of the sidebands and that by means of a low-pass filter with linear phase shift and

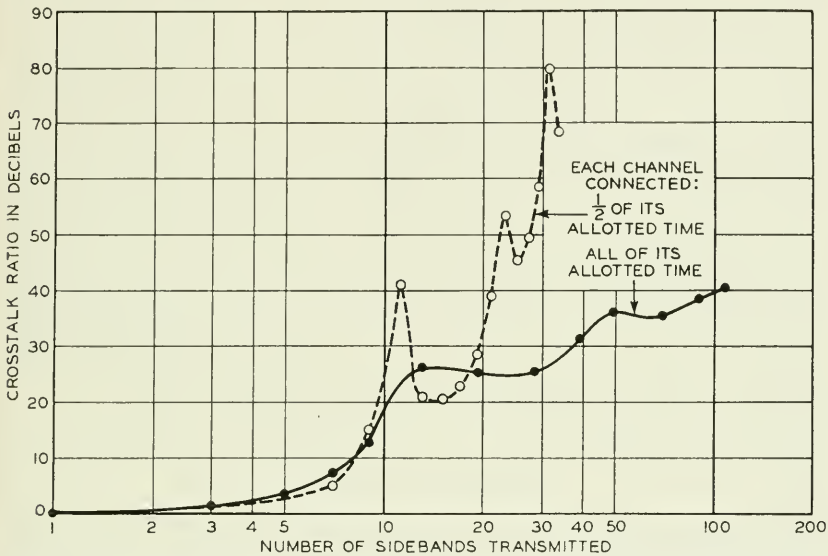


Fig. 2—Crosstalk between adjacent channels of ten-channel time division multiplex system with on-and-off switching. No attenuation or phase distortion in pass band of line

uniform attenuation in its pass-band all other sidebands are removed from the line. The expression (23) then becomes:

$$\frac{Y_{jk}}{Y_{jj}} = \frac{1 + 2\sum_{m=1}^{M} \cos 2m\pi(k - j)/N}{1 + 2M} = \begin{cases} 1, & k = j \\ \dfrac{\sin (2M + 1)\pi(k - j)/N}{(2M + 1) \sin \pi(k - j)/N}, & k \neq j \end{cases} \tag{24}$$

In particular, if

it follows that
$$\left.\begin{aligned} M &= (N-1)/2, \\ \\ Y_{jk}/Y_{jj} &= 0, \; k \neq j \end{aligned}\right\} \tag{25}$$

Thus there exists in theory a system employing sidebands on zero frequency and the first $(N - 1)/2$ harmonics of the switching frequency, in which multichannel transmission is possible without interchannel interference. Since the required condition (25) may also be written $N = 2M + 1$, an odd number of channels is obtained. Since N sidebands are transmitted, the band width used is the same as the minimum required for N single sideband amplitude modulation channels on a frequency discrimination basis. Sidebands produced on higher harmonics in the time division process must be removed by filtering.

It is to be noted that since it is equality of the N sideband contributions which is important and the amount of each contribution is determined by the transmission characteristic of the line as well as the transmitting and receiving switching processes, it would be theoretically possible to make up for sideband irregularities by equalizing the line. However, the equalization required in the line would be of "stairstep" type rather than smoothly varying with frequency since an error in the value of one harmonic of the switching function produces the same error throughout the entire range occupied by the pair of sidebands associated with that harmonic.

General Switching Functions with Crosstalk Suppression and Minimum Band Width

The above discussion based on the properties of a commutator has led us to an ideal switching function which is, except for an unimportant proportionality factor,

$$F_j(t) = 1 + 2 \sum_{m=1}^{(N-1)/2} \cos m[qt - (j-1)2\pi/N], \; N \text{ odd} \tag{26}$$

This type of switching is approximately realizable with synchronized commutators having contact widths very narrow in comparison with the spacing between contacts. For a 3000-cycle speech band, the minimum switching rate would be 6000 cycles per second. Such a speed would be difficult to obtain with ordinary mechanical means but would be feasible with rotating electron beams.

The concept of combining detected contributions from a number of sidebands in proper phase to give in-phase addition of desired components and cancellation of unwanted ones leads to a generalization of the switching processes over those possible with synchronized commutators. We note that the switching functions $F_j(t)$ of (2) and $G_k(t)$ of (9) are analogous to carrier waves applied to a product modulator, and an electrical analogue of time division may be realized therefore by applying signal and a suitable carrier to a product modulator. Phase shifts in the carrier supply circuit may be made to serve the same purpose as the angular displacements between commutator segments. It is thus of interest to examine various other possible forms of the function $F_j(t)$ which are suitable for multiplex transmission and investigate methods by which they can be realized.

We note that (26) is suitable for an odd number of channels because it makes use of the direct signal component (or sideband on zero frequency) in addition to the paired sidebands on harmonics of the switching frequency. It seems reasonable to expect that systems for even numbers of channels can be devised using only upper and lower sidebands on harmonics and omitting the signal itself. Complete information for the separation of N channels should be contained in any set of N sidebands; hence we should not be forced to start with the sidebands of lowest frequency, but be able to use other sets with a more suitable place in the spectrum or with better equalization of amplitudes.

We shall derive an expression for a quite general switching function meeting the desired conditions of freedom from crosstalk and economy of band width for an even number of channels by assuming the following forms for A_{mj} and θ_{mj} in (3),

$$A_{mj} = \begin{pmatrix} A, \; n \leqq m \leqq n + N/2 - 1 \\ 0, \; m < n, \text{ or } m > n + N/2 - 1 \end{pmatrix} \tag{27}$$

$$\theta_{mj} = (j - 1)\,(m + h)\,\psi + \alpha \tag{28}$$

The switching function assumed contains $N/2$ harmonics and hence will produce N sidebands. The values of n, h, α and ψ are first assumed to be arbitrary. At the receiving end, a switching function similar except for a time displacement t_0 will be assumed. That is, in (9), we take

$$B_{mk} = \begin{pmatrix} B, \; n \leqq m \leqq n + N/2 - 1 \\ 0, \; m < n \text{ or } m > n + N/2 - 1 \end{pmatrix} \tag{29}$$

$$\Phi_{mk} = (k - 1)\,(m + h)\psi + (mq + \nu)\tau + \alpha \tag{30}$$

Transmission over the line is assumed to be of the distortionless form obtained by setting $Z_0(i\omega) = Z_0$, a constant, in (20). Substituting (27)–(30) in (14), we then calculate

$$Y_{jk} = \frac{NABZ_0e^{-it_0\omega_j}}{4} \begin{cases} 1, & j = k \\ \dfrac{2 \sin N(j - k)\psi/4 \cos (4n + 4h + N - 2)(j - k)\psi/4}{N \sin (j - k)\psi/2}, & j \neq k \end{cases} \tag{31}$$

Freedom from interchannel interference is obtained if the numerator vanishes and the denominator does not vanish for all unequal integer values of j and k less than N. We note that this can be accomplished by setting

$$N\psi = (4n + 4h + N - 2)\psi = 2\pi, \tag{32}$$

in which case the numerator becomes $2 \sin (j - k)\pi/2 \cos (j - k)\pi/2 = \sin (j - k)\pi = 0$. Solving (32) for h and ψ, we find

$$\left.\begin{aligned} h &= -(2n - 1)/2 \\ \psi &= 2\pi/N \end{aligned}\right\} \tag{33}$$

The denominator thus becomes $N \sin (j - k)\pi/N$, and since the largest possible value of $j - k$ is $N - 1$, the denominator cannot vanish for $j \neq k$. The value of n remains arbitrary; hence we may start with any harmonic we please. The value of α is also immaterial. The general form of switching function thus derived is, omitting the constant of proportionality:

$$F_j(t) = \sum_{m=n}^{n+N/2-1} \cos [(\nu + mq)t - (j - 1)(2m - 2n + 1)\pi/N + \alpha], \tag{34}$$
$$n = 0, 1, 2, \cdots, N \text{ even}$$

Dependence of the phase angle upon the initial harmonic may be avoided in certain special cases; for example, if we set $n = rN/2$, where r may be zero or any positive integer, we obtain the result that a particular switching function which satisfies the required conditions is:

$$F_j(t) = \sum_{m=rN/2}^{(r+1)N/2-1} \cos [(\nu + mq)t - (j - 1)(2m + 1)\pi/N + \alpha], \tag{35}$$
$$r = 0, 1, 2, \cdots, N \text{ even}$$

Specific methods of realizing satisfactory switching functions will now be examined.

Plus-and-Minus Switching with Commutator

Instead of opening and closing the circuit between each individual channel and the line, the commutator may be designed to reverse the polarity of the connections on alternate contacts with a given channel. An appropriate switching function for alternate polarity reversal is:

$$F_1(t) = \begin{Bmatrix} 0, -(2N - x)T/2N < t < -xT/2N \\ A, -xT/2N < t < xT/2N \\ 0, xT/2N < t < (2N - x)T/2N \\ -A, (2N - x)T/2N < t < (2N + x)T/2N \end{Bmatrix} \tag{36}$$

The relation of $F_j(t)$ to $F_1(t)$ is as specified by (15). By the usual Fourier series expansion, we then find

$$F_1(t) = \sum_{m=0}^{\infty} \frac{4A}{(2m+1)\pi} \sin \frac{(2m+1)x\pi}{2N} \cos (2m+1)qt/2 \tag{37}$$

where $q = 2\pi/T$. In the limiting case in which the duration of contact time is made small compared with the interval between contacts, the expression (37) combined with (15) gives the general result:

$$F_j(t) = \frac{2Ax}{N} \sum_{m=0}^{m} \cos \left[\frac{(2m+1)qt}{2} - \frac{(2m+1)(j-1)\pi}{N} \right] \tag{38}$$

By comparing this function with (35), setting $\nu = q/2$, we note that plus-and-minus switching with small contact time can be made to give freedom from crosstalk and economy of band width in a. time division multiplex system provided that we select the sidebands on $N/2$ adjacent odd harmonics of $q/2$, *beginning with* $m = rN/2$, i.e., *the* $(rN + 1)^{st}$ *harmonic of* $q/2$, *where* r *may be zero or any positive integer.* It is necessary that even harmonics of $q/2$ be suppressed, which will be the case if the positive and negative switching contacts are identical except for change in sign.

Vestigial Sideband Systems

It is seen that the use of a commutator results in the production of a large number of sidebands, and that for economy of bandwidth in the transmission medium it is desirable to select only a definite group by filtering. The requirements on the filter are not easily met since there should be linear phase shift and uniform attenuation throughout the range of frequencies occupied by the desired sidebands and a high degree of suppression of all other frequencies. The requirements may be simplified somewhat by using the principle of vestigial sideband transmission which allows the attenuation to increase gradually within the desired band and compensates for energy removed from wanted sidebands by allowing a corresponding amount of energy to be transmitted from adjacent sidebands which would normally be suppressed. Linear phase shift is required for all frequencies at which appreciable transmission exists.

Figure 3 shows the curve of admittance vs. frequency required for on-and-off and plus-and-minus switching. Referred to the mid-band admittance as unity, the admittance is reduced to one-half (six *db* loss) at the frequencies $rNq/2$ and $(r + 2)Nq/2$ which are the nominal upper and lower cutoffs. N is an even integer and r is zero or any positive integer. The admittance curve has odd-symmetry about the cutoff frequencies—that is, if at a frequency x cycles below a cutoff frequency, the admittance has the value a, it must be $1 - a$ at a frequency x cycles above the cutoff. The nominal

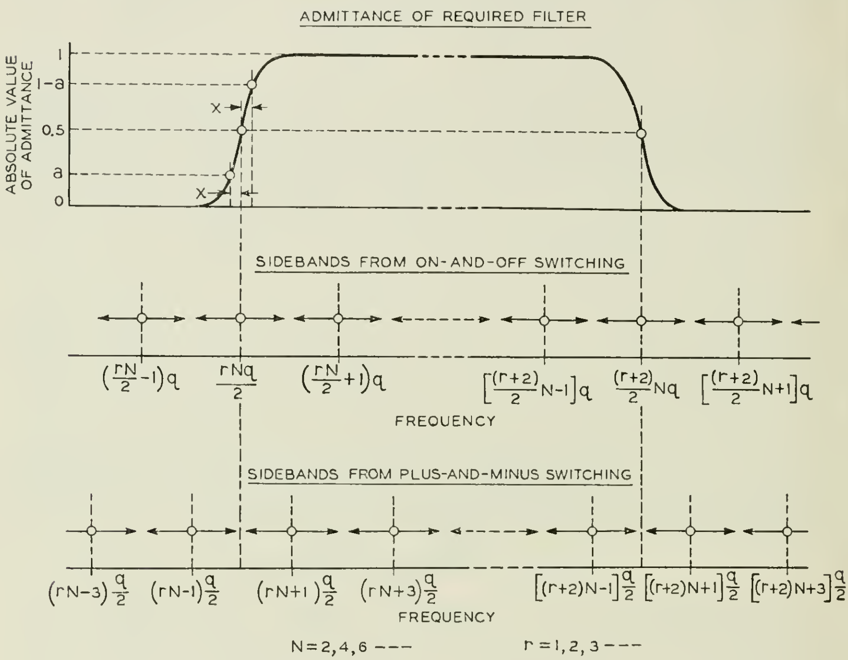


Fig. 3—Vestigial sideband transmission in time division multiplex systems

band transmitted in the case of on-and-off switching consists of the upper sideband on the harmonic $rNq/2$, the lower sideband on the harmonic $(r + 2)Nq/2$, and all intervening sidebands. In the case of plus-and-minus switching the upper and lower sidebands on frequencies $(rN + 1)q/2$ to $[(r + 2)N - 1]q/2$ inclusive are transmitted. Impairment of the nominal band by the filter is made up by transmitting the appropriate parts of sidebands outside the nominal range. It is easily verified that either of the systems depicted in Fig. 3 satisfies the required conditions for multiplex transmission without interchannel interference when the sidebands produced by a given signal have equal amplitudes over the range utilized.

The vestigial method is required only when strong sidebands very near the desired ones must be removed. Modifications of the time division process exist in which vestigial filters are unnecessary because very little energy appears at unwanted side frequencies. It is clear that if we regard the problem as one of producing certain sidebands on carriers of definite phases, we are not restricted to commutating devices only but may make use of general modulator technique. Further details concerning specific circuit arrangements are described in *U. S. Patent* 2,213,938, W. R. Bennett; and *U. S. Patent* 2,213,941, E. Peterson. As a general guide the following table of carrier phases for an N-channel system (N even) is furnished:

Table of Phase Shifts for N-Channel System

($N/2$ Carrier Frequencies Required)

Carrier Frequency		ν	$\nu + q$	$\nu + 2q$	$\nu + 3q$	$\cdots$
Phase Shift In Carrier	Channel 1	0	0	0	0	$\cdots$
	2	π/N	$3\pi/N$	$5\pi/N$	$7\pi/N$	$\cdots$
	3	$2\pi/N$	$6\pi/N$	$10\pi/N$	$14\pi/N$	$\cdots$
	4	$3\pi/N$	$9\pi/N$	$15\pi/N$	$21\pi/N$	$\cdots$

Transmission Requirements

Practical success of a time division multiplex system requires the maintenance of a satisfactory ratio of wanted signal to crosstalk. In order to accomplish this, the transmission link must preserve the amplitude and phase relations of a group of sidebands. A physical picture of the relations involved may be obtained from Fig. 4, which is drawn for the particular case of a 5-channel system of the on-and-off switching type. For this example the theory previously developed shows that five sidebands of equal amplitude are sufficient, namely—the signal itself (which may be regarded as a sideband on a carrier of zero frequency), the upper and lower sidebands on the switching frequency and on the second harmonic of the switching frequency. If we take the phases of the switching fundamental and its second harmonic as applied to the first channel as a reference, the proper phases of fundamental and second harmonic respectively for the other four channels are given by the following table:

Channel Number	Fundamental Phase	Second Harmonic Phase
2	72°	144°
3	144°	288°
4	216°	432°
5	288°	576°

In Fig. 4, we have assumed a single-frequency signal component as input to the first channel. If the line has distortionless attenuation and phase characteristics, the five resulting side frequencies are received in the first channel as the five in-phase vectors of equal amplitude shown in (a). Reception in the second channel is shown by (*b*) in which the vector 1 represents the directly transmitted signal component (or sideband on *d-c*), 2 and 3

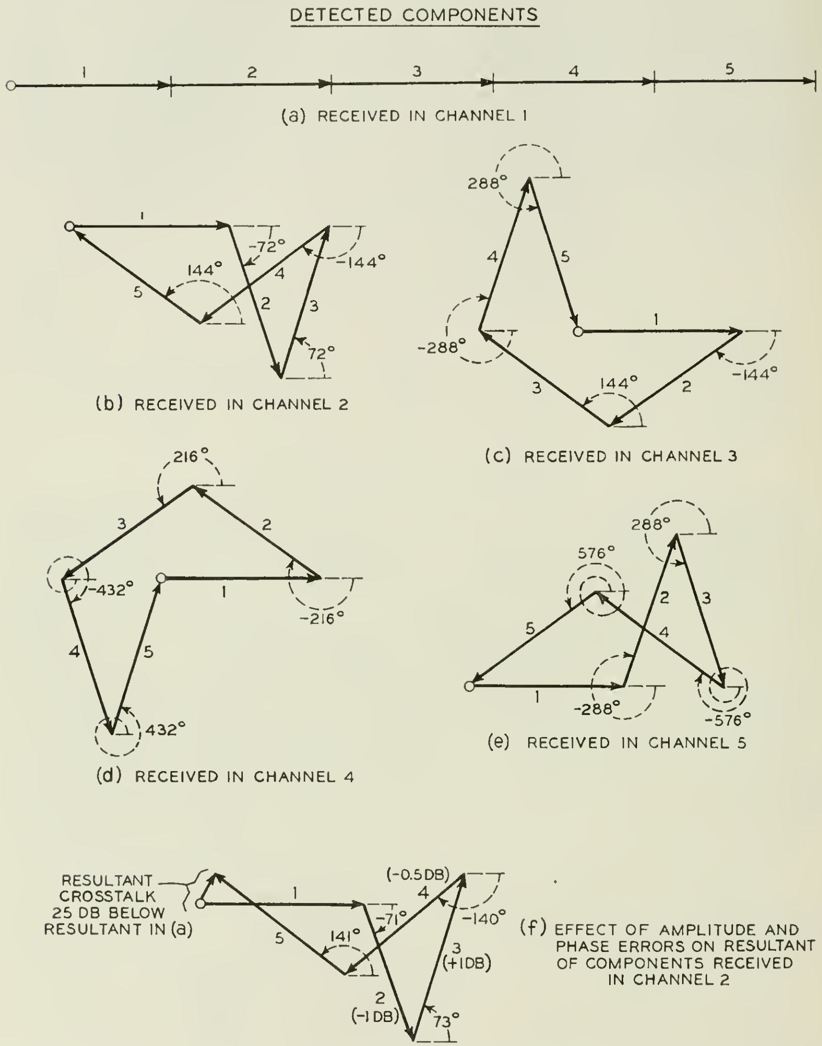


Fig. 4—Graphical representation of operation of time division multiplex system. Signal transmitted in channel 1 of 5-channel 5-sideband system

represent the detected components from the upper and lower side frequencies associated with the fundamental switching frequency, and 4 and 5 the components resulting from the upper and lower side frequencies of the second harmonic. Components 2 and 3 are shifted $-72°$ and $+72°$ respectively and 4 and 5 are shifted $-144°$ and $+144°$ in relation to the phase of component 1. As shown in (*b*), the five vectors combine in the form of a closed polygon giving a resultant of zero amplitude. Similar vector diagrams for reception in the third, fourth, and fifth channels are shown in (*c*), (*d*), and (*e*). The appropriate diagrams for transmission in channels 2, 3, 4, and 5 and receiving in any channel can be obtained from (*a*) – (*e*) by cyclic permutation of the channel numbers, i.e., transmission in 1 and reception in 2 corresponds to transmission in 2 and reception in 3, etc.

Production of crosstalk by phase and amplitude distortion in the transmission medium is illustrated by (*f*), Fig. (4), which shows the resultant component received in channel 2 when signal is transmitted in channel 1 and an imperfect line is used to connect the transmitting and receiving terminals. The vector 1 is taken as the reference amplitude and phase. The gain characteristic of the line is assumed to be one *db* low at the side frequency producing vector 2, one *db* too high for vector 3, 0.5 *db* low for vector 4, and with no error for vector 5. The phase curve is assumed to depart from a straight line by $-1°$, $-1°$, $-4°$, $+3°$ at the side frequencies from which components 2, 3, 4, 5 respectively are derived. The vector polygon fails to close and the resultant represents an unwanted signal received in channel 2 at a level 25 *db* below the wanted signal received in channel 1.

We may make an estimate of the accuracy of the equalization required in the general case by writing the transfer impedance $Z(i\omega)$ in the form:

$$Z(i\omega) = \rho(\omega) Z_0 e^{-it_0\omega - i\beta(\omega)} \tag{39}$$

where $\beta(\omega)$ represents the departure of the phase shift from a straight line and the variation from flat gain is given by

$$g(\omega) = 20 \log_{10} \rho(\omega) \tag{40}$$

The expression (39) may be rewritten as:

$$Z(i\omega) = [1 + z(i\omega)] Z_0 e^{-it_0\omega}, \tag{41}$$

where

$$z(i\omega) = \rho(\omega) e^{-i\beta(\omega)} - 1 \tag{42}$$

If we assume that the switching function is of the general form (34), we calculate from (14) the general relation:

$$Y_{jk} = K \begin{cases} N + \sum_{m=n}^{n+\frac{N}{2}-1} (z[i(\nu + mq + \omega_j)] + z^*[i(\nu + mq - \omega_j)]), \ j = k \\ \sum_{m=n}^{n+\frac{N}{2}-1} (z[i(\nu + mq + \omega_j)]e^{-i(j-k)(2m-2n+1)\pi/N} \\ \qquad + z^*[i(\nu + mq - \omega_j)]e^{i(j-k)(2m-2n+1)\pi/N}), \ j \neq k \end{cases} \tag{43}$$

where K is a complex constant of proportionality. The case of $j = k$ which gives transmission within the channel contains a variation with signal frequency caused by the summation of the departures from ideal transmission at the N sideband frequencies. This term presumably will be unimportant if the transmission characteristic is sufficiently good to meet crosstalk requirements; hence we may neglect the z and z^* terms in the case of $j = k$ and write the ratio of interference to desired signal as:

$$C_{jk} = \frac{Y_{jk}}{Y_{jj}} = \frac{1}{N} \sum_{m=n}^{n+\frac{N}{2}-1} (z[i(\nu + mq + \omega_j)]e^{-i(j-k)(2m-2n+1)\pi/N} + z^*[i(\nu + mq - \omega_j)]e^{i(j-k)(2m-2n+1)\pi/N}) \tag{44}$$

The crosstalk ratio will in general vary with the signal frequency. The requirement would logically be based on the total interference power weighted in accordance with the interfering effect at individual frequencies. Thus we might set

$$X_{jk} = S_j \int_{\omega_a}^{\omega_b} W_{jk}(\omega_j) \mid C_{jk} \mid^2 d\omega_j, \tag{45}$$

where X_{jk} represents the weighted interference power received in the k^{th} channel when a reference signal wave of mean power S_j is applied to the j^{th} channel. The limits of integration ω_a and ω_b are the lowest and highest signal frequencies used. The function $W_{jk}(\omega_j)$ represents the proper weighting with frequency of the interference and takes into account the distribution of the interfering signal and the relative importance of the different interfering frequencies.

Equation (45) is sufficient for computation of interchannel interference introduced by the line when the transmission characteristics of the line are known. A more valuable result, however, would be the expression of the required line characteristics in terms of the allowable interference. In general this would require some specification of the nature of the departures from the ideal characteristic. Except perhaps for systems with very few channels, it seems reasonable to assume that the departures are distributed fairly uniformly throughout the frequency range transmitted by the line, and hence that for purposes of estimating requirements we may replace $| C_{jk} |^2$ in (45) by its average value over the band. We may then write (45) in the form:

$$\overline{| C_{jk} |^2} = \frac{X_{jk}}{S_j \int_{\omega_a}^{\omega_b} W_{jk}(\omega_j)\, d\omega_j} \equiv U_{jk} \tag{46}$$

The value of the right-hand member, which we have designated by the symbol U_{jk}, is either known or can be determined for the particular type of signal. Hence our problem is reduced to finding the allowable departures in transmission which keep the mean square absolute value of C_{jk} from exceeding a prescribed maximum value.

We note that C_{jk} is the sum of N complex quantities, each of which is restricted to a range of values determined by the transfer impedance of the line in an individual band of frequencies. A convenient simplification may be made by regarding the N complex quantities as N independent chance variables. This is tantamount to assuming that departures in one band do not affect departures in any other band; the assumption is not strictly true, but should lead to no important error. We may then make use of the following theorem[8]: If

$$\zeta = b_1 z_1 + b_2 z_2 + \cdots + b_n z_n, \tag{47}$$

where $z_1, z_2, \cdots z_n$ are n independent complex chance variables and $b_1, b_2, \cdots b_n$ are complex constants,

$$\overline{| \zeta |^2} = \overline{| b_1 |^2}\, \overline{| z_1 |^2} + \cdots + \overline{| b_n |^2}\, \overline{| z_n |^2} \tag{48}$$

Application of this theorem to (44) gives

$$\begin{aligned} \overline{| C_{jk} |^2} &= \frac{1}{N^2} \sum_{m=n}^{n+\frac{N}{2}-1} \left(\overline{| z[i(\nu + mq + \omega_j)] |^2} + \overline{| z^*[i(\nu + mq - \omega_j)] |^2} \right) \\ &= \overline{| z |^2}/N, \end{aligned} \tag{49}$$

if the average square of the absolute value of the departure is the same in all bands and is equal to $\overline{| z |^2}$, which we shall define as the average squared absolute value of the departure for the entire line band used.

From (42) and (40),

$$\begin{aligned} | z(i\omega) |^2 &= 1 - 2\rho(\omega) \cos \theta(\omega) + \rho^2(\omega) \\ &= 1 - 2 \cdot 10^{g(\omega)/20} \cos \theta(\omega) + 10^{g(\omega)/10} \end{aligned} \tag{50}$$

[8] R. S. Hoyt, *B. S. T. J.*, Vol. XII, No. 1, Jan. 1933, p. 64.

Since it seems certain that g and θ must remain small to make the system operative, we investigate the nature of (50) when expanded in powers of g and θ. The leading terms are:

$$|z(i\omega)|^2 = \frac{(\log_e 10)^2}{400} g^2(\omega) + \theta^2(\omega) + \cdots \tag{51}$$

Hence for g and θ small, we have independent of any correlation which may exist between g and θ,[9]

$$\overline{|z(i\omega)|^2} = \left(\frac{\log_e 10}{20}\right)^2 \overline{g^2(\omega)} + \overline{\theta^2(\omega)} \tag{52}$$

Let

$$\sigma_1^2 = \overline{g^2(\omega)}, \qquad \sigma_2^2 = \overline{\theta^2(\omega)} \tag{53}$$

Then from (46), (49), (52),

$$U_{jk} = \left[\left(\frac{\log_e 10}{20}\right)^2 \sigma_1^2 + \sigma_2^2\right] \Big/ N \tag{54}$$

In (54) σ_1 is the r.m.s. departure of the gain in *db* from a constant and σ_2 is the r.m.s. departure of the phase shift in radians from a straight line. If σ_2 is expressed in degress instead of radians, (54) becomes

$$U_{jk} = 10^{-3}\,(13.25\,\sigma_1^2 + .3046\,\sigma_2^2)/N \tag{55}$$

The total interference received in any one channel is the sum of the individual contributions from the other $N - 1$ channels. The addition factor required to express the total in terms of the interference from one channel depends on the nature of the individual loads. Thus if the probability that any one channel is transmitting a signal wave is τ, the average total interference power received in one channel is

$$X = \tau(N - 1)X_{jk} = US_j \int_{\omega_a}^{\omega_b} W_{jk}(\omega_j)\, d\omega_j, \tag{56}$$

where

$$U = (N - 1)\tau U_{jk} = \frac{(N - 1)\tau}{N} 10^{-3}(13.25\sigma_1^2 + .3046\sigma_2^2) \tag{57}$$

For large values of N, the ratio $(N - 1)/N$ approaches unity, and the average interference becomes independent of the number of channels. The average interference may not be the most significant quantity, however. For example, if there is a considerable probability that all channels are

[9] This method of avoiding any assumption concerning correlation of attenuation and phase was suggested by Dr. T. C. Fry.

carrying energy simultaneously, as would be the case if the channels were subdivisions of a common signal band, the peak value of interference would probably be of more significance than the average value.

It is convenient to let

$$H = 10 \log_{10} \frac{S_j}{X} \tag{58}$$

$$F = -10 \log_{10} \int_{\omega_a}^{\omega_b} W_{jk}(\omega_j)\, d\omega_j \tag{59}$$

H is the ratio expressed in *db* of mean signal power in one channel to the total interference power received in one channel, and F is the weighting factor expressed in *db*. From (56),

$$U = 10^{-(H-F)/10} \tag{60}$$

Equation (57) may be written in the form,

$$\frac{\sigma_1^2}{a^2} + \frac{\sigma_2^2}{b^2} = 1, \tag{61}$$

where

$$\left.\begin{aligned} a &= 8.69 \sqrt{\frac{NU}{(N-1)\tau}}\ db \\ b &= 57.3 \sqrt{\frac{NU}{(N-1)\tau}}\ \text{degrees} \end{aligned}\right\} \tag{62}$$

Without the numerical factors, a and b are expressed in nepers and radians respectively.

If we regard σ_1 and σ_2 as variables, (61) determines a family of ellipses in which a and b are the semi-axes. By assigning values to N, τ, and $H - F$ we may thus represent the requirements on gain and phase variation by elliptical boundaries in the $\sigma_1\sigma_2$-plane. Figure 5 shows such a diagram constructed for a large number of channels each active one-fourth of the time and with flat weighting. In terms of the symbols above, we have set $N/(N-1)$ equal to unity, $\tau = 1/4$, and $F = 0$. Gain and phase variations included within a particular ellipse produce average interference power less than the amount designated on the boundary in terms of *db* down on mean power in one channel. The requirements imposed on both gain and phase variation are considerably more stringent than the corresponding requirements for carrier systems using frequency discrimination and employing comparable band widths.

Requirements on linear transmission of the line are, of course, not the only considerations involved in a comparison of time division multiplex

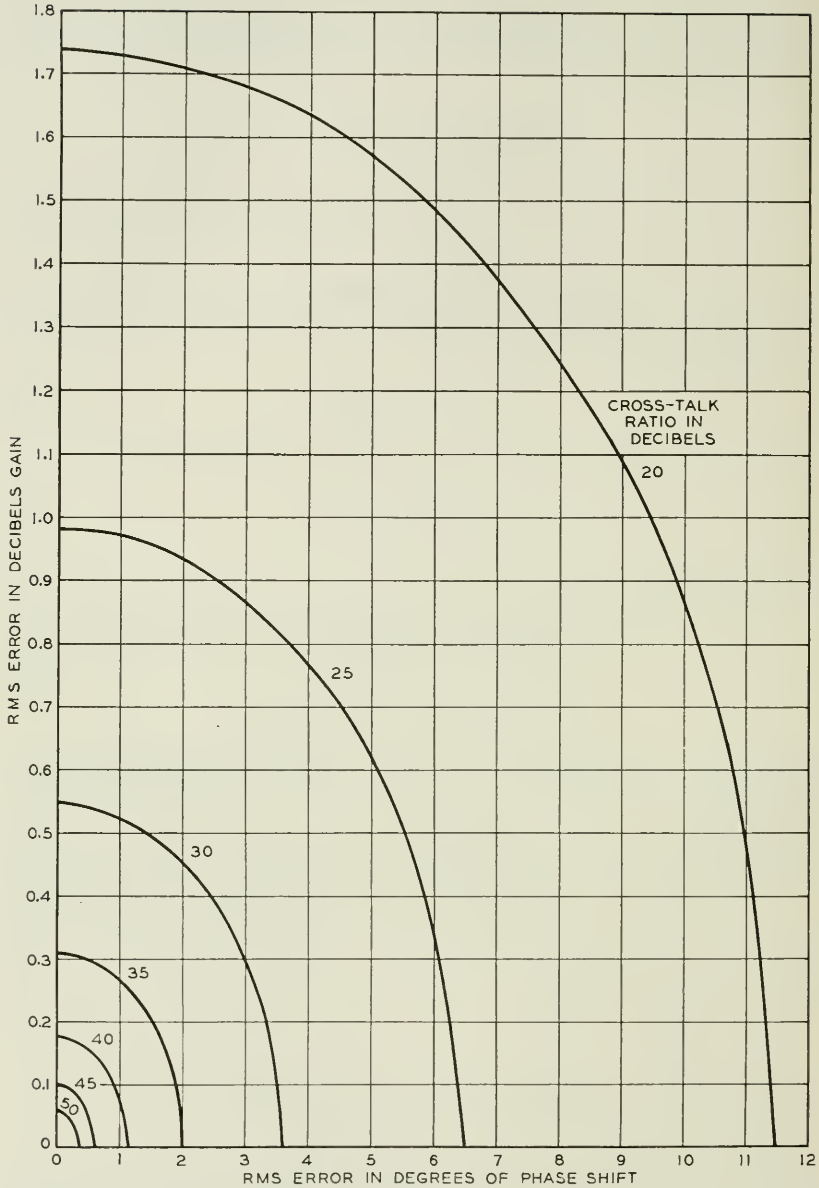


FIG. 5—Gain and phase requirements for transmission of time division multiplex signals. Each channel active 25% of time

methods with competing methods of superimposing channels. Other aspects to be considered are the synchronization of transmitting and receiving switching processes, the effects of non-linearities in the line, and the sensitivity of the system to external interference. It is thought, however, that the severe restrictions imposed on phase and attenuation characteristics when economy of band width is required form the weakest feature of the method and will in many cases provide the primary criterion for judging its availability in the solution of particular problems. Conversely, if the crosstalk requirements of the system are sufficiently mild to enable the transmission problem to be solved, the other problems also become relatively simplified.

Further References

A. M. Nicolson, *U. S. Patents,* 1,962,610; 2,021,743; 1,951,524; 1,956,397; 2,007,809.

Callahan, Mathes, and Kahn, "Time Division Multiplex in Radio-Telegraphic Practice," *Proc. I. R. E.*, 26 (1938), 1, pp. 55–74.

H. Raabe, "Untersuchungen an der wechselzeitigen Mehrfachübertragung (Multiplexübertragung)," *E. N. T.*, 16 (1939), 8, pp. 213–228.

Steady State Delay as Related to Aperiodic Signals

By R. V. L. HARTLEY

The concepts of phase and envelope delay, as applied to any linear system, rather than only to a medium, are discussed. Criteria are set up for the time of occurrence of that part of an aperiodic signal which corresponds to a small segment of the spectrum. The original spectrum of the signal gives the time of entry and this spectrum as modified by the phase characteristic of the system gives the time of exit.

If the amplitude is constant over the segment, it is shown that when the criterion is the time of maximum envelope of the disturbance, the aperiodic delay is identical with the envelope delay. When it is the time of maximum absolute value, the delay depends on the signal spectrum, the phase shift of the system, and the envelope delay, but not on the phase delay.

If the amplitude vaies rapidly with frequency, the component of an aperiodic disturbance which corresponds to a narrow segment of the spectrum persists so long that the resulting over-lapping of neighboring segments makes their interpretation difficult.

IN THE earlier applications of steady state theory to transmission problems the emphasis was placed on the variation of amplitude with frequency. The use of long loaded lines made it necessary to take account of phase distortion[1] as well. With the development of telephotography and television[2], the phase characteristic was found to provide a useful index for predicting the overlapping of adjacent picture elements. For these purposes it has been found convenient to express the phase characteristic in terms of phase or envelope delay. These may be called "steady state delays" since they are defined and measured in terms of sinusoidal disturbances of adjustable frequency. However, the signals for which they are intended to furnish an index are aperiodic in nature. It seemed worthwhile, therefore, to examine more closely the relations existing between "aperiodic delays," defined in terms of such signals, and steady state delays.

Let us first review the development of the concepts of steady state delay. Early in the study of the propagation of sinusoidal waves a distinction was made between phase and group velocity. If we fix on a particular distance of transmission the ratio of this distance to each of these two velocities may be interpreted as a delay associated with the transmission. In the

[1] For discussion and references see "Phase Distortion and Phase Distortion Correction," S. P. Mead, *B. S. T. J.*, Vol. VII, p. 195, 1928.

[2] Symposium on Television, *B. S. T. J.*, Vol. VI, p. 551.

communication art, these delays have been called phase and envelope delay, respectively. If the medium exhibits dispersion they vary with frequency. Let us fix our attention on the conditions throughout the medium at a particular instant during the transmission of a sinusoidal disturbance. We may determine the total change of phase in passing from the input to the output. This may be more than a single cycle. If now we divide this phase shift by the frequency, expressed in the same angular units, we get the time which will be required for the phase at the input to progress to the output, or the phase delay. Also it may readily be shown that the derivative of this phase shift with respect to frequency is equal to the envelope delay as defined above in terms of the group velocity. The simplest treatment of this is based on the consideration of two sinusoidal waves of equal amplitude and slightly different frequencies.

While these delays can be easily interpreted for most media, difficulties arise in the case of those substances which exhibit anomalous dispersion. Here, in the neighborhood of certain frequencies, the phase shift varies rapidly with frequency, and often appears to be discontinuous. Actually the apparent discontinuity is a region of very rapid decrease of phase with frequency, which leads to a negative value of envelope delay. In the same region the transmission varies rapidly with frequency, and selective reflection occurs at the entering boundary. This effect can be explained in terms of resonance in the elements which make up the fine structure of the medium.

The next step was to dissociate the idea of delay from that of velocity in a medium, and associate it with a steady state transfer characteristic between any two points of a linear system. This would permit its application to all sorts of complicated networks in which uniform propagation cannot be readily visualized. Here two types of characteristic are to be distinguished. One, which is associated with what might be called "damped" systems, exhibits a reasonably gradual variation of both phase shift and attenuation with frequency. This is the analog of a medium having normal dispersion. The other, which is associated with "resonant" systems, exhibits the phenomena associated with anomalous dispersion. In the case of filters and hollow wave guides these resonances give rise to regions of high attenuation and reactive impedance, which are the analogs of the regions of high absorption and selective reflection at the boundary of a medium. In applying the idea of delay to networks then, we can expect the results to agree with our intuitions only so long as we keep away from the critical frequencies of resonant systems.

In computing or measuring the phase shift of a system, at a single frequency, the result is indeterminate so far as the addition of multiples of 2π is concerned. This does not affect the envelope delay, which depends only on the derivative, and so this type of delay can be generalized directly to include the transfer characteristics of arbitrary networks. To give an exact meaning to phase delay some convention would have to be adopted for determining what, if any, multiple of 2π is to be added to the computed phase for the frequency in question. Apparently no such convention has been agreed upon which is of general application. For damped networks which transmit frequencies down to zero, it is customary to assume the phase shift to be zero at zero frequency, and, for higher frequencies, to add multiples of 2π so that the phase shift varies continuously with frequency. If, then, B is the computed phase shift, between $-\pi$ and π, we may represent the continuously varying phase shift by $B + 2m\pi$, where m is the number of discontinuities in B which have been eliminated in passing from zero to the frequency in question. The phase delay may then be defined as

$$D_p = \frac{B + 2m\pi}{\omega}. \tag{1}$$

Any similar convention for resonant systems would be less simple, and since, as will appear below, phase delay has little bearing on aperiodic signals, it seems unwise to attempt to formulate such a convention here.

In contrast with steady state delay, let us now examine the delay of an aperiodic signal. If the signal is transmitted without distortion the concept of delay of the signal as a whole is simple. If, because of distortion, the sent and received signals are different we may still agree upon some recognizable feature of each as determining its time of occurrence. If the distortion is considerable the delay may vary greatly with the distinguishing characteristic chosen. For example, if it depends on the behavior of components of high frequency the delay may be quite different from what it is if it depends on those of low. In the first case the result would be little affected if, before transmission, the signal were sent through a high-pass filter and, in the second, if it went through a low-pass filter. In each case we measure a delay associated with a disturbance which comprises only those Fourier components of the signal which occupy a particular limited range of frequency. We may carry this idea farther and make use of a very narrow band-pass filter. By varying the mid-frequency of this band we obtain a delay which is a function of frequency. Its value, at any frequency, is the delay, as defined by our convention, of a disturbance which corresponds to that part of the spectrum of the signal which is in the immediate neighborhood of the frequency in question. Our problem then is to find recognizable features of a disturbance of this kind such that, when they are used as criteria of delay, the result can be related directly to the phase or envelope delay as defined in terms of periodic disturbances.

Compared with the pair of equal sinusoids used in the derivation of envelope delay, this disturbance differs in that, in any finite range of frequency, there are an infinity of sinusoids, the amplitudes of which need not all be the same. For simplicity, we assume the actual filter to be replaced by an idealized one in which there is no distortion within the band and no transmission outside it. If the signal as a whole be represented by a Fourier integral, we may obtain the desired disturbance, for an angular frequency, ω_1, by integrating from $\omega_1 - \delta$ to $\omega_1 + \delta$. The disturbance may be represented by

$$f(t) = \text{real part of } M \int_{\omega_1-\delta}^{\omega_1+\delta} \exp\left[-\alpha + i(\omega t - \theta)\right] d\omega, \tag{2}$$

where M is a constant dependent on the magnitude of the signal and α and θ are functions of frequency and position which describe the spectrum of the signal at various points in the system.

The first step is to perform the indicated integration and express the resulting function of time in a convenient form. For this we let

$$\epsilon = \omega - \omega_1 .$$

Since we are interested only in small values of ϵ we may replace α by

$$\alpha = \alpha_1 + \alpha_1'\epsilon,$$

where α_1 and α_1' are the values of α and $\dfrac{\partial\alpha}{\partial\omega}$ at ω_1. Similarly,

$$\theta = \theta_1 + \theta_1'\epsilon.$$

We define an instant, T_e, by

$$T_e = \theta_1', \tag{3}$$

and a time, τ, by

$$\tau = t - T_e . \tag{4}$$

Substituting these in (2) and performing the integration, we get

$$f(t) = \text{real part of } 2M \exp\left[-\alpha_1 + i(\omega_1\tau - (\theta_1 - \omega_1\theta_1'))\right] \frac{\sinh(-\alpha_1' + i\tau)\delta}{(-\alpha_1' + i\tau)} .$$

If we introduce the angles,

$$\beta = \text{arc tan} \frac{-\alpha_1'}{\tau},$$

and

$$\gamma = \text{arc tan} \frac{\tanh(-\delta\alpha_1')}{\tan \delta\tau},$$

and take the real part, we get

$$f(t) = 2M \exp(-\alpha_1) \frac{[(\cosh \delta\alpha_1' \sin \delta\tau)^2 + (\sinh \delta\alpha_1' \cos \delta\tau)^2]^{\frac{1}{2}}}{(\alpha_1'^2 + \tau^2)^{\frac{1}{2}}}$$

$$\cos(\omega_1\tau - (\theta_1 - \omega_1\theta_1') + \beta - \gamma). \tag{5}$$

Let us consider first the extreme case where the spectrum of the signal is

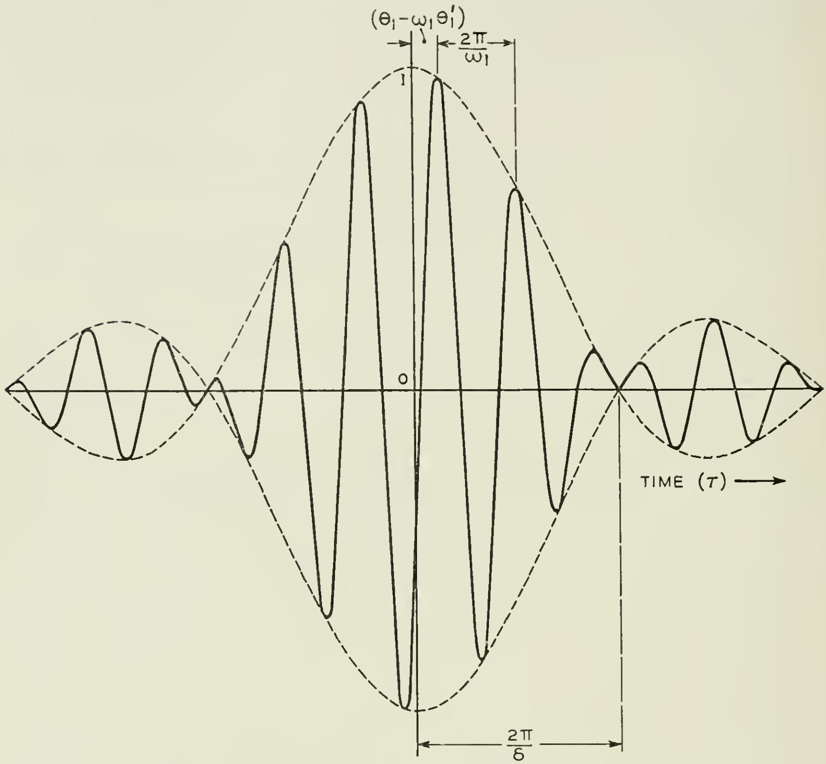


Fig. 1—Elementary disturbance corresponding to a narrow segment of the spectrum

uniform in amplitude in the neighborhood of ω_1, so that α_1' is zero. Then

$$f(t) = 2\delta M \exp(-\alpha_1) \frac{\sin \delta\tau}{\delta\tau} \cos(\omega_1\tau - (\theta_1 - \omega_1\theta_1')). \tag{6}$$

Here the amplitude includes a constant factor which is proportional to the bandwidth, 2δ, and to the magnitude, $M \exp(-\alpha_1)$, at the frequency, ω_1, and a function of time, a plot of which is shown in Fig. 1. This function consists of a sinusoidal wave of frequency, ω_1, the amplitude of which varies with time, the envelope being symmetrical about the instant, $T_e = \theta_1'$, at which it is a maximum. T_e, the time of maximum envelope, is then a unique instant which is suitable for defining the time at which the disturbance occurs. It is determined solely by the slope of the phase frequency curve for the spectrum.

The instant, T_e, may be interpreted, in accordance with the principle of stationary phase, as the one at which the sinusoidal components of (2) are most nearly in the same phase, and so have the least destructive inter-

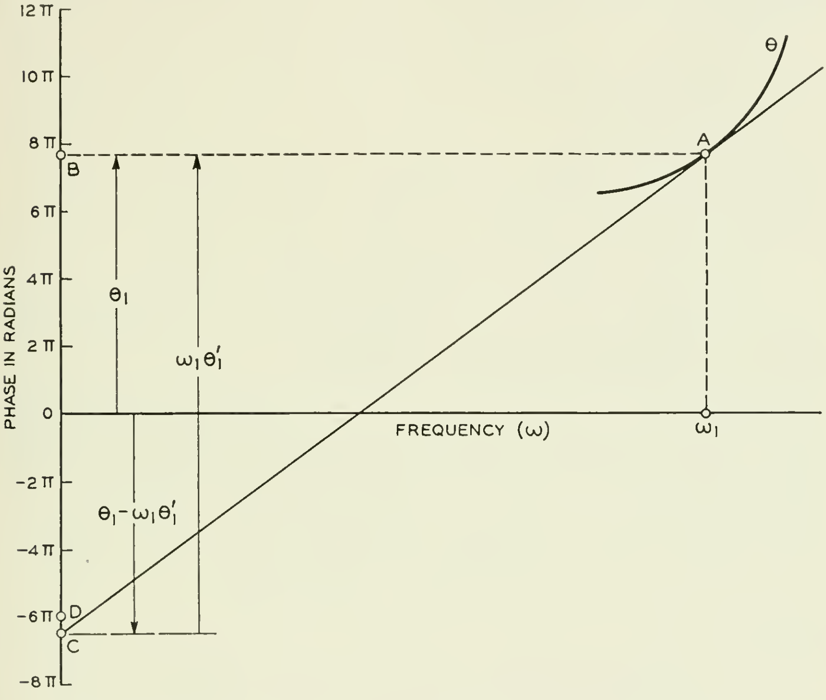


Fig. 2—Graphical representation of the phase of an elementary disturbance

ference. This condition will hold when the instantaneous phase angle is changing least rapidly with frequency, that is, when

$$\frac{\partial}{\partial \omega} (\omega t - \theta) = 0,$$

from which

$$t = \theta_1'.$$

The angle, $\theta_1 - \omega_1\theta_1'$, in (6), gives the phase of the wave at the instant, T_e, when its envelope is a maximum. The interpretation of this angle will be aided by the geometrical construction of Fig. 2 which is similar to that employed for phase and group velocity[3]. The abscissae are values of ω and the ordinates are values of phase in radians. A portion of the function, θ, in the neighborhood of ω_1 is shown. The distance, OB, is θ_1. The slope of the tangent, CA, to the curve at A is θ_1'. The distance, CB, is $\omega_1\theta_1'$. Consequently, OC, or the intercept of this tangent on the phase axis, is $\theta_1 - \omega_1\theta_1'$. If, as shown in the figure, the absolute value of this intercept is greater than π, we may transform (6) to a form in which the angle is less than π, by the substitution

$$\varphi = \theta_1 - \omega_1\theta_1' + 2n\pi, \tag{7}$$

where n is an integer and

$$|\varphi| < \pi.$$

In Fig. 2, n is 3, and φ is the distance DC. (6) then becomes

$$f(t) = 2\delta M \exp(-\alpha_1) \frac{\sin \delta\tau}{\delta\tau} \cos(\omega_1\tau - \varphi),$$

and φ is the ordinary phase lag of the sinusoid, relative to an origin of time given by the instant of maximum envelope.

We may choose as the instant at which the disturbance occurs, not T_e, at which the envelope is a maximum, but T_a, at which the instantaneous value of the function has its maximum absolute value. Since δ is small compared with ω_1, this will occur very nearly at the smallest absolute value of τ for which $\cos(\omega_1\tau - \varphi)$ is ± 1. This will occur for

$$\tau = \frac{\varphi}{\omega_1}, \quad \text{when} \quad -\frac{\pi}{2} < \varphi < \frac{\pi}{2},$$

and for

$$\tau = \frac{\varphi \pm \pi}{\omega_1} \quad \text{when} \quad -\pi < \varphi < -\frac{\pi}{2} \quad \text{or} \quad \frac{\pi}{2} < \varphi < \pi.$$

From (4), (3) and (7),

$$T_a = \frac{\theta_1 + k\pi}{\omega_1},$$

where k is an integer such that

$$-\frac{\pi}{2} < \Psi = \theta_1 - \omega_1\theta_1' + k\pi < \frac{\pi}{2}.$$

The significance of this can be seen from Fig. 3. Here, in addition to the θ curve of Fig. 2, there are plotted a series of curves whose ordinates differ

[3] Lamb, "Hydrodynamics," Cambridge U. Press 1916, p. 371.

from it by multiples of π. In so far as any one purely sinusoidal component of the disturbance is concerned, values of phase determined by those curves which differ by an even multiple of π would be indistinguishable. Those differing by an odd multiple would represent a reversal of sign. Let us

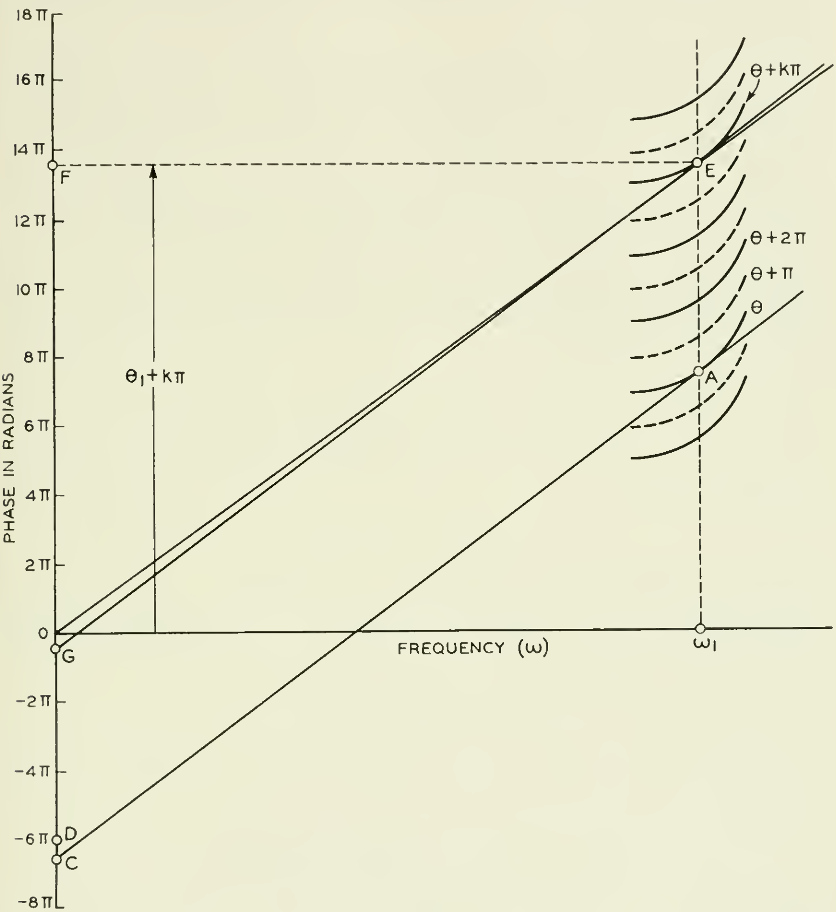


Fig. 3—Graphical representation of the time of maximum absolute value

now select that curve for which the tangent at ω_1 intersects the phase axis nearest the origin, and call it $\theta + k\pi$. Since, for the case drawn,

$$|DC| < \frac{\pi}{2},$$

$$k = 2n.$$

If it were greater we should have

$$k = 2n \pm 1.$$

It is then obvious that the time of maximum absolute value, T_a, is given by the slope of the line OE. It differs from T_e by the difference in slope of the lines OE and GE.

We have then deduced from the spectrum of the disturbance its time of occurrence in terms of two definitions of the latter. The next step is to compare these times for the input and output and determine the corresponding delays. Let us consider first the case where the attenuation is independent of frequency, so that α_1' is zero in the output signal also. We may then confine our attention to the phase, θ. Let us represent its value at the input by b, and the phase shift of the system by B. Then at the output θ will be equal to $b + B$. If we take the time of occurrence as determined by the maximum envelope, these times at the input and output are

$$T_{e_0} = b_1',$$

$$T_{e_1} = b_1' + B_1'.$$

The delay is then

$$D_e = T_{e_1} - T_{e_0} = B_1',$$

which is by definition the envelope delay of the system.

If we take the time of occurrence based on the maximum absolute value, we have, at the input,

$$T_{a_0} = \frac{b_1 + k_0\pi}{\omega_1},$$

where

$$-\frac{\pi}{2} < \Psi_0 = b_1 - \omega_1 b_1' + k_0\pi < \frac{\pi}{2}.$$

At the output,

$$T_{a_1} = \frac{b_1 + B_1 + (k_0 + k_3)\pi}{\omega_1},$$

where

$$-\frac{\pi}{2} < \Psi_3 = b_1 + B_1 - \omega_1(b_1' + B_1') + (k_0 + k_3)\pi < \frac{\pi}{2}.$$

The delay,

$$D_a = T_{a_1} - T_{a_0} = \frac{B_1 + k_3\pi}{\omega_1}.$$

While there is a superficial similarity between this and the phase delay (1), it is of little real significance; m, in (1), is determined by the aggregate increase in phase shift with frequency, while k, is determined mainly by the rate of increase at ω_1. An example of a situation in which the two delays are very different, is furnished by a wave guide when the frequency only just exceeds the cutoff. The phase delay is then almost zero while the rate of change of phase shift with frequency is very large.

Thus the delay based on maximum absolute value depends on both the envelope delay and the phase shift of the system, but not on the phase delay. There remains to examine this dependence in more detail. The value of k_3 depends on the spectrum of the signal as well as the characteristic of the system. It is of interest to see if it can be replaced by a quantity derived from the system characteristic alone. The most obvious thing to try is a delay which is derived from the phase shift of the system in the same way that the time of absolute maximum is derived from that of the signal spectrum. This would be

$$D_S = \frac{B_1 + k_2\pi}{\omega_1},$$

where

$$-\frac{\pi}{2} < \Psi_2 = B_1 - \omega_1 B_1' + k_2\pi < \frac{\pi}{2}.$$

The difference between this and the aperiodic delay based on absolute value is

$$D_S - D_A = \frac{\pi}{\omega_1}(k_2 - k_3),$$

$$= \frac{1}{\omega_1}(\Psi_0 + \Psi_2 - \Psi_3).$$

Since $k_2 - k_3$ is either zero or an integer and $|\Psi_3|$ is less than $\frac{\pi}{2}$, if

$$-\frac{\pi}{2} < \Psi_0 + \Psi_2 < \frac{\pi}{2},$$

$$D_S - D_A = 0.$$

If

$$-\pi < \Psi_0 + \Psi_2 < -\frac{\pi}{2},$$

$$D_S - D_A = -\frac{\pi}{\omega_1}.$$

If

$$\frac{\pi}{2} < \Psi_0 + \Psi_2 < \pi,$$

$$D_S - D_A = \frac{\pi}{\omega_1}.$$

Thus the delay as derived from the system characteristic alone may be identical with the aperiodic delay based on maximum absolute value or it may differ from it by $\pm\frac{\pi}{\omega_1}$, that is by half a period. Which condition holds depends on the interrelation of the phase functions which characterize the signal spectrum at the input and the transmission of the system, and not on either of these functions alone.

If the attenuation is not uniform, α_1' cannot be neglected and the expression for the output signal becomes more complicated. Both the amplitude and phase in (5) vary with time in a manner which depends on the value chosen for δ. The expression becomes fairly simple, however, for the case where α_1' is very large, as in anomalous dispersion and in highly resonant systems. Then, even when δ is small, we may assume that

$$\cosh(\delta\alpha_1') = \exp(\pm\,\delta\alpha_1'),$$

$$\sinh(\delta\alpha_1') = \pm\exp(\pm\,\delta\alpha_1'),$$

according as $\alpha_1' \gtrless 0$.

The amplitude factor in (5) then becomes

$$\frac{M\exp(-\alpha_1 \pm \delta\alpha_1')}{(\alpha_1'^2 + \tau^2)^{\frac{1}{2}}}.$$

Here the exponent is equal to the value of α at that edge of the segment of the spectrum where the amplitude is greatest. The amplitude is symmetrical about $\tau = 0$, that is, about $t = \theta_1'$, at which point it has its maximum value. Hence the instant of maximum envelope is still given by the slope of the phase, frequency curve, as when α_1' is small. However, the maximum is now extremely flat and its sharpness no longer depends directly on δ. Over the range of values of τ for which $\tau^2 << \alpha_1'^2$, the amplitude is sensibly constant. When $\tau = \pm\alpha_1'$, it is reduced to $\frac{1}{\sqrt{2}}$ times its maximum. For $\tau^2 >> \alpha_1'^2$, it varies inversely as $|\tau|$.

To investigate the oscillating factor of (5) we note that now

$$\gamma = \pm\delta\tau \pm \frac{\pi}{2},$$

where the sign of $\delta\tau$ depends on that of α_1' and that of $\frac{\pi}{2}$ does not. The oscillating factor then is

$$\cos\left[(\omega_1 \mp \delta)\tau - (\theta_1 - \omega_1\theta_1') - \eta\right],$$

where

$$\eta = \arctan\frac{\alpha_1'}{\tau} \pm \frac{\pi}{2}. \tag{8}$$

The frequency, $(\omega_1 \mp \delta)$, is that of the edge of the segment of the spectrum where the amplitude is relatively very large. The phase differs from that for small values of α_1' by a quantity η which is an ambiguous function of the time τ. This ambiguity may be removed if we assume that the phase varies continuously and that, for very small values of τ, the amplitude has the same sign as the spectrum component corresponding to an infinitesimal value of δ. As τ increases through zero, $\arctan\frac{\alpha_1'}{\tau}$ changes discontinuously from $\mp\frac{\pi}{2}$ to $\pm\frac{\pi}{2}$ according as $\alpha_1' \lessgtr 0$. To avoid a similar discontinuity, in η we say that the sign of $\frac{\pi}{2}$ in (8) is to be taken opposite for positive and negative values of τ. If we make it $\pm$ for $\tau < 0$, and $\mp$ for $\tau > 0$, according as $\alpha_1' \gtrless 0$, then η is zero in the neighborhood of $\tau = 0$. Since the amplitude factor is always positive, this corresponds to a spectral component of positive amplitude. If we make the sign of $\frac{\pi}{2}$ $\mp$ for $\tau < 0$, and $\pm$ for $\tau > 0$, η becomes $\pm\,\pi$, which is the equivalent of a negative amplitude. Hence a knowledge of the spectral component of frequency ω_1 enables us to determine the sign in (8). For large values of (τ), η reduces to $\pm\frac{\pi}{2}$.

Here we have assumed the amplitude of the input signal to be independent of frequency. If this is not the case the same conditions hold at the input as have just been discussed for the output of a resonant system.

The main conclusion to be drawn from the foregoing is that when the amplitude is changing rapidly with frequency, the component of an aperiodic disturbance which corresponds to a narrow segment of the spectrum persists for a considerable period so that there is much overlapping of the contributions of neighboring segments. It is therefore difficult to deduce the nature of the disturbance at any particular time from any narrow region of its spectrum. For the same reason it is difficult to associate the delay experienced by an aperiodic signal with the steady state characteristic of a network when the attenuation of the latter is changing rapidly with frequency.

The net result of our study then is that steady state phase delay has no direct relation to the particular types of delay of an aperiodic signal which we have chosen to investigate. When the amplitude does not change rapidly with frequency, envelope delay is identical with the delay produced in the maximum value of the envelope of a disturbance corresponding to that part of the signal spectrum which is in the immediate neighborhood of the frequency in question. The envelope delay, together with the phase shift, determines the delay in the maximum absolute value of this disturbance, subject to an uncertainty of half a period. This uncertainty depends on the particular combination of signal spectrum and system characteristic. When the amplitude does change rapidly with frequency, the envelope delay still gives the delay in the maximum value of the envelope. However, this maximum is so flat that the interpretation of the results is very difficult.

Engineering Requirements for Program Transmission Circuits*

By F. A. COWAN, R. G. McCURDY and I. E. LATTIMER

Present-day program networks are reviewed from the standpoints of engineering, design, and operation as developed to meet the needs of the broadcasters. The factors requiring consideration in the further development of program networks in anticipation of future needs are also discussed. The presentation of the paper is supplemented by a demonstration of the quality obtainable by transmission over various types of telephone facilities.

Introduction

THE growth of radio broadcasting to the magnitude of a major national industry within the last twenty years has been accompanied by the development of a nation-wide system of wire-line networks interconnecting hundreds of broadcasting stations. Papers have been presented before this Institute from time to time[1,2,3] describing the types of plant used for these networks and discussing important features of their design and operation. With these twenty years of experience as a background, it should now be of interest to review how the various requirements of broadcasting have influenced the development of the networks and to consider some of the factors which have determined the point to which transmission and operating features have so far been carried.

Simply stated, broadcasting is a means by which sounds originated at one place are reproduced simultaneously to large numbers of listeners distributed over wide areas. The simplest possible radio broadcasting system would consist of a microphone, a radio broadcast transmitter and some radio receiving sets. Such a system could serve only the listeners within the comparatively limited service area of the transmitter. To serve the whole nation many transmitters must be established about the country. Furthermore the most desirable sources of program are not usually in the neighborhood of the transmitter to which a particular listener can tune, since talent tends to be concentrated in certain parts of the country, and special events of interest may occur anywhere. To give a true country-wide service so that every listener can hear the programs he enjoys wherever they may

* Presented at A.I.E.E. Winter Convention, Philadelphia, Pa., January 27-31, 1941. Published in *Electrical Engineering*, Transactions section, April 1941.

[1] For all numbered references, see list at end of paper.

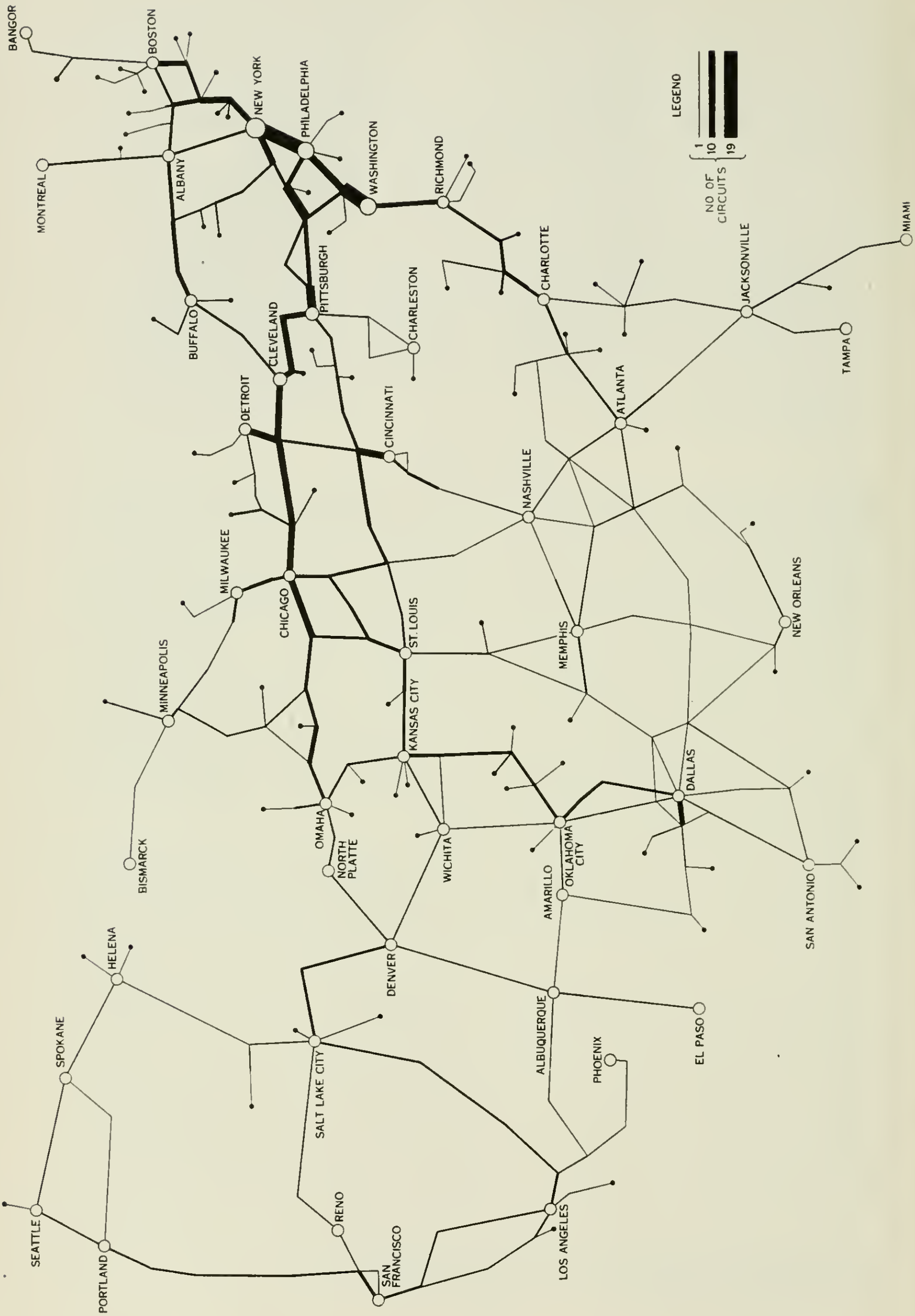


Fig. 1—Major network circuits in the United States

originate, a supplementary transmission system must be provided, interconnecting the many studios and broadcasting stations. The wire networks that perform this function comprise the subject matter of this paper.

The present extent of the wire-line facilities which are associated with the major portion of these networks is indicated in Fig. 1. The width of the lines on this chart has been made proportional to the numbers of circuits in the various sections. The total length of these circuits is in excess of 110,000 miles, and it is not unusual for a program originating at some point on a network to traverse more than 7,000 miles of circuit before being broadcast by the most remote station.

The requirements which the program networks must meet are in the final analysis determined largely by the needs of the broadcasters. The objective of a program network service is to meet these needs in as complete and prompt a manner as possible consistent with reasonable cost. With this objective in mind, it is necessary in planning the plant to consider not only the day-to-day needs, but the possible future needs as well. The importance of this may be appreciated when it is considered that plant provided today for program transmission service will need to be adaptable to the service requirements ten or twenty years hence. As a result of such planning, cables and equipment installed five, ten, and fifteen years ago meet present-day requirements, and, with some rearrangements, will take care of those likely to develop tomorrow.

The detailed planning of program transmission circuits requires consideration of:

1. The numbers of circuits likely to be required, section by section, over each route;
2. The provisions for reliability, flexibility, operation, and supervision essential to a high-grade network service;
3. The transmission requirements, or electrical characteristics, necessary to achieve a natural reproduction of the program.

These three general classes of requirements will be considered in order.

Number of Circuits Required

The circuits which have been established on a full-time basis for continuing use form the backbone of the program networks. Even for these circuits, however, permanence is relative since frequent extensions and rearrangements are made to meet changing requirements of the broadcasters. Aside from these fulltime circuits there are intermittent requirements occasioned by special events and other short-period needs of the broadcasters, some of which involve networks almost as extensive as the full-time networks. In addition reliability of service requires provision for rerouting the networks in the event of trouble. Figure 2 shows the year-by-year growth in the operated mileage of program circuits for the period 1926 to 1940. Of the more than 110,000 miles of circuits shown for 1940, about 45,000 miles have been provided for the short-period services and as stand-by facilities for protection. In addition to these, there are still other circuits, normally assigned to other services, which are arranged to be readily adaptable to program service to supplement the reserve facilities maintained on a full-time basis.

The time interval necessarily accompanying any extensive construction project makes it necessary to engineer plant considerably in advance of actual service requirements to meet, not only the expected growth, but also the changes in network routing. Figure 3 shows for two typical sections

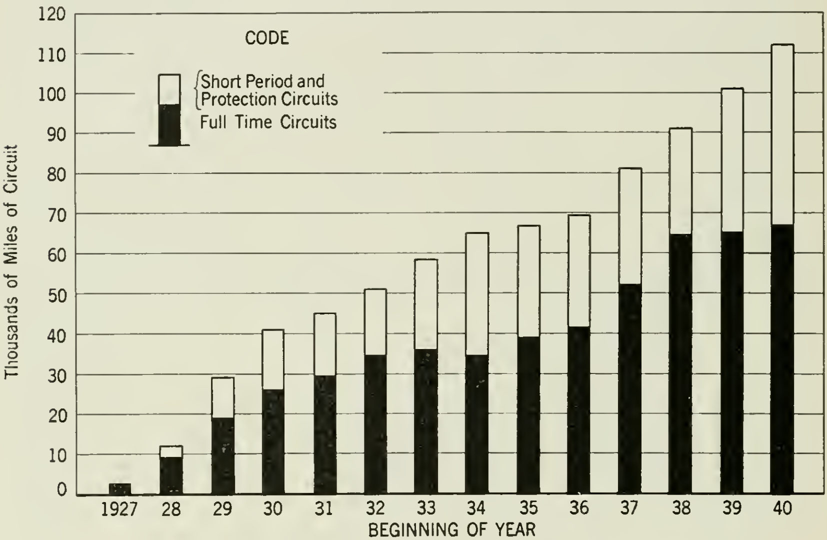


Fig. 2—Growth in mileage of major network circuits

along major routes the variations in requirements for full-time network circuits resulting from growth and rearrangements required by the broadcasters. While, in planning to meet these rapid variations in circuit requirements, advantage can be taken of some latitude which exists in the choice of routes for occasional services and protection facilities, the task of balancing the provision of circuits against requirements is an entertaining and at times difficult one for the circuit engineer.

Operating Requirements

Considering for a moment the variety of programs originating at many different points that can be heard on any home radio set in the course of an evening without once changing the tuning, it will be apparent that minute-to-minute rearrangements of an established interconnecting network must be possible. For example, studios have to be changed from receiving to originating, sections of the network have to be made to transmit first in

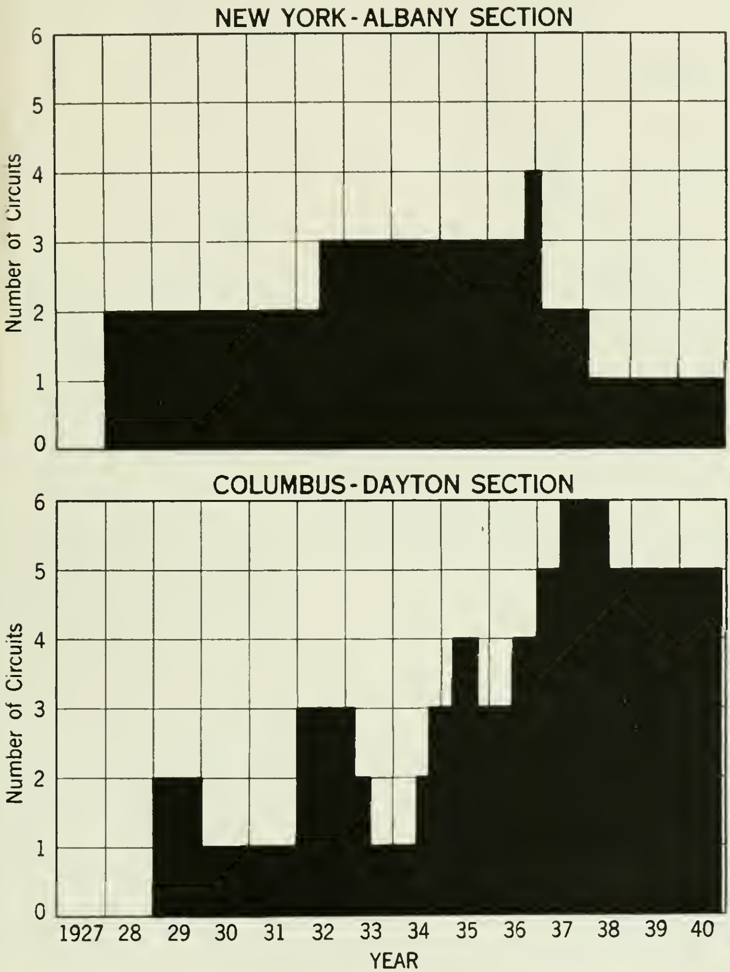


Fig. 3—Variations in full-time program circuits

one direction and then in the other, and branches have to be connected and disconnected. These changes in the network have to be made in the few seconds elapsing between the close of one program and the start of the next on the receipt of selected cue words or sounds. Even during the course of a single program, switches or reversals may have to be made to change the originating point temporarily. To provide for these rapid changes, special operation and special switching and reversing equipment are required at many points along the network. Much of this equipment is under remote control from selected points.

The greater portion of the switching of program circuits is done at about 25 points throughout the country on the major networks. On the average more than 25,000 switching operations per month are performed at these 25 points. During the busy hours of any typical evening there may be something over 500 men on duty at all of the offices about the networks.

At points where switching requirements are simple, the switching equipment consists merely of a few keys. At the larger points where the switching requirements are complex, the switching equipment consists of elaborate relay and control arrangements. These are so designed that it is possible to set up in advance the circuit combinations required for the ensuing program period without disturbing the programs in progress. The actual switching operation takes place at the instant the monitoring attendants signal the the receipt of the last of selected cues, and not before then. This type of arrangement affords a maximum of protection against error, as it is possible to check the presetting for the next switch or make a last minute change if necessary any time before the switch has been made.

Figure 4 shows a picture of such a switching arrangement in use at Omaha, Nebraska for one broadcasting company. At this point 13 circuits used in various trunk and branch sections of two networks are connected to the switching equipment. These are grouped in various combinations to take care of as many as five simultaneous programs. A maximum of five cues might, therefore, be involved in a switch at this point.

The operation and maintenance of the networks are carried out by a special organization under centralized authority and trained in the application of uniform methods and procedures found by experience to be productive of best results. Transmission is monitored continuously at strategic points about the networks. In order to facilitate the activities of this group many thousands of miles of intercommunicating telephone and telegraph circuits are provided full time for their use.

A picture of a monitoring position in the program transmission office at Washington, D. C. is shown in Fig. 5. It will be noted that the monitoring attendant is using an individual headset. This is of a special high fidelity type and is used to avoid the confusion that would result from attempting to monitor a number of different programs simultaneously with loud-speakers. Loud-speakers are available, however, for supplementary checks of quality whenever required.

Accurate transmission measuring equipment is necessary at the various operating points about the networks to insure satisfactory transmission maintenance results.

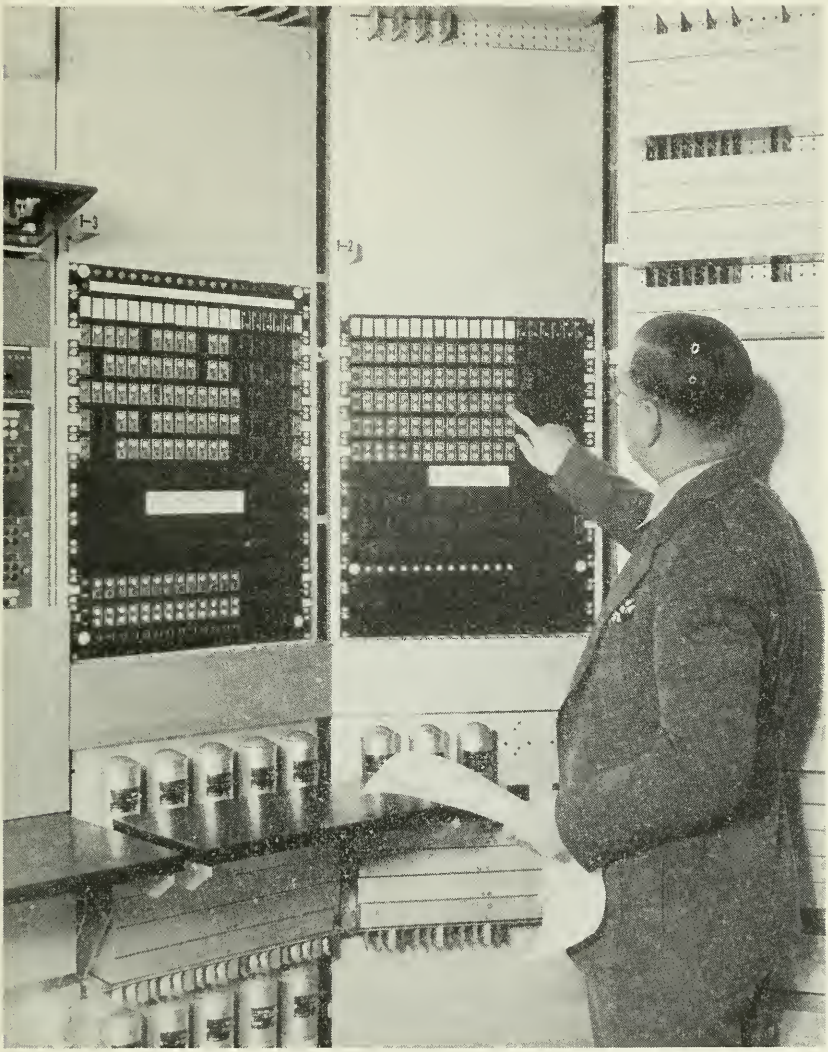


Fig. 4—Switching panel at Omaha, Nebraska

Transmission Requirements

The general transmission requirement for a broadcasting system is that the program material be transmitted with a high degree of naturalness. Although the exact determination of the transmission characteristics which would accomplish this involves many considerations it will be assumed in this discussion that an ideal transmittion system is one in which the sound waves impressed on the listeners' ears in the home are an exact replica of the sound waves striking the microphone in the distant studio. Limitations inherent in the human ear, in the program material to be transmitted, and in the usual listening conditions, however, make such ideal transmission unnecessary. In expressing the requirements for satisfactory transmission, frequency range, attenuation distortion, delay distortion, nonlinearity and noise are used as indices of quality.

Fig. 5—Monitoring position at Washington, D. C.

Before taking up the transmission requirements of a program circuit, it is important to consider further the fundamental factors that are involved in fixing the characteristics considered desirable for the entire system. According to Harvey Fletcher,[4] the zone of audibility of the average normal human ear for pure or single frequency sounds is the area within the curve of Fig. 6. The abscissas represent frequency and the ordinates show the range of intensity recognizable as sound, between the lower limit or threshold of audibility and the upper limit where the sensation of pain is felt. It is seen that the extreme frequency range shown on the chart is from about 20 to 20,000 cycles per second. This range is for young people. It is considerably less for middle-aged and elderly people, and varies with individuals.

In addition to the limitation of the ear there is the fact that there is little energy present in most program material in the extremes of this range, particularly in the upper frequencies. The energy versus frequency spectra of music and other forms of program have been published elsewhere.[5] Figure 7 shows the frequency range which must be transmitted for a number of instruments, speech, and certain noises, so that competent observers cannot detect any impairment.[6] For whole orchestras, experiment has

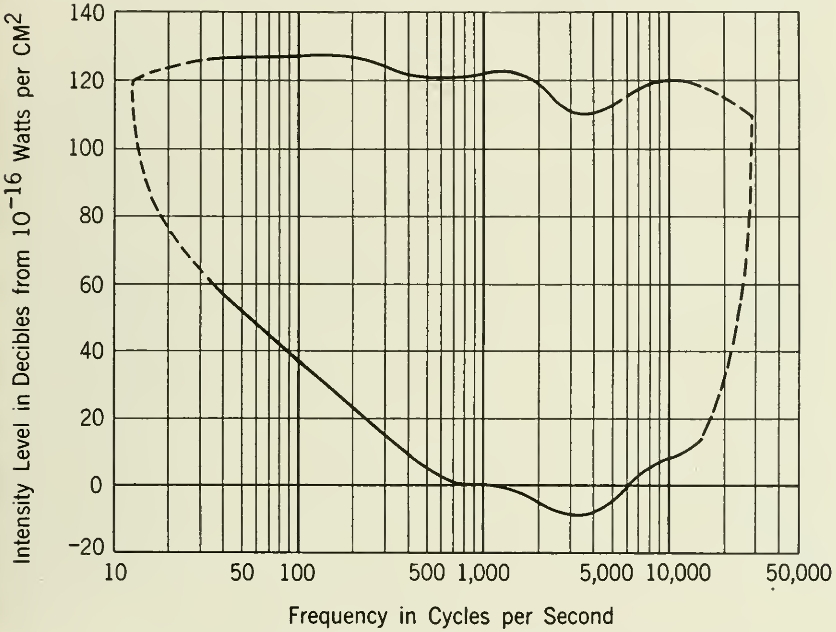


Fig. 6—Limits of audible sound

shown that the elimination of frequencies below 40 and above 15,000 cycles per second is undetectable.[4] It the upper limit of the transmitted frequencies is lowered from 15,000 cycles, the impairment is at first barely detectable but increases at an accelerating rate. When the limit is materially lower than 8,000 cycles, the loss is readily apparent to many people.

Another important consideration is volume range—that is the difference between the maximum and minimum levels of the program. The ordinates of Fig. 6 show that for part of the frequency range, the ear can respond to a range of intensities of more than 120 decibels, with perhaps 100 decibels as a mean. However, the following considerations show that the volume range which the transmission system needs to accommodate is considerably narrower than the intensity range to which the ear can respond.

In the first place, the range of program volumes to which the ear can respond is much less than the range of single-frequency intensities shown by the curve. Program waves are in general very irregular in shape, and even at constant volume contain large and small peaks differing in amplitude by many decibels. The range between the volume at which the highst peaks reach the maximum instantaneous intensity which the ear can tolerate and

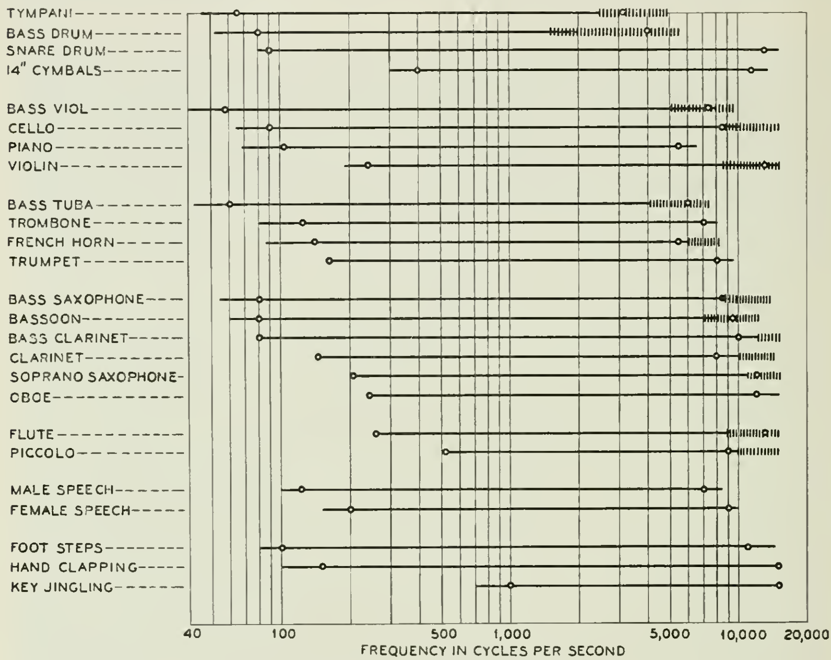


Fig. 7—Audible frequency ranges of music, speech, and noise

Observers voted for entire band by ratio of 60 to 40 over band shown by extremities of lines, and by ratio of 80 to 20 over band indicated by circles. Broken lines show range of noise accompanying music.

the volume at which the smallest peaks are just above the threshold of audibility is therefore less by a number of decibels than the intensity range of the ear as measured by single frequencies.

In the second place, the volume range of the usual program material has definite limits. Measurements have shown that a large symphony orchestra produces a maximum volume range of about 70 decibels.[4] The volume range of most other types of program is considerably less than this, for example, being only about 25 to 30 decibles for dance music and as little as about 15 decibels for much of the dialogue of actors in radio drama.

In the third place, the usual listening conditions impose a definite limit on the useful volume range. The loudest passages in the music of a symphony orchestra correspond to a sound level of about +95 decibels at a point, say one-third the way back in an auditorium, but most people in their homes prefer a level which is lower than this by 5 to 10 decibels. Figure 8

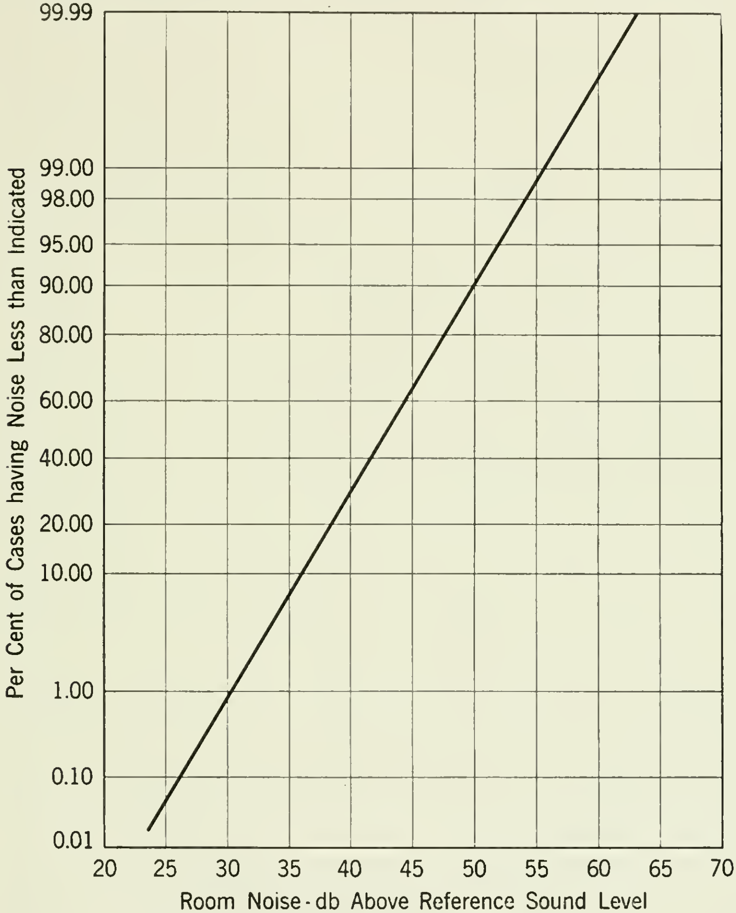


Fig. 8—Residence room noise with radio sets silent
Average 43 db. Standard deviation 5.5 db.

shows the results of an extensive survey[7] of acoustic room noise in homes. It will be noted that the average noise level is +43 decibels on the sound-level scale, and that few homes are quieter than +30 decibels even in the suburbs. The signal-to-noise range inherent in the listening conditions, and allowing nothing for the noise contributed by the transmission system or the room noise where the program is being produced, is therefore seen to be somewhere between 45 and 65 decibels. There is, therefore, no advantage to the listener in providing a permissible volume range materially wider than this in the transmission system.

The above discussion applies primarily to the transmission of symphonic and similar high-grade program material. Much program material is less exacting in its requirements, but on the other hand some sound effects such as the tearing of paper require the reproduction of higher frequencies for complete naturalness.

With these broad considerations in mind, the requirements of high-quality program circuits may now be taken up. As has been noted, the program network is but one part of the over-all broadcasting system which, in addition, includes microphones and studio equipment, radio transmitters, and the home receivers with their loud-speakers. It may be taken as a goal for the program networks that their transmission be nearly enough distortionless so that the over-all performance in regard to naturalness of reproduction will not be limited by them.

To meet such a requirement for short program circuits having only one or two sections is not difficult technically and does not in general require costly types of plant. However, the vast country-wide program networks are made up of many sections of circuits in tandem, which as mentioned before may total in some cases as much as 7,000 miles. This makes it necessary to design and operate the individual circuits to very close limits so that the cumulative discrepancies in the whole network will not exceed tolerable values; and to consider carefully the types of plant employed lest by virtue of sheer numbers of units involved, the total cost be out of line with the over-all grade of service being given the listner. These two conflicting factors are important ones in the consideration of transmission requirements for networks. The determination of the practical working characteristics of program networks involves a consideration not only of the physical and cost factors discussed above but also of such other factors as cost of studios, broadcast transmitters and receivers, and the limitations of the frequency allocations of broadcast stations.

From the standpoint of frequency band the consideration of all factors has resulted in the major present-day program networks being set up to transmit a frequency band with an upper limit of about 5,000 cycles. All program facilities installed in the last ten years or so, however, have been designed to be adaptable to the future transmission of frequencies up to 8,000 cycles. Operation on an 8,000-cycle basis, however, requires the release of additional frequency space now occupied by other services in much of the plant and a general readjustment of the program-circuit characteristics. In 1933, experimental wire circuits were set up between Philadelphia and Washington to transmit frequency bands up to 15,000 cycles. These were employed in a demonstration of stereophonic transmission and reproduction of music.[4] Studio-transmitter loops have been provided to transmit wider frequency bands than the 5,000 cycles currently provided on the nation-wide networks. At the present time, many of the studio-transmitter loops are being set up to transmit bands up to 15,000 cycles. A demonstration will be given at the close of the paper of the transmission of programs over cable circuits about 1,200 miles in length with frequency bands extending to 15,000, 8,000, and 5,000 cycles. The 5,000-cycle circuit is of the type in present commercial use. The 8,000-cycle circuit is of a type to which much of the present program plant can readily be modified. The 15,000-cycle circuit consists of a standard carrier system to which has been added program terminal equipment now under development.

In the consideration of transmission requirements for program circuits other than nominal frequency band, the variation in performance with length and type of circuit is important, since the factors tending to impair transmission are in the nature of small amounts of distortion or noise which accumulate over the length of the circuit. If these effects varied in some definite manner with length, transmission requirements could be fixed on that basis. However, good engineering practice frequently requires choosing for the various sections of a long circuit, different types of facilities whose contributions to the total effects are not in proportion to their length. Even the determination of the maximum permissible distortion and noise on a circuit is influenced by outside factors such as are involved in the broadcasters weighing operating flexibility and cost against the frequency of occurrence of unfavorable network routings and the number of stations affected. For example, in order to secure operating flexibility with a minimum of total network mileage, most of the networks employ the so-called "round robin" principle for a part of the network. In this arrangement the circuit follows a route from station to station forming a continuous loop which returns to its starting point. This naturally results in increased circuit mileage between the program source and the more distant listeners with an attendant increase in undesired transmission effects. For these reasons no exact or specific transmission requirements can be stated for even the over-all performance of program transmission service.

Volume Range

The permissible volume range for a program circuit is determined by the maximum volume which can be transmitted as limited by nonlinear distortion or crosstalk, and the minimum volume which can be transmitted without impairment from the noise present on the circuit.

In connection with their design the various types of program circuit are subjected to listening tests in which the transmission of program over a long loop of the circuit is compared with transmission over a local distortionless circuit. Each type of circuit thus is rated as to the maximum volume it can transmit without noticeable distortion. The highest volume which can be permitted without excessive crosstalk into other program or message telephone circuits is also investigated, and whichever limit is the lower determines the maximum allowable working volume for service. The range between the maximum permissible volume and the noise level on very long lengths of the present program circuits is about 45 or 50 decibels, except under some conditions on certain open-wire sections. On the individual links making up the long circuit, the range is 10 or 20 decibels greater than this.

Attenuation and Delay Distortion

Another important consideration is the amount of attenuation and delay (or phase) distortion to be permitted within the transmitted frequency band. It is the practice to equip program circuits with adjustable attentuation equalizers. By means of these once the desired frequency band has been chosen the deviation in attenuation at any frequency within that band, compared with that at 1,000 cycles, can be adjusted within close limits. On very long circuits, however, experience has shown that even with automatic regulating features and careful operation residual variations which may amount to several decibels may develop as a result of changing temperature and other conditions. These variations are kept within tolerable limits by readjustment of the equalizers from time to time.

Associated with the attenuation distortion is another effect detrimental to program quality, namely, differences in time of transmission for different frequency components of the signal. In practice, circuits tend to have a lower velocity of transmission near the edges of the frequency band than in the middle portions. This results in frequency components near the edges being delayed as compared to the middle portions of the band. This difference in time of transmission is called delay distortion of the circuit. Careful listening tests have shown that it becomes noticeable if, at the highest transmitted frequency, the delay is more than eight milliseconds greater than at 1,000 cycles, and if, at 100 cycles, it is more than about 15 milliseconds greater than at 1,000 cycles. It is controlled by careful attention to the design of loading systems, amplifiers, repeating coils, and all other elements of the circuit. Since such small amounts of over-all delay distortion are detectable and since networks frequently have 100 or more amplifiers in tandem between an originating point and the broadcasting stations on the more distant portions of the networks, it is necessary that the delay distortion of all individual components of a network be held to exceedingly close limits. Accumulations of residual delay distortion which cannot be entirely eliminated in design are reduced by the use of delay equalizers along the circuits when they are set up.

Conclusion

From this discussion it is seen that the program networks are comprised of many parts, each of which must meet exacting requirements in order that over-all results will be satisfactory. It is seen that equally important with transmission are the requirements for plant flexibility, adequate reserves, uniform practices, and centralized supervision of the networks.

The features discussed have been those found desirable for present-day network service. As indicated earlier, consideration of the needs of the future as well as those of the present is an essential feature of the design and engineering of the plant for program-network service. As a result of having done this it will be possible to provide with present plant, and with new plant currently being installed, adequate network facilities as the broadcasting art develops toward higher standards of performance. With the past experience as a guide, it appears that there should be no fundamental difficulty in meeting all reasonable requirements, always remembering that in the long run, requirements and costs bear definite relations to each other.

Acknowledgment

The authors of this paper gratefully acknowledge the assistance in its preparation of many of their associates, especially Mr. W. E. Bloecker, Mr. D. K. Gannett, and Mr. G. S. Bibbins.

References

1. "Telephone Circuits for Program Transmission," F. A. Cowan, *A.I.E.E. Transactions*, Vol. 48, no. 3, July 1929, pages 1045–9.
2. "Long Distance Cable Circuit for Program Transmission," A. B. Clark and C. W. Green, *Bell System Technical Journal*, Vol. 9, no. 3, July 1930, pages 567–94.
3. "Wide Band Open Wire Program System," H. S. Hamilton, *Electrical Engineering*, Vol. 53, no. 4, April 1934, pages 550–62.
4. "Symposium on Wire Transmission of Symphonic Music and Its Reproduction in Auditory Perspective," *Electrical Engineering*, Vol. 53, no. 1, January 1934, pages 9–32, 214–19; *Bell System Technical Journal*, Vol. 13, no. 2, April 1934, pages 239–308:
   "Basic Requirements," Harvey Fletcher.
   "Physical Factors," J. C. Steinberg and W. B. Snow.
   "Loud Speakers and Microphones," E. C. Wente and A. L. Thuras.
   "Amplifiers," E. O. Scriven.
   "Transmission Lines," H. A. Affel, R. W. Chesnut and R. N. Mills.
   "System Adaptation," E. H. Bedell and Iden Kerney.
5. "Absolute Amplitude and Spectra of Certain Musical Instruments and Orchestras," L. J. Sivian, H. K. Dunn, and S. D. White. *Journal Acoustical Society of America*, Vol. 2, January 1931, pages 330–71.
6. "Audible Frequency Range of Music, Speech and Noise," W. B. Snow, *Journal Acoustical Society of America*, Vol. 3, July 1934, pages 155–6.
7. Two papers: "Room Noise at Telephone Locations," D. F. Seacord, *Electrical Engineering*, Vol. 58, no. 6, June 1939, pages 255–7, and Vol. 59, no. 6, June 1940, pages 232–4.

Abstracts of Technical Articles by Bell System Authors

Notes on the Time Relation between Solar Emission and Terrestrial Disturbances.[1] Clifford N. Anderson. Although the correlation between general solar activity and terrestrial disturbances is quite evident, the association of individual storms with specific sunspot groups has never been very satisfactory. Disturbances sometimes have occurred when no sunspots were visible and at other times large sunspots have been unaccompanied by any abnormal disturbances. A possible explanation of such anomalies may lie in longer transit times for the disturbing solar emission than is usually assumed. Some indication is given in this paper that these transit times may range from periods as short as only one or two days to as much as three months. The corresponding velocities for the above transit times are of the order of 2000 and 20 kilometers per second.

Curves show the approximate relation between the angle of emission, velocity, day of emission, and the days intervening between the passage of a spot through the central meridian of the sun and the corpuscular encounter with the earth.

The Effect of the Earth's Curvature on Ground-Wave Propagation.[2] Charles R. Burrows and Marion C. Gray. Curves are presented for the rapid calculation of the ground wave for radio propagation over a spherical earth of arbitrary ground constants, antenna heights, and polarization.

Based on the pioneering work of G. N. Watson, a rigorous theory of the propagation of electromagnetic waves round a spherical earth has been developed in the past twenty years. Watson developed his method in detail only in the limiting case of an earth of infinite conductivity, but his work has since been extended by various authors to cover other values of the earth's conductivity. Theoretically, therefore, solutions are available for any values of the earth's constants (dielectric constant and conductivity) and for either vertically polarized or horizontally polarized waves. In practice, unfortunately, the computations required are lengthy and involved, and for the most part the recent theoretical papers have confined their calculations to a few specific values of the earth's constants. The present paper attempts to summarize the results so far obtained in a manner

[1] *Proc. I.R.E.*, November 1940.
[2] *Proc. I.R.E.*, January 1941.

that will make them more easily available to the practical engineer, and to fill the gaps in these results by developing a series of curves from which the field for any values of the earth's constants may be read, with all the accuracy that could be expected in engineering practice.

Electrical Breakdown of Anodically Oxidized Coatings on Aluminum: A Means of Checking Thickness of Anodized Finishes.[3] K. G. Compton and A. Mendizza. The existing methods for determining the thickness of anodically produced oxide coatings on aluminum are relatively few and are almost entirely of a destructive nature. It is a fairly well established fact that, within the thickness limits normally encountered in practice, the voltage breakdown is a linear function of thickness of oxide film. The authors have endeavored to utilize this fact in developing a test method for determining the thickness of coatings produced under known and controlled conditions with practically no injury to the finish. Data are given which show the relationship between breakdown resistance, anodizing time, thickness of coating, current density and sealing of anodically oxidized polished commercially pure aluminum. Statistical data for the values obtained are also given, indicating the good reproducibility of the breakdown values. By calibrating a particular anodic process, satisfactory results may be obtained in a relatively short time and often without destroying or marring the article. Since the oxide coating is not entirely homogeneous it is necessary to obtain a fairly large number of readings for every test condition. The authors have found that approximately twenty-five readings are usually sufficient and can be made in a relatively short time. Although only one of the many anodizing possibilities has been investigated, the applicability of this method of evaluating the thickness of oxide coatings may be extended to all commercial treatments.

Ultrasonic Absorption and Velocity Measurements in Numerous Liquids.[4] Gerald W. Willard. By means of ultrasonic light-diffraction phenomena the velocity and absorption of sound in some forty transparent liquids were measured in the frequency range of 6 to 30 Mc. Among the list of materials studied are mixtures of liquids in varying proportions, several solutions of solids in liquids, and a non-liquid jell. A novel-construction glass-to-metal-to-quartz cell made possible the study of highly solvent liquids. Velocity values were obtained from measurements of the diffraction spectra spacing. Absorption values were obtained by measurement of the sound radiator voltages required to produce certain color transmission effects at measured distances from the sound radiator. The use of a mercury arc light-source

[3] *A.S.T.M. Proc.*, Vol. 40, 1940.
[4] *Jour. Acous. Soc. Amer.*, January 1941.

enhanced the necessary color effects. The relation between sound beam width (in the optical direction) and light transmission was studied. In general, the values of velocity obtained were found to be independent of frequency, and the absorption to be proportional to frequency squared and unrelated to calculated viscous and thermal losses. A simple calculation is proposed for estimating absorption errors caused by sound beam diffraction and spreading. These apply as well to absorption measured in other methods than here used.

Contributors to this Issue

W. R. Bennett, B.S., Oregon State College, 1925; A.M., Columbia University, 1928. Bell Telephone Laboratories, 1925–. Mr. Bennett has been engaged in the study of the electrical transmission problems of communication.

F. A. Cowan, B.S. in Electrical Engineering, Georgia School of Technology, 1919. American Telephone and Telegraph Company at Atlanta, Ga., 1920; Long Lines Engineering Department, Special Service Group, New York, 1922; Division Transmission Engineer, Division No. 1, New York, 1926; Engineer of Transmission, Long Lines Department, 1928; Engineering Department, American Telephone and Telegraph Company, New York, 1940–. Mr. Cowan is Transmission Engineer and as such is responsible for assisting the Associated Companies in connection with telephone and telegraph transmission and protection matters.

R. V. L. Hartley, A.B., Utah, 1909; B.A., Oxford, 1912; B.Sc., 1913; Instructor in Physics, Nevada, 1909–10. Engineering Department, Bell Telephone Laboratories, 1913–. Mr. Hartley took part in the early radio telephone experiments and has since been associated with research on telephony and telegraphy at voice and carrier frequencies. As Research Consultant he is concerned with general circuit problems.

Irving E. Lattimer, B.E.E., University of Michigan, 1913. Central Union Telephone Company, Engineering Department, Chicago, 1913–14; Chicago Telephone Company, Engineering Department, 1914–18. American Telephone and Telegraph Company, Long Lines Engineering Department, New York, 1918–. Mr. Lattimer is Transmission Methods Engineer.

R. G. McCurdy, B.S., University of California, 1913. Pacific Gas and Electric Company, 1913; Joint Committee on Inductive Interference (California), 1913–1916. American Telephone and Telegraph Company, 1916–1934; Bell Telephone Laboratories, Inc., 1934–. Assistant Director of Transmission Development, 1937–1939; Director of Transmission Engineering, 1939–.

S. O. Rice, B.S. in Electrical Engineering, Oregon State College, 1929; California Institute of Technology, 1929–30, 1934–35. Bell Telephone Laboratories, 1930–. Mr. Rice has been concerned with various theoretical investigations relating to telephone transmission theory.

W. W. Werring, M.E., Cornell University, 1922; Instructor, Mechanical Engineering Laboratory at Cornell, 1920–22. Engineering Department, Western Electric Company, 1922–25; Bell Telephone Laboratories, 1925–. Mr. Werring was at first associated with the Apparatus Analysis Laboratory and went with the Materials Department on its formation. His earlier work there was in the field of insulating materials and plastic molding. More recently he has been concerned with the application of welding, x-rays, optical and photographic methods, gauging and precision measurements and the engineering features of standardization of dimensions and tolerances.

VOLUME XX JULY, 1941 NUMBER 3

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

50c per copy *$1.50 per Year*

THE BELL SYSTEM TECHNICAL JOURNAL

Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.

SUBSCRIPTIONS

Subscriptions are accepted at $1.50 per year. Single copies are 50 cents each. The foreign postage is 35 cents per year or 9 cents per copy.

The Bell System Technical Journal

Vol. XX *July, 1941* *No. 3*

Industrial Mathematics*

By THORNTON C. FRY

The report consists of three major sections. The first discusses mathematical specialists in industry, calls attention to the essentially consultative character of their work, and makes some observations regarding the education, employment and supervision of this type of personnel.

The second section deals, not with the work of these specialists, but with the uses to which mathematics is put at the hands of industrial workers in general, the various ways in which it contributes to the economy and effectiveness of research, and the kinds of mathematics that are most used. A number of illustrations are given, together with brief surveys of the utilization of mathematics in four important industries: communications, electrical manufacturing, petroleum and aircraft.

The third section is devoted to statistics, which touches industrial life at rather different points, and hence could not conveniently be included in the general discussion.

Introduction

MATHEMATICAL technique is used in some form in most research and development activities, but the men who use these techniques would not usually be called mathematicians.

Mathematicians also play an important role in industrial research, but their services are of a special character and do not touch the development program at nearly so many points.

Because of this contrast between the ubiquity of mathematics and the fewness of the mathematicians, this report is divided into sharply differentiated parts. Under "Mathematicians in Industry" an attempt is made to explain what sort of service may be expected of industrial mathematicians, and to develop some principles of primary importance in employing and managing them. An attempt is also made to appraise future demand for men of this type, and to discuss the sources from which they can be drawn. Under "Mathematics in Industry" appear brief surveys of the extent and character of the utilization of mathematics in a few special industries, and examples of specific problems in the solution of which mathematical methods have been necessary or advantageous.

* This discussion of the part which Mathematics might play and to a certain extent is playing in industry was prepared for the National Research Council *Survey of Industrial Research*, a survey undertaken at the request of the National Resources Planning Board. The document which the survey produced has been published as "Research—A National Resource, Part II, Industrial Research" and is available through the Government Printing Office.

In these two sections mathematics is interpreted broadly to include not only the fundamental subjects, algebra, geometry, analysis, etc., but also their manifestations in applied form as mechanics, elasticity, electromagnetic theory, hydrodynamics, etc. Statistics, however, touches industrial activity in a rather different way, and is therefore discussed separately under a third heading, "Industrial Statistics and Statisticians."

One observation which will be made in more detail later is worthy of mention here, because of the present and prospective scarcity of suitably trained industrial mathematicians. Though the United States holds a position of outstanding leadership in pure mathematics, there is no school which provides an adequate mathematical training for the student who wishes to use the subject in the field of industrial applications rather than to cultivate it as an end in itself. Both science generally, and its industrial applications in particular, would be advanced if a group of suitable teachers were brought together in an institution where there was also a strong interest in the basic sciences and in engineering.

Mathematicians in Industry

What is a Mathematician?

If every man who now and then computes the average of a set of instrumental readings or solves a differential equation is a mathematician, there are few research workers who are not. If, on the other hand, only those who are primarily engaged in making additions to mathematical knowledge are mathematicians, there are almost none in industry. Neither definition is sound. The first is absurd; the second not closely related to the essential nature of mathematical thought. This report adopts a definition based upon the character of the man's thinking rather than the ultimate use to which his thinking is put.

Some men would be called mathematicians in any man's language; others physicists or engineers. These *typical* men are differentiated in certain essential respects:

The typical mathematician feels great confidence in a conclusion reached by careful reasoning. He is not convinced to the same degree by experimental evidence. For the typical engineer these statements may be reversed. Confronted by a carefully thought-out theory which predicts a certain result, and a carefully performed experiment which fails to produce it, the typical mathematician asks first, "What is wrong with the experiment?" and the typical engineer, "What is wrong with the argument?" Because of this confidence in thought processes the mathematician turns naturally to paper and pencil in many situations in which the engineer or physicist would resort to the laboratory. For the same reason the mathematician in his "pure" form delights in building logical structures, such as topology or abstract algebra, which have no apparent connection with the world of physical reality and which would not interest the typical engineer; while conversely the engineer or physicist in his "pure" form takes great interest in such useful information as a table of hardness data which may, so far as he is aware, be totally unrelated to any theory, and which the typical mathematician would find quite boring.

A second characteristic of the typical mathematician is his highly critical attitude toward the details of a demonstration. For almost any other class of men an argument may be good enough, even though some minor question remains open. For the mathematician an argument is either perfect in every detail, in form as well as in substance, or else it is wrong. There are no intermediate classes. He calls this "rigorous thinking," and says it is necessary if his conclusions are to be of permanent value. The typical engineer calls it "hair splitting," and says that if he indulged in it he would never get anything done.

The mathematician also tends to idealize any situation with which he is confronted. His gases are "ideal," his conductors "perfect," his surfaces "smooth." He admires this process and calls it "getting down to essentials"; the engineer or physicist is likely to dub it somewhat contemptuously "ignoring the facts."

A fourth and closely related characteristic is the desire for generality. Confronted with the problem of solving the simple equation $x^3 - 1 = 0$, he solves $x^n - 1 = 0$ instead. Or asked about the torsional vibration of a galvanometer suspension, he studies a fiber loaded with any number of mirrors at arbitrary points along its length. He calls this "conserving his energy"; he is solving a whole class of problems at once instead of dealing with them piecemeal. The engineer calls it "wasting his time"; of what use is a galvanometer with more than one mirror?

In the vast army of scientific workers who cannot be tagged so easily with the badge of some one profession, those may properly be called "mathematicians" whose work is dominated by these four characteristics of greater confidence in logical than experimental proof, severe criticism of details, idealization, and generalization. The boundaries of the profession are perhaps not made sharper by this definition, but it has the merit of being based upon type of mind, which is an attribute of the man himself, and not upon such superficial and frequently accidental matters as the courses he took in college or the sort of job he holds.

It is, moreover, a more fundamental distinction than can be drawn between, say, physicist, chemist and astronomer. That is why the mathematician holds toward industry a different relationship than other scientists, a relationship which must be clearly understood by management if his services are to be successfully exploited.

The Place of the Mathematician in Industrial Research

The typical mathematician described above is not the sort of man to carry on an industrial project. He is a dreamer, not much interested in things or the dollars they can be sold for. He is a perfectionist, unwilling to compromise; idealizes to the point of impracticality; is so concerned with the broad horizon that he cannot keep his eye on the ball. These traits are not weaknesses; they are, on the contrary, of the highest importance in the job of finding a system of thought which will harmonize the complex phenomena of the physical world, that is in reducing nature to a science. The job of industry, however, is not the advancement of natural science, but the development, production and sale of marketable goods. The physicist, the chemist, and especially the engineer, with their interest in facts, things and money are obviously better adapted to contribute directly to these ends. To the extent that the mathematician takes on project responsibility, he is forced to compromise; he must specialize instead of generalize; he must deal with concrete detail instead of abstract principles. Some mathematicians cannot do these things at all; some by diligence and self-restraint can do them very well. To the extent, however, that they succeed along these lines they are functioning not as mathematicians but as engineers. As mathematicians their place in industry is not to supply the infinite attention to practical detail by which good products, convenient services, and efficient processes are devised; their function is to give counsel and assistance to those who do supply these things, to appraise their everyday problems in the light of scientific thought, and conversely to translate the abstract language of science into terms more suitable for concrete exploitation.

In other words, the mathematician in industry, to the extent to which he functions as a mathematician, is a consultant, not a project man.

Qualifications Necessary for Success as an Industrial Mathematician

The successful industrial mathematician must not only be competent as a mathematician; he must also have the other qualities which a consultant requires:

First, though his major interests will necessarily be abstract, he must have sufficient interest in practical affairs to provide stimuli for useful work and to reconcile him to the compromises and approximations which are necessary even in the theoretical treatment of practical problems. This usually means that the type of mathematician who could not do a good engineering job if he turned his hand to it will not get on very well in an industrial career.

Second, he must be gregarious and sympathetic. If he shuts himself off from his associates, much of his thinking will have no bearing on their needs and that which does will exert less influence than it might. If he does not translate his thoughts into their language, they will miss the significance of much of his work and he will have but a limited clientele.

Third, he must be cooperative and unselfish. A man cannot be at once consultant and competitor to his associates. Self-seeking attempts to gain credit for his contributions to the industry will inevitably alienate his clientele. There are two reasons for this: In the first place a mathematician's appraisal of mathematical work, even if made from a detached point of view, is heavily weighted on the side of its fundamental scientific significance, whereas its industrial value should be judged on very different grounds and can best be appraised by the engineer. In the second place, the engineer in charge of a project can give credit without embarrassment for help received; it is to his credit to have known where help was to be had. The same story told by another, and particularly by the consultant himself, has an entirely different flavor.

Fourth, he must be versatile. Jobs change, and even the same job may give rise to questions which require very different mathematical techniques.

Fifth, he must be a man of outstanding ability. No one wants the advice of mediocrity. Among industrial mathematicians there is no place for the average man.

Employment and Supervision

Perhaps the greatest hazard in hiring mathematicians for industry arises from the fact that the employment officer is not often a judge of mathematical ability. Paradoxically, however, his mistakes are not usually made in judging mathematical aptitude, since general scholastic rating is an unusually trustworthy index of mathematical ability. But because of a feeling of incompetence bred by his lack of mathematical lore, he spreads the mantle of charity over other characteristics with regard to which he should trust his own judgment. If, for example, the applicant gives an incoherent account of the problems on which he has been working, the interviewer excuses it on the ground of his own lack of mathematical training, an excuse which would be quite adequate if the circumstances demanded that he meet the applicant on the applicant's ground. What he overlooks is that the applicant has failed to meet him on his own ground; has failed, in other words, to display the essential ability to translate his thoughts into the language of his hearer. Or perhaps a personality defect is excused on the ground that "after all, he will be working by himself and won't have to meet people," whereas in fact the real value of a consultant comes not in what he does at his desk, but in how much of it gets through to his associates. The applicant who is boastful or pushing or querulous should not be hired on the general theory that "all mathematicians are queer."

High standards in all such matters, and an interest in practical things as well, are as important as technical mathematical ability. These are stiff specifications, and the men to fill them are not to be found in every market place. They are, however, the requirements implicit in the nature of the job and no good can come from failing to recognize them.

After the right man is hired, he is not a difficult person to supervise if his function as a consultant to the rest of the staff is kept clearly in mind. The broad objectives must be to avoid barriers which would tend to deter his associates from seeking his services, and to assure that his work is justly appraised and fairly compensated.

The three barriers most likely to arise between him and his associates are jealousy, red tape and unavailability.

Jealousy is unavoidable if the man himself is self-seeking; once such a man is hired trouble is inevitable. But the man is not always to blame. A generous and cooperative recruit will be spoiled by an atmosphere too highly charged with progress reports, or by a salary policy which bases revisions upon the dollar value of the last year's work. Actually the "progress" which is significant to management will be far more accurately appraised by his colleagues than by himself, hence his reports have little value except as they give him an opportunity to review and criticize his own activities. If too much emphasis is placed upon them, even this value will be lost and they will be written in the spirit of making a case for himself, which is exactly the spirit most certain to breed jealousy. Similarly, a salary policy based on dollar returns is essentially unjust, for the money value of various bits of theoretical work has almost no correlation with the scientific acumen which they require. This does not mean that a mathematician's pay should, in the long run, be independent of the dollar value of his services. It means only that whether he gets a raise this year, and how big it shall be, should properly be based on the size, character and satisfaction of his clientele, and not upon the commercial importance of the questions they saw fit to bring him last year.

Red tape is easily avoided by avoiding it. No engineer, whatever his rank in the organization, ought ever need permission to consult a mathematician in the company's employ, and the mathematician in turn ought not need a specific work order or expense allowance before giving his advice. In this respect he should be on the same basis as the free-lance investigators who are to be found in most large research laboratories, and who are generally known as staff engineers.

Unavailability is a more serious matter. It is well recognized that in industrial research the urgent job always tends to take precedence over the important one. Left to themselves, fundamental studies give way to the detailed development "which ought to go into production next month." Mathematical studies are no more susceptible than other fundamental research to such interruptions, but the effect upon the career of the mathematician may be more far-reaching, for as soon as he is assigned an urgent project of special character his availability as a consultant ceases or at best is temporarily impaired. If his value to the industry is greater as a project man than as a consultant this need not be a cause for regret; but to turn a good mathematician into a poor engineer, or an irreplaceable mathematician into a replaceable engineer, is unfortunate for both employer and employee.

The Mathematical Research Department of the Bell Telephone Laboratories

In the Bell Telephone Laboratories men of this type have been grouped together as a separate organization unit. They have no more specific function than to be helpful to their associates in other parts of the Laboratories. No engineer is obliged to consult them about any phase of his work; no particular jobs come to them by reason of prerogative; conversely, there is no sort of help which an engineer or physicist may not seek from them if he so desires. No routine need be complied with in advance in order to secure their services, and no report is required afterwards, though written reports are frequently prepared when needed for scientific record. The expense of the group is distributed broadly over the activities of the Laboratories, not charged to specific jobs. Every effort is made to maintain a spirit of service among the members of this group, and though responsibility for engineering projects occasionally descends upon them, it is regarded as an undesirable necessity to be avoided whenever possible and liquidated at the earliest opportunity.

The group has functioned successfully for a number of years. Its members are respected by their engineering associates, and like their jobs. Information regarding their activities reaches management almost entirely through spontaneous acknowledgments made by the engineers they assist. These expressions of appreciation are generous, but rather erratic in that they concentrate attention first on one man, then on another, as the genius and training of the individual happen to click with the important job of the moment. This has not affected the morale of the group adversely, probably because a serious effort is made to avoid erratic salary revisions in which the man who is at the moment in the limelight benefits at the expense of others who are doing equally good but less conspicuous work.

From the standpoint of the men, the principal advantages of being associated together instead of distributed through the engineering departments, is the stimulus of contact with men of like interests. From the standpoint of management, the advantages are wider availability, greater flexibility in matching the talents of the man with the requirements of the job, and a more uniform appraisal of ability because of supervision by a man of adequate mathematical background.

So far as is known, mathematicians have not been organized into separate administrative groups in other industries. In most laboratories their numbers have been thought too small to make such an arrangement feasible, and they have been treated as staff engineers distributed throughout the various general departments. It is believed, however, that there are a few industries in which this arrangement could be introduced with profit at this time, and that it has sufficient merit to justify its adoption wherever possible.

The Mathematician in the Small Laboratory

What has been said above relates primarily to conditions in large industries. The qualifications for success in the small industry are not dissimilar, though the relative emphasis to be placed upon them is somewhat different. Matters of personality (gregariousness, unselfishness, etc.) are not quite so important, because they are offset to some extent by the friendly coherence of the small group. On the other hand, a strong interest in things as well as ideas, and the ability to translate from the language of concrete experience to that of abstract thought and conversely, take on even greater importance. As Dr. H. M. Evjen, himself a worker in a small laboratory, says:

"In order to be of optimum value, the mathematician must keep in close touch with realities. In a sufficiently large organization, employing both theoretical and experimental men, the best results, therefore, can be obtained only by the closest cooperation between the two groups. In smaller organizations, employing —for instance—only one scientifically qualified man, it is difficult to say whether this man should be of the theoretical or the experimental type. If he is a theoretical man, no success can be expected unless he is willing to roll up his sleeves and get his feet firmly planted on the ground. In fact, even if he has highly qualified experimental assistants, he should not feel averse to 'getting down in the dirt.' Secondhand information is always of inferior quality. . . .

The mathematician not only is useful as an auxiliary to whom the practical man can turn with special problems. A properly trained mathematician, with a sufficiently broad vision, can be very much more useful as an active participant in the industrial problems. Due to his training in exact thinking he should be better able to see through the maze of intricate details and discover the fundamental problems involved."

Number Employed

The number of mathematicians employed in communications, electrical manufacturing, petroleum and aircraft is estimated at about 100. The number employed in other places is no doubt somewhat less, but it is probably not an insignificant part of the whole, since mathematicians are found here and there in some very small industries. For example, the Brush Development Company with a total engineering force of only 17, has found it desirable to supplement this group with a man hired specifically as a consultant in mathematics.

It is perhaps not too wide of the mark to estimate the total number at 150, not including actuaries and statisticians.

This number can be checked in another way. The membership list of the American Mathematical Society lists 202 men with industrial addresses. Of these, 102 are in financial and insurance firms and are presumably statisticians. The remaining 100 names are those of industrial employees with mathematical interests strong enough to belong to an organization devoted exclusively to the promotion of mathematical research. Some of these are not mathematicians by the definition adopted in this report. On the other hand, there are also 158 names for which only street addresses are given, some of whom are known to be industrial mathematicians. Balancing these uncertainties against one another, and remembering that many industrial mathematicians find little profit in belonging to an association devoted primarily to pure mathematics, the estimate given above does not appear unreasonable.

Future Demand

The appraisal of future demand is even more speculative than the estimation of present personnel. Two statements, however, seem warranted: (1) The demand for mathematicians will never be comparable to that for physicists, chemists or engineers. (2) It will certainly increase beyond the number at present employed.

The first statement is justified by the fact that physicists, chemists, and other experimental workers deal directly with the natural laws and natural resources which it is the business of industry to exploit, whereas mathematicians touch these things only in a secondary way.

The second statement would perhaps be granted on the general ground that throughout the whole of industry, research is becoming more complex and theoretical, and hence the value of consultants in general, and of mathematical consultants in particular, must increase. It is not necessary, however, to rely solely on such general considerations. Direct evidence exists in certain industries, notably aircraft,[1] where many of the major research problems are generally recognized to be more readily accessible to theoretical than experimental study, and in certain others, such as industrial chemistry,[2] where one may reasonably assume that modern molecular physics will soon begin to play an important part in determining speeds of reaction. There is also the general alertness of executives to the dollar value of a theoretical framework in planning expensive experiments, and the gradually changing attitude toward mathematics that stems from it. As Dr. W. R. Burwell, Chairman of the Brush Development Company, writes:

"There is a definite trend toward a greater use of mathematics in industry which is somewhat commensurate with the trend toward the acceptance of research and development departments as necessary adjuncts to successful businesses. It is becoming more and more generally recognized that mathematics is not only a necessary tool for all engineers, physicists and chemists who make any pretense of going beyond strictly observational methods and experimental solutions to their problems but that it is also performing an important function as the recording medium for those generalizations which lay the foundation for the advances of scientific knowledge. . . .

Even in an organization as small as ours, the use as a consultant is really important and we are constantly having instances where the mathematician because of his training is serving as an interpreter of mathematical and physical theories, sometimes influencing the direction of experimental work and sometimes eliminating the need for it."

If, therefore, the estimate of 150 mathematicians in industry at present is realistic, it may not be too wide of the mark to forecast several times that number a decade or so hence.

Source of Supply

Based on these estimates, a demand for new personnel of the order of 10 a year may be predicted. This number sounds small; but if we reiterate that mediocrity has no place in the consulting field, and that

[1] See pages 31–34.
[2] See pages 30–31.

these 10 must be *exceptional* men, it does not seem unreasonable to ask where they may be found.

Most mathematicians now in industry were trained as physicists or as electrical or mechanical engineers, and gravitated into their present work because of a strong interest in mathematics. Few came from the mathematical departments of universities. As scientists they are university trained, but as mathematicians they are self-educated.

Their training has not been ideal. Industrial mathematics is being carried on by graduates of engineering or physics not so much because of the value of that training as because of the weakness of mathematical education in America. The properly trained industrial mathematician should have, beyond the usual courses of college grade, a good working background of algebra (matrices, tensor theory, etc.), some geometry, particularly the analytic sort, and as much analysis as he can absorb (function theory, theory of differential and integral equations, orthogonal functions, calculus of variations, etc.). These should have been taught with an attitude sympathetic to their applications, and reinforced by theoretical courses in sound, heat, light and electricity, and by heavy emphasis upon mechanics, elasticity, hydrodynamics, thermodynamics and electromagnetic field theory. He should understand what rigor is so that he will not unwittingly indulge in unsound argument, but he should also gain experience in such useful but sometimes treacherous practices as the use of divergent series or the modification of terms in differential equations. He should have enough basic physics and chemistry of the experimental sort to give him a realistic outlook on the power as well as the perils of experimental technique. By the time he has acquired this training he will usually also have acquired a Ph.D. degree, but the degree itself is not now, and is not likely to become, the almost indispensable prerequisite to employment that it is in university life.

There is nowhere in America a school where this training can be acquired. No school has attempted to build a faculty of mathematics with such training in mind. Hence industry has had to make such shift as might be with *ersatz* mathematicians culled from departments of physics and engineering. To make matters worse, a student with strong theoretical interests who enrolls in physics these days is almost certain to spend most of his time on modern mathematical physics, which insists almost as little upon fidelity to experience and experiment as does "pure" mathematics, from which it differs more essentially in matters of language and rigor than of general philosophic attitude. At the moment, therefore, engineering schools must be looked upon as the most hopeful sources of industrial mathematicians.

Historically it is easy to explain how this situation came about. Fifty years ago America was so backward in the field of mathematics that there was not even a national association of mathematicians. A quarter of a century later it was just coming of age in mathematics and was properly, if not indeed necessarily, devoting its entire attention to improving the quality of instruction in the "pure" field. The first faint indications that industrial mathematics might someday become a career had indeed begun to appear, but they were not impressive enough to attract the attention of university executives.

Today we lead the world in pure mathematics, and perhaps also in that other field of mathematics which has somehow come to be known as modern physics. We have strong centers of actuarial and statistical training. But in the field of applied mathematics which is the particular subject of this report, we stand no further forward than at the turn of the century, and far behind most European countries.

A quarter of a century ago it would have been difficult to find suitable teachers. Just now it could be done, primarily because a number of European scholars of the right type have been forced to come here, and a few others have developed spontaneously within our own borders. There are perhaps half a dozen of them, but they are so scattered, sometimes in such unpropitious places, as to have little influence on the development of industrial personnel.

It is unfortunate that no university with strong engineering and science departments has seen fit to bring this group together and establish a center of training in industrial mathematics. We have estimated a demand of about 10 *exceptional* graduates per year. If that estimate is even remotely related to the facts, such a department would have a most important job to do.

Mathematics in Industry

Subjects Used

As Dr. H. M. Evjen, Research Physicist of the Geophysical Section of the Shell Oil Company, remarks:

> "Higher mathematics, of course, means simply those branches of the science which have not as yet found a wide field of application and hence have not as yet, so to speak, emerged from obscurity. It is, therefore, a temporal and subjective term."

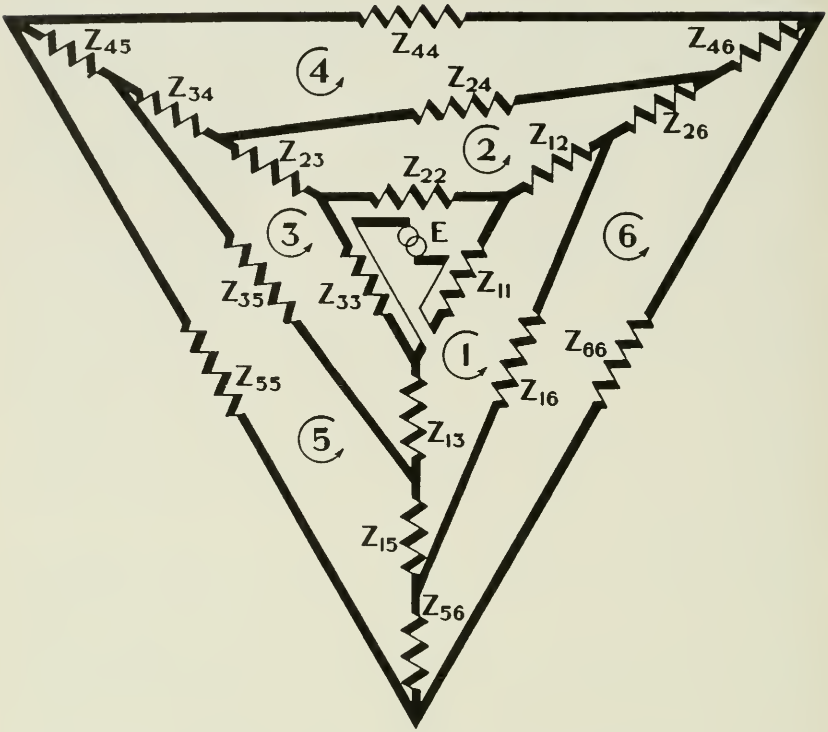
If this is accepted as a definition of higher mathematics—and it is a valid one for the pure science as well as for its applications—it follows automatically that industry relies principally upon the lower branches. What it uses much, ceases by the very muchness of its use to be high. The theory of linear differential equations, for example, is a subject by which the average well-trained engineer of 1890 would have been completely baffled. The well-trained engineer of 1940 takes it in his stride and regards it as almost commonplace. The well-trained engineer of 1990 will certainly regard as equally commonplace the theory of analytic functions, matrices and the characteristic numbers (Eigenwerte) of differential equations, which today are thought of as quite advanced.

With this as a background, there need be no apology associated with the statement that such simple processes as algebra, trigonometry and the elements of calculus are the most common and the most productive in modern industrial research. They frequently lead to results of the greatest practical importance. The single sideband system of carrier transmission, for example, was a mathematical invention. It virtually doubled the number of long distance calls that could be handled simultaneously over a given line. Yet the only mathematics involved in its development was a single trigonometric equation, the formula for the sine of the sum of two angles.

Next in order of usefulness come such subjects as linear differential equations (e.g., in studying the reaction of mechanical and electrical systems to applied forces, the strains in elastic bodies, heat flow, stability of electric circuits and of coupled mechanical systems, etc.); the theory of functions of a complex variable (particularly in dealing with potential theory and wave transmission, propagation of radio waves and of currents in wires, gravitational and electric fields as used in prospecting for oil, design of filters and equalizers for communication systems, etc.); Fourier, Bessel, and other orthogonal series (in problems of heat flow, flow of currents in transmission lines, deformation and vibration of gases, liquids and elastic solids, etc.); the theory of determinants (particularly in solving complicated linear differential equations, especially in the study of coupled dynamical systems); and the like.

Less frequently we meet such subjects as integral equations, which has been made the basis of one version of the Heaviside operational calculus, and which has also been used in studying the seismic and electric methods of prospecting for oil; matrix algebra, which has been applied to the study of rotating electric machinery, to the vibration of aircraft wings, and in the equivalence problem in electric circuit theory; the calculus of variations, in improving the efficiency of relays; and even such abstract subjects as Boolean algebra, in designing relay circuits; the theory of numbers, in the

DETERMINANTS



$$D = \begin{vmatrix} Z_1 & -Z_{12} & -Z_{13} & O & -Z_{15} & -Z_{16} \\ -Z_{12} & Z_2 & -Z_{23} & -Z_{24} & O & -Z_{26} \\ -Z_{13} & -Z_{23} & Z_3 & -Z_{34} & -Z_{35} & O \\ O & -Z_{24} & -Z_{34} & Z_4 & -Z_{45} & -Z_{46} \\ -Z_{15} & O & -Z_{35} & -Z_{45} & Z_5 & -Z_{56} \\ -Z_{16} & -Z_{26} & O & -Z_{46} & -Z_{56} & Z_6 \end{vmatrix} ; Z_j = \sum_{k=1}^{6} Z_{jk}$$

Driving point impedance in mesh j = $Z_{(jj)} = \dfrac{D}{D_{jj}}$

Transfer impedance between mesh j and mesh k = $Z_{(jk)} = \dfrac{D}{D_{jk}}$

(D_{jk} = the first minor of the element Z_{jk} in D)

Many properties of the complicated networks studied at Bell Telephone Laboratories are most conveniently expressed by means of determinants. Above are shown a six-mesh network; its "circuit discriminant", D; and some formulae which illustrate how simply the properties of the system can be found from D. Note that, since $Z_{jk} = Z_{kj}$, D is symmetrical.

design of reduction gears, and in developing a systematic method for splicing telephone cables; and analysis situs, in the classification of electric networks.

Least frequently of all, but by no means never, the industrial mathematician is forced to invent techniques which the pure mathematician has overlooked. The method of symmetric coordinates for the study of polyphase power systems; the Heaviside[3] calculus for the study of transients in linear dynamical systems; the method of matrix iteration in aerodynamic theory;[4] much of the technique used in the design of electric filters and equalizers—these may stand as illustrative examples.

The student of modern mathematics will be impressed at once by two aspects of this review: first, by the heavy emphasis on algebra and analysis, and the almost complete absence of geometry beyond the elementary grade; second, the complete absence of the specific techniques which play such a large rôle in modern physics and astrophysics. It is not easy to say just why advanced geometry plays no larger part in industrial research; however, the fact remains that it does not.[5] As regards modern physics, one may perhaps extrapolate from past history and infer that what is now being found useful in interatomic physics will soon be needed in industrial chemistry. In making this extrapolation, however, it is well to bear in mind that the physics in question is for the most part a mental discipline, its connection with the world of reality still ill-defined and incompletely understood. Therefore it may not prove to be as quickly assimilable into technology as have other disciplines whose symbols could be more immediately identified with experience.[6]

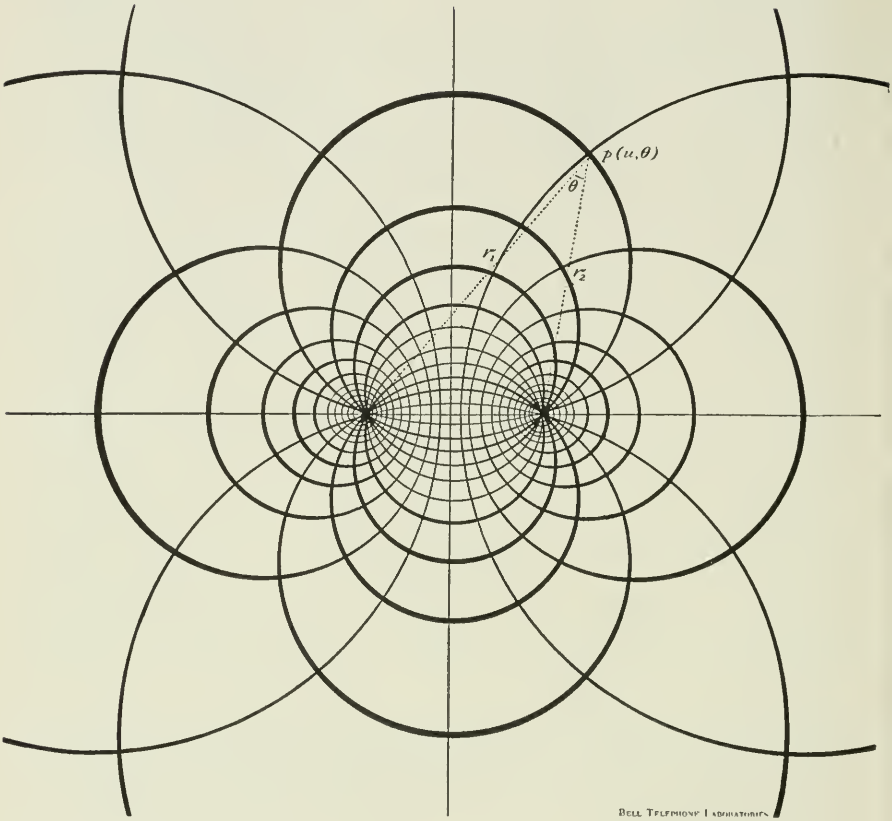
[3] Heaviside was not himself an industrial employee, but the reformulation of his work in terms of integral equations, and its interpretation in terms of Fourier transforms were both carried out in America by industrial mathematicians.

[4] This method was developed in The National Physical Laboratory of England, in the course of studies which in America would probably have been undertaken by a government or industrial laboratory.

[5] Mr. Hall C. Hibbard of the Lockheed Aircraft Corporation comments on this remark as follows: "It is possible that the usefulness of this principle of mathematics has been overlooked to a large extent in certain fields where it might be applied to advantage. In particular, that phase of engineering known as 'lofting,' which deals with the development of smooth curved surfaces, might offer an interesting field for certain types of advanced geometry. Practically all of this work is now done by 'cut and try' methods and the application of mathematics would no doubt save a great deal of time. The same thing is true in the field of stress analysis, where a great deal of time is absorbed in determining the location and direction of certain structural members. It is even possible that the application of vector analysis technique would greatly simplify certain forms of structural analysis, particularly space frameworks. The lack of application of geometry in these fields is probably due to the wide gap that exists between the mathematician and the 'practical' designer and draftsman. Advanced geometry might also turn out to be a very useful tool in connection with problems that we are now encountering in the forming of flat sheet into surfaces with double curvature, an operation that is extensively employed in aircraft manufacture."

[6] In this connection, see the quotation from Dr. E. C. Williams on pages 30–31.

Bicircular Coordinates



$$(x + \coth u)^2 + y^2 = \operatorname{csch}^2 u;\ x^2 + (y - \cot\Theta)^2 = \csc^2\Theta$$
$$u = \log (r_2/r_1)$$

Using the bicircular system of coordinates facilitates finding the distribution of electric charge on two parallel conductors, and thence their capacity. Rotating the bicircular system about the vertical axis generates a toroidal coordinate system which facilitates determining the capacity of a torus.

Finally, we must remark upon two facts: (1) that approximate solutions of problems, and hence methods of iteration (successive approximation), play a much more conspicuous role in applied mathematics than in the pure science; (2) that the highly convenient assumption that linear approximations to natural laws (such as Hooke's law and Ohm's law) are sufficiently exact for practical purposes is less often true than formerly was the case, so that nonlinear differential equations are of great importance to the modern engineer.

Types of Service Performed by Mathematics

Leaving aside the important but rather trite observation that mathematics is a language which simplifies the process of thinking and makes it more reliable, and that this is its principal service to industry, we may distinguish certain less inclusive, but perhaps for that reason more illuminating, categories of usefulness.

First: It provides a basis for interpreting data in terms of a preconceived theory, thus making it possible to draw deductions from them regarding things which could not be observed conveniently, if at all.

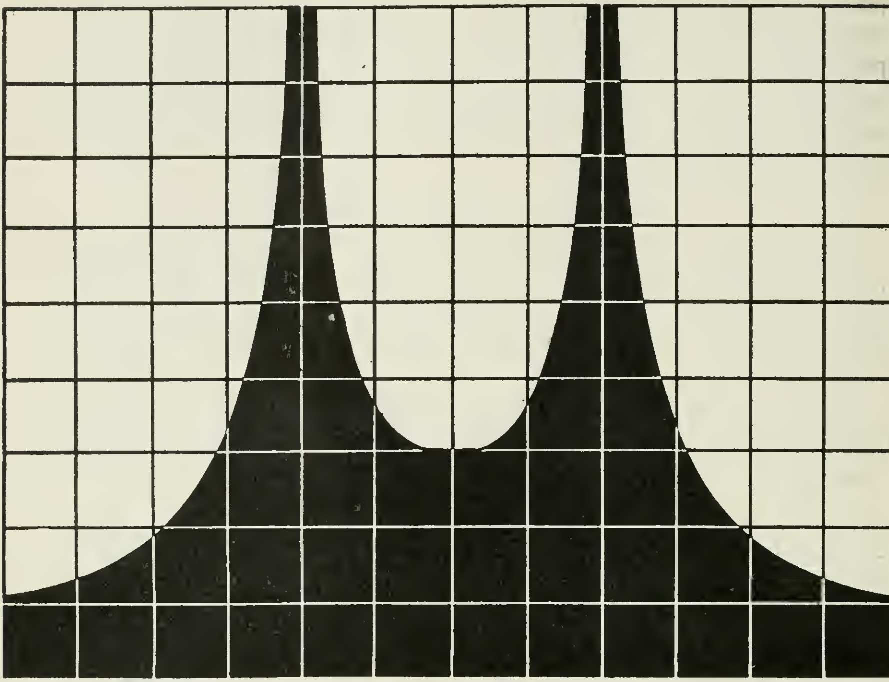
(*a*) An illustration is the standard method for locating faults on telephone lines. Mathematical theory shows that a fault will affect the impedance of the line in a way which varies with frequency, and that the distance from the place of measurement to the fault can be deduced at once from the frequencies at which the impedance is most conspicuously affected. This is obviously much more convenient than hunting the fault directly.

(*b*) A second illustration is the mapping of geological strata by means of measurements made upon the surface of the earth. One method extensively employed uses a large number of seismographs, each of which records the miniature earthquake shock produced at its location by a charge of dynamite set off at a known place. A theory of reflection and refraction similar to that used in geometrical optics shows that certain observable characteristics of these records are related to the depth and tilt of the underground layers, and hence enables the situation of these layers to be plotted. By this means the location of the highest point of an oil-bearing stratum can be found, and the most favorable position for drilling determined.

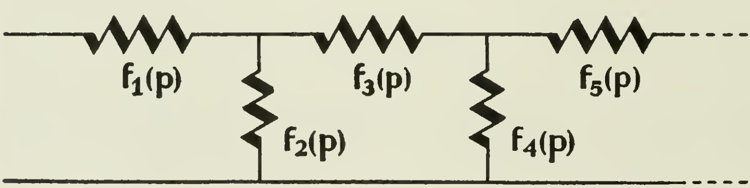
Underground geology is also studied by means of gravity, electrical or magnetic measurements upon the surface. In this case the basic theory is that of the Newtonian potential field, and the interpretation of the data leads into the subject of inverse boundary value problems, which is still insufficiently understood. Enough progress has been made in several geophysical laboratories, however, so that the gravity method is now being widely used, and the electrical methods appear promising for some applications.

Second: When data are incompatible with the preconceived theory, a mathematical study frequently aids in perfecting the theory itself. The

CONTINUED FRACTIONS



$$Z = f_1(p) + \cfrac{1}{\cfrac{1}{f_2(p)} + \cfrac{1}{f_3(p) + \cfrac{1}{\cfrac{1}{f_4(p)} + \cdots\cdots}}}$$



A mathematical method of systematically designing a circuit of predetermined impedance has been developed in Bell Telephone Laboratories. The given impedance, as a function of frequency, is expanded in a Stieltjes continued fraction, whose terms give the electrical constants of the desired network.

classical illustration in pure science is the discovery of the planet Neptune. The motion of the planet Uranus was found to be inconsistent with the predictions of the Newtonian theory of gravitation, if the solar system consisted only of the seven planets then known. Mathematical investigation indicated, however, that if an eighth planet of a certain size was assumed to be moving in a certain orbit, these discrepancies disappeared. Upon turning a telescope to the spot predicted, the new planet was found.

An illustration comes from the aircraft industry. I quote it from a report sent me by Mr. C. T. Reid, Director of Education of the Douglas Aircraft Company:

(*c*) "The behaviour of airplanes with 'power on' did not check closely enough with stability predictions which had been made without consideration of the effects of the application of power; therefore, a purely mathematical analysis of the longitudinal motion of an airplane was carried out, involving the solution of three simultaneous linear first-degree differential equations. The results led to the development of equations for dynamic longitudinal stability with 'power on' which enable the aerodynamicist more accurately to predict the stability characteristics of a given design. 'Power-on' dynamic longitudinal stability is an important design criterion in aircraft construction."

(*d*) Another illustration arises in communication engineering. Theoretical studies had established the fact that vacuum tubes would spontaneously generate noise because of the discrete character of the electrons of which the space current is composed. The theory predicted how loud this noise would be in any particular type of vacuum tube, a most significant result since it established a limit to the weakness of signals which could be amplified by this type of tube. The predictions of the theory were supported by experimental data so long as the tubes were operating without appreciable space charge. But it was found that when space charge was present the noise level fell far below the predicted minimum. In this case the missing factor in the theory was immediately obvious, but an understanding of the mechanism by which the reduction was affected and its incorporation into the theory in a workable form, required an extensive and difficult mathematical attack.

Third: It is frequently necessary in practice to extrapolate test data from one set of dimensions to a widely different set, and in such cases some sort of mathematical background is almost essential.

An example of this kind of service, concerned with the theory of arcs in various gases, is furnished me by Mr. P. L. Alger, Staff Assistant to the Vice President in Charge of Engineering, of the General Electric Company:

(*e*) "An example of this kind of problem is that of the theory of arcs in various gases. It has been experimentally known that the duration, stability and voltage characteristics of electric arcs in different gases and under different pressures vary very widely. The behaviour of such arcs is of great importance, both in

Elliptic Integrals

$$H=\frac{4I}{r}\left[\frac{1}{1-k^2}\int_0^{\pi/2}\sqrt{1-k^2\sin^2\lambda}\,d\lambda-\int_0^{\pi/2}\frac{d\lambda}{\sqrt{1-k^2\sin^2\lambda}}\right]$$

Some simple engineering problems require advanced mathematics in their solution. This is true, for example, in the computation of the magnetic field outside the spiral grid of a vacuum tube, a problem of interest to Bell Telephone Laboratories. If the grid is closely coiled, the current can be treated as a continuous cylindrical sheet, of radius a. Then the component of the magnetic field parallel to the axis of the grid at a distance r from the axis is given by the above function of two Elliptic Integrals whose "modulus" is $k=a/r$.

welding and in the design of circuit breakers and other protective devices. Recently a mathematical theory has been developed which relates the arc phenomena to the heat transfer characteristics of different gases. This theory has given excellent correlation between the known experimental results, and has enabled very useful predictions of performance under new conditions to be made. The theory has been applied in the design of high voltage air circuit breakers, which are of important commercial value, and it is also greatly curtailing the time and expense necessary to develop many other devices in which arc phenomena are of importance."

A second example, furnished me by Mr. Reid, has to do with the interpretation of wind-tunnel data in aerodynamics:

(*f*) "Here it is obviously impracticable to perform full-scale tests of such parts as wings or fuselage, much less of entire aircraft, and the extrapolation from the results of wind-tunnel measurements to the full-scale characteristics of airplanes must be based on theoretical considerations."

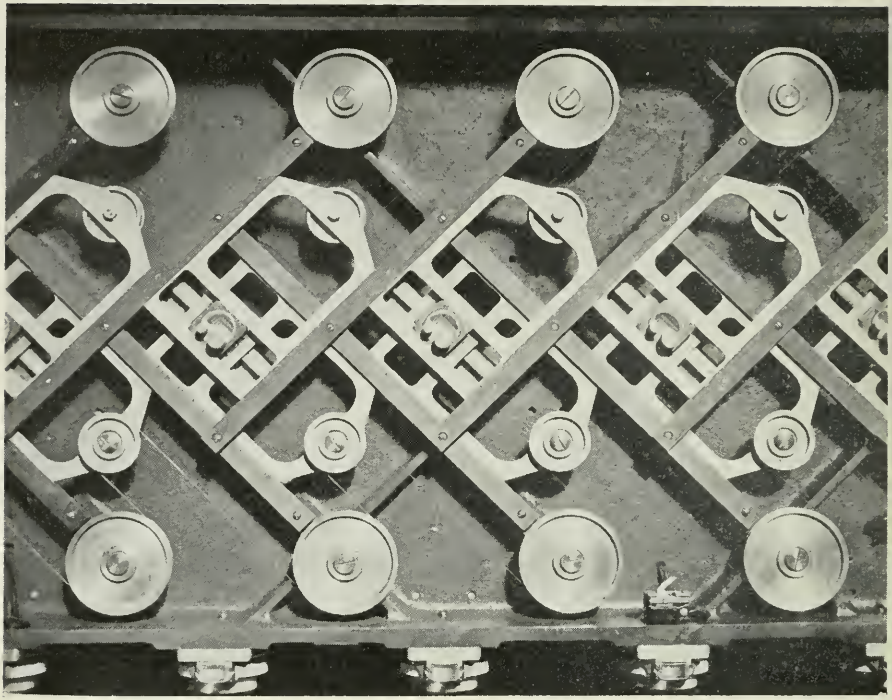
Fourth: Mathematics frequently aids in promoting economy either by reducing the amount of experimentation required, or by replacing it entirely. Instances of this kind are met everywhere in industry, not only in research activities, but in perfecting the design of apparatus and in its subsequent manufacture as well.

Mr. Alger describes in general terms one situation frequently met in research activities as follows:

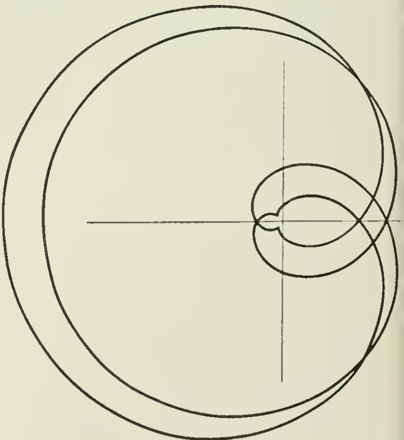
"The first type of problem is one in which there are so many different independent dimensions of a proposed shape to be chosen, or in general so many independent variables, that it is hopeless to find the optimum proportions by experiment. The truth of this can readily be seen when it is realized that the number of test observations to be made increases exponentially with the number of variables. If 10 points are required to establish a performance curve for one variable, 1,000 observations will be required if there are 3 independent variables, and a million if there are 6 variables."

As an illustration he cites the following problem:

(*g*) "An example of this kind of problem is that of designing a *T* dovetail to hold the salient poles in place on a high speed synchronous generator. A large machine of this type may have 10 or more laminated poles carrying heavy copper field coils, each assembled pole weighing several tons and traveling at a surface peripheral speed of 3 miles a minute. The centrifugal force on each pound of the pole then amounts to approximately 500 pounds. The problem of designing dovetails to hold these poles in place, even at over speed, is, therefore, one of great importance and technical difficulty. For each such dovetail, there are 7 different dimensions which may be independently chosen. While empirical methods have enabled satisfactory results to be obtained in some cases, application of mathematics has recently enabled marked improvements in dovetail



THE ISOGRAPH



The Isograph was developed in Bell Telephone Laboratories to find mechanically the complex roots of polynomials of high degree. Let the polynomial to be factored be $p(z) = \sum_{o}^{n} a_j z^j$ or $\sum_{o}^{n} a_j r^j \cos j\Theta + i\sum_{i}^{n} a_j r^j \sin j\Theta$ if $z = r(\cos\Theta + i\sin\Theta)$. The isograph maps the complex values of $p(z)$ as the variable describes the circle $|z| = r$. This graph loops the origin once for each root smaller in absolute value than r. The number of roots between trial values of r is determined by counting loops, and by interpolation a value of r is found for which the graph passes through the origin. This value of r and the corresponding value of Θ define the real and imaginary parts of a root.

designs to be made. Generally speaking, these improvements have permitted an overall strength increase of 20 per cent to be obtained under steady stresses, and much higher gains to be made under fatigue stress conditions; while at the same time the certainty of obtaining the desired results on new designs has been very greatly enhanced."

A second example was brought to my attention by Mr. L. W. Wallace, Director of the Engineering and Research Division of the Crane Company:

(*h*) "A pipe fitting weighing several hundred pounds and intended for high pressure service had a neck of elliptical cross-section. As originally designed, the thickness of the casting was intentionally not uniform, the variations having been introduced empirically to strengthen it where strength was supposed to be most needed. A redesign carried out on the basis of the theory of elasticity showed the distribution of metal to be inefficient and resulted in a new casting in which the weight was reduced by half, while at the same time the bursting strength was doubled. The method used in arriving at this result is an interesting illustration of sensible mathematical idealization. The casting was regarded as an elliptical cylinder under hydrostatic pressure. As the stresses for this idealized structure were already known, the design problem reduced at once to the simple matter of establishing thicknesses sufficient to withstand these stresses."

Another example from the field of geophysical prospecting is furnished by Mr. Eugene McDermott, President of Geophysical Service, Inc.:

(*i*) "A specific case of mathematical research in instrument design was recently encountered. The instrument in question was intended for the measurement of gravity. After the machine had been completely built it was found to be unexplainably inaccurate. After weeks of trial and error it was turned over to a mathematician to try to find the trouble. He soon showed by simple trigonometry that the axis of the instrument would have to be located on its pivot with an accuracy which is not attainable. He also pointed out a means of avoiding this feature by a relatively simple change in design, and this appears to have remedied the trouble."

Another illustration from the petroleum industry, but this time concerned with the production of oil rather than prospecting for it, comes from Dr. E. C. Williams, Vice President in Charge of Research of the Shell Development Company:

(*j*) "The petroleum industry has one important problem not found in other fields; it has to do with oil production from the ground. A mathematical problem arising from this subject is the following: The oil-gas mixture underground flows under pressure through porous media; with a certain spacing of wells, determine the most economical way to recover this mixture. This is sometimes equivalent to asking: 'In what way can the largest fraction of the oil be obtained over a certain period of time?' Simplified problems of this kind have been solved by potential theory methods, since classical hydrodynamics becomes too involved, and in the general problems where the flow constants vary with liquid-gas composition, etc., partial differential equations are found which can be solved by approximate methods. On the basis of the solution of this mathematical problem, aided by extensive laboratory determinations of the required constants, one is able to find the best of several ways of producing from a given oil field."

As a final example under the heading of economy, we may mention the flight testing requirements imposed upon the aircraft industry by the Civil Aeronautics Authority. Of these, Mr. E. T. Allen, Director of Flight and Research of the Boeing Aircraft Company, says:

(*k*) "It was formerly required that each type of transport plane must be tested at all the altitudes at which it was intended to be flown, and at all flying fields where it was expected to be used. The cost of such testing was extremely high. A mathematical study of steady flight performance has, however, identified the basic parameters and established their relations to one another. This has made possible a scientific interpretation of flight test data taken at any suitable location convenient to the aircraft factory, and a reliable conclusion therefrom as to the performance to be expected under other conditions. This has greatly reduced both the cost and the time necessary to establish performance figures."

Fifth: Sometimes experiments are virtually impossible and mathematics must fill the breach. An example comes to me from Mr. Hall C. Hibbard, Vice President and Chief Engineer of the Lockheed Aircraft Corporation:

(*l*) "An unfortunate phenomenon that must be dealt with in aircraft design is a type of violent vibration which may be set up in the wings if the plane is flown too fast. It is known as flutter, and is highly dangerous, since the vibrations may be of such intense character as to cause loss of control or even structural failure. The technical problem is therefore to be sure that the critical speed at which flutter would occur is higher than any at which the craft would ever be flown. It is a phenomenon with respect to which wind tunnel experimentation is difficult, and flight testing very dangerous. It has been the subject of a number of mathematical investigations, the results of which have reached a sufficiently advanced stage that they are now being used to predict the critical speeds and flutter frequencies of aircraft while still in the design stage. Even more important, the mathematical investigation of this problem points the way to modifications of design which will insure that flutter cannot occur in the usable speed range."

Telephony provides a second example:

(*m*) The equipment in an automatic telephone exchange must be capable of connecting any calling subscriber with any called subscriber. It consists of several stages of switches, each of which can be caused to make connection with a number of trunks which lead in turn to switches in the next succeeding stage. Enough switches must be provided so that only a very small proportion of subscribers' calls will fail to be served immediately. Since the demands made by the subscribers fluctuate from moment to moment, the number of switches required depends in part upon the height to which the crests occasionally rise in this fluctuating load. It is also influenced, however, by the way the trunks are arranged, by the order in which the switches choose them, and by many other factors. Experimental appraisal of the effect of these various factors is impossible, both because it would be very costly, and because it would be exceedingly slow. Mathematically, however, they have been studied by the theory of *a priori* probability,[7] which is used not only in determining how much apparatus to install in a working exchange, but also in comparing the relative merits of alternative arrangements while in the development stage.

Sixth: Mathematics is frequently useful in devising so-called crucial experiments to distinguish once for all between rival theories. A famous example in the field of physics was the study of the refraction of starlight near the sun's disc, which afforded a means of deciding between Newtonian and relativistic mechanics. In this case, mathematical investigation showed that the result to be expected was different according to the two theories, and astronomical observations confirmed the prediction of relativistic mechanics. In the industrial field, an example of this kind comes to me from Dr. Joseph A. Sharpe, Chief Physicist in the Geophysical Laboratory of the Stanolind Oil and Gas Company:

(*n*) "As an example of the second sort of use of analysis there is the case of our study of 'ground-roll,' the large amplitude, low frequency surficial wave which caused so much grief in the early days of seismic reflection prospecting when filters were not used as extensively as at present. We hope to use our study of this wave motion as an aid to a better understanding of the properties of the surficial layers of soil and their effects on the reflected waves in which we are primarily interested.

Two views on the ground-roll are current, although neither is based on very much observation, and this of an uncontrolled sort. One view states that the ground-roll is an elastic wave. Analysis predicts that this wave will have a certain velocity in relation to the velocities of other waves, that it will have a certain direction of particle motion and relation of maximum horizontal to maximum vertical component of displacement, that it will attenuate with distance according to a certain law, that it will attenuate with depth in a certain way, and that its velocity will follow a certain dispersion law. The second view maintains that the 'ground-roll' is a wave in a viscous fluid, and analysis predicts a behavior which is similar in certain cases, and different in others, to that of the elastic wave. Having the predictions of the analysis at hand, we are enabled to devise a group of observations, and the special equipment for their prosecution, which will provide crucial tests of the two hypotheses."

Seventh: Mathematics also frequently performs a negative service, but one which is sometimes of very great importance, in forestalling the search

[7] *Not* statistics, which is *a posteriori* probability. This is one of the few cases in industry where the *a priori* theory finds application.

for the impossible; for many desirable objectives in industry are as unattainable as perpetual motion machines, and frequently the only way to recognize the fact is by means of a mathematical argument.

(*o*) A certain type of electric wave filter which is usually referred to as an "ideal" filter would be very useful if it could be produced. However, it has been shown mathematically that such a structure would respond to a signal before the signal reached it; in other words, that it would have the gift of prophecy. Since this is absurd, it follows that no such filter can be built, and consequently no one tries to build it.

Still another example from the field of communication deals with the design of feedback amplifiers.

(*p*) In practice, any amplifier is intended to handle signals in a given frequency band. For various reasons, it is preferable not to have it amplify disturbances outside this band, and hence its gain characteristic is made to drop off as rapidly as possible outside the limits of the useful band. It has been shown theoretically, however, that the gain cannot decrease at more than a certain rate, which can easily be computed, without causing the amplifier to become unstable. As a matter of fact, the allowable rate at which the gain may fall is often surprisingly low, and a great deal of design effort would be wasted in the attempt to obtain an impossible degree of discrimination if the theoretical limitations were unknown.

Eighth: Finally, mathematics frequently plays an important part in reducing complicated theoretical results and complicated methods of calculation to readily available working form. So many and so varied are the services falling in this category that it is difficult to illustrate them by means of examples. We arbitrarily restrict ourselves to two, chosen primarily for the sake of variety. The first comes from Mr. Hibbard:

(*q*) "In aircraft design the metal skin, though thin, contributes a large part of the structural strength. Nevertheless, such thin metallic plates will buckle or wrinkle after a certain critical load is exceeded. Beyond this point the usual structural theories can not be applied directly and it is therefore necessary to introduce new methods of attack to predict the ultimate strength of the structure. These stiffened plates are difficult to deal with theoretically, but by interpreting the effect of the stiffeners as equivalent to an increase in plate thickness or a decrease in plate width, the calculations can be brought within useful bounds."

The reduction of electric transducers to equivalent T or Π configurations, the interpretation of the elastic reaction of air upon a microphone as equivalent to an increase in the mass of its diaphragm, the postulation of an "image current" as a substitute for the currents induced in a conducting ground by a transmission line above it, and a host of other common procedures could be cited as similar instances of simplification based upon more or less valid mathematical reasoning.

The second example is furnished by Dr. E. U. Condon, Associate Director of the Research Laboratories of the Westinghouse Electric and Manufacturing Company:

(*r*) "In the manufacture of rotating machinery it is of extreme importance to have the rotating parts dynamically balanced, in order to reduce to a minimum the vibration reaction on the bearings which unbalance produces. Theory shows the phases and amplitudes of the bearing vibrations produced by excess masses located at various places on the rotor; conversely, by solving backward from observed vibration data, one can compute what correction is needed to eliminate the unbalance. Recently a most valuable machine has been developed which not only measures the unbalance, but also automatically shows what correction should be made, thus eliminating the necessity for these calculations.

The rotor to be balanced is whirled in bearings on which are mounted microphones that generate alternating voltages corresponding to the vibrations of the bearings. These voltages are fed into an analyzing network, which automatically indicates the correction needed in order to achieve dynamic balance. In some cases the output of the balancing machine has been arranged to set up a drilling machine so it will automatically remove the right amount of metal at the right place. These machines are finding application in the manufacture of small motors, of automobile crankshafts, and in the heavy rotors of power machines."

In the same class would come the isograph, by means of which the complex roots of polynomials can be located; the tensor gauge which registers the principal components of strain in a stressed membrane without advance knowledge of the principal axes; and slide rules for a great variety of special purposes such as computations with complex numbers, the calculation of aircraft performance, aircraft weight and balance, and the like. Perhaps we ought also include in the same category the use of soap-bubble films for the study of elastic stresses in beams, the use of current flow in tanks of electrolyte for the study of potential fields, and the use of steel balls rolling on rubber membranes stretched over irregular supports as a means of studying the trajectories of electrons in complicated electric fields. These are all mechanical methods for saving mathematical labor, but they are more than that, for they all rest upon a foundation of mathematical theory. They are, in fact, examples of the use of mathematics to avoid the use of mathematics.

Mathematics in Some Particular Industries

Communications

The communication field is the one in which mathematical methods of research have been most freely used. This is due partly to the fact that the transmission of electric waves along wires and through the ether follows laws which are particularly amenable to mathematical study; partly also to the fact that so much of the research has been centralized in a single laboratory, thus bringing together a large number of engineers into a single compact group, and justifying the employment of consultative specialists. Most important of all, however, is the fact that there are two devices—vacuum tubes and electrical networks—without which modern long-distance telephony would be impossible; and one of these, the electrical network, is and has been since its earliest days almost entirely a product of mathematical research. Mathematics has thus been as essential to the development of nation-wide telephony as copper wire or carbon microphones.

Number of Mathematicians. The Mathematical Research Department of the Bell Telephone Laboratories contains 14 mathematicians. Perhaps an equal number of men scattered through various engineering departments should also be classified as mathematicians according to the definition adopted for this report. Say a total of 25 or 30 for the Bell Laboratories, a few more for the Bell System as a whole, and perhaps 40 or 50 for the entire communication field including the companies interested in radio and television. A few of these men carry on a considerable amount of experimentation, but their significant work is theoretical.

In addition, there is a much larger number of men who use mathematical methods extensively in their daily work, but whose mental type is not that which we have described as mathematical, and who are therefore not included in the numbers quoted above. This is true in particular of the engineers who have the responsibility for designing networks.

Uses of Mathematics. Mathematical activity is most intense: (1) in designing wave filters and equalizers, (2) in studying transmission by wire and ether, the concomitant problems of antenna radiation and reception, inductive interference between lines, etc., (3) in studying various problems related to the standard of service in telephone exchanges, such as the amount of equipment required, the probability of delays and double connections, the hunting time of switches, etc., (4) in providing a rational basis for the design of instruments, such as transmitters and receivers, vacuum tubes, television scanning devices, etc., (5) in developing efficient statistical methods for the planning and interpretation of experiments, and for controlling the quality of manufactured apparatus.

Future Prospects. During the last 20 years the number of men employed in communication research has increased with great rapidity, but this rapid expansion appears to be about over. A large increase in the mathematical personnel of the industry therefore appears unlikely. It seems inevitable that the problems will increase in complexity, and that theoretical methods will become increasingly important, but it is believed that this trend will be matched by progressively better trained engineering personnel, rather than by an increased number of mathematicians. Indeed, unless the qualifications of the mathematicians rise progressively with those of the engineers, it may turn out that less rather than more will be employed.

Electrical Manufacturing

Substantially all the research in the power fields is carried on by a few electrical manufacturers. The power companies usually accept and exploit such equipment as the manufacturers supply, and contribute to improved design principally through their criticisms of past performance. Many of their engineers, however, are individually active in the invention and development of improved equipment.

Number of Mathematicians. The number of mathematicians in the industry is smaller than in communications, and is not easy to estimate because their work is less segregated from other activities. The total number who would here be rated as mathematicians is probably about 20.

As in communications, some are engaged partly in experimental work. There are some, however, whose relationship as consultants is clearly recognized, and there is evidence that management is becoming increasingly conscious of the nature and value of their services.

Uses of Mathematics. Mathematical activity is most intense: (1) in studying structural and dynamic problems, such as the strain, creep and fatigue in machine parts, vibration and instability in turbines and other rotating machinery, etc., (2) in appraising the evil effects of suddenly applied loads, lightning or faults upon power lines, and their associated sources of power, and devising methods to minimize these effects, (3) in studying system performance, particularly the most effective or economical location of proposed new equipment, and the evaluation of performances of alternative transmission or distribution systems, (4) in refining the design of generators, motors, transformers and the like, so as to improve their electrical efficiency and reliability, and in similar improvement of the thermal efficiency of turbines, (5) in the design of miscellaneous instruments and apparatus.

Statistical methods are being introduced into manufacturing and research, but are not yet utilized to the same extent as in telephony.

Future Prospects. The amount of money spent on development in these industries is gradually increasing, and as in other fields the problems are becoming more complex. Hence a slow increase in the number of mathematicians seems probable, with rising standards in the qualifications required, not only as to mathematical training, but as to temperament and personality as well.

The Petroleum Industry

The petroleum industry consists of many producing units of various sizes, highly competitive in character, and surrounded by a number of consulting service organizations, all of which are small. The larger producing companies—and within their resources, the service units also—maintain research laboratories. They tend to be secretive about the developments which take place in these, sometimes to a surprising degree. Hence there is much duplication of effort, particularly in such matters as the design of instruments for geophysical prospecting, and in methods of interpreting the data derived from them.

Number of Mathematicians. The industry employs more mathematicians than is generally appreciated, some of them men of very considerable ability. The total of first-rank men is perhaps 15 or 20. Due to the small size of the individual research staffs, however, most of these men carry considerable project responsibility along with their theoretical work. This is the normal state of affairs in small groups: the abnormality is the lack of contact with, and stimulus from, similar men in other companies.

Uses of Mathematics. Petroleum research extends in three directions: prospecting for oil, producing it, and refining it.

There are five recognized methods of prospecting: gravity, seismic, electric, magnetic and chemical. In the first four, important mathematical problems arise in designing sufficiently sensitive instruments and in interpreting data. The fifth requires the use of statistical methods.

Research on methods of producing a field has led to a few mathematical studies of underground flow, and would undoubtedly give rise to others if the results of these studies could be profitably applied. However, since the rate at which oil is brought to the surface is almost entirely determined by law, and the same is indirectly true of well location also, mathematical consideration of the subject is largely sterile, at least so far as American oil fields are concerned.

The third activity—refining—is essentially a chemical industry. Hence the following remarks by Dr. E. C. Williams, Vice President in Charge of Research of the Shell Development Company, presumably apply not only to the petroleum business, but to manufacturing chemistry in general:

"The two chief problems in chemistry are (aside from the identification of substances): The calculation of chemical equilibrium and the calculation of the rates of attainment of these equilibria. The first problem, involving thermodynamics and statistical mechanics, is rather well understood and usually, by very simple computations, information sufficiently accurate for industrial application, at least, can be found. Frequently, when several equilibria are possible simultaneously, complicated equations arise, but we rarely solve them directly, but rather set up tables of the dependent variable (the per cent conversion possible) as a function of the independent variables (temperature, pressure concentration). The sources of these data, however, are numerous and at times require complicated mathematics, as in the calculation of thermodynamic properties from spectroscopic data via quantum statistics.

The situation is much less favorable in the calculation of the rates of chemical reactions. A semi-empirical method, based on quantum mechanics, has been applied with a little success to some of the simplest reactions taking place in the gas phase, but virtually no progress has been made in the more important field of heterogeneous reactions (reactions of gases on surfaces, for example). We may say that no satisfactory mathematical theory for such calculation exists at the present time. Some progress is being made, but we are far from being able to predict a suitable catalyst for any desired reaction. For the present we are happy to be able to account for observations made on some simple reactions."

Future Prospects. It is inconceivable that research in the industry will not continue at at least its present level. Hence more, rather than less mathematical work will probably be undertaken in prospecting and in refining. A demand of moderate proportions should exist for able mathematicians with a suitable background of geology and classical physics for the geophysical work, and of physical chemistry and molecular physics in the chemical field.

Aircraft Manufacture

The aircraft industry also consists of a number of independent units, and is highly competitive. It is a new industry in which rapid technical development and rapid increase in size has been the rule. It has depended primarily upon government-supported laboratories and, to a lesser extent, upon the universities for its research, and has busied itself with the exploitation of that research in the advancement of aircraft design. No unit of the industry has had or, for that matter, now has a research laboratory, in the sense in which the words would be used in older and larger businesses, but the beginnings of research departments have appeared, and individual researchers and research projects are clearly recognizable.

Numbers of Mathematicians. Some men in the engineering departments of these companies should undoubtedly be classed as mathematicians, but it is impossible to make even an approximate estimate of their number. It is possible, however, to cite pertinent information which bears on the importance of mathematics to the industry.

The design of a modern four-engine transport plane requires about 600,000 hours of engineering time up to the point where complete working drawings have been prepared. About 100,000 hours are spent on mathematical analysis of structures, performance, lift distribution and stability. Most of this work is routine, but some is fundamental in character, as is evident from several of the examples mentioned earlier in this report.

Of 670 men in the engineering department of one of the larger companies, about 25 have mathematical training beyond that usually obtained by engineers, and 10 or so of these are using this advanced training to a significant extent.

Uses of Mathematics. In designing an airplane, five factors are of particular importance. These may be used to indicate the directions in which mathematical research may be expected.

(1) Performance (that is, pay-load, range, speed, climbing rate, etc.)

In the past, forecasts of performance have been based almost entirely on empirical data. Mathematical methods of estimation are now being developed from hydrodynamic theory, however, and are being used to an increasingly greater extent.

(2) Lift and Drag (i.e., the force variation over the wings)

This is the principal objective in the aerodynamic design of the wing. The technique of prediction rests on two supports: wind tunnel experiments and airfoil theory, by means of which experimental data are interpreted and applied. For example, airfoil theory suggests the shape of airfoil to avoid unfavorable pressure distributions and is leading to improved wing sections. This part of aircraft design is already highly mathematical, but a number of fundamental problems still remain unsolved. For example, the theory is still unable to predict stall, and too little is known about optimum shapes or about turbulence, though the recently developed statistical theory of turbulence has contributed to the understanding of the airflow over an airplane and resulted directly in a decrease in airplane drag and consequent improvement in performance.

(3) Stability (inherent steadiness of motion)

The stability of an airplane in flight is inherent in its aerodynamic design and quite distinct from its control or maneuverability. The theory of "small oscillations" has been successfully applied to rectilinear flight. More recently the problem of predicting the response of an airplane to control maneuvers has used the Heaviside operational calculus. Current problems of dynamical stability in which applied mathematicians are interested are the behavior of an airplane when running on the ground and the behavior of seaplanes when running on the water (porpoising).

(4) Structural Safety

Very precise appraisal of structural strength is required in aircraft design. In most industries inaccuracy can be compensated by increased factors of safety, but the pay-load of an airplane is so small a proportion of its total weight that slight increases in factors of safety would seriously reduce its carrying power or even make it unable to get off the ground. Mathematical methods have always been used in this phase of aircraft design in so far as they were available. The standard technique is first to design a part on the basis of calculated strength, then build and test it, and if the tests do not agree with predictions, revise the design and build and test the modified part. This process is continued as many times as necessary to attain a satisfactory result. It is slow and expensive. Theoretical methods are now reliable enough that the majority of structural tests confirm predictions with sufficient accuracy to require no revision. However, new problems constantly present themselves—the introduction of pressurized cabins recently gave rise to several—and hence continual mathematical study is required. A beginning has also been made in the use of the principles of probability in setting up structural loading factors.

(5) Flutter

We have already commented upon the impracticability of studying this phenomenon by any means other than the mathematical. The general equations are complicated, and have only been solved by making important simplifying assumptions. The results are serviceable for check purposes, but need further elaboration. The importance of the problem increases progressively as more efficient planes are designed, and the necessity for an adequate mathematical theory is becoming critical.

Future Prospects. It appears inevitable that from motives of economy the industry will rely increasingly upon theoretical methods of design, and that mathematics will play a larger part in the future than at present. It is also probable that for competitive reasons the various companies will supplement government research by fundamental studies of their own. Furthermore, in view of the present fragmentary state of aerodynamic theory, it would not be surprising if part of the research effort was devoted to the improvement of the basic theory itself.

The reliability of these predictions is, of course, conditioned by the financial prospects of the industry. Just now, war orders are causing abnormal inflation of earnings; when these cease, retrenchment will be inevitable. The industry is not highly mechanized, however, and hence its present cycle of inflation does not imply so large an expenditure for plant as would be true in most manufacturing fields. For this reason, the period of deflation may prove to be one of large war profits in the bank, but insufficient orders to occupy the time of many competent technical men whom the management would be reluctant to let go. If this should occur, an almost explosive development of research may take place.

Whether the development is explosive or not, however, it is probable that the industry will soon become one of the largest employers of industrial mathematicians.

Industrial Statistics and Statisticians

The subject of statistics enters the business world at points quite distinct from those touched by the rest of mathematics. Moreover, the types of business activity to which it most frequently applies—insurance and finance, economic forecasting, market surveys, elasticity of demand against price, benefit and pension plans, etc.—belong to the field of economics which is the subject of a separate report, and need not be touched on here.

There are certain other respects in which statistical theory could be of great service in industry, but they have been exploited to only a limited extent. This report must therefore point out these hopeful fields rather than record achievements in them.

Statisticians in Industry

By "statistician" we mean a person versed in and using the mathematical theory of statistics, not one who collects, charts and scrutinizes factual data. In the business world the word is more often used in the latter sense.

There is a very great difference between the number of statisticians in industry, and the number of men interested in some form of statistics. How great the discrepancy is will be clear from a comparison of the membership of the American Statistical Association, which devotes itself to the application of statistics in its broadest sense, and of the American Institute of Mathematical Statistics, which confines itself narrowly to the development of statistical technique. The former lists 277 names with industrial addresses; the latter only 10.

Statistics in Industry

Dr. W. A. Shewhart, Research Statistician of the Bell Telephone Laboratories, has delineated broadly and succinctly the field in which statistics may be expected to find application as follows:

"Since inductive inferences are only probable, or, in other words, since repetitions of any operation under the same essential conditions cannot be expected to give identical results, we need a scientific method that will indicate the degree of observed variability that should not be left to chance. Hence it appears that the use of mathematical statistics is essential to the development of an adequate scientific method, and that mathematical statistics may be expected to be of potential use wherever scientific method can be used to advantage."

More specifically, there are five recognizable types of industrial engineering activity in which statistical theory either is, or should be used.

(*a*) In studying experimental data to determine whether the observed variations should be regarded as accidental or significant. An example is found in the field of geochemical prospecting. The surface soil overlying regions in which there is oil contains a higher proportion of hydrocarbons and waxes than occur in other locations. Chemical analysis of surface soil therefore affords a means of prospecting for oil. Mr. Eugene McDermott writes:

"In the geochemical method, it was found necessary to determine between samples showing significantly high analysis values, and those which were normal values. These normal sample values, of course, had considerable variation between themselves, due to analysis and in larger part sampling errors. After examining these data for a long period of time, it was decided to approach the problem statistically. This disclosed at once that areas surveyed could be divided into positive (having significant values, and hence favorable from the standpoint of petroleum possibilities), negative (no significant values and unfavorable for petroleum) and marginal (indeterminate). The latter case is always the most difficult one in surveying, and while we are now able to recognize it, further work is needed to fully interpret it. This kind of mathematics is being applied at the present moment, and bids fair to solve the problem."

(*b*) In planning the kind of experiments from which such data arise. Whether variations are or are not significant depends in no small degree upon the fashion in which the data were taken. Consideration of the experiment in advance from a statistical point of view often results in economy of procedure, or even points the difference between a trustworthy and a meaningless result.

The following example is quoted from an address by Dr. R. H. Pickard, Director of the British Cotton Research Association:

"To illustrate the advantage of good experimental design I may refer to some experiments carried out at the Shirley Institute to find the effect of various treatments on a quality of cloth. This quality varies considerably at different parts of the same piece of cloth, and in order to measure the effect of the treatments the tests are repeated systematically so that the variations are 'averaged out.' Some of the natural variation, however, is systematic, and by adopting a 'Latin Square' arrangement of treatments on the cloth (such as is much used in agricultural yield trials), these systematic variations are eliminated from the comparison, and in the instance quoted the result was to reduce by one-half the number of tests necessary for a given significance as compared with a random arrangement."[8]

To the extent to which biology becomes an important element in industrial research—and it would appear to be on the point of doing so in such fields as food manufacturing—it can be expected that the type of statistical work listed under (*a*) and (*b*) will rapidly increase.

(*c*) In laying out an inspection routine. Manufacturing inspection frequently yields data which are best interpreted statistically, either because only spot-checks are taken, or because the method of inspection gives measurements which are themselves subject to accidental fluctuation. In such cases statistical theory is of great advantage in setting up an effective and economical inspection program. It is being so used in certain industries, notably in electrical manufacturing and textiles, but the potential field of usefulness is far from covered.

The following example is quoted from an address by Mr. Warner Eustis, Staff Officer on Research of the Kendall Company:

"Surgical sutures are twisted strands of sheep intestine, which has been slit lengthwise. . . . After a stated number of days a sewing with such material, implanted in the body during a surgical operation, will be digested and disappear as the healing processes progressively take up the load originally held by the suture. . . . Here is a product which it is impossible to test in any way without destroying the product, especially as each suture is sealed in an individual, sterilized tube. Our final product tests must all be conducted by breaking open a sterile tube and testing the product therein. The quality appraisal of such a product naturally rests upon probability, rather than upon an actual testing of each item. Due to the nature of such a product, in which a single failure may destroy human life, the need for accurate quality appraisal is superlative."[9]

(*d*) In the control of manufacturing processes. Inspection is not merely a means of discarding bad product; it is also a means of detecting trouble in the factory. This is obvious in the extreme cases when the product is

[8] "The Application of Statistical Methods to Production and Research in Industry," by R. H. Pickard, Supplement to the *Journal of the Royal Statistical Society*, Vol. 1, No. 2, 1934, pp. 9-10.

[9] "Why the Kendall Company is interested in Statistical Methods," by Warner Eustis, *Proceedings of the Industrial Statistics Conference* held at M. I. T., Cambridge, Mass., Sept. 8-9, 1938, pp. CXLIII-CXLIV, published by Pitman Publishing Corporation.

unusually bad. By the use of suitable routines set up in accordance with statistical theory, the day-to-day results of inspection can be used to detect incipient degradation in the process of manufacture which might otherwise escape notice. This procedure is used extensively by the Western Electric Company in assuring uniform quality in many items of manufacture, and to a lesser extent in other industries. Of it, Mr. J. M. Juran, Manufacturing Engineer of the Western Electric Company, says:

"Too frequently we have seen an inspection group grow lax in vigilance until a complaint from the customer wakes them up. They promptly swing the pendulum a full stroke in the opposite direction, and the factory groans in its effort to meet the now unreasonable demands. A sound and steady control, like a sound currency in commercial relations, gives factory foremen a feeling of confidence and gives the consumer a feeling that control is being exercised before the product reaches him."[10]

(*e*) In writing rational specifications. Obviously, if such a procedure helps the manufacturer to assure uniform quality, it is also of value to the purchaser of his products. Hence the subject of statistics enters into the writing of the buyer's specifications. It has been so used to a limited extent in the Bell System in connection with telephone apparatus, and by the United States Government in the purchase of munitions. However, it must still be rated as a relatively undeveloped field. Of it, Captain Leslie E. Simon, Ordnance Department of the United States Army, says:

"Statistical methods have proved to be a powerful tool in the critical examination of some ammunition specifications prior to final approval. Their use, either directly or indirectly, is almost essential in determining a reasonable and economic standard of quality through the method of comparing the quality desired with that which can be reasonably expected under good manufacturing practice. In like manner, the statistical technique renders a valuable service in framing the acceptance specification. Through its use the quantity and kind of evidence which will be accepted as proof that the product will meet the standard of quality can be clearly expressed in a fair, unequivocal and operationally verifiable way."

Conclusion

It is perhaps unusual to conclude a survey of this sort by stating the impressions which it has made upon its writer. In the present instance, however, the element of self-education has been so large that these impressions may summarize the report better than any more formal recapitulation. They are:

[10] "Inspectors' Errors in Quality Control," by J. M. Juran, *Mechanical Engineering*, Oct., 1935, pp. 643–644.

(1) Because of its general significance as the language of natural science, mathematics already pervades the whole of industrial research.

(2) Its field of usefulness is nevertheless growing, partly through the development of new industries such as the aircraft business, and partly through the incorporation of new scientific developments into industrial research, as in the application of quantum physics in chemical manufacturing and statistical theory in the control of manufacturing processes.

(3) The need for professional mathematicians in industry will grow as the complexity of industrial research increases, though their number will never be comparable to that of physicists or chemists.

(4) There is a serious lack of university courses for the graduate training of industrial mathematicians.

(5) Management, which is already keenly alive to the importance of mathematics, is also rapidly awakening to the value of mathematicians and the peculiar relationship which they bear to other scientific personnel.

This last observation is not trivial. There was a day when, in engineering circles, mathematicians were rather contemptuously characterized as queer and incompetent. That day is about over. Just now, an attitude more commonly met is one of amazed pride in pointing to some employee who "isn't like most mathematicians; he gives you an answer you can use, and isn't afraid to make approximations." As the proper function of the industrial mathematician becomes better understood, these proud remarks will no doubt cease. Those who are adapted to the job will be taken for granted; the others will be recognized as personnel errors and not mistaken for the professional type. Perhaps the present report may speed this day. If so, it will have been a service to the profession and to industry.

The Transmission Characteristics of Toll Telephone Cables at Carrier Frequencies

By C. M. HEBBERT

IN THE design of a new telephone transmission system a knowledge of the characteristics of the medium over which the waves are to pass is, of course, a prerequisite. What painstaking experimentation is necessary to accumulate such knowledge, however, what voluminous data are involved, what minutiae of detail, and what extremes of accuracy, are things far less obvious.

Recent papers have described a new 12-channel carrier telephone system for operation over cable pairs.[1] For this system a knowledge of the maximum cable losses is needed in order to determine the necessary repeater gains. Accurate data on the insertion loss slope versus frequency are required so that compensating equalizers can be designed to give uniform transmission over the frequency band. In order to design a regulating system to compensate for the variations in attenuation which result from changes in cable temperature, precise knowledge of these variations as a function of frequency is essential. It is necessary to know the impedance of the cable pairs in order that the amplifier impedance may be matched to it, thereby avoiding reflections which would aggravate cross-talk effects. For various purposes, e.g., testing the cables, designing the coils to balance out crosstalk, etc., it is also necessary to know the fundamental parameters (resistance, inductance, capacitance and conductance) or so-called primary constants of the pairs. The velocity of transmission also plays a part in determining the characteristics of the channels. In addition to all these transmission characteristics, it is, of course, essential to know the crosstalk couplings between different pairs. This subject has been treated elsewhere[2], however, and is not considered herein.

In order that the cable carrier systems may be applied in the plant without requiring extensive transmission measurements on each individual carrier pair in each repeater section, it is important that the differences in the transmission characteristics between different pairs be known. The problem therefore becomes one of statistical analysis. In most cases the

[1] "A Carrier Telephone System for Toll Cables," C. W. Green and E. I. Green, *B.S.T.J.*, Vol. 17, January 1938, page 80. "Experience in Applying Carrier Telephone Systems to Toll Cables," W. B. Bedell, G. B. Ransom and W. A. Stevens, *B.S.T.J.*, Vol. 18, October 1939, page 547.

[2] "Crosstalk and Noise Features of Cable Carrier Telephone System," M. A. Weaver, R. S. Tucker and P. S. Darnell, *B.S.T.J.*, Vol. 17, January 1938, page 137.

effects involved are cumulative with distance and the accuracy involved in the determination of the various characteristics is therefore set by the maximum distance over which the system is designed to operate. For a distance of 4000 miles the total loss at the top frequency of 60 kc. will be approximately 16,000 db, the attenuation difference between the top frequency of 60 kc. and the bottom frequency of 12 kc. nearly 6000 db if the cable is at about the average temperature, 55°F. The range of variation in loss with temperature, assuming aerial cable over the whole distance, will be about ±8 per cent of the total at 60 kc. It is desired to correct these frequency differences in loss and variations with temperature so accurately that individual channels will be constant to within ±2 db.

Prior to the beginning of experimentation with cable carrier systems limited use had been made, in connection with carrier systems operated over open-wire lines at frequencies up to 30 kc., of conductors in relatively short entrance and intermediate cables. The available data, however, were quite inadequate for the cable carrier problem. Accordingly, an extensive series of tests was undertaken. Reels of standard toll cable were placed in temperature controlled rooms where the extreme temperature variations of the mid-west could be substantially duplicated (the actual laboratory temperatures ranged from just below 0° F to 120° F) and measurements were made to determine the changes in the parameters of the cable accompanying these wide temperature variations at frequencies from 1 kc. to 100 kc. and higher in some cases. Certain of the tests even studied the effect of varying the humidity content of the cables. Further measurements were then made on suitable lengths of pairs in actual commercial cables in which carrier systems were to be installed. These results corroborated and extended the data from the laboratory measurements; the subsequent operation of equalizers, regulators, etc., based upon these data, showed no essential discrepancies.

The present paper, after referring to the types of toll cables employed for the new carrier systems, outlines the methods employed in determining their characteristics both in the laboratory and in the field, summarizes these characteristics for typical 19-gauge cable at frequencies up to 100 kc. and finally extends them to frequencies as high as 700 kc. for 16 and 19-gauge cables.

Types of Carrier Toll Cables

The type K carrier system has been designed so that it may be applied to existing cables, thus in many cases avoiding the installation of expensive new cables. Most of the standard toll cable in the Bell System contains chiefly 19-gauge paper insulated conductors in "multiple twin quads," i.e., two conductors are twisted together to form a pair and two pairs twisted together to make a quad. The nominal capacitance of a pair is .062 mf. per mile. There are various twist lengths of both pairs and quads in a given cable as well as cables ranging in size from 12-quad cable to the oversize 19-gauge cable containing 225 quads. Some type K is operating over "paired" cable, i.e., cable in which only the wires of each pair are twisted together.

Operation in two directions is accomplished by using either a separate cable for each direction or a single cable with two groups of conductors separated by a layer shield. This avoids serious near-end crosstalk effects which would result from the large level difference existing between opposite directions at a repeater point.

Methods of Measurement

As mentioned above, 250-foot reels of standard toll cable were placed in a special room which could be accurately maintained at any desired temperature from about zero to 120 degrees, Fahrenheit, and measurements made for various frequencies and temperatures. For the most part, these consisted of open-circuit admittance and short-circuit impedance measurements on part of the pairs in the cable at temperatures of about zero, 30, 50, 90 and 120 degrees F., over the frequency range from 4 to 100 kc. From these measurements computations could then be made to determine the resistance, inductance, capacitance and conductance as well as the attenuation, phase constant and characteristic impedance of this type of cable at the different temperatures and frequencies. Detailed data on frequency and temperature variations of these quantities are given below. Most of these data were obtained from measurements made on 16- and 19-gauge pairs in a typical reel-length of standard toll cable. The temperature is difficult to maintain at a constant level and d-c. resistances of certain pairs were measured at frequent intervals during the process in order to get accurate temperature readings by comparing with resistance-temperature curves of these pairs. Thermocouples were also attached to the cable at various points along its length and sheath temperatures determined from them. After stabilizing the room temperature as closely as possible, the variations in cable temperature took place slowly enough to be allowed for in the computations.

After the selection of the Toledo-South Bend route for a trial installation, further measurements were made on certain of the pairs in this cable. The test sections, extending out of Lagrange, Indiana, were made about 10 miles long in order to obtain the averaging effect of length. For this distance it was not possible to use open and short-circuit measurements as was done in the laboratory, and a substitution method[3] was devised (Fig. 1).

[3] This was devised by H. B. Noyes and will be described by him in a paper in the *Bell Laboratories Record*.

This consists essentially of first measuring the input and output a-c. currents at the two ends of the test pair by means of thermocouple-milliammeter arrangements and then immediately sending d-c. over another pair (called reference pair in Fig. 1) built out to a convenient fixed d-c. resistance, the same for all measurements, and adjusting resistance networks at both ends of the line until the meter readings are the same as for the a-c. Suitable charts then enable readings of attenuation (insertion loss) corresponding to the d-c. (and therefore also to the a-c.) readings to be made very rapidly.

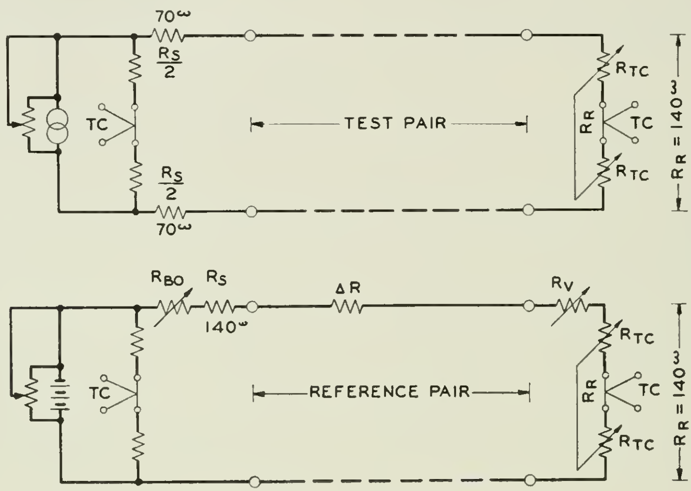


Fig. 1—Simplified attenuation measuring circuit

Toll Cable Characteristics Below 100 Kc.

Primary Line Parameters

The four primary line parameters, R, L, G and C—series resistance, series inductance, shunt leakage and shunt capacitance—are of the same sort for all kinds of transmission lines, but the relative importance of the various elements changes considerably with frequency and the type of structure considered. The old name primary "constants" is obviously a misnomer, and it is simpler to speak of them as line "parameters," since this does not necessarily imply anything regarding their constancy or inconstancy under various conditions.

The "true" or distributed values of these parameters are usually obtained from measurements of the open circuit admittance, $G' + j\omega C'$, and the short-circuit impedance, $R' + j\omega L'$ of the actual pair in a short length of cable, by means of the following formulas:

$$R \text{ (resistance)} = R'(1 - \tfrac{2}{3}\omega^2 L'C') + \cdots$$
$$L \text{ (inductance)} = L'(1 - \tfrac{1}{3}\omega^2 L'C' \cdots) + \tfrac{1}{3} R'^2 C' + \cdots$$
$$G \text{ (conductance)} = G'(1 - \tfrac{2}{3}\omega^2 L'C' \cdots) - \tfrac{1}{3} R'\omega^2 C'^2 + \tfrac{2}{5} R'\omega^4 L'C'^3 \cdots$$
$$C \text{ (capacitance)} = C'(1 - \tfrac{1}{3}\omega^2 L'C' + \tfrac{2}{3} R'G' \cdots) \quad (1)$$

These formulas give accuracies within one per cent for reel-lengths of 500 feet or less and frequencies up to 100 kilocycles for 19-gauge cables having a capacity of .062 mf per mile. All the curves of R, L, G, C herewith are based on true values obtained from such computations.

Resistance

The quantity R, series resistance in ohms per mile, has a large variation with frequency produced by the well-known phenomenon called skin effect and another large increment, resulting from the closeness of the wires in cables, known as the proximity effect.[4-7]. The magnitude of the proximity effect varies with the diameter of the conductors as well as with their separation. The curves in Fig. 2 show the increment in resistance resulting from skin effect and the total increase including proximity effect as computed for a pair of wires separated by various multiples of their diameters. The abscissa, B, in Fig. 2 is a sort of universal parameter used in data on skin effect so that a single curve will suffice for various gauges. If f is frequency in cycles per second and R_0 is the d-c. resistance for 1000 feet of the wire (not a 1000-foot loop), the parameter B is given by the equation

$$B = \sqrt{f/R_0} \doteq \sqrt{f/8} \quad (2)$$

for 19-gauge wire so that $B = 80$ corresponds to 51,200 cycles. According to the curves at $B = 80$ (51.2 kc), the skin effect increases the a-c. resistance to about 12 per cent more than the d-c. resistance.

For a separation of two diameters between centers of the wires of a pair ($k = .25$) the proximity effect adds another 6 per cent to the resistance ratio making the total a-c. resistance about 1.18 times the d-c. resistance at 51.2 kc. If the wires are closer together ($k = .4$) the a-c. resistance is computed to be about 1.30 times the d-c., which is about the ratio actually measured. The effects caused by the presence of the adjacent pair in a

[4] J. R. Carson, "Wave Propagation over Parallel Wires—The Proximity Effect," *Phil. Mag.*, Vol. 41, April 1921, pp. 607–633.
[5] A. E. Kennelly, F. A. Laws and P. H. Pierce, "Experimental Researches on Skin Effect in Conductors," *A.I.E.E. Trans.*, Vol. 34, Part 2, 1915, pp. 1953–2021.
[6] A. E. Kennelly and H. A. Affel, "Skin Effect Resistance Measurements of Conductors at Radio Frequencies," *I.R.E. Proc.*, Vol. 4, No. 6, Dec. 1916, pp. 523–574.
[7] Günter Wuckel, "Physics of Telephone Cables at High Frequencies," *EFD* 47, (Nov. 1937) pp. 209–224.

quad, the surrounding wires and the lead sheath are not included in these computations.

These values assume a temperature of 20° Centigrade (68° Fahrenheit) but if the temperature varies, so also does the resistance. Figure 3 shows the a-c. temperature coefficient of resistance and its variation with temperature for 19-gauge pairs in ohms per ohm per degree, Fahrenheit, i.e.,

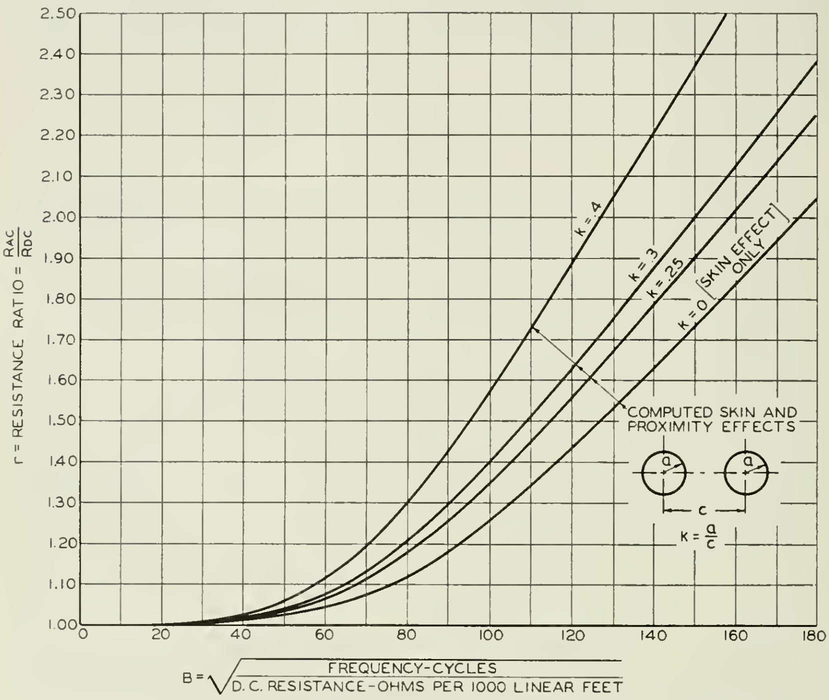


Fig. 2—Skin effect and proximity effect on a.-c. resistance of toll cable pairs

$$A = \frac{1}{R_1}\frac{dR}{dt} \tag{3}$$

where A is the a-c. temperature coefficient of resistance of copper at t_1 degrees, Fahrenheit. The a-c. resistance R at temperature t is given by the formula

$$R = R_1\,[1 + A(t - t_1)] \tag{4}$$

where R_1 is the a-c. resistance at temperature t_1 degrees, Fahrenheit. The coefficient A decreases with increasing frequency, but not indefinitely; it approaches 1/2 the d-c. coefficient as its asymptotic limit with frequency.

The normal variation of air temperatures in the middle western part of the country is from about 1° Fahrenheit, (−17° Centigrade) to plus 109° Fahrenheit (43° Centigrade). Extremely hot summers like that in 1936, which was preceded by a severe winter, show even higher temperatures and there are occasional periods in mid-western winters when the temperature hovers continuously around −30° F., for a week or two. These temperatures are almost the temperatures assumed by open wires, but wires inside a lead sheath like those in an aerial cable are subjected to much higher than air temperatures in hot weather when exposed to direct sunlight in the absence of wind. Some observations made at Lagrange, Indiana, in 1936

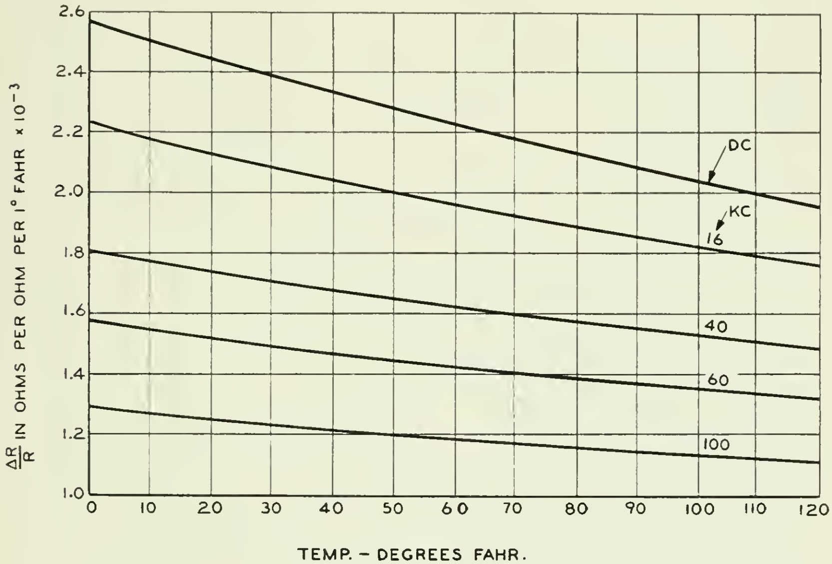


Fig. 3—Resistance-temperature coefficient—$\frac{\Delta R}{R}$ in ohms per ohm per 1° fahr.-19 gauge cable

on d-c. resistance of cable pairs showed that temperatures in the ten miles of cable averaged about 122° Fahrenheit (50° Centigrade) when the air temperature read on thermometers was about 104° Fahrenheit (40° Centigrade). Similar data taken at Chester, New Jersey, showed temperatures in the cables 20° to 25° Fahrenheit higher than the air temperature on hot bright days.

The actual observed cable temperature range in that season (1936) as indicated by the d-c. resistance measurements was thus from −4° Fahrenheit (−20° Centigrade) to 122° Fahrenheit (50° Centigrade). In terms of a-c. resistance changes, this amounts to a resistance change of about 20 ohms per mile at 50 kc., the resistance at the lower temperature being about 96 ohms per mile and at the higher temperature about 116 ohms per mile. This amounts to about ± 10 per cent variation from the mean.

In addition to the wide annual variations, there are daily variations of as much as 50° Fahrenheit at times, that is, almost half as much as the normal annual variation. The practical importance of these large resistance changes lies in their large contribution to changes in attenuation as will be brought out more fully in connection with variations of attenuation with temperature.

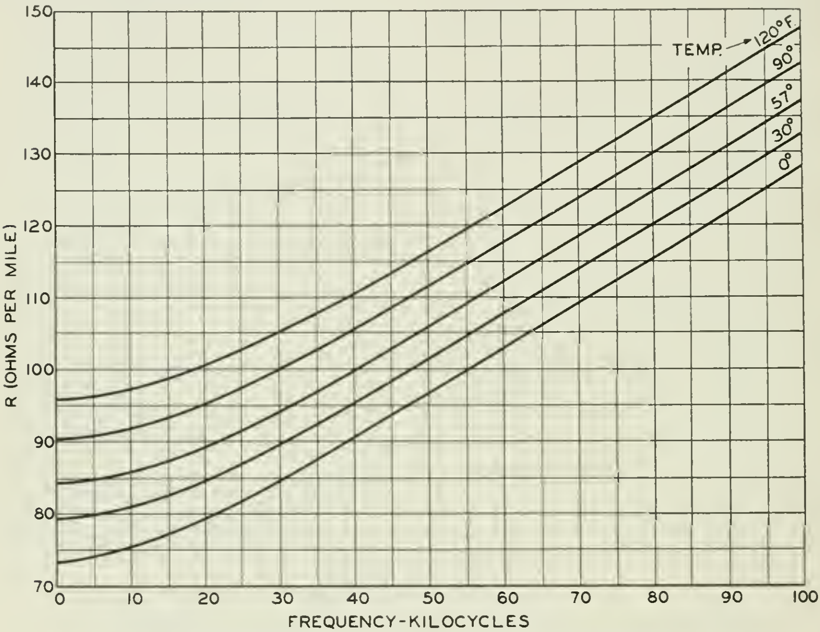


Fig. 4—Resistance per mile vs. frequency—19 gauge pairs

Underground and buried cable are, of course, not subjected to such wide annual variations and daily variations are almost entirely eliminated by the attenuation of heat changes by the soil. Cable in ducts usually lies well below the freezing line and this depth at the same time protects it from the summer's heat. The normal range for cable in ducts is from about freezing to about 70 degrees, F. Cable buried only a foot or so underground would have a considerably larger annual temperature range but a great deal of such cable is buried two to three feet deep.

Curves in Fig. 4 show the actual a-c. resistance variation with frequency and in Fig. 5 are shown temperature variations of resistance at typical frequencies for 19-gauge toll pairs in a reel-length of standard toll cable.

In addition to the variations with frequency and temperature there are the initial differences between pairs on account of manufacturing processes,

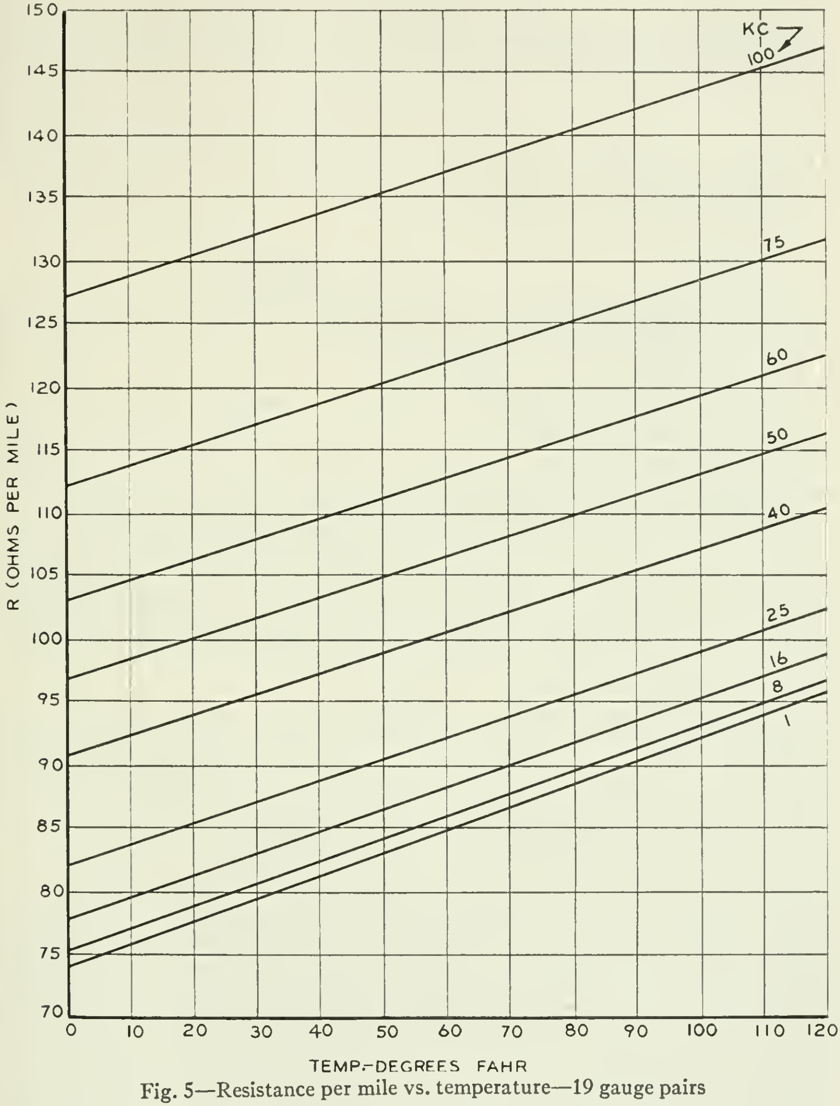


Fig. 5—Resistance per mile vs. temperature—19 gauge pairs

etc. One such source of variation in resistance of pairs is the difference in wire diameters caused by wear of the dies used in drawing wire.[8] The

[8] John R. Shea and Samuel McMullan, "Developments in the Manufacture of Copper Wire," *B.S.T.J.*, VI (April 1927), pp. 187–216.

permissible variation in diameter for ordinary toll cable is $\pm$ 1 per cent which means a d-c. resistance variation of about $\pm$ 2 per cent.

Still another cause of resistance variation is the presence of small quantities of impurities in the copper which show up as a reduction of as much as 2 per cent in the conductivity. This causes trouble in calibrating temperature-resistance curves in the laboratory setup.

Finally, in a single reel length the outside pairs are longer than the inside pairs. The total pair-to-pair variation in resistance from the average of the reel caused by all these factors amounts to about $\pm$ 3 per cent with a standard deviation of about 1.5 per cent.

Inductance

The inductance of a circuit formed by two parallel wires closely spaced relative to their length is

$$L = 0.64374\left[2.3026 \log_{10} \frac{2D}{d} + \mu\delta\right] \times 10^{-3} \text{ henrys per loop mile} \qquad (5)$$

where d, the wire diameter, and D, the separation of the wires, are measured in the same unit; μ is the permeability, and δ is a frequency factor.

As is well known, the tendency of alternating currents to concentrate on the surface of a wire reduces the magnetic flux within the wire and decreases the internal inductance of the wire. This internal inductance is given by the term $\mu\delta$ in Equation (5). In like manner, the "proximity effect" produces a concentration of current density in the adjacent portions of the two wires of a pair.

Another term might well be added to formula (5) to represent this proximity effect. The procedure outlined by J. R. Carson on pages 625 and 626 of the *Philosophical Magazine* paper[4] of 1921 has been carried out with the results given in an Appendix to this paper. Formula (11a) of the Appendix gives the ratio, K, of the a-c. inductance of the pair (less the "geometric inductance") to the a-c. inductance of a wire with concentric return, which is given by a well-known formula (7a in the Appendix). It will be seen that the factor introduced by proximity effect decreases with frequency but is asymptotic to a definite value, depending upon the separation of the two wires, as the frequency increases indefinitely. Similar curves are given in an extensive study of the mutual inductance of four parallel wires of a quad by R. S. Hoyt and Sallie Pero Mead[9]. Their theoretical studies agree closely with experimental values given by R. N. Hunter and

[9] Ray S. Hoyt and Sallie Pero Mead, "Mutual Impedances of Parallel Wires," *B.S.T.J.*, XIV (1935), pp. 509–533.

R. P. Booth[10] who made measurements on 18-gauge and 20-gauge pairs in various arrangements and on a 55-foot length of 19-gauge quadded cable.

Overall inductance variations of 19-gauge pairs with frequency and temperature are shown by the curves of Figs. 6 and 7.

The magnitude of inductance variations from pair to pair in reel lengths of cable is about ± 3 per cent from the mean with a standard deviation of about 1.5 per cent.

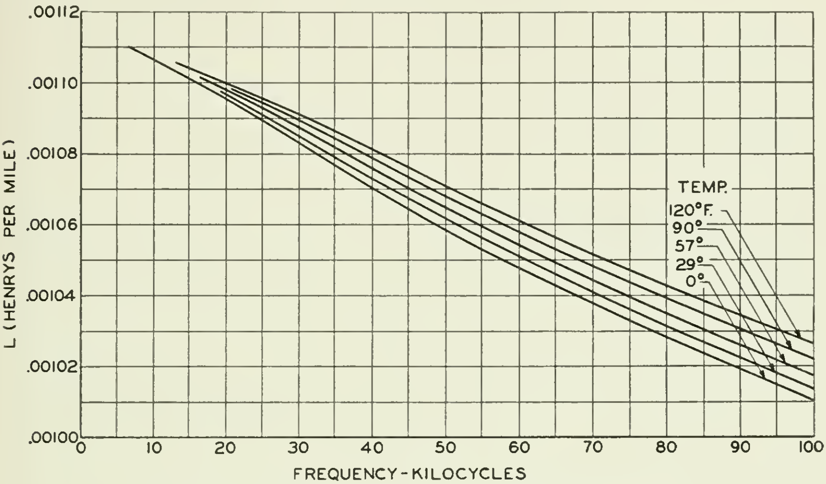


Fig. 6—Inductance per mile vs. frequency—19 gauge pairs

Capacitance

The usual formula for capacitance of two parallel wires in space, separated by a distance negligible compared with their length, is

$$C = \frac{0.019415}{\log_{10} \dfrac{2D}{d}} \times 10^{-6} \text{ farads per loop mile} \tag{6}$$

Conditions in a cable are vastly different from those assumed in this formula which assumes that the two wires are at a great distance from other wires and from the ground. In the cable, pairs are twisted and, in addition, other wires are very near and the sheath is effectively at ground potential, resulting in a considerable modification of the capacitance. Moreover, the formula assumes that the wires are in air, which has a dielectric constant almost equal to unity. (1.00059 at 0° Centigrade) The dielectric constant of

[10] R. N. Hunter and R. P. Booth, "Cable Crosstalk—Effect of Non-Uniform Current Distribution in the Wires," *B.S.T.J.*, XIV (1935) pp. 179–194.

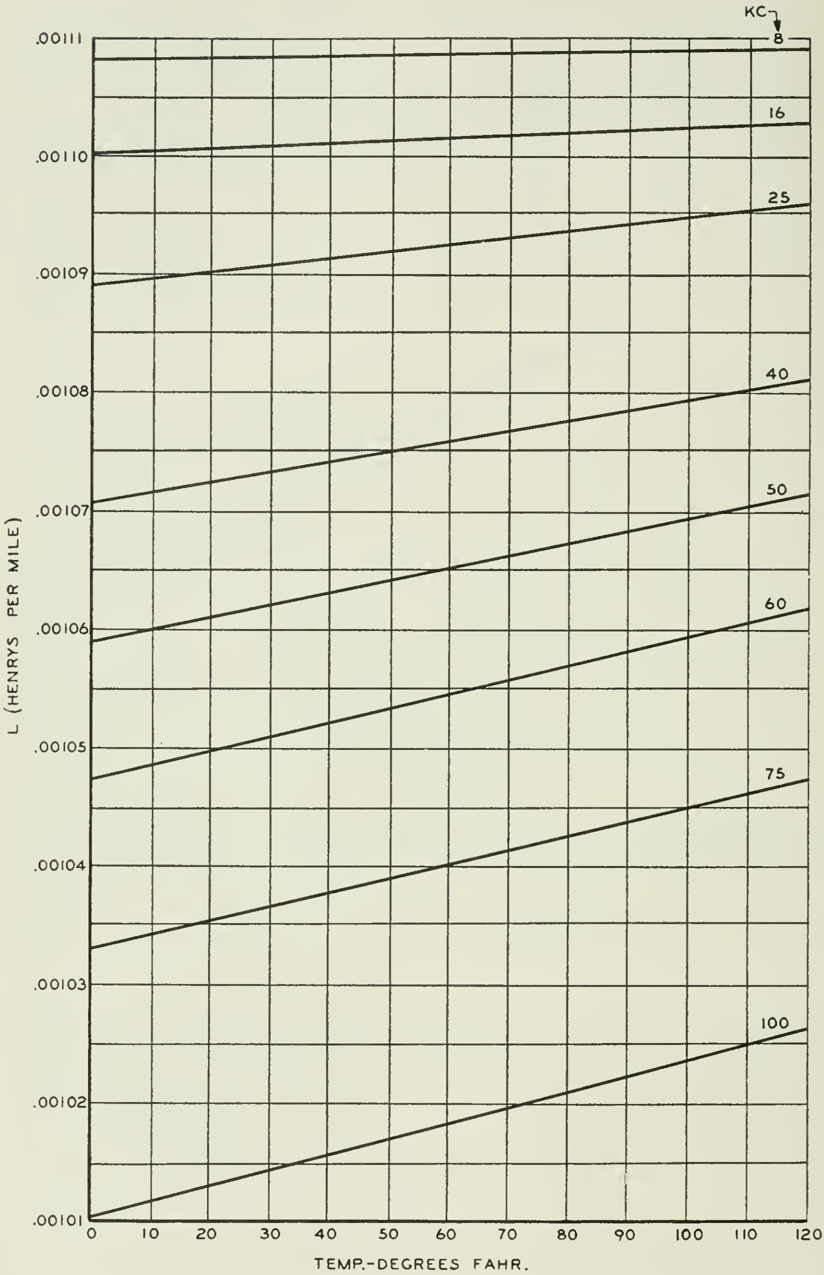


Fig. 7—Inductance per mile vs. temperature—19 gauge pairs

the paper in cables varies from 1.7 to 1.9 depending upon the amount of air and impurities contained in the paper. Of the space around the wires inside the sheath about 40 per cent is filled with paper and the remaining 60 per cent with air. Frequency and temperature of the cable affect the true dielectric constant in a complicated way. Slight amounts of moisture remaining in the cable even after drying affect the dielectric constant and the capacitance as well as the leakage conductance and introduce further changes in the frequency-temperature characteristics.

Results of an extensive study of the dielectric constant were given by E. J. Murphy and S. O. Morgan in a series of recent papers[11]. They point out (first paper, p. 494; second paper, p. 641) that a dielectric may be thought of as an assemblage of *bound charges*, that is charged particles which are so bound together that they are not able to drift from one electrode to the other under the action of an applied electric field of uniform intensity. But the applied field disturbs the equilibrium of the forces acting on the bound charges and they take up new equilibrium positions, thereby increasing their potential energy when the applied field is removed. Then when the applied field is removed, some of this energy is dissipated as heat in the dielectric. If the applied field is alternating, the bound charges swing back and forth with certain amplitudes and the sum of the product of the amplitude by the charge extended over all the bound charges in a unit volume determines the *dielectric constant* of the material. The energy dissipated as heat by the motions of the bound charges is the *dielectric loss*, which is proportional to the a-c. conductivity after the d-c. conductivity has been subtracted from it.

Considering the fact that positive and negative charges will be displaced in opposite directions and such a motion constitutes an electric current, there is thus what is called a *polarization current* or *charging current* flowing while the polarization (or displacement of charges) is being formed. If the current alternates too rapidly for the polarization to form completely before the field reverses its direction, the magnitude of the dielectric polarization and the dielectric constant will be reduced. The result of this lag, therefore, is that the dielectric constant (and likewise the capacity) decreases with increasing frequency. This is the phenomenon known as anomalous dispersion from its relation to the anomalous dispersion of light, i.e., at visible frequencies.

A further important concept in dielectric theory is that the molecules of all dielectrics except those in which the positive and negative charges are symmetrically located, possess a permanent electric moment characteristic

[11] E. J. Murphy and S. O. Morgan, "The Dielectric Properties of Insulating Materials," *B.S.T.J.* XVI (1937) pp. 493–512; XVII (1938) pp. 640–669; XVIII (1939) pp. 502–537. These are referred to as "First Paper," etc.

of those molecules. These polarized molecules are called *dipoles* and when an electric field is applied the dipole axes tend to line up in the direction of the applied field. It is probable that for a combination dielectric such as the paper and air in cables with possible traces of moisture, in spite of oven-drying, the dipoles constitute only part of the charges. The frequency also is too low, in most of the data, to emphasize the effects due to dipoles. The paper-air combination introduces another slowing up of the polarization process on account of *interfacial polarization.* Maxwell showed that if the dielectric in a condenser consisted of two layers of materials having different constants, the capacity depends upon the charging time because of time required in charging the interface between the two dielectrics. For a-c. this means a decreasing capacity with increasing frequency and, since there

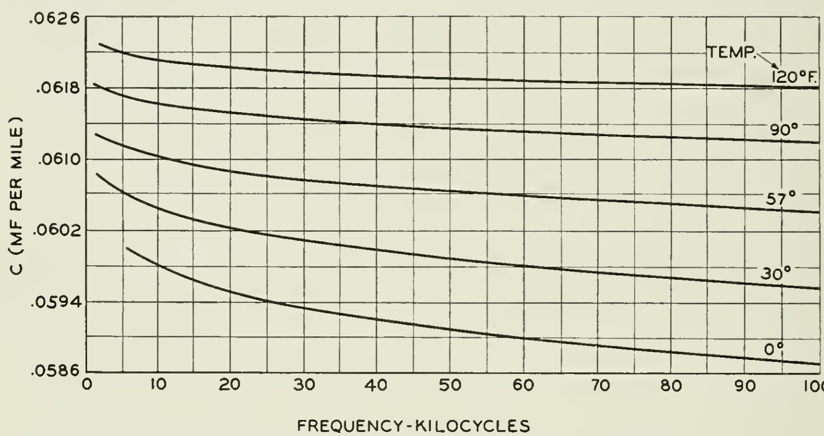


Fig. 8—Capacitance per mile vs. frequency—19 gauge pairs

are effectively an immense number of interfaces between paper and air in the cable, this effect must be of some importance.

Increasing the temperature increases the thermal energy of the molecules and their consequent thermal motion which helps maintain the random orientation of the molecules. Thus, the thermal motion opposes the action of the electric field in maintaining the alignment of the dipoles so that as the temperature rises, the polarization is reduced. But in the cable there are unequal expansions of the copper and the lead sheath which may act to increase the internal pressure as the temperature rises, increasing dielectric densities as well as bringing the wires closer together.

The final result of all these effects on the capacitance of the cable pairs is shown by the curves of Fig. 8, which give the 19-gauge capacitance-frequency relations for several temperatures. Figure 9 shows the variation of capacitance with temperature for several frequencies. The largest change

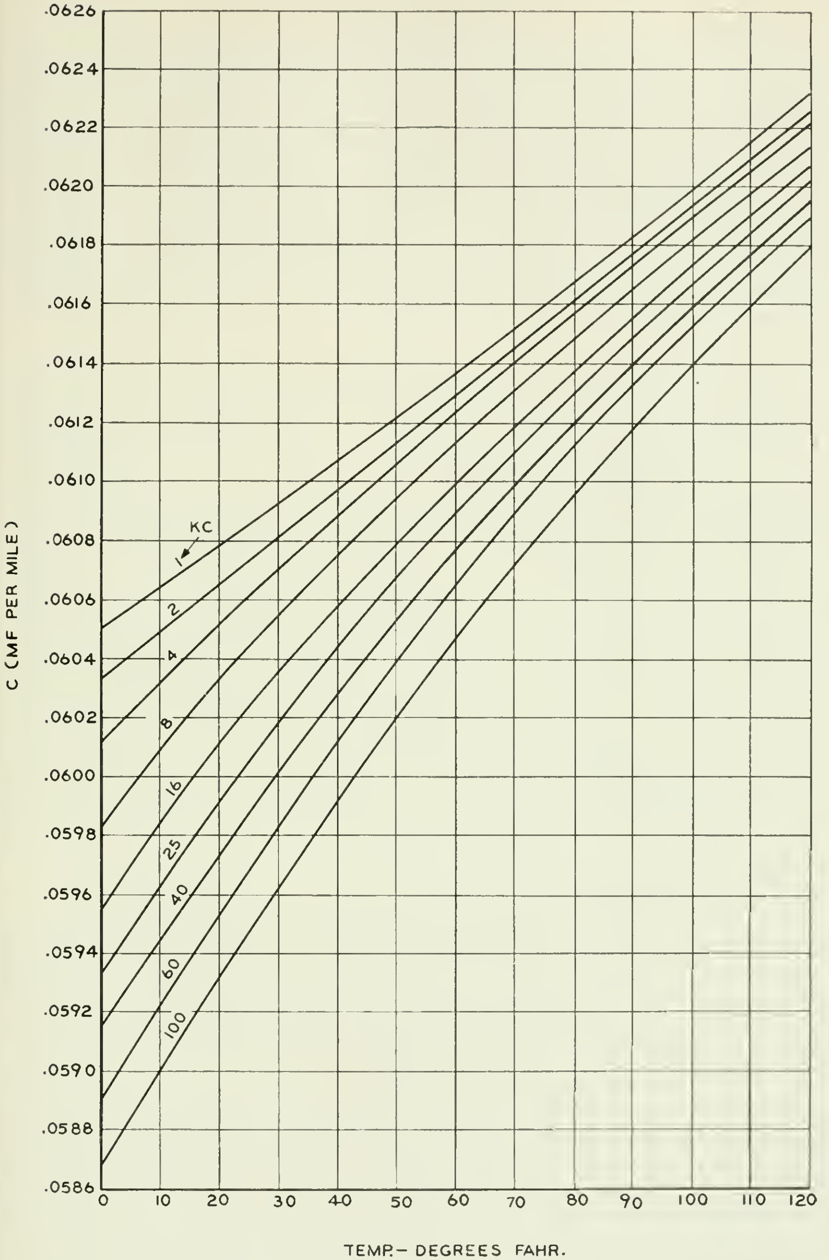


Fig. 9—Capacitance per mile vs. temperature—19 gauge pairs

shown is at 100 kc. and amounts to about 6 per cent increase for 120 degrees increase in temperature.

Leakage Conductance

The variation in the dielectric constant of the insulating layers of paper is further reflected in the leakage conductance, G. This is probably the most inconstant of the parameters and is a function of separation of the wires and their diameters, as well as frequency and temperature and the nature of the dielectric. Humidity, if present, is a highly important contributor to high leakance, but in properly dried cables the humidity is not very great. It will be seen in the later discussion of attenuation and the factors affecting it that conductance is a much less important factor, relatively, than it is for open-wire lines.

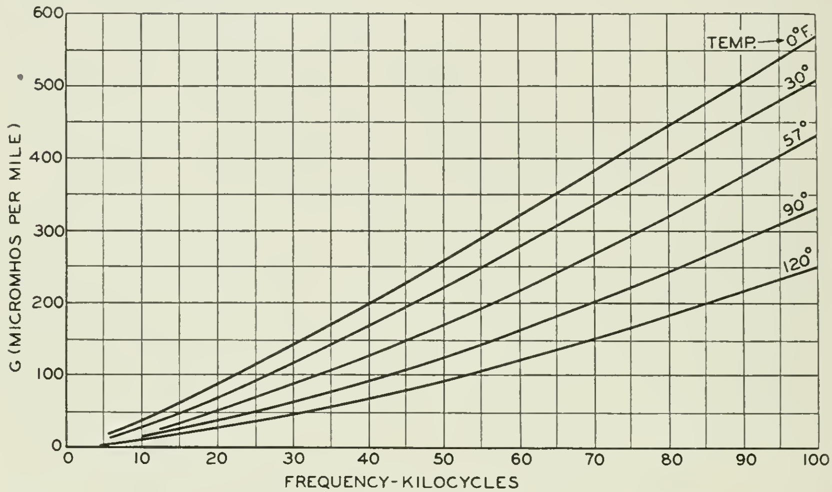


Fig. 10—Conductance per mile vs. frequency—19 gauge pairs

The curves of Fig. 10 show the variation of conductance of 19-gauge pairs with frequency at five temperatures from zero to 120° F. When plotted on log-log paper these curves are nearly linear, showing that conductance varies with frequency approximately according to a formula of the type

$$G = aF^k \tag{7}$$

where a is about $.0001 \times 10^{-6}$ and k is about 1.33 for the 57° data. F. B. Livingston in a paper[12] on conductance in cables stated that for the data there given k averaged about 1.3.

The range of conductance from pair to pair in a reel is about ± 11 per cent from the average and the standard deviation about 5.5 per cent.

[12] F. B. Livingston, "Conductance in Telephone Cables," *Bell Laboratories Record*, Vol. XVI (Dec. 1937) p. 141.

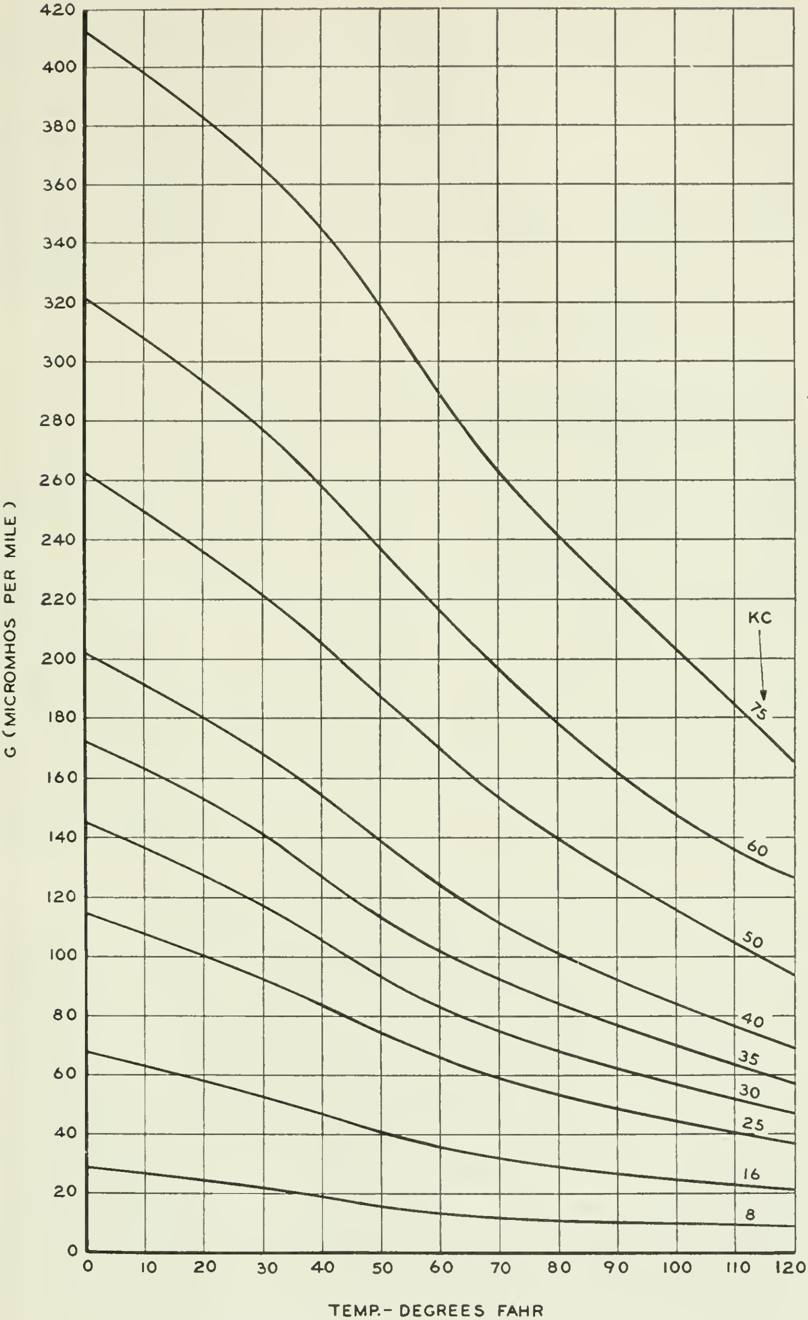


Fig. 11—Conductance per mile vs. temperature—19 gauge pairs

As might be expected, temperature has a great deal to do with the value of G. Variations with temperature shown by the curves in Fig. 11 may be expressed for small temperature ranges by the equation

$$G = G_1 [1 + k(t - t_1)] \tag{8}$$

where G_1 is the value of G at the temperature t_1 and k is the temperature coefficient of leakage conductance. Curves of k based on measurements on a 61-pair, 16-gauge cable are given in Fig. 12 in the neighborhood of 70 degrees Fahrenheit. It will be noticed that k is negative below 1200 kc. but at high frequencies the coefficient increases rapidly from its minimum value reached at about 500 kc.

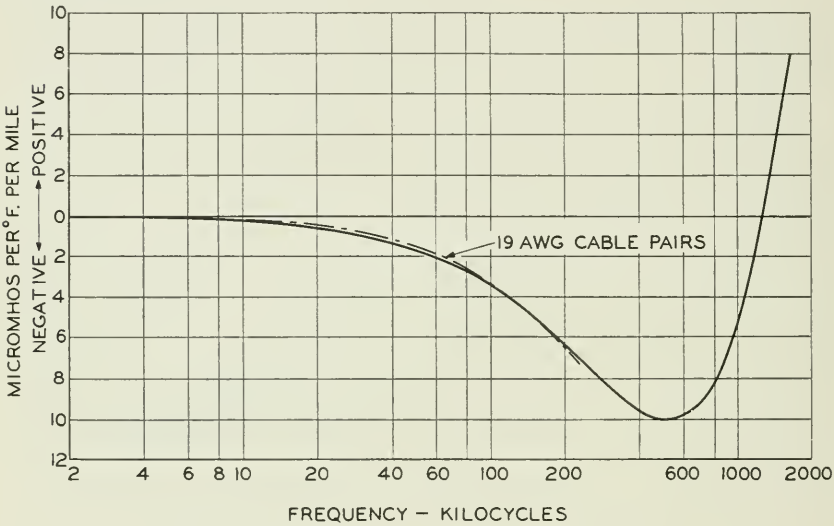


Fig. 12—Conductance-temperature coefficient; micromhos per mi. per 1°F. 16 AWG 61-pair paper insulated cable at 70°F.

It was mentioned above that moisture in the cable has a pronounced effect on the conductance. To drive out excess moisture during manufacture the reels of paper covered cable (or cores) are placed in vacuum driers[13] and then stored in a room maintained at about 110° F. and at a relative humidity of 1/2 of 1 per cent or less until the cables are covered with their lead-antimony sheaths. The lead presses are adjacent to the ovens and the cable is fed through the wall directly into the press so that it emerges at the opposite side covered with the sheath. This procedure minimizes the amount of moisture entering the paper of the cables after they have been dried. The practical measure of the moisture content and the effectiveness of the drying

[13] C. D. Hart, "Recent Developments in the Process of Manufacturing Lead-Covered Telephone Cable," *B.S.T.J.*, VII (1928) pp. 321–342.

is the value of the quantity $G/2C$ = conductance divided by twice the capacitance, both measured at the room temperature in the factory. The quantity $G/2C$ is used because it is the coefficient in the leakage component of attenuation as explained in connection with the formula (12) below for high-frequency attenuation. The average value of $G/2C$ for 1000 cycles at 70° F. is about 8.3. This quantity increases with frequency and at the same time decreases with temperature in the same way G changes, since the capacitance changes are relatively so much smaller than the conductance changes.

Layer to Layer Variations of Primary Parameters

The values of the primary parameters vary from inside layers to outside layers of cables, in addition to variations mentioned under specific parameters above. There are three basic reasons for this variation with location in the cable. The first is that the length of an outside pair is usually considerably greater than the length of an inside pair. Unusual twisting of the inside layers might make up for this difference but in the ordinary construction this is not done. This increase in length amounts to as much as 1 or 2 per cent and is reflected at once in the d-c. resistance as well as in the a-c. parameters.

The second reason is that, particularly in the outside layer, the sheath being made of lead-antimony, has electrical properties considerably different from the properties of copper wires. The large size of the sheath relative to the wires is an important factor. High-frequency currents in the wires near the sheath produce fields cutting the sheath which affects the fields in a different fashion from the way adjacent copper wires affect the field of a pair of conductors near the center.

The third reason, closely allied to the second, is that the conductors in the core of the cable are surrounded by a practically symmetrical mass of copper conductors and paper plus the sheath, while conductors in any other layer are surrounded by an unsymmetrical arrangement of conductors and paper.

A fourth factor is the variation in the amount of space allowed pairs in the core by the pressure of the outside layers.

The magnitude of these effects is indicated by the curves of Fig. 13, showing layer-to-layer variations in per cent for Resistance, Inductance, Conductance and Capacitance. Such large variations would be of considerable importance were it not for the fact that in the process of splicing pairs are made to pass, in effect, from inside to outside of the cable and vice versa. A long study of these variations will be found in the paper by Wuckel[7]. The effects of splicing together sections slightly different in their character-

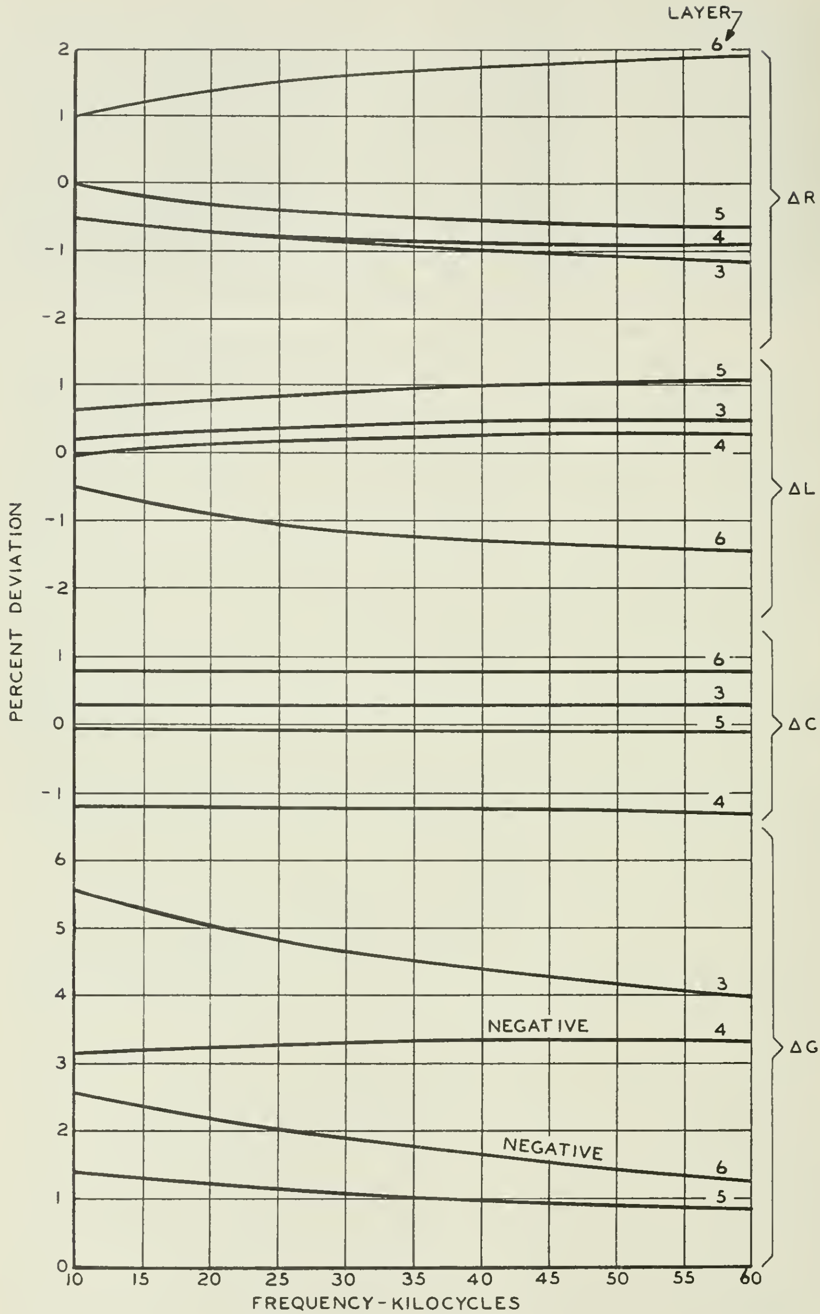


Fig. 13—Percentage deviations of layer average values of R, L, G, C from grand average—19 gauge pairs

istics were given for impedance, attenuation and delay distortion by Pierre Mertz and K. W. Pfleger[14].

Closely allied to these effects is the possibility of temperature differences across the section of cable in actual installed cables. Splicing usually takes care of this, too, but there are traces of such a lag in cases where the pairs remain in the outer part of the cable for a long distance and then pass to the inner group for the remaining part of the line. No such cross-sectional variation entered into the laboratory measurements as the temperatures were sufficiently well maintained close to given desired values.

Attenuation

The propagation constant γ is given by the familiar formula

$$\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)}$$
$$= j\omega\sqrt{LC}\sqrt{\left(1 + \frac{R}{j\omega L}\right)\left(1 + \frac{G}{j\omega C}\right)} \quad (9)$$

The real part, α, is the attenuation in nepers and the imaginary part, β, is the phase in radians. Expressing the attenuation in terms of reals, gives

$$2\alpha^2 = \sqrt{(R^2 + \omega^2 L^2)(G^2 + \omega^2 C^2)} - (\omega^2 LC - RG)$$
$$= \omega^2 LC\sqrt{(1 + R^2/\omega^2 L^2)(1 + G^2/\omega^2 C^2)} - (\omega^2 LC - RG) \quad (10)$$

In cables, $G/\omega C$ is small as compared to unity, in which case (10) may be reduced to the approximate form

$$2\alpha^2 = \omega^2 LC\sqrt{1 + R^2/\omega^2 L^2} - (\omega^2 LC - RG) \quad (11)$$

The formula for β^2 is the same as for α^2 except for the sign of the last two terms in (10) and (11), that is, the sign in front of the parenthesis is $+$ instead of $-$.

By expanding the square root term in (10) and using only first order terms in the expansion, another approximate form, frequently found useful in checking high-frequency values, is obtained, viz.,

$$a \doteq \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}} = \left(\frac{R}{2L} + \frac{G}{2C}\right)\sqrt{LC} \quad (12)$$

(Terms neglected in this approximation all include powers of ω^2 in their denominators and so become negligible at high frequencies.) The first term is commonly called the "resistance component of attenuation" and represents series losses. The second term represents shunt losses and is called the "leakage component of attenuation". The quantity $\sqrt{L/C}$, as is well known, represents the nominal characteristic impedance of the circuit.

[14] Pierre Mertz and K. W. Pfleger, "Irregularities in Broad-Band Wire Transmission Circuits," *B.S.T.J.*, XVI (Oct. 1937) pp. 541–559.

The shapes of attenuation-frequency curves at high, low and intermediate temperatures are shown by the curves of Fig. 14. These curves do not appear to be strikingly different in shape but more detailed study of the variations with temperature show the rate of change (decibels per degree per mile) to vary with frequency according to the curve of Fig. 15. The frequency of maximum rate of variation depends upon the gauge, as does the

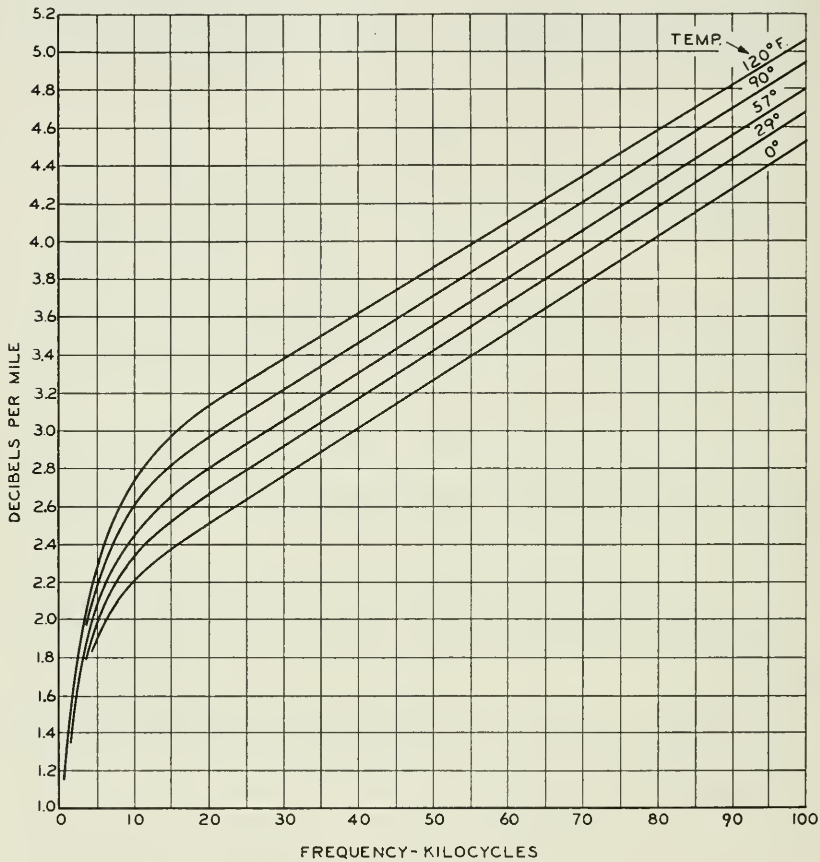


Fig. 14—Attenuation; decibels per mile—19 gauge pairs

actual rate of change. If the curves are plotted as attenuation coefficients (db per db per degree, Fahrenheit) with the abscissa

$$B = \sqrt{\frac{\text{frequency}}{R_{dc} \text{ per 1000 ft}}},$$

the same as for skin and proximity effects in Fig. 2, the peaks are brought together as indicated in Fig. 16.

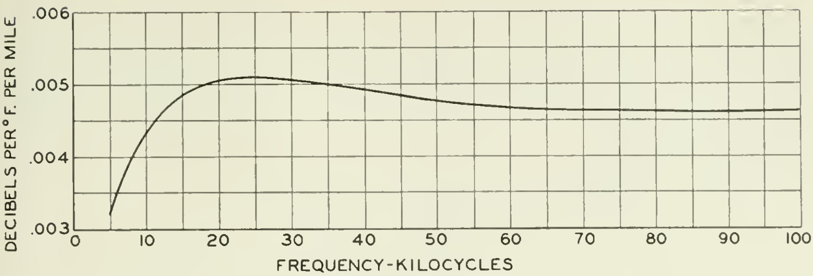


Fig. 15—Temperature variation of attenuation; decibels per degree Fahrenheit per mile vs. frequency—19 gauge pairs

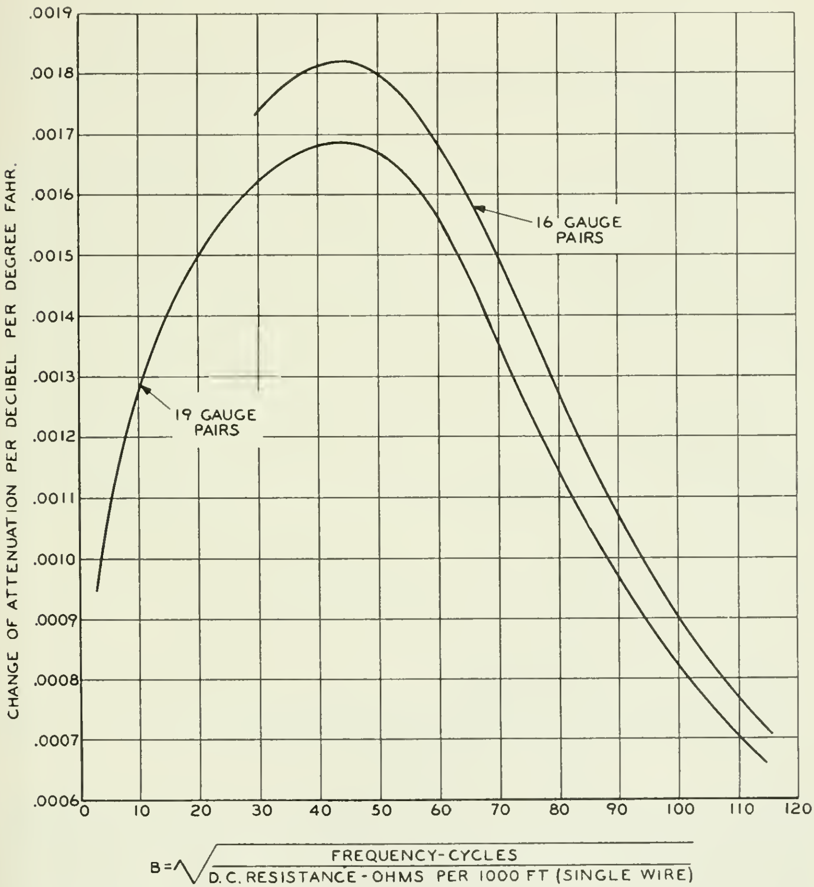


Fig. 16—Temperature coefficient of attenuation, α. $A = A_{50}[1 + \alpha(T - 50)]$

The change in attenuation with temperature may be formulated in various ways as a function of its component variables, R, L, G, C, the most obvious way being to take the partial derivatives of one of the equations (9), (10), or

(11) with respect to R, L, G and C, in order to get a differential expression $d\alpha$ in terms of dR, dL, dC and dG. This may then be interpreted as a change with temperature or a manufacturing variation, or a variation of attenuation from pair to pair in the cable. This procedure applied to (10) results in the formula

$$\begin{aligned}4\alpha\cdot d\alpha = {} & \left[G + R\sqrt{(G^2+\omega^2C^2)/(R^2+\omega^2L^2)}\right]dR \\ & + \left[R + G\sqrt{(R^2+\omega^2L^2)/(G^2+\omega^2C^2)}\right]dG \\ & - \left[\omega^2C - \omega^2L\sqrt{(G^2+\omega^2C^2)/(R^2+\omega^2L^2)}\right]dL \\ & - \left[\omega^2L - \omega^2C\sqrt{(R^2+\omega^2L^2)/(G^2+\omega^2C^2)}\right]dC \qquad (13)\end{aligned}$$

Curves of Fig. 17 show the components of the temperature variation produced by changes in R, L, G and C for standard 19-gauge cable.

A better formula from the point of view of equalizer design results from applying Taylor's series expansion to equation (9) and taking the real part of the resulting expressions. In this method, the variables are taken to be LC, R/L and G/C which effectively reduces the number of variables by one. There is a further advantage which appears in equation (14), below, namely, that the coefficient of the per cent variation in LC is just 1/2 the attenuation constant α and this means, therefore, only a slight addition to the basic equalizer which matches the curve for α vs frequency. There are thus added only two new types of temperature equalizers, one for R/L and one for G/C correction. Since equation (10) is already in the real form, it is more straightforward to expand it by Taylor's series and use the required number of terms. The formula thus obtained is naturally the same as that obtained from (9) and is as follows:

$$\begin{aligned}\Delta\alpha = {} & \frac{\alpha}{2}\frac{\Delta(LC)}{LC} + \frac{1}{2}\frac{R}{\omega L}\sqrt{\frac{\alpha^2+\omega^2LC}{1+R^2/\omega^2L^2}}\frac{\Delta(R/L)}{R/L} \\ & + \frac{1}{2}\frac{G}{\omega C}\sqrt{\frac{\alpha^2+\omega^2LC}{1+G^2/\omega^2C^2}}\frac{\Delta(G/C)}{G/C} \\ & - \frac{1}{8}\frac{R^2}{\omega^2L^2}\frac{(R/\omega L)\sqrt{\alpha^2+\omega^2LC}+\sqrt{\alpha^2-RG}}{\sqrt{(1+R^2/\omega^2L^2)^3}}\left[\frac{\Delta(R/L)}{R/L}\right]^2 + \cdots \qquad (14)\end{aligned}$$

Application of this formula gives slightly different values for the temperature-attenuation coefficient at different parts of the temperature range for most frequencies. This means that the change of attenuation with temperature is not quite linear. The nonlinearity is so small below 100 kc. that it has not been measured with any certainty on lengths of cable varying from 500 feet up to 10 miles, but on long cable carrier circuits corrections for it may become necessary.

The formula has another use, however, in determining the effects of small manufacturing variations on the probable attenuation of cables made up of the resulting product, as well as computing fairly accurately the attenuation of all the pairs in a layer by means of the average values of the constants and the departures of the values of the constants of the individual pairs from the average values. Actual attenuation variations of pairs in a reel

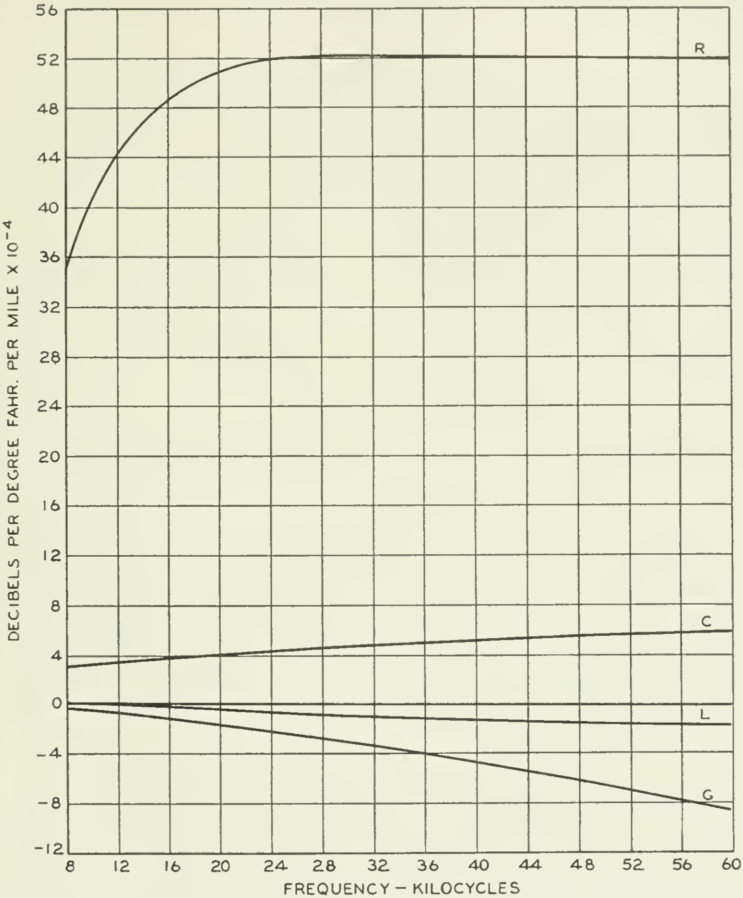


Fig. 17—Analysis of variation of attenuation with temperature; variations due to the components R, L, G, C—19 gauge pairs

are about ± 5 per cent and the standard deviation of the variations is about 2.5 per cent.

Practically, since $G/\omega C$ is small, formula (11) may be used in the Taylor series expansion with the variables R/L, LC and RG. This gives the formula

$$4\alpha \cdot \Delta\alpha = \frac{RC}{\sqrt{1 + R^2/\omega^2 L^2}} \cdot \Delta(R/L) + \Delta(RG) + \omega^2\left(\sqrt{1 + R^2/\omega^2 L^2} - 1\right)\Delta(LC) \quad (15)$$

Components of $\Delta\alpha/\Delta T$ computed by formula (15) are given in Fig. 18 It is evident that changes in R/L are responsible for most of the change in attenuation since the small changes in attenuation introduced by ΔLC and ΔRG tend to annul each other over most of the frequency range shown. This is to be expected from the approximate high-frequency formula (12) in which $G/2C$ is much smaller than $R/2L$.

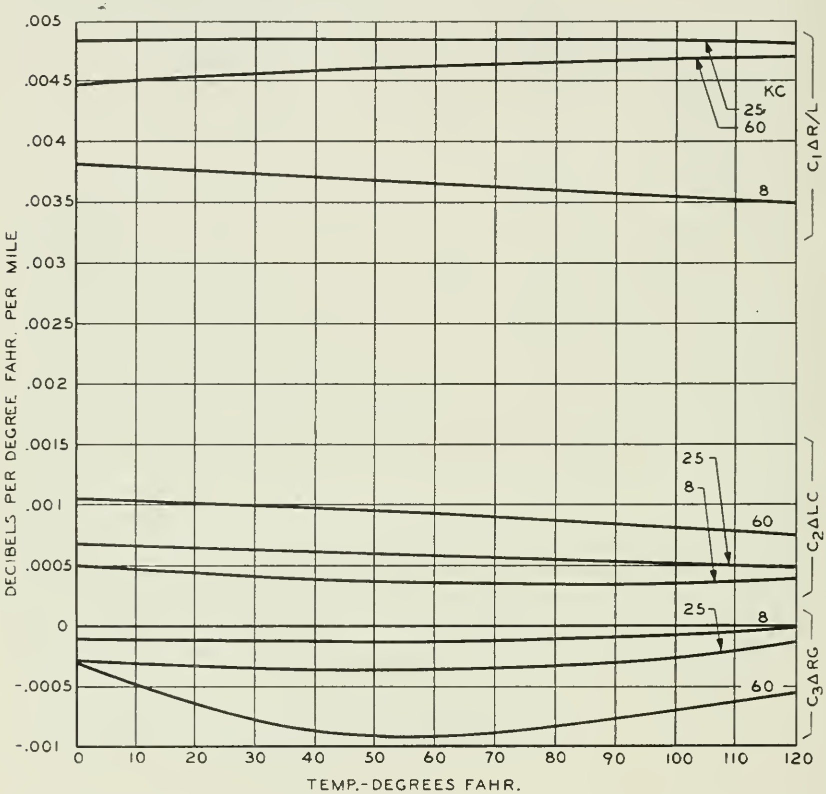


Fig. 18—Components of $\frac{\partial\alpha}{\partial T}$ = DB per °F. per mile = $C_1\Delta R/L + C_2\Delta LC + C_3\Delta RG$ vs. temperature—19 gauge pairs

Phase Change and Velocity

As pointed out above, merely changing the signs of the last parenthetical expression in equations (10) and (11) gives corresponding formulas for phase angle in radians. Fortunately, the phase change is nearly, though not quite, linear with frequency (Fig. 19). The velocity, $V = \omega/\beta$, for 19-gauge pairs is about 105,000 miles per second at 10 kc. and increases slowly to about 125,000 mps. at 100 kc. At high frequencies the internal inductance is small and, if the inductance L is expressed in abhenries and capacitance C in abstat-microfarads, then $V^2 = 1/LC = 1/k$, where k is the dielectric constant[15].

Impedance

The characteristic impedance is

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{16}$$

which has a large reactive component at low frequencies as shown by the curves of Fig. 20, based on the same reel-length measurements as the curves

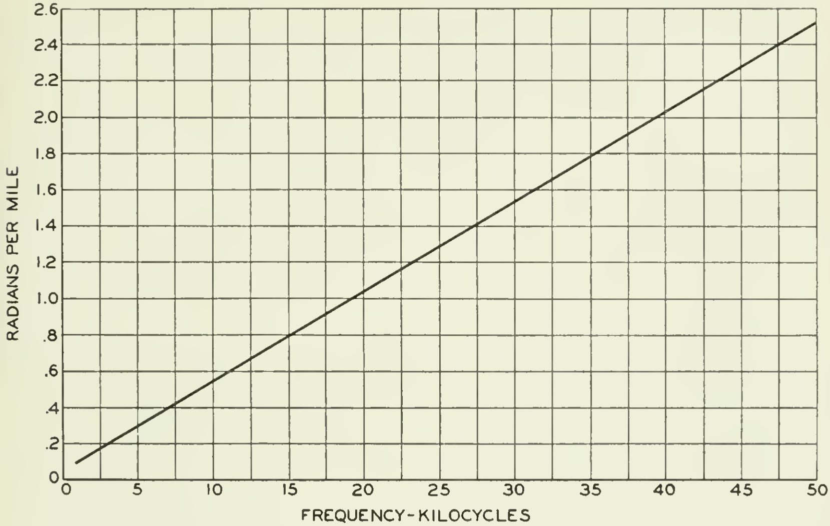


Fig. 19—Phase; radians per mile—16 gauge pairs 36°F.

for R, L, G and C in Figs. 3 to 11. In actual long cables the curves are irregular with frequency as a consequence of small irregularities along the line[13]. (See, for example, Fig. 26.) There are also small variations with temperature; for the resistance component about ± 1.5 per cent from the average for the temperature range zero to 120° F. at 10 kc., and about ± 1 per cent at 100 kc. The reactive component varies ± 10 per cent at 10 kc. over the same temperature range*.

[15] G. H. Livens, "The Theory of Electricity," p. 456 and p. 539.

* K. Simizu and I. Miyamoto, "Effect of Temperature on the Non-Loaded Carrier Cable, *Nippon Elec. Comm. Eng.*, May 1939, p. 596–599. Give similar data on the variation of parameters and attenuation for spiral-four cable at frequencies 0-30 kc. and temperatures 0–50°C. They do not specify the length measured but state that the wire diameter was 1.5 mm. From their d-c. resistance data the length appears to have been about 160 feet.

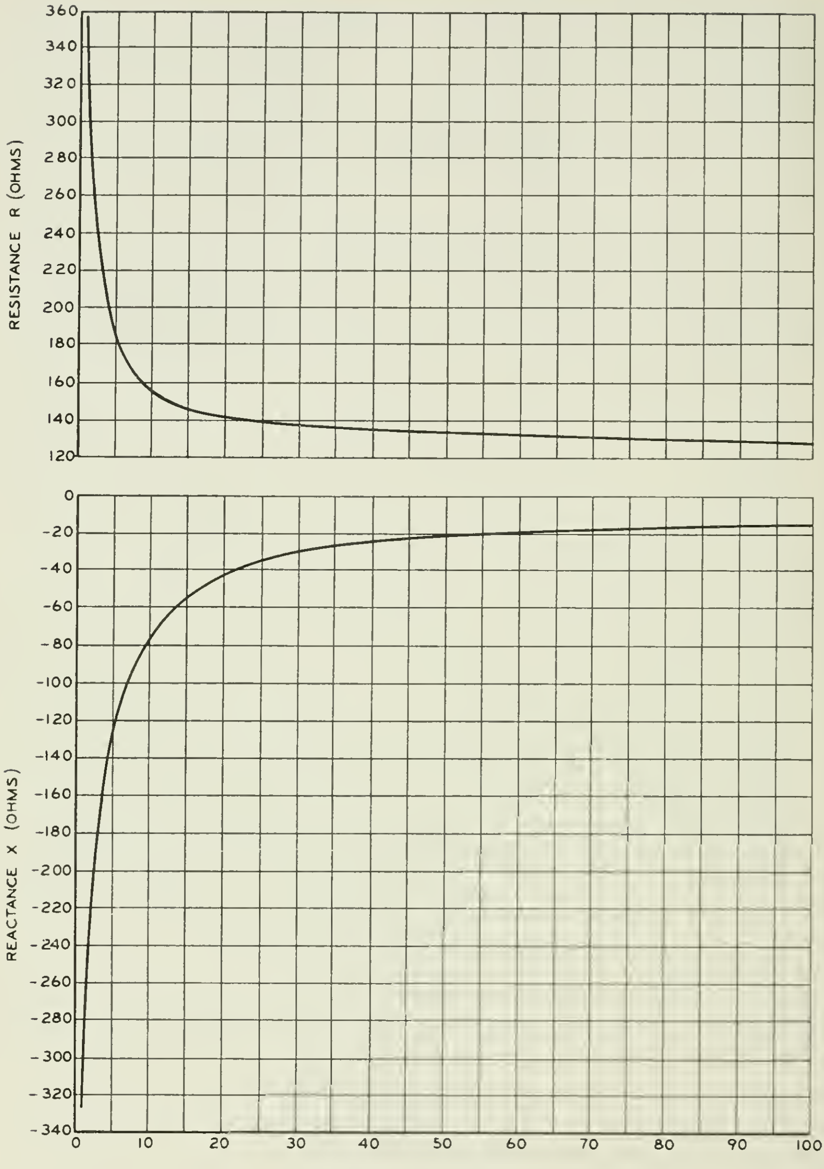


Fig. 20—Characteristic impedance, $Z = R - jx$; temperature 90°F.—19 gauge pairs

Characteristics of Toll Cable Above 100 Kc.

The preceding discussion has dealt largely with the characteristics of toll cables up to 100 kc. However, some measurements have been made extending to much higher frequencies. In the laboratories measurements were made on 16 and 19-gauge pairs in reel-lengths at frequencies up to about 3000 kilocycles. Field data at frequencies from 100 kc. to 2000 kc. were obtained on 16-gauge and 19-gauge pairs in cables about 3 miles long at

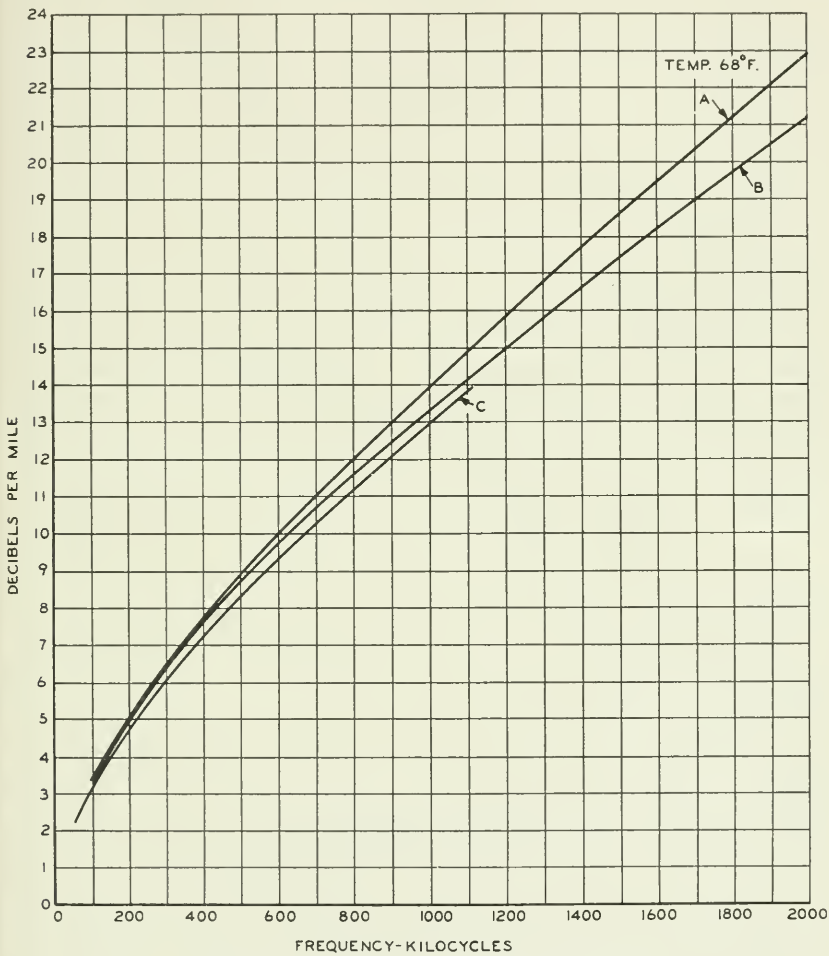


Fig. 21—Attenuation of 16 gauge cable pairs

A, on reel, 247 feet; B, fourteen reels, 1.3 miles; C, aerial cable, 3.6 miles, Ticonderoga, N. Y.

Ticonderoga, New York. A third set of data was obtained from measurements on 7000 feet of a special type (61-pair) of 16-gauge cable on reels in the laboratory under controlled temperature conditions. Figure 21 shows the attenuation values to 2000 kc. obtained in the three sets of 16-gauge

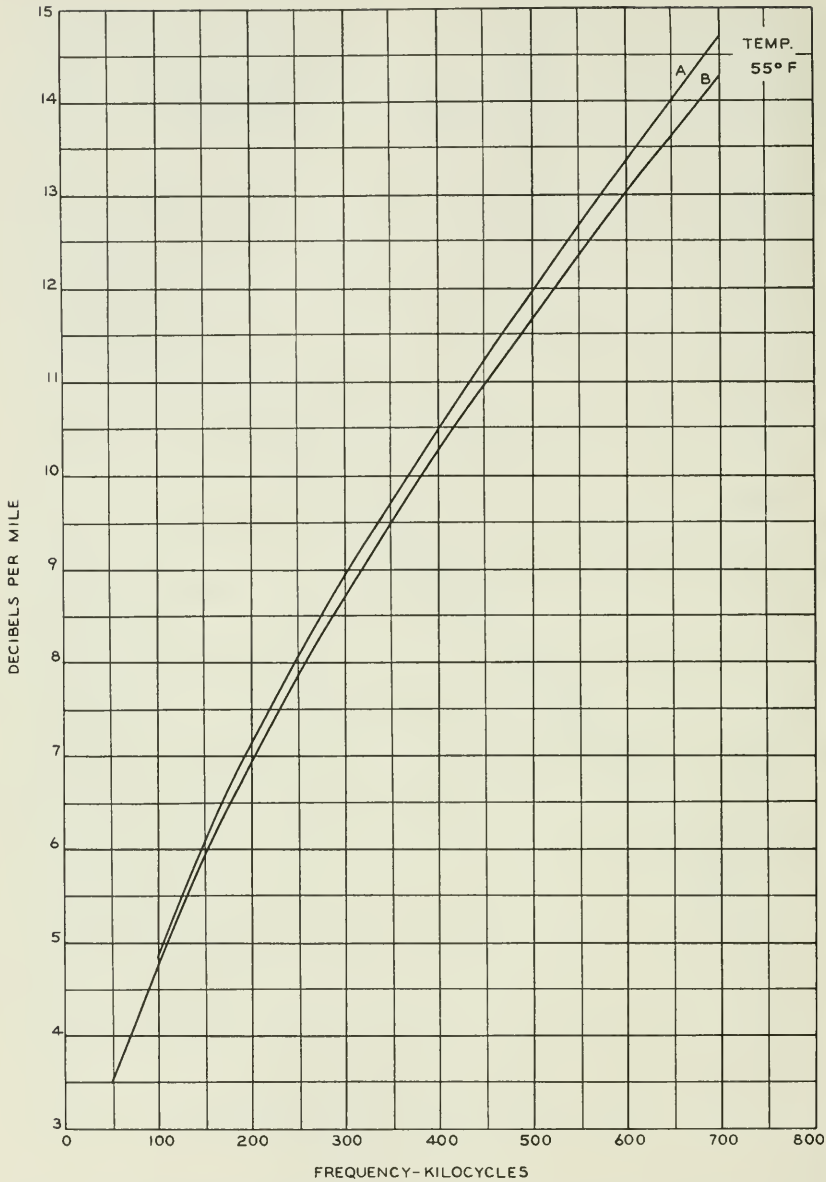


Fig. 22—Attenuation of 19 gauge cable pairs
A, on reel, 247 feet Bell Tel. Labs., Inc.; B, aerial cable, 3.6 miles, Ticonderoga, N. Y. Curves A and B show average of 10 pairs.

data at 68 degrees Fahrenheit. Figure 22 shows results on 19-gauge pairs at 55 degrees from field and laboratory data up to 700 kc.

The curves of the change in attenuation per degree F. per mile (db/1°F./mi) as shown by Figs. 23 and 24 are highly dependent upon the temperature, showing that at these high frequencies the attenuation is decidedly nonlinear with temperature in the toll cables.

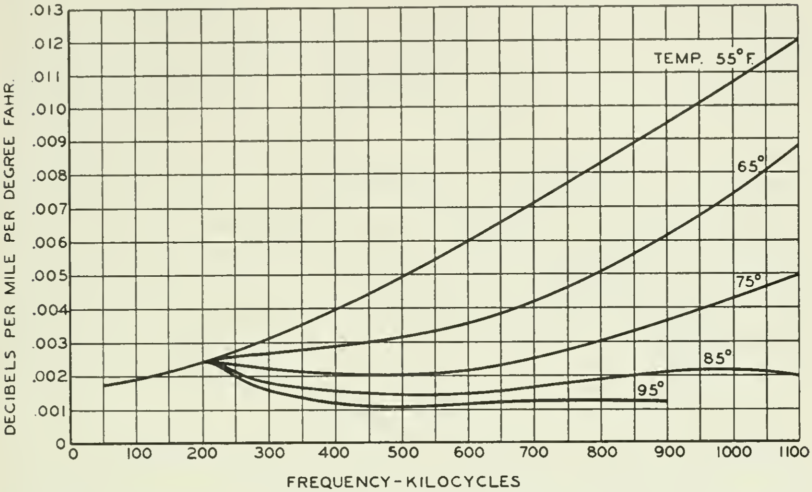


Fig. 23—Variation in attenuation at different temperatures for 1°F. change in temperature; aerial cable—16 gauge pairs

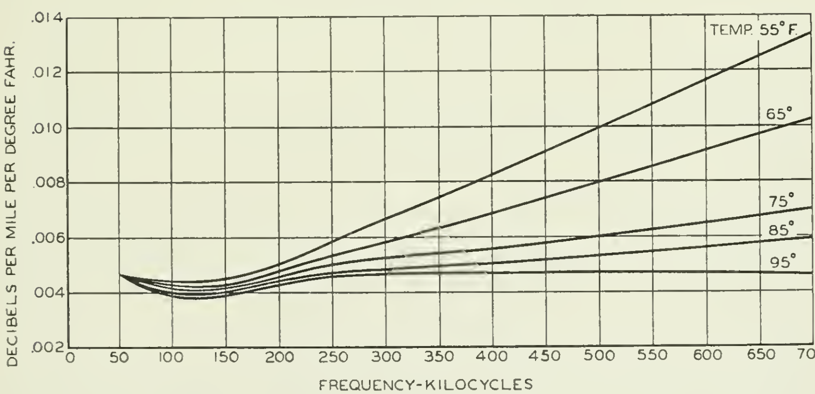


Fig. 24—Variation in attenuation at different temperatures for 1°F. change in temperature; aerial cable—19 gauge pairs

Toll Entrance Cable

The insertion losses measured between 125-ohm resistances on various lengths of 13, 16 and 19-gauge toll entrance cables at Denver, Colorado, are shown in Fig. 25. The data have been reduced to a per-mile basis by

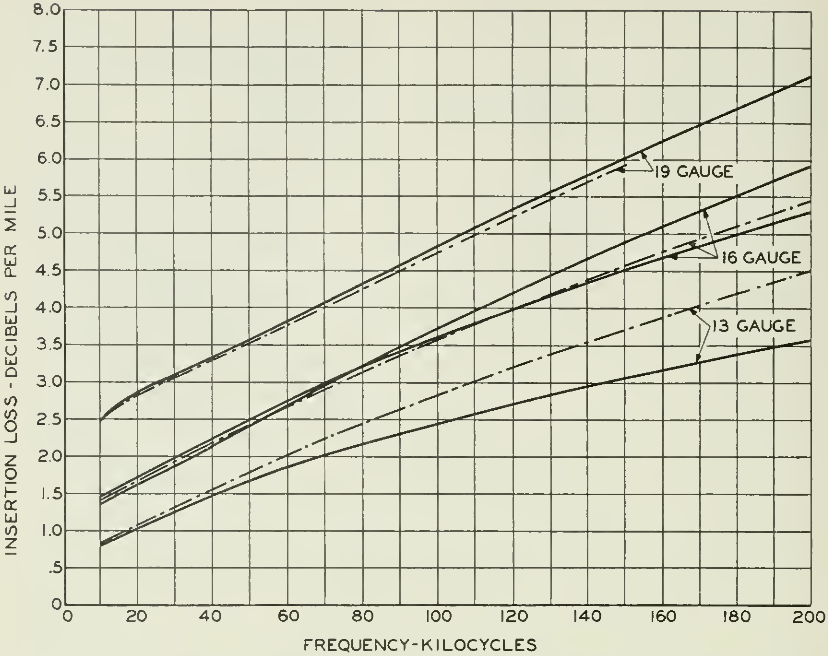


Fig. 25—Carrier frequency loss* of toll entrance cable; non-loaded, quadded—temperature 60°F., approx.

*Insertion loss between 125-ohm resistances

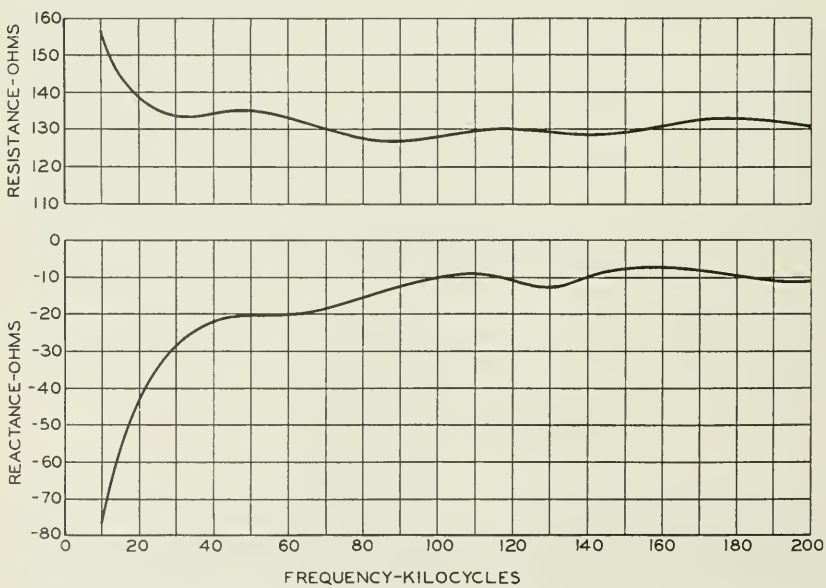


Fig. 26—Carrier frequency impedance of toll entrance cable, Denver, Colo.—19 gauge, quadded, non-loaded—terminated in 125 ohms

direct division of the measured attenuations by the lengths. This is, of course, not strictly accurate, but the errors are very small at these frequencies. This is nonloaded cable and frequencies measured were from 10 kc. to 200 kc. The values check closely the values shown in the previous figures for frequencies below 100 kc. The 13-gauge figures are the first of such data given herein but the first laboratory measurements on reel-lengths (begun in 1921) included reels of 13-gauge cable, and curves of 13-gauge attenuation and impedance were given in a paper[16] by E. H. Colpitts and O. B. Blackwell.

Corresponding data on impedance show the values given on Fig. 26. The wavy characteristic of these curves, as mentioned in the section on Impedance above, is caused by small irregularities in the pairs, particularly differences between pairs in successive reel-lengths giving rise to reflection currents at certain frequencies[14]. In new construction smoother impedance characteristics can be obtained when it is important to do so, by close control of the product during manufacture, followed by suitable splicing methods.

Acknowledgment

It is practically impossible to name all my associates in the Bell Telephone Laboratories whose work has been drawn upon in assembling these data, but I am especially indebted to Mr. Pierre Mertz and Mr. E. I. Green for their helpful suggestions and continued encouragement.

APPENDIX

Wave Propagation over Two Parallel Wires: The Proximity Effect—Inductance*

In his paper[4] on the Proximity Effect, J. R. Carson carried out the detailed computations for the ratio C of the a-c. resistance of two parallel wires to the a-c. resistance of a wire when the return conductor is concentric. He gave a formal expression for the impedance (equation 64 of his paper), viz.,

$$R + iX = 2Z + ipL \tag{1a}$$

This simple equation is complicated by the fact that Z and L are given by two complex expressions involving Bessel functions and the set of harmonic coefficients of the Fourier-Bessel expansion for the axial electric force in

[16] E. H. Colpitts and O. B. Blackwell, "Carrier Current Telephony and Telegraphy," *Jour. A. I. E. E.* XL, Feb. 1921, pp. 205–300.

* This work, done under the direction of Mr. J. R. Carson, was completed in April, 1922. For the general theory of wave propagation on parallel conductors see a paper by Chester Snow, "Alternating Current Distribution in Cylindrical Conductors," Proc. Int. Math. Congress, Toronto (1924) Vol. II, pp. 157–218.

one of the wires and the separation of the wires. The Bessel functions are of order zero to infinity and the argument, b, is given by

$$b = ia\sqrt{4\pi\lambda\mu i p} \tag{2a}$$

where

a = radius of the wire in cm.
λ = conductivity of wire in c.g.s. units
μ = permeability of wire in c.g.s. units
p = 2π times the frequency in cycles per second
i = $\sqrt{-1}$

The separation comes in by way of the quantity k, the ratio a/c of the radius to the interaxial separation of the wires, and a function s which can be expressed as a continued fraction in k^2, viz.,

$$s = \cfrac{1}{1 - \cfrac{k^2}{1 - \cfrac{k^2}{1 - \cdots}}} \tag{3a}$$

which results in

$$s = \frac{1}{1 - k^2 s} \tag{4a}$$

from which

$$s = \frac{1 - \sqrt{1 - 4k^2}}{2k^2} \tag{5a}$$

as given by Carson's equation (38).

The actual expression for $R + iX$ is as follows:

$$\begin{aligned} R + iX &= 2Z + ipL \\ &= -4ip \log ks + 2Z_0\left[1 + \sum_{n=1}^{\infty} (-ks)^n h_n J_n/J_0\right] \end{aligned} \tag{6a}$$

where

$$\begin{aligned} Z_0 &= R_0 + iX_0 \\ &= \frac{2p}{b}\frac{u_0 v_0' - u_0' v_0}{u_1^2 + v_1^2} + i\frac{2p}{b}\frac{u_0 u_0' + v_0 v_0'}{u_1^2 + v_1^2} \end{aligned} \tag{7a}$$

which is the impedance of a wire with concentric return expressed as usual[17] in terms of the ber and bei functions related to the Bessel functions by the formula

[17] Russell, "Alternating Currents," Edition 1904, Vol. I, p. 370.

$$u_n + iv_n = J_n\,(b\,i\,\sqrt{i}) \tag{8a}$$

and primes denote derivatives with respect to b.

Substituting Z_0 from (7a) in (6a) and carrying out the algebraic processes involved gives finally

$$R + iX = 2\,R_0\,C + i\,(-4\,p\log ks + 2\,K\,X_0) \tag{9a}$$

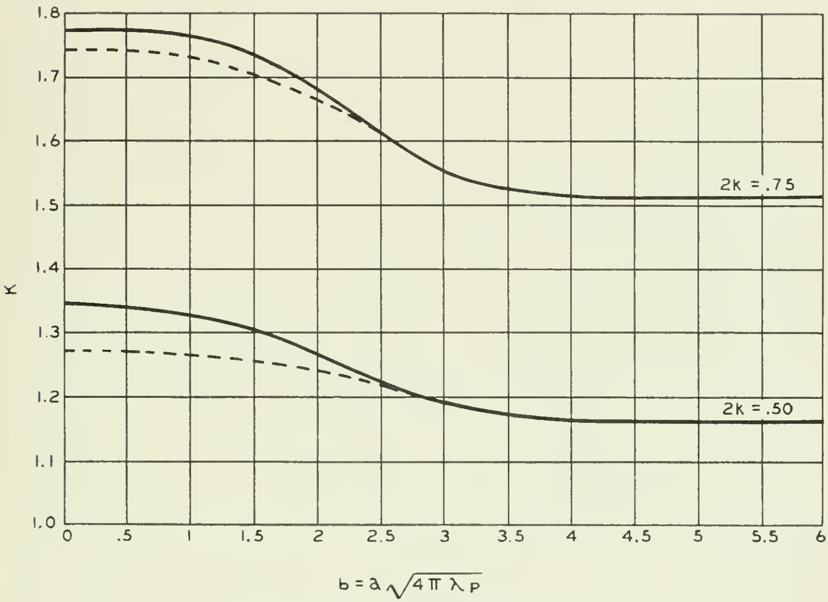


Fig. 27—Values of correction factor, K
For $2k = .75$ and $.50$

where

$$C = 1 + \frac{4p}{bR_0}\sum (k^2s^2)^n w_n + \frac{4p}{bR_0}\sum n(k^2s)^{n+1} gw_n$$
$$-\frac{4}{b}\frac{X_0}{R_0}\frac{u_1^2+v_1^2}{u_0^2+v_0^2}\sum n(k^2s)^{n+1}\frac{u_{n-1}v'_{n-1} - u'_{n-1}v_{n-1}}{u_{n-1}^2+v_{n-1}^2} \tag{10a}$$

and

$$K = 1 + 2\sum (k^2s^2)^n\,\frac{u_{n-1}v'_{n-1} - u'_{n-1}v_{n-1}}{u_{n-1}^2+v_{n-1}^2}\cdot\frac{u_0'^2+v_0'^2}{u_0u'_0+v_0v'_0}$$
$$+\frac{4p}{b^2X_0}\sum n(k^2s)^{n+1} bg\,\frac{u_{n-1}v'_{n-1} - u'_{n-1}v_{n-1}}{u_{n-1}^2+v_{n-1}^2} + 2w_n\,\frac{u_0u'_0+v_0v'_0}{u_0^2+v_0^2}$$

$$= 1 + \frac{4p}{bX_0} \sum (ks)^{2n} \frac{u_{n-1} v'_{n-1} - u'_{n-1} v_{n-1}}{u^2_{n-1} + v^2_{n-1}}$$

$$+ \frac{4p}{bX_0} \sum n(k^2 s)^{n+1} g \frac{u_{n-1} v'_{n-1} - u'_{n-1} v_{n-1}}{u^2_{n-1} + v^2_{n-1}}$$

$$+ \frac{4}{b} \frac{u^2_1 + v^2_1}{u^2_0 + v^2_0} \sum n(k^2 s)^{n+1} w_n \tag{11a}$$

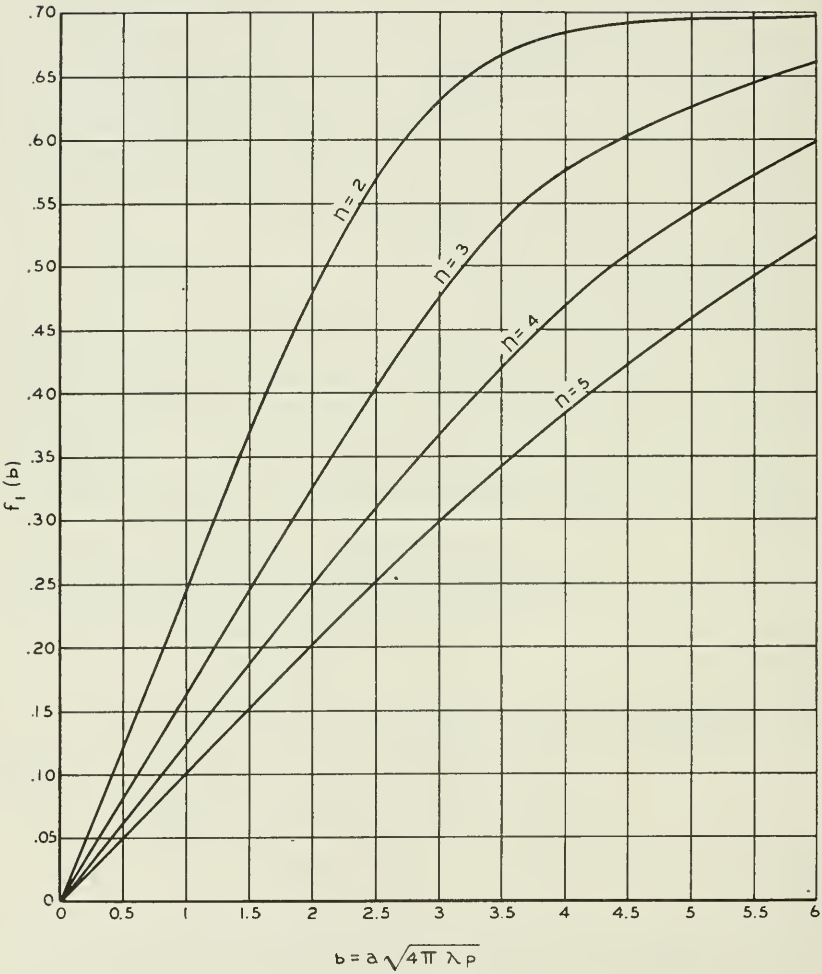


Fig. 28—Values of auxiliary functions

$$f_1(b) = \frac{u_{n-1} v'_{n-1} - u'_{n-1} v_{n-1}}{u^2_{n-1} + v^2_{n-1}}$$

In these formulas

$$g = \frac{2}{b}\frac{u_0'v_0 - u_0v_0'}{u_0^2 + v_0^2} = -\frac{R_0}{p}\frac{u_1^2 + v_1^2}{u_0^2 + v_0^2} \tag{12a}$$

$$w_n = \frac{u_nv_n' - u_n'v_n}{u_{n-1}^2 + v_{n-1}^2} \tag{13a}$$

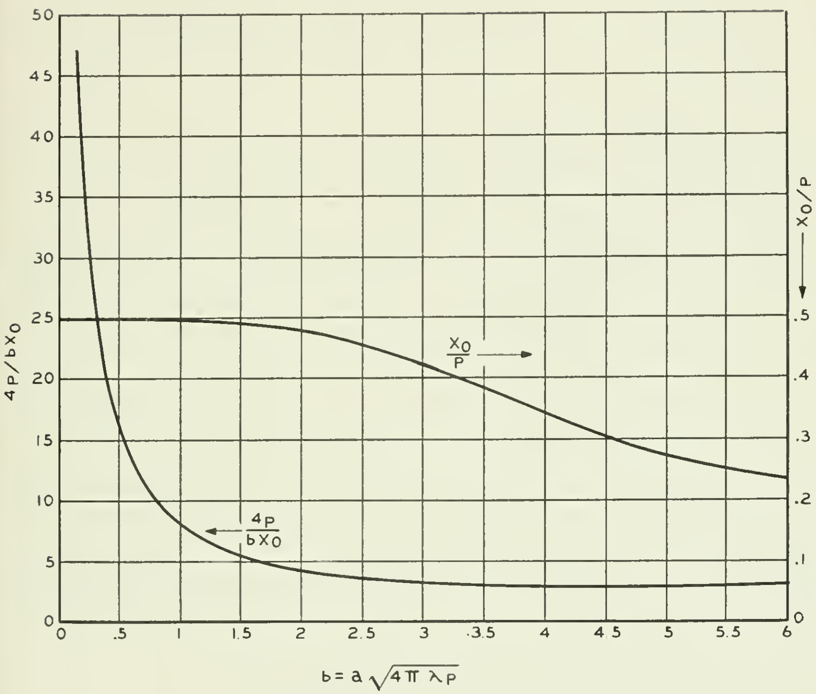


Fig. 29—Values of functions

$$\frac{4p}{bX_0} \quad \text{and} \quad \frac{X_0}{P}$$

The curves of figures 28–30 show the auxiliary functions vs b and figure 27 the correction factor K. The dotted curve for K is computed from Mie's formula[18] (14a) for small b. Two values of $2k$ are shown, .75 and .50, respectively.

G. Mie[18] gave formulas for small and large values of b, as follows:

For small b,

[18] G. Mie, *Annalen der Physik*, Vol. II, (1900) pp. 201–249.

$$L = 1 - 4 \log k - l_n - l'_n \tag{14a}$$

where $l_n = .417\ b^4/16 - .003\ b^8/256$

$$l'_n = b^4(1.33k^2 - .917k^4 - .652k^6 - .496k^8 \cdots)/16$$
$$- b^8(.633k^2 - 1.354k^4 + .539k^6 + .584k^8 \cdots)/256$$

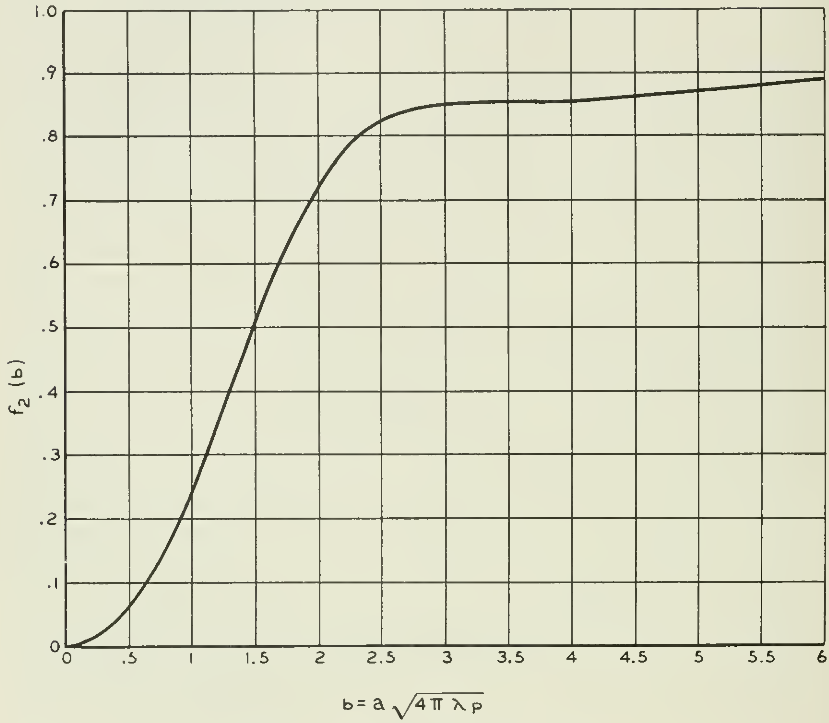


Fig. 30—Values of auxiliary functions

$$f_2(b) = \frac{u_1^2 + v_1^2}{u_0^2 + v_0^2} = \frac{ber'^2 + bei'^2}{ber^2 + bei^2}$$

For large b,

$$L = -4 \log ks + \frac{2\sqrt{2}}{b} C_m - \frac{3}{2b^2} \tag{15a}$$

where $C_m = s/(2 - s)$.

Some Analyses of Wave Shapes Used in Harmonic Producers

By F. R. STANSEL

Analyses by Fourier's Series have been made of waves consisting of sinusoidal, rectangular and trapezoidal pulses and also waves of the type found in multivibrator circuits. The method of increasing harmonic content by modulating a wave with a submultiple is treated mathematically.

THE heterodyne method of frequency comparison requires, except in the case of the comparison of nearly identical frequencies, the generation of harmonics of either the unknown, or of the standard frequency or of both. These harmonics may be generated directly in the modulator which produces the difference frequency, or "beat note", or may be generated in an entirely separate circuit before the frequency is applied to the modulator. An example of the latter is the multivibrator circuit often used in connection with a frequency standard to produce a series of harmonics of this standard frequency.

The design of harmonic generators for frequency measuring equipment presents a different problem from the design of equipment for producing a single harmonic such as doubler or tripler stage in a radio transmitter. In the latter case the amplitude of the one harmonic and the efficiency are of primary importance. In frequency measuring equipment, although a large amplitude of each harmonic is desirable, it is of greater importance that each harmonic within the range to be used, which may be up to the 100th or 150th harmonic or even higher, be present and that the amplitude of nearby harmonics be of the same order of magnitude. Unless the latter conditions are met, there is a danger that the beats obtained with a weak harmonic will either be entirely overlooked or mistaken for a higher order modulation product.

The generation of harmonics is usually accomplished by the distortion of the wave shape in some nonlinear circuit element such as a vacuum tube. One such harmonic generator consists of a vacuum tube biased so that there is no output for a portion of the cycle. The plate current of such a tube may be approximated by a sine wave shaped pulse such as shown in Fig. 1. Any such periodic wave can be resolved into its harmonic components[1] and in the case of this wave the amplitude of the nth harmonic is found to be

[1] This and the subsequent analyses were made by application of Fourier's Series. See I. S. Sokolnikoff and E. S. Sokolnikoff, "Higher Mathematics for Engineers and Physicists," Chapter VI.

$$h_n = \frac{A}{n\pi(1 - \cos b/2)} \left[\frac{\sin (n - 1)b/2}{n - 1} - \frac{\sin (n + 1)b/2}{n + 1} \right] \tag{1}$$

in which A is the amplitude of the pulse and b the pulse width as shown in Fig. 1.

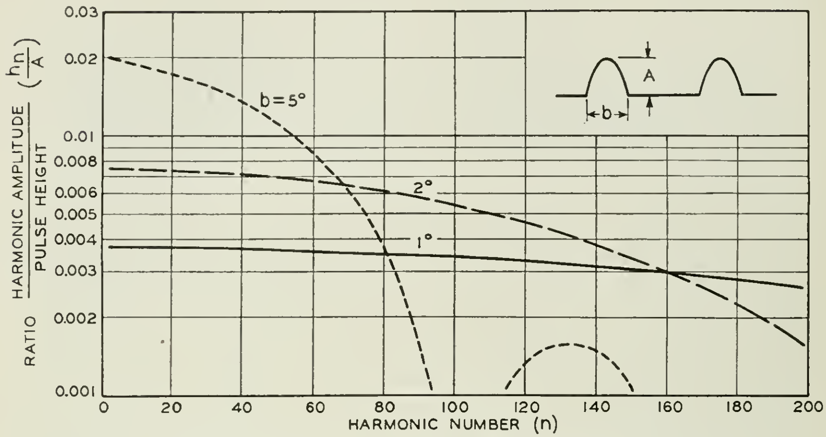


Fig. 1—Harmonic content of a wave consisting of sinusoidal pulses

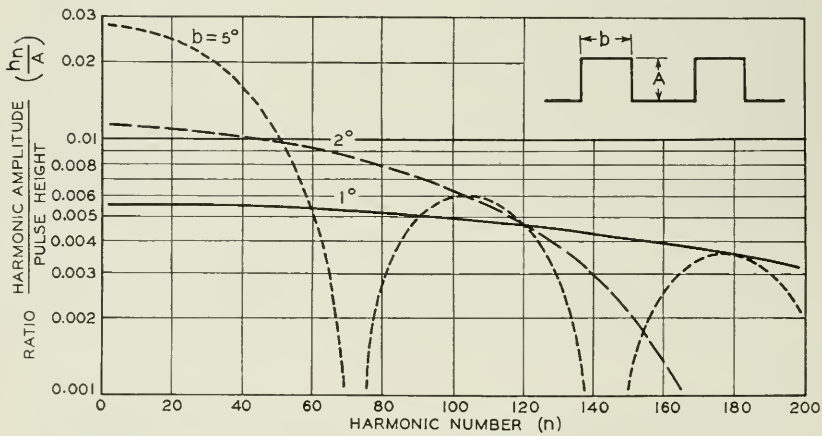


Fig. 2—Harmonic content of a wave consisting of rectangular pulses

The form of this expression immediately suggests that for some harmonics the terms

$$\frac{\sin (n - 1)b/2}{n - 1} - \frac{\sin (n + 1)b/2}{n + 1}$$

may become equal to zero causing these harmonics to vanish. That this is the case is shown in the curves of Fig. 1 in which the harmonic amplitudes are plotted against n for pulse widths of 5°, 2° and 1°. With a 5° pulse harmonics in the vicinity of the 105th and again the 150th become negligibly small. For a shorter pulse width the amplitude of the lower harmonics decreases but all harmonics up to beyond the 200th are present.

The wave shown in Fig. 1 can only be considered as a first approximation of the plate current in such a harmonic generator as it implicitly assumes that the tube is linear to cut-off. More frequently sufficient excitation is placed on the grid of the tube to saturate it and the resulting current wave may better be represented by a series of rectangular pulses such as shown

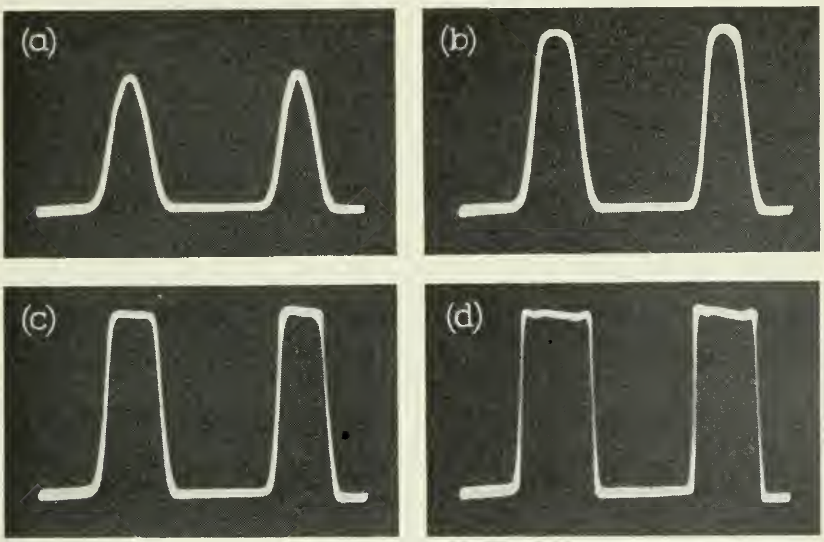


Fig. 3—Oscillograms of the plate current of a vacuum tube showing the transition from sinusoidal to rectangular pulses as excitation is increased

(a) Excitation 6 volts
(b) Excitation 8 volts
(c) Excitation 10 volts
(d) Excitation 20 volts

in Fig. 2. This transition from sine wave pulses to rectangular pulses as the grid excitation is increased is shown in the series of oscillographs in Fig. 3.

The analysis of a wave consisting of rectangular pulses such as the one in Fig. 2 shows the amplitude of the nth harmonic to be

$$h_n = \frac{2A}{n\pi} \sin \frac{nb}{2} \tag{2}$$

From this equation it is seen that certain of the harmonics are not present as the expression (2) becomes equal to zero whenever

$$n = \frac{2\pi}{b} m \tag{3}$$

$$m = 1, 2, 3, 4, \cdots$$

Thus for a rectangular pulse of 5° ($\pi/36$ radians) pulse width the 72nd, 144th, 216th, etc. harmonics vanish, and harmonics in the vicinity of these missing harmonics have lower amplitudes as can be seen from the curves of Fig. 2.

As the pulse width of a rectangular wave increases, the number of harmonics which vanish increases. For a pulse width of 90° every fourth harmonic is missing. For a pulse width of 180°, the familiar square wave,

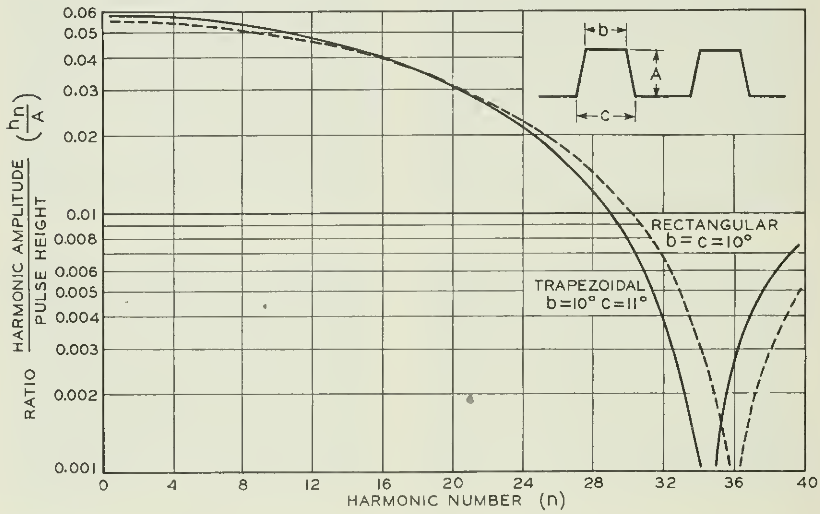


Fig. 4—Comparison of the harmonic content of waves consisting of rectangular and o trapezoidal pulses

every even harmonic vanishes and the wave contains only odd harmonics. As the pulse width is increased beyond 180° the number of harmonics increases and it can be shown that a wave having a pulse width greater than 180° will have the same harmonic content[2] as a wave of pulse width (360°—b). Thus for a large harmonic content it is desirable to have a wave having either extremely narrow pulses or pulses lasting nearly 360°.

True rectangular pulses are never obtained in practice. One common type of distortion in such pulses when obtained by the "limiter" action of a vacuum tube consists in the pulses having sloping rather than vertical sides. The sloping sides arise from the fact that the pulses are essentially sine waves

[2] This statement is correct for absolute magnitude of the harmonics only. Certain of the harmonics in the two waves will be 180° out of phase.

with their tops chopped off. The analysis of a pulse of the dimensions shown in Fig. 4 shows that the amplitude of the nth harmonics is given by the expression

$$h_n = \frac{4A}{n^2\pi(c - b)}\left[\cos\frac{nb}{2} - \cos\frac{nc}{2}\right] \tag{4}$$

In order to show better the relationship between a wave of rectangular pulses and one of trapezoidal pulses, consider the ratio of the nth harmonic for these two waves. From (2) and (4)

$$\frac{h_n \text{ for trap. pulse}}{h_n \text{ for rect. pulse}} = \frac{2(\cos nb/2 - \cos nc/2)}{n(c - b)\sin nb/2} \tag{5}$$

Substituting $c - b = \delta$ and expanding $\cos nc/2 = \cos\,(nb/2 + n\delta/2)$, the right hand side of (5) becomes

$$\frac{2}{n\delta}\left[\frac{\cos nb/2}{\sin nb/2} - \frac{\cos nb/2 \cos n\delta/2}{\sin nb/2} + \sin n\delta/2\right] \tag{6}$$

For small values of $n\delta/2$, that is for trapezoidal waves whose base is only slightly wider than the top, $\cos n\delta/2$ may be replaced by unity and $\sin n\delta/2$ by $n\delta/2$. The first two terms then cancel and the approximation

$$\frac{h_n \text{ for trap. pulse}}{h_n \text{ for rect. pulse}} \cong 1 \tag{7}$$

is obtained showing that a slight slope in the sides of the pulse has only a second order effect on the harmonic content of the wave.

The curve in Fig. 4 shows the harmonic content of a rectangular wave having a pulse width of 10° compared with that of a trapezoidal wave having a pulse width of 10° at the top and 11° at the bottom. For lower harmonics the amplitudes are nearly the same, but in the vicinity of the 36th harmonic there is an essential difference. For the rectangular pulse, the 36th harmonic vanishes, while the trapezoidal pulse has a minimum at a somewhat lower value of n and all harmonics have finite values.[3] This is shown in Table 1 which tabulates the amplitude of the harmonics in this case.

A second form of distortion in rectangular pulses is the rounding of the corners at both the top and the bottoms of the pulse. This type of distortion is more difficult to analyze and while no complete analysis has been made the effect of such distortion is known to be, in general, to reduce the amplitude of the higher harmonics.

[3] In discussing the curves in Fig. 1 thru 5 it must be remembered that while these are drawn as solid lines, the lines have a meaning only for integral values of n. Fractional values of n are meaningless.

From the examination of these cases it is evident that in the design of a harmonic generator of the type here considered the decision as to the pulse

TABLE 1
Harmonic Content of Rectangular and Trapezoidal Pulses Shown in Figure 4

Harmonic	$\frac{\text{Harmonic amplitude}}{\text{Pulse height}} = \frac{h_n}{A}$	
	Rectangular	Trapezoidal
Fundamental	.0555	.0581
2	.0554	.0580
3	.0550	.0576
4	.0545	.0570
5	.0540	.0563
10	.0488	.0505
15	.0411	.0416
20	.0314	.0307
25	.0209	.0191
30	.01061	.00811
32	.00681	.00411
33	.00501	.00226
34	.00325	.000489
35	.001901	−.001085
36	0	−.00276
37	−.00180	−.00412
38	−.00291	−.00557
40	−.00545	−.00791

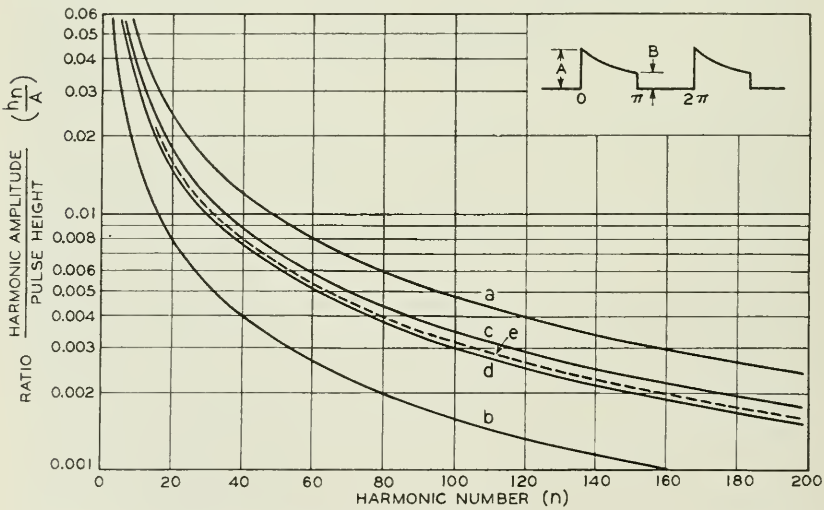


Fig. 5—Harmonic content of multivibrator wave
(a) Odd harmonics $\tau = 1/2$
(b) Even harmonics $\tau = 1/2$
(c) Odd harmonics $\tau = 1/10$
(d) Even harmonics $\tau = 1/10$
(e) All harmonics $\tau = 0$

width must be based on the type of service to which it is to be put. If only a few harmonics are required, a considerable gain in the amplitudes of the harmonics can be obtained by using a wider pulse width. When a wide range of harmonics is required, the band width must be greatly reduced to avoid blank intervals in the frequency spectrum.

A second type of harmonic generator is the multivibrator. The output wave of such a harmonic generator has a shape similar to that shown in Fig. 5. The current pulse lasts for a complete 180° rising abruptly to the peak value, then falling more or less exponentially to a lower value and finally breaking abruptly to zero. Assuming an exponential decay this wave will be found to contain the following harmonics

$$h_n = \frac{A(1-\tau)}{\sqrt{n^2\pi^2 + (\ln \tau)^2}} \text{ for even harmonics} \tag{8}$$

$$h_n = \frac{A(1+\tau)}{\sqrt{n^2\pi^2 + (\ln \tau)^2}} \text{ for odd harmonics} \tag{9}$$

Except for small values of n, the $(\ln \tau)^2$ term is negligible and these equations can be written

$$h_n = \frac{A(1-\tau)}{n\pi} \text{ for even harmonics} \tag{10}$$

$$h_n = \frac{A(1+\tau)}{n\pi} \text{ for odd harmonics} \tag{11}$$

In all of the above equations $\tau = B/A$, the ratio of the amplitude at the end to the amplitude at the beginning of the pulse.

The curves in Fig. 5 show the harmonic content of such a wave for $\tau = \frac{1}{2}$ and $\tau = \frac{1}{10}$. In the first case the amplitudes of the odd and even harmonics differ by approximately 9.5 db while in the second case the amplitudes are not greatly different. The dotted curve shows the limiting condition which all harmonics approach as τ approaches zero, that is as the current at the end of the pulse approaches zero.

The analysis of such a pulse except assuming a linear rather than exponential decay yields the following equations

$$h_n = \frac{A\,(1-\tau)}{n\pi} \text{ for even harmonics} \tag{12}$$

$$h_n = \frac{A}{n\pi}\sqrt{(1+\tau)^2 + \frac{4(1-\tau)^2}{n^2\pi^2}} \text{ for odd harmonics} \tag{13}$$

As n becomes large the second term under the radical becomes small and (13) becomes

$$h_n = \frac{A\ (1 + \tau)}{n\pi} \text{ for odd harmonics} \tag{14}$$

Equations (12) and (14) are identical with (10) and (11) showing that in harmonic generators of this type the harmonic content of the output wave is primarily a function of the initial and final values of the current rather than of the shape of the decay curve.

All of the foregoing curves show that the amplitudes of the higher harmonics are quite small so that in many applications some method of increasing their amplitudes may be required. This can be accomplished by the use of tuned amplifiers. An alternative method is to modulate a standard frequency wave with a lower derived frequency.

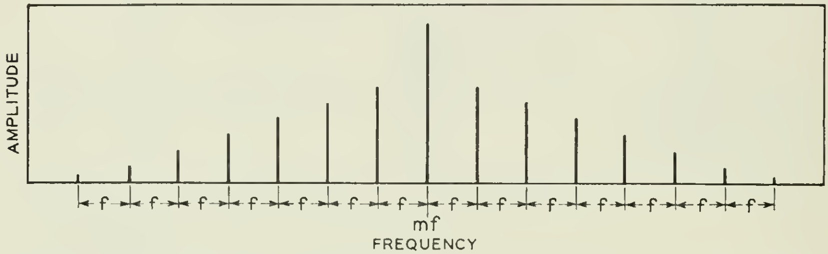


Fig. 6—Frequency spectrum of wave of frequency mf modulated by a series of pulses of frequency f

Assume a standard frequency of the form

$$A \cos m\omega t$$

This wave is completely modulated by a rectangular wave of frequency $\omega/2\pi$ and pulse width b. The modulated wave will then be of the form

$$I = A[1 + Kf(t)] \cos m\omega t \tag{15}$$

As shown previously the modulating wave is of the form

$$f(t) = \frac{b}{2\pi} + \sum_{n=1}^{\infty} \frac{2}{n\pi} \sin \frac{nb}{2} \cos n\omega t \tag{16}$$

For 100 per cent modulation $K = 1$. Since

$$\cos (m\omega t) + \cos (n\omega t) = \tfrac{1}{2} \cos (m + n)\ \omega t + \tfrac{1}{2} \cos (m - n)\ \omega t \tag{17}$$

the modulated wave is

$$\begin{aligned} i_p = \frac{Ab}{2\pi} \cos m\omega t &+ \sum_{n=1}^{\infty} \frac{A}{n\pi} \sin \frac{nb}{2} \cos (m + n)\ \omega t \\ &+ \sum_{n=1}^{\infty} \frac{A}{n\pi} \sin \frac{nb}{2} \cos (m - n)\ \omega t \end{aligned} \tag{18}$$

The frequency spectrum of this wave is shown in Fig. 6. The original standard frequency $m\omega/2\pi$ is present and on either side above and below $\omega/2\pi$ cycles apart are additional components. The rate at which the amplitude of these frequencies dies out depends on the modulating pulse width and is equal to half the amplitude of the corresponding harmonic in Fig. 2.

If the standard frequency is not a pure wave but contains harmonics each of these harmonics will be modulated by the rectangular pulses, that is the function (16). The result will be a series of frequency spectra similar to the one in Fig. 6, each centered at one of the harmonics of the standard frequency. By proper choice of the frequency of the modulating wave these spectra may be made to overlap giving a continuous series of harmonic of the modulating frequency with much larger amplitudes than can be obtained from a straightforward harmonic generator. As an example, a one-megacycle wave heavily modulated with 100 kc was found to give strong 100 kc harmonics up to well over 35 mc.

Acknowledgment

These analyses were originally made in connection with a research program at the Polytechnic Institute of Brooklyn. The author wishes to thank Professor E. Weber of this Institute for verifying the derivation of these equations.

Forces and Atoms: The World of the Physicist*

By KARL K. DARROW

ONE of the signs whereby a physicist may be known is a fondness for putting dots upon blackboards. This is not an irrational habit, but a symbolic practice. It is a symbol of his manner of regarding the world as a multitude incredibly enormous of particles incredibly small. The dots stand for the particles, and the bare regions of the blackboard for the empty spaces between them. The habit has not indeed been universal. Many a thinker has preferred to consider the world as a continuum, a solid or jelly or fluid; and we shall see that this alternative has always been very near in the background, even when the "atomists" were at their most triumphant. Let me however defer this other idea, and derive as much as possible from the notion of particles in a void.

But when the dots are set down on the otherwise clean board with regions of black emptiness between, the story is far from completed. It is, in fact, only begun, for the major part is yet to be written: the account of the forces among the particles. Though these last be separated from each other by spaces apparently empty, yet they are not unconscious of each other, for each of them is subject to a force—the resultant of many forces, due to all rest.

One might attach an arrow to each dot, to signify the strength and the direction of the force which acts upon it. One might draw wandering curves all over the board, to intimate at every point the direction and strength of the force which a particle would feel, were it to be at that point. This is accepted practice, but it would be worth the doing only if our assumptions and our ambitions were much more specific than for the present they are. Perhaps at least the blackboard should be smeared with a uniform coating of chalk, to signify that a particle in space is not left entirely to itself, but feels the influence of the others. Among our not-so-distant ancestors there seems to have been a psychological need for a gesture of the sort; they talked about space as though it were filled with a "medium" or "aether", because it seemed wrong to them to say that space is empty if the particles which wander in it are subject to forces. Our generation has nearly lost the need, whether of an aether to occupy actual space or of a

* Opening lecture of a course on "Nuclear Physics and Theory of Solids" delivered in the Spring semester of 1941, during the author's tenure of the William Allan Neilson chair at Smith College.

smear of chalk to symbolize it on the blackboard. Let us say that space is empty and leave the blackboard black between the dots, without deeming ourselves deprived of the right of saying that the particles exert and suffer forces on and from one another.

What is there to be said about the forces? A very great deal! for most of theoretical physics is made up of beliefs or ideas about the forces, augmented by the mathematical operations—very hard and very long-winded, in far too many cases—required for making the ideas really useful. So great a programme is indicated by that sentence, that I am wasting words in adding that it will not be fulfilled in one or two lectures, nor in the whole of the course. Only the most general of statements can be made in what follows. Of these I lay down at once the first, which is negative and self-evident:

The forces cannot be purely repulsive. For if they were, all of the particles would rush off into the uttermost depths of space, and we should have no model at all for a universe which, with all its faults, does manage at least to stick together.

Therefore there must be attractive forces, and these by and large must overpower the repulsive ones, if any such there be.

But need there be any repulsive forces at all? (Let the sophisticated reader now forget for a little that there are electrical forces which are repulsive, so that he may enquire with an open mind as to whether such could be avoided.) At first, it may not seem so; and one may invoke the great authority of Newton, who is often thought to have contented himself with assigning to all bodies the power of attracting one another with the force of gravity. He did not so content himself, and we shall learn this shortly. For the moment, let it be remembered that forces of attraction unopposed would tend to draw all of the particles of the universe into a single compact clump. If the volume of each particle were infinitely small, so also would be that of the ultimate clump; if the volume of each particle were irreducible below a certain minimum—but we shall ere long find what *that* idea can involve us in! Briefly, there must be something to oppose the attractive forces. To call this something by the name of "force", or even to call it by any single name, would be to limit it unduly. So to the second general statement I give the form:

There must be attractive forces, but there must also be antagonists to them.

If someone wanted a particular problem of the theory of physics identified to him as the profoundest, the problem of these antagonists might well be selected as such.

There is indeed one famed and spectacular case, which makes one antagonist clear. It is the case of the heavenly bodies: the planets revolving around the sun, the satellites around the planets. Why does not the moon fall onto the earth and the earth fall into the sun? Newton's laws of motion tell the answer. The antagonist is *motion*—or, to speak more precisely, momentum—or, to speak yet more precisely, angular momentum. If two particles attract one another but are moving with a relative motion which is not along the line that joins them, they never will meet. However great the attraction between them, it cannot draw them together. Attraction can do no more than constrain them to swing in permanent orbits around their common centre of mass. Therefore,

The celestial bodies exhibit to us a system kept stable by the attraction of gravity, with motion for the antagonist thereto.

However natural this statement may now seem, it is by no means an idea inborn in the human mind. There was an era when it was believed that motion dies out of itself, unless continually sustained by a never-ceasing stimulus. Were motion to die out of itself, it could not be an eternal antagonist to gravity. Newton cleared the way for the new idea by abolishing the old one.

May we now assume that the ultimate particles of the world act on each other by gravity alone, with motion as the sole antagonist to keep the universe from gathering into a single clump?

The answer to this question is a forthright and irrevocable NO.

That the answer should be *no* is not at all surprising to this generation, which is familiar with other forces than gravity, the electromagnetic forces especially. Those who underrate the prowess of our forerunners may feel surprise on hearing that the negative answer was quite as apparent to Newton. No apology is ever needed for quoting verbatim what Newton wrote in English, though it is a dangerous act for the quoter, whose writing must suffer by contrast with the simple elegance of the seventeenth century. Incurring the danger, I cite from the *Opticks* (a book of which the name falls decidedly short of the scope):

"The attractions of gravity, magnetism and electricity reach to very sensible distances, and so have been observed by vulgar eyes, and there may be others which reach to so small distances as hitherto escape observation. . . . The parts of all homogeneal hard bodies which fully touch one another stick together very strongly. And for explaining how this may be, some have invented hooked atoms, which is begging the question; and others tell us that bodies are glued together by rest, that is, by an occult quality, or rather by nothing; and others, that they stick together by conspiring motions, that is, by relative rest among themselves[1]. I had rather infer from their cohesion, that their particles attract one another by some force, which in immediate contact is exceedingly strong, at small distances per-

[1] These remarks seem to be aimed at Lucretius, or else at the Greeks from whom Lucretius took some of his ideas.

forms chemical operations, and reaches not far from the particles with any sensible effect."

Further along in the *Opticks* we read:

"Thus Nature will be very conformable to herself and very simple, performing all the great motions of the heavenly bodies by the attraction of gravity which intercedes those bodies, and almost all the small ones of their particles by some other attractive and repelling powers which intercede the particles."

With the powerful aid of Newton we have now distinguished between the attractive force of gravity and another attractive force, for which I retain the old-fashioned name "cohesion". I give another basis of distinction, one which could not have been found until in the mid-nineteenth century the equivalence of heat with mechanical work was established. Consider a piece of solid or liquid matter, and put the question: how much work must be done to tear its atoms apart and dissipate them into the infinite reaches of space, if the only force whereby they act on one another is the attraction of gravity? The question is answerable, if it is known how massive the atoms are and how far apart (on the average) they are. These things are known. The result of the computation is to be compared with the amount of work which is actually expended—in the form of heat—when the solid or liquid is volatilized into vapor. It is found that only about the billionth part of a millionth part of the heat so spent is devoted to "breaking down the gravitational bond", to doing work against the attraction of gravity which is overcome when the atoms are dispersed[2]. All the rest is required for overcoming that more intimate force of cohesion.

Gravity now is pushed into the background, and sinks into the relative insignificance which may be gauged from the fact that in the endless speculations of physicists and chemists as to how matter is built up and joined together, it is completely left out. The force which dominates the planets, which makes a hill so hard to climb and a height so dangerous to fall from—how amazing that it should be trivial, compared with others which the flame of the gas-jet vanquishes as the water boils out of the kettle! Trivial of course by comparison only, and at small distances, not at great; or to phrase the situation better, it is the force of cohesion which is trivial at great distances, gigantic at small. This is the contrast which is implied by the technical terms of physics, "long-range forces" versus "short-range forces".

[2] The computation for mercury was made by my colleague Dr. L. A. MacColl, on the basis most favorable to gravity: by assuming mercury to be a continuum, or in other words, to be made up of infinitesimal atoms infinitely close together—an assumption giving the greatest possible value to the work required for spreading the mercury through infinite space, if gravity be the only restraint. The latent heat of vaporization of mercury is found by experiment to be $1.88.10^{16}$ times this value. Thus the contrast mentioned in the text is not contingent upon knowledge of the mass and spacing of the atoms, though the knowledge is available if wanted.

Gravity is long-range, because it falls away gently with increase of distance; cohesion is short-range, because it falls away precipitately. We shall soon be meeting with other examples of either character.

One other fact to illustrate the short-range quality of the cohesive forces: When a kettle of water is boiling away on the stove, the amount of heat consumed in dispersing the first cubic inch that departs is the same as is spent in dispersing the second, and the third, and each of the others down to and including the last. This could not be so, if the particles were drawn together by important long-range forces; for then each cubic inch would be easier to drive off than that which last preceded it into the vaporous state, since there would be less of the liquid remaining behind to attract it.

The celestial bodies—useful as they have been in showing us the laws of motion—have therefore served us badly by hinting that gravity is the sole attractive force, a hint which is quite misleading. In another important respect they fail to give us a lead: they show us no examples of collision. Collision, more commonly known as impact, is one of the most important of earthly phenomena, as it is one of the most uncomfortable. The apple which fell in the orchard of Newton, and inspired him with the law of gravitation, may have been a legendary apple; if it was real, we may be sure that it ended its fall in a collision—ended its fall, not its existence. It did not pass through the globe and pop out of the ground in the Antipodes; it did not instantly merge with the grass or the soil of the orchard; it bounced and rolled a little, perhaps, and then lay quietly pressing against the earth, entire and whole. The earth was impenetrable to the apple, as the apple to the earth.

We do not even have to look to impact, to be taught this lesson about the impenetrable. Not less impressive than the fact that the piece of iron sticks together, is the fact that it does not shrink. For any particular choice of temperature and pressure, it has a particular volume which is its own. Work or heat must be expended to dilate it or tear it apart altogether, but also work must be expended to make it denser.

Having ascribed to attractive forces the fact that it takes heat—or let me say henceforward, energy—to vaporize a piece of matter solid or liquid, we now ascribe to repulsive forces the fact that it takes energy to squeeze the piece. The forces must be short-range—still more short-range than are the cohesive forces, inasmuch as these come into play to capture the atoms and hold them together, before those get their opportunity of crying "hold, enough!" They must be very potent, for the most terrific pressures which have been achieved by man do not avail to squeeze the most compressible solid into half of its original volume. Why talk of artificial pressures? everywhere in the globe of the earth, except within a hundred miles of the surface, the pressure is greater by far than any of them; and yet, the average density of the earth is less than double that of its superficial crust.

We have imagined that as two atoms approach each other, the gravitational force between them rises gently, the cohesive force remaining undetectable till they come very close together, when at some critical distance it begins a sharp and sudden rise which quickly carries its value far over that of gravity. Now we are to conceive of yet a third force, repulsive, undetectable till they come still closer together, then at a lesser critical distance entering on a sharper more sudden rise which rapidly carries its value far over those of both of the other two.

This essential and powerful force has no name of its own. This is because it is usually described in words not conveying directly the notion of force. What we have now encountered is the concept of the incompressible atom, the particle of irreducible volume—the doctrine that the atoms are to be pictured not as infinitely small like the points of geometry, but as hard impenetrable elastic pellets, minute indeed but not inconceivably so. This is a doctrine frankly expressed by many a thinker of the past, who perhaps was more unwilling than we to receive uncritically that difficult dogma of the point of infinite smallness. Hearken again to Newton: "It seems probable to me that God in the beginning formed matter in solid, hard, massy, impenetrable, moveable particles . . . incomparably harder than any porous bodies compounded of them; even so very hard, as never to wear or break in pieces; no ordinary power being able to divide what God himself made one in the first creation."

The completely unsqueezable atom corresponds to a force of repulsion which passes suddenly from zero to an infinite strength at a certain critical distance. The critical distance is the "radius of the atom." Reversely the idea of a force of repulsion rising rapidly indeed, but always continuously, as two particles draw nearer—this corresponds to a squeezable atom, without a definite radius. Solids and liquids in bulk are compressible, and this seems to rule out the former idea, which anyhow is more drastic than one likes to accept. It is not ruled entirely out, for there may be interstices among the particles, and the shrinkage entailed by pressure may be ascribed to the atoms so setting themselves that the cavities lessen in size. However, this does not seem adequate, and it is better to accept a compressible atom and make it share with the cavities the responsibility for the shrinkage. Then there is also the fact that solids expand when warmed. This is ascribed to the atoms dancing around with the heat, and so we approach a new situation in which repulsion and motion are allied as the two antagonists to cohesion.

Instead of exploring this situation further, let us ask whether there is a difference between the concept of the more-or-less squeezable atom and that of the force-field curiously devised which I have been describing? Formally, there is not. But in respect of the path which the mind next tries to follow, there is a difference, and a great one.

The compressible atom being accepted, one asks, of what is it made? and finds that one is thinking of a continuous substance, elastic and dense. One who is trying to become a thoroughgoing atomist is hardly pleased to discover a continuum at the base of the theory. The displeasure would not be long-lasting, if by assigning a few simple qualities to the continuum one could arrive at the right numerical values for things that can be measured—if one could infer, for instance, that the continuum by its nature divides itself into globules of just the same radii as the structure of crystals demands for the atoms. We are to meet in nuclear physics with a calculation singularly like this—but in general, the feat has not been done. It is not an adequate retort to say that the thoroughgoing atomist is obliged to assign to his atoms the sizes and the masses which they actually have, without giving any deeper reason. He manages to avoid the question; it becomes imperious, when the continuum is brought upon the scene. The road to success may lie by way of the continuum, but it is a road that has not been successfully trodden.

The force-field around the point-particle being accepted, one asks, why this so curious force-field? An inverse-square field would seem so natural as not even to ask for further explanation (but this is probably because the human mind has had two and a half centuries for getting accustomed to it). This combination of a short-range attraction with a repulsion still shorter in range cries out from explanation. Could one but somehow reduce it all to inverse-square forces, one would be more contented. This road seems impassable, but already it has been trodden—built and trodden—to splendid successes. Therefore I lay aside the compressible atom scooped out of a continuum, mentioning that even now we have not heard the last of it. Two stages of preparation are now required.

First, I must take more care henceforward in using the words "atom" and "particle". Hitherto I have used them interchangeably; from this moment on, "atom" is to have one meaning and "particle" another. Of the two, it will be "atom" which comes the closer to meaning what both words have meant up to now. Atom will attract atom by the force of cohesion; atom will repel atom by the nameless short-range force. The atoms in their turn will be made up of more elementary particles, bearing such names as "nucleus" and "electron". As to the forces between them,—that is the topic to which we are coming.

Second, I must introduce at long last the forces which the reader has so long been missing from this discourse: the electromagnetic.

Of these, it is the "electrostatic" force which stationary charges exert on one another which concerns us the most. Newton spoke of it in one of the passages which I have just been citing, but the pleasure was denied him of knowing how it resembles gravity. Both follow the law of the inverse-square; yet two centuries were to elapse between the years when Newton proved this for the one and Coulomb for the other. The electrostatic force is broader though than gravity, for it includes an attraction and a repulsion. There are two categories of charge, the positive and the negative: any charge repels those of its own category, attracts those of the other.

This entry upon the scene of a long-range repulsion modifies the prospects of a successful picture of the world as a congeries of particles, and seems at first glance to brighten them greatly. Dismiss gravity—forget about cohesion—put the question: in an imaginary universe made up of electrified particles some positive and some negative, acting on one another by electrostatic forces only, is it possible to have stability with all of the particles standing still?

Again the answer is no. This is not, however, too disappointing: we are accustomed to motion as the antagonist of gravity in the celestial case; shall we not now introduce it to be an ally to the electrostatic repulsion, the two of them conjointly being the antagonists of the attraction?

Now with real surprise and disappointment, one stands confronted again by the ruthless negative answer. The past revives: I have said that a pre-Newtonian philosopher would scarcely have accepted motion as the deathless antagonist to gravity, because he would have believed that motion dies out of itself. Well, the motion of an electrified particle *does* die out of itself —so says the electromagnetic theory. A proviso must here be inserted for correctness' sake, though it does not alter the situation. Uniform motion does not tend to die out—but uniform motion is useless to our ambitions. The orbital motion of a planet, the swing of a pendulum,—on these the theory must be built; but these are accelerated motions; and accelerated motions destroy themselves, when the moving body is electrified. Their energy passes into light, and the body sinks to rest. Aristotle was avenged in the nineteenth century on those who sneered at him; for what he had believed of motion generally, was in effect what they believed of the motion of electricity. Still, as nearly everyone knows, there is, after all, an electrical theory of matter; the elementary particles are deemed to be electrified, and the forces between them are deemed to be electromagnetic.

How is all this to be reconciled? By a statement which is the prelude to the final one—provided, that is, that all works out as well as physicists now hope, and provided also that we avert our eyes from the phenomena called "nuclear". Having imagined the elementary particles as points possessed of mass and bearing charges, and acting upon one another by electromagnetic forces, we are to treat their motions by the method of quantum mechanics, and not by the method of classic mechanics.

I will not pretend that this is a slight innovation, nor try to represent it as anything less than a great and difficult revolution in some of our most cherished habits of thought. Concepts formerly sharp, even those of position and motion themselves, become hazy; there are pitfalls and labyrinths; the mathematical technique is novel and hard. Yet in the picture of the universe as now presented, there are particles possessed of charge and mass; there are electromagnetic forces between the particles; there is motion of the particles; there is radiation, which it is just barely permissible to disregard in an outline like this one, and which I am disregarding; and outside of the realm of "nuclear" phenomena, there is nothing else. The stability of the world, that is to say, of the picture, is assured by attractions and repulsions electrical in nature, and by motion, with radiation playing an essential part.

The hydrogen atom appears before the eye of the mind as a system of a nucleus and an electron: two particles of known, equal and opposite charges, of known unequal masses, attracting one another by electrostatic force. The force draws them together, but there is kinetic energy and there is motion, and so they stay apart. It takes a definite amount of energy to separate them, and the theory derives its actual value very exactly from a basic principle. Any other atom appears before the mind as a system of a nucleus and two or more electrons. The nucleus bears a positive charge, the electrons are negative; the nucleus attracts the electrons, but they repel one another; there is motion; between the attraction and the motion and the repulsion, there is stability. A molecule is a system of two or more nuclei positively charged and two or more electrons negatively charged, and the same three qualities hold the balance. A tangible piece of metal is an enormous multitude of nuclei and electrons, these latter enjoying a very wide variety of motions, some moving almost as freely as though the metal were a vacuum: again the balance is held, the metal tending neither to shrink nor to explode.

All this is a programme for the explanation of Nature; and it is a programme which has been largely fulfilled—wherefore this lecture and a portion of the course. Not everything has been explained, nor ever will be. Quite apart from the phenomena called nuclear, there are countless things and happenings on earth which are so complicated, that they may well obey our fundamental laws without ever giving us the chance to prove it. If we should apply our assumptions to them and start to work out the consequences, it would take a geological era to finish the job. Perhaps all phenomena of life are of this type. The most that can be asked for is, that the theory should deal capably with all the phenomena for which it cannot reasonably be claimed that they are so complex as to defy any theory. I do not allege that our theory of massive particles, electromagnetic forces and quantum mechanics has done even this. It has, however, done a great deal, so much that it takes a rather skeptical physicist to deny it in the realms to which it lays claim.

In the light of this theory, let us consider the situation of the several forces.

Gravity remains apart and inaccessible, one of the ultimate forces, quite probably a quality of space as Einstein has proposed.

The *electromagnetic forces* remain ultimate, not explained in terms of anything else, united among themselves by the theory of relativity, responsible for the incessant passage of energy to and fro between matter and light which is one of the major features of the world. The ionization of atoms, the generation and the absorption of light, show us these forces at work within the atoms, holding together the electrified particles of which the atoms are made, balanced by motion and by their own dual character of attractions and repulsions.

Cohesion, and the chemical forces which bind atoms into molecules and grade insensibly into cohesion, and the nameless *repulsive* force which holds the balance to them and led many to the concept of the more-or-less-compressible atom: these are derivable from the electromagnetic forces between the elementary particles whereof the atoms are made up. I repeat: *derivable from the electromagnetic forces, with the aid of quantum mechanics,*—without which aid they would not have been derived. In the literature one finds incessant reference to "exchange forces"; these are not a novel category, but a step in the derivation.[3] Here are the fields of research where work is the most active. The theory of chemical forces, which some call "quantum chemistry", is well advanced; the theory of metals, not so well. Much earlier and much more often than we like, do we impinge on the class of phenomena, for which it can all too reasonably be claimed that they are so complex as to defy the theorist probably for all time. Yet there are many simple ones which have brilliantly been explained, and there is satisfaction on the whole—until one raises the eyes and looks ahead: for the nuclear phenomena are still before us.

As a prelude to these we may view the electron itself. Hardly have we begun to "look narrowly" upon it, before we see the spectre rising up of that old antithesis between the point-atom and the atom carven out of a continuum; nor is it long before the spectre grows more frightful than it was in the earlier case. If the point-electron is adopted, all the old conceptual

[3] There is also a strange quality of Nature bearing in quantum-mechanics the name of "the exclusion-principle of Pauli," which to some extent resembles a repulsive force acting between similar particles such as electron and electron or proton and proton, under very special conditions.

troubles return in the company of a new one. The intrinsic energy of this point-particle is infinite—so says the electromagnetic theory; the mass must therefore be infinite—so says the relation of Einstein of which I will presently show the power. If from this alternative we rebound to that of a globule of continuous electrical fluid, the old difficulties come back in the company of another new one. The parts of the globule of negative electricity repel each other, so our electron-model turns out to be a high-explosive bomb. The reader if he wishes may seek in Lorentz' "Theory of Electrons", a classic of some thirty years ago, the details of a scheme for preventing the electron from exploding by means of nonelectrical forces—a surrender, therefore, of the viewpoint that the ultimate forces are electrical.

Leaving these difficulties still unmastered, I turn to nuclear physics. This is a term which covers two fields: on the one hand, the structure and the qualities of atom-nuclei; on the other, some remarkable attributes of electrons, which they display either when they have tremendous energies, or under conditions which it takes tremendous energies to create. "Tremendous" energies are enjoyed by electrons fresh from radioactive substances, are obtained from the cyclotron and the electrostatic generator, and are found at their extremest in the cosmic rays. Of these attributes the only one which I will mention is mortality.

Mortality: this is a very obnoxious attribute for an elementary particle. All atomists heretofore have devised their atoms specifically to be immortal, to be *the* immortal things, to be the one thing permanently changeless under the flux of phenomena. But the electron is mortal, subject to birth and to death. Electrons are born in pairs, a positive and a negative springing together into existence. Electrons die in pairs, a positive uniting with a negative and the two of them passing out of existence.

These are not exactly cases of something coming out of nothing and something turning into nothing. Energy, mass and momentum are all conserved. Corpuscles of light disappear where and when an electron-pair is born, are born where and when a pair of electrons vanishes. So far as can be told, the corpuscles of light possess just the energy, just the mass and just the momentum which is destined to go to the nascent electrons or to be left unpossessed by those about to die. Now I have to admit my fault in not elevating earlier the corpuscles of light to a parity with the electrons and the atoms. They have the singular attribute of moving always with the same speed (when in a vacuum); they do not collide with one another, or rather such collisions have not been detected, though collisions with electrons are known; and they suffer from mortality, very much more so than do electrons. (Positive electrons are so rare, that negative electrons enjoy an almost perfect security.) Immortality is reserved for energy and mass and momentum. Now we feel ourselves swerving again toward a continuum-theory. The ground is slippery, and I step hastily from it into the last section of this lecture, into nuclear physics proper.

All of the theory of nuclei is firmly grounded on one basic statement, which is this: the masses of all nuclei are *nearly* integer multiples of a common unit, this being *slightly less than* the mass of the lightest among them.

Here is a statement bitterly disappointing! the little word "nearly" and the three little words "slightly less than" conjointly make a bright hope stillborn. Were it not for those words, we should already have joyously leaped to the conclusion that all nuclei are clusters of a single kind of fundamental particle, different clusters differing only in how many of the particles they comprise. The conclusion is so tempting that one is quite unable to resist it, hoping against hope that the words of frustration can somehow or other be cancelled. Soothing the reader with this veiled assurance, I adopt the conclusion.

The conclusion itself must be tempered at once, for there is a second basic statement coequal with the first: the charges of all nuclei are integer multiples of a common unit of charge. No pernicious adverbs here! this statement is an exact one, to the best of our knowledge and belief. The common unit of charge, as nearly everyone knows, is equal to the electron-charge and positive in sign.

The conclusion would still be sound, if the charges of all nuclei were proportionate to their masses (we should merely attribute an equal charge to every particle). Definitely this is not so, being most strikingly denied by the fact of "isobars": there are nucleus-types agreeing in mass, disagreeing in charge. We seek the next simplest assumption, and find that it suffices: Two types of fundamental particles—equal in mass—the one of them charged positively, the other neutral—each nucleus to be distinguished by two integers, one being the number of the charged component particles of the cluster, the other the number of the neutrals—"proton" and "neutron" for the names of the two.

This is the beginning of the programme for nuclear theory. Having taken the first step by writing it down, we enter upon the second—and find ourselves on the very road which our ancestors trod when atomic theory was new, facing the same ascents, the same passes and the same morasses. The long-range forces—the short-range forces—the cohesion—the repulsion—the more-or-less-incompressible particle—the troubles of the concept of the point-particle—the countervailing troubles of the continuum carven into globules—the dream of reducing everything to long-range forces and motion holding each other in balance—every one of these rejoins us on our journey. The mighty difference is, that the road still ends in the darkness, and the dream is still a dream. Therefore it is that the language of nuclear theorists wanders about in the most disconcerting way, so that often in a single article the wording in one place will be intelligible only to a few hundred (if so many) of the most advanced of specialists, and in another will sound like the voice of Newton speaking out of the *Opticks*, only in a much more cumbrous manner.

In the atomic world we have already seen how gravitation is neglected, being pushed into the background by the electromagnetic forces and the cohesions and repulsions derivable from these. Now in their turn the electromagnetic forces must recede into the background. This sounds extraordinary. Have we not all been told of the supreme importance of nuclear charges? Have we not been taught that by its charge a nucleus attracts electrons and organizes them into a family about itself and so creates an atom,—an atom which coheres with others, so that the world as we know it is organized by the charges of nuclei? All this is true, and very important from our viewpoint—but not so important, it seems, from the viewpoint of a nucleus. To this little cluster of protons and neutrons, the mass is more important than the charge; the total number of its component particles is more important than the number of protons separately or the number of neutrons separately; the cohesive forces are more important than the electrical. Perhaps a nucleus cares little about its charge, and nothing at all about the swarm of electrons which that charge coerces to swirl about it like a cloud of flies, though if it were not for those swirls the world would be barren and dead.

The cohesive forces certainly overpower the electrical. We are in no doubt of this, for the electrical forces are repulsive. Newton had gravity available for binding his atoms together; it was of the right type but inadequate, so he gave it cohesion as an ally. The electrostatic force between proton and proton is a repulsion, so to bind such particles together the Newtons of nuclear physics must overcome it with cohesion as an adversary. How greatly it is overcome is shown in much the same sort of way, as I followed when invoking the vaporization of solids to show how greatly the cohesion of atom with atom surpasses gravity. It is possible (at the end I will mention how) to compare the amount of energy required for tearing apart a cluster of two protons and a neutron with that required for tearing apart a cluster of two neutrons and one proton. The two amounts differ by only a few per cent; and more surprising yet, the former is the greater! Though the first-named of the clusters contains the inherent explosive power of two protons trying to drive themselves apart by the long-range repulsion, it is stuck tighter together than the other, which contains nothing of the sort. As a minor detail this shows that the cohesive forces depend to some extent on whether the particles are neutrons or protons; but the major conclusion is, that the cohesive forces are the masters.

Are they short-range or long-range? By calling them "cohesive" I have already committed myself, but correctly. There is an argument quite similar to the second which I drew from the vaporization of liquids. Think again of the kettle of water boiling away on the stove. It takes as much energy from the flame to disperse the last cubic inch of water that goes as it does to drive off the first, despite the fact that the first is exposed to all the long-range forces of attraction exerted on it by all the other cubic inches remaining in the kettle, and the last is not. Therefore the long-range forces which act between atom and atom are trivial, and cohesion is a force exerted by the atoms on their near neighbors only. Think now of the cluster of protons and neutrons which is a nucleus—a massive one by choice, built of two hundred particles or more. Imagine it taken to pieces by detaching one particle after another. I admit that this precise experiment is beyond the art of the physicist, but for a certain reason—the one which I have already promised to give, and will give at the end—he is as confident of its result, as he ever is of the result of any experiment which he has not actually performed. The result is, that it takes *roughly* as much energy to remove a particle when there are two hundred left behind to pull it back, as when there are but a dozen left behind, or any number in between. Therefore the long-range forces which act between the fundamental particles are minor, and the intra-nuclear cohesion is a short-range force.

I have carefully made these last statements rather weaker than their analogues for the water boiling away. The amount of energy required for taking away a particle does depend to some extent on the number left behind, and the long-range forces are therefore minor but not trivial. If the long-range forces are attractive, the binding-energy of a particle—this is the shorter name which is given to the "energy required for taking away a particle"[4]—must be greater, the greater the size of the cluster, *i.e.*, the greater the mass of the nucleus. Now for nuclei of some fifty particles or more, the contrary is the case. Therefore the long-range force, or the major one if there are more than one, is a repulsion. We already know of one long-range repulsion, to wit, the electrostatic force between proton and proton. Is this the force in question? The answer is oddly difficult to give with assurance, but at present is believed to be *yes*.

If the answer is definitely *yes*, then the electrostatic force has after all one role of supreme importance in nuclei. It fixes their maximum size and their maximum charge, therefore limits the number of chemical elements, and may indeed be all that prevents the universe from caving together into a single lump of protons and neutrons with the electrons fluttering help-

[4] It ought strictly to be called the "unbinding-energy" or "binding lack-of-energy," since it is given as positive when energy must be contributed to the system in order to detach the particle.

lessly around it. So long have the chemists been on the search for new elements, and so completely have they searched, that we may believe them when they say that apart from the works of the "atom-smashers," no nucleus exists having more than 238 particles altogether, 92 of which are protons. Even the atom-smashers or (as I should rather call them) the transmuters, for all the wonder and power of their art, have not forced the total number of protons upward by more than two or the total number of particles altogether upward by more than one. Moreover all of the two dozen or so most massive nuclei known are subject to explosion—to explosions quite terrific, some of them spontaneous, others touched off by what seems a very minor cause. It may therefore be taken as nearly certain that there is an upper limit to the size of nuclei, and probable that it is electrostatic force that sets the limit.

Now we come down to the short-range repulsion. Such a one there must be, for again we can rehearse the ancient argument. A piece of iron does not shrink into a point; therefore the iron atoms must either exert a force of repulsion or else be more-or-less compressible pellets. A nucleus does not shrink into a point, but offers an impenetrable front, measurable though small, to an oncoming neutron; therefore the nuclear particles—but why repeat the words?

Shall we interpret neutrons and protons alike as systems of particles still smaller, acting on one another by electromagnetic forces, to be treated by quantum mechanics? Alas, if there is one surety in this field, it is that we cannot play quite the same game twice. Quantum mechanics may not be used up (some think that it is) but the electromagnetic forces certainly are. In this direction we have as yet no leadership.

Shall we then adopt the compressible globule or the point-particle with a curious field of force surrounding it? Though the language of nuclear theorists verges sometimes on the former, it is the latter practice which is common—a fact which will hardly surprise the reader. In the specialized literature, one finds many a speculation and (what is of more moment) many an inference about the force-field which is drawn pretty directly from reliable data. As a rule the inferences are expressed in language very different from the phrases of this lecture: "interaction" is used instead of "force-field," and there are queer and slightly comic technical terms such as "potential-well." When you read of a "rectangular potential-well," interpret that what I have been calling the "cohesive force" becomes suddenly enormous at a certain specific radius; when of an "error-well" (!) understand that the cohesive force increases rapidly according to a certain law with decline of distance; when of a "Coulomb interaction" realize that it is the inverse-square force-field of the electrostatic repulsion between proton and proton. Of these interactions I will give only two facts: first, that the short-range attraction is confined within a very few times 10^{-13} cm of the centre of the proton or neutron, whereas the cohesive attraction of atom for atom spreads over a radius a hundred thousand times as great; second, that the three short-range attractions of proton for proton, neutron for neutron and neutron for proton are nearly the same.

Shall we adopt the force-fields as given to us by experiment, with some plausible assumptions added (for one cannot as yet do without them) and operate on them by the procedures of quantum mechanics, hoping to arrive at (say) values of binding-energies compatible with the data? This is the present, or perhaps I should say the recent, programme of nuclear theory. If one reads the theoretical papers of any one year out of the last ten, one may readily get the impression that success is just around the corner. But if one reads the papers of two or more years and takes note of the rapid changes, the prospect does not look quite so rosy—nor when one overhears the conversations of the theorists themselves. I will not conduct the reader down the paths which are as yet so tortuous and hazy; it will be better to fill in the picture with a few of the many remaining details.

Mass was the first of properties (along with hardness) to be assigned to the elementary particles; the second was charge; to these have lately been added angular momentum and magnetic moment. It is difficult to say when the idea of a spinning atom was first propounded (one recalls the vortices in a continuous fluid which Kelvin introduced as one of the most brilliant of all attempts to contrive a continuum and atoms as a part of it) but easy to fix the time when the idea of the spinning electron became so definite and sharp, as to be successfully used in explaining crucial data; this was 1925. The electron, the proton and the neutron all have equal angular momentum; its amount, common to these three which at present claim most strongly the rank of *elementary* particle, is one of the universal constants. When protons and neutrons are assembled in a nucleus, their axes of spin all point in an identical direction, though not by any means necessarily in the same sense in that direction. It is possible for a nucleus to have zero angular momentum, through half of its particles setting themselves in the one sense and half in the other; the lightest nucleus for which this happens is the alpha-particle, composed of two protons and two neutrons. The magnetic moments of the three elementary particles are very far from equal, that of the electron being some seven hundred times as great as that of the proton, which in turn is half again as great as that of the neutron. One of the tragedies of theoretical physics occurred in this connection. A principle of quantum mechanics had been proposed, superbly capable of serving as a basis for most of the incomplete principles which had already so well justified themselves in atomic physics, and including among its parts the actual values of the angular momentum and magnetic moment of the electron. Its empire would have been extended, had the ratio of the magnetic moments of proton and electron been equal to the reciprocal of the ratio of the masses of these two—actually the former ratio is too great by a factor of 2.78. This contretemps has led many to deny the title of "elementary" particle to the proton; while as for the neutron, the fact that it lacks an apparent electric charge while nevertheless displaying a magnetic moment leaves it also open to suspicion.

Few readers of these pages will be unaware that electrons are observed proceeding out of nuclei: it may well be a source of wonderment that they are denied a residence in these assemblages of protons and neutrons only. This is of course another example of the mortality of the electron. Having observed that it is subject to birth and to death, should we be deterred from supposing that it is born as it quits the nucleus from which it comes? This rhetorical question gives a false impression of the course of history. There was indeed an era when electrons were believed to inhabit nuclei, when nuclei were regarded as assemblies of protons and electrons only. It ended in 1932; but the observation of the birth and the death of electrons did not ensue for yet another year. What happened in 1932 was the discovery of the free neutron. Only when this particle had been discovered did a physicist (Heisenberg) think it worth while to begin to develop in detail the theory that the components of nuclei are protons and neutrons and no other particles but these.

Now I bring this article to a close by fulfilling my promise to speak of Einstein's relation between energy and mass, which on the one hand has been rigorously tested in the realm of nuclear physics, and on the other has extended that realm.

The relation may be worded in several ways; I will employ the shortest: *energy has mass.*

Now imagine an assemblage of particles sticking together. "Sticking together" is not the dignified phrase of a physicist; such a one would say, more abstractly but more exactly, that energy must be given the particles to take them apart. But energy has mass; therefore the mass of the assemblage must be augmented, when they are taken apart. Therefore the mass of the interconnected assembly is less than the sum of the masses of the particles when free.

Now with a single stroke this principle does away with what otherwise would have been a quite unsurmountable obstacle to the doctrine that all nuclei are made up of protons and neutrons. For "proton" and "neutron" are not merely the names of hypothetical particles whereof nuclei are made up; they are also the names of the two lightest of nuclei. These two lightest of the nuclei are so massive, that it could not possibly be said that the other nuclei are made up of them, were it not for the detraction of mass which occurs when they are bound up together. This deficit of mass corresponds to the unbinding-energy or, badly called, the binding-energy of which I earlier spoke. The binding-energy is the amount of energy which must be supplied to the nucleus, to break it up into protons and neutrons. The deficit of mass—the difference between the actual mass of the nucleus, and the masses of all of its neutrons and protons dispersed into freedom—is related to the binding-energy by Einstein's relation.

I have said that this relation has been tested in the realm of nuclear physics, and has served also to extend that realm. The possibility of testing arises from the fact that in certain cases the physicist is able to convert a system of two nuclei into a system of two other nuclei, the masses of all four being known. This seems a somewhat pedantic way of expressing the well-known fact that in performing an act of transmutation, the physicist causes one nucleus as "projectile" to impinge upon another as "target," whereupon the two merge and two others spring apart from the scene of the merger. The masses of the two initial nuclei do not as a rule add up to the same precise sum as the masses of the two final nuclei. But if to the first pair of masses we add that of the kinetic energy of the projectile, and if the second pair is augmented by that of the kinetic energies of the final nuclei—why, then, the equation balances, and Einstein's relation is justified.

As for the extensions of the realm of nuclear physics, or let me rather say, the realm of physics generally: no fewer than three have been stressed in these few pages. First, mass could not be conserved in the birth or the death of electron-pairs, were not the energy of the electrons accompanied by its mass when it passes out of or into the form of radiant energy. Then, we should not so soon have known that the system of two protons and one neutron requires less energy to unbind it, than the system of two neutrons and one proton; this was deducible from the masses of these two nuclei, before it was attested by the discovery that the former changes spontaneously into the latter. Then, we should not have the evidence that the binding-energy of the individual particle lessens, as the number of particles remaining behind in the nucleus increases; for this is a statement derived from observations on the masses of the nuclei.

So all seems well with the model of the nucleus as a system of protons and neutrons, and the particle-theory stands triumphant. Yet notice at what a price this triumph has been bought! Of all the attributes of the fundamental atom, of the elementary particle, constancy of mass was the earliest and the most firmly accepted. The elementary particle was a bit of immutable mass, set forever apart from change. Now it turns out that when the particle adheres to another, some of its mass departs. What has departed is not perished and gone. It is known sometimes to have passed into radiant energy, sometimes into energy of motion, sometimes into that mingling of the two which is known by the name of heat. Changelessness has ceased to be the quality of the atom, remaining that of the mass and the energy of the world as a whole. Immortality has gone from the atom back into the continuum. This is as good a place as any to step out from the incessant alternation, never yet ended and probably endless, between the particle and the continuum as the basis of thought about physics.

Abstracts of Technical Articles by Bell System Authors

Tentative Standards for Wood Poles Become Approved American Standards.[1] RICHARD C. EGGLESTON. The six American Tentative Standards covering specifications for wood poles, several of which were approved by the American Standards Association as tentative in 1931 and the rest in 1933, have now been reviewed by the ASA committee and approved by the ASA as full American Standards. In reviewing the standards, the committee found that the general principles of the standard requirements have been universally recognized as a satisfactory basis for the selection of poles. Covering as they do northern white cedar poles, western red cedar poles, chestnut poles, southern pine poles, lodgepole pine poles, and Douglas fir poles, the standards represent a rational uniform standardization system for the six major pole timbers of the United States.

The standards establish practical limits that can be applied economically in the production of poles for general use, but they are intended also to be flexible enough to cover the purchase of poles of high quality for special purposes. At the same time, it is not desired that they should be so restrictive that any considerable quantity of usable poles produced under normal production practices would be labeled substandard because of the specification restrictions.

The standard specifications include material requirements for shape and straightness of grain, limit defects such as knots, checks, insect damage and decay, and define the minimum quality of acceptable poles. In the standards, departures from straightness are held within practical limits for ordinary use. Decay and the presence of wood-rotting fungi are generally prohibited. Definite limitations on knots are set, and fire-killed poles are acceptable only by special agreement between producer and purchaser.

The standard dimensions now included with the specifications in one standard for each type of pole, were based on recommended fibre stresses contained in the American Standard for Ultimate Fiber Stresses of Wood Poles (05a-1930). They were approved as American Standards from their inception, and until they were included with the specifications in the present American Standards they were considered as separate standards. These standard dimensions have all been prepared according to the same principles for all types of poles. The sizes at six feet from the butt in all six standards have been so fixed with respect to ground-line resisting moments,

[1] *Industrial Standardization,* June 1941.

that, for any given class and length of pole, all six species are equal in strength. In calculating these six-feet-from-butt sizes, distance from the butt to the ground line for any given pole was assumed, by definition, to be as shown in the column in the tables headed "Ground Line Distance from Butt." The equality-in-strength principle holds good, however, for any reasonable depth of set required.

Approval of the six standards at this time followed a policy adopted by the Standards Council of the American Standards Association in April 1939. At that time the Standards Council decided to withdraw approval of standards having a tentative status, and requested the reconsideration of such standards with the idea of either discarding them or of advancing them to American Standards.

Equilibrium Relations in the Solid State of the Iron-Cobalt System.[2] W. C. Ellis and E. S. Greiner. There are important transformations in the solid in the iron-cobalt system. One of these originates from the A_3 transformation in iron. Cobalt in the binary system at first raises the A_3 transformation to a maximum in the region of 45 weight per cent cobalt. Further additions decrease the temperature of transformation which rapidly approaches room temperature in the region of 80 weight per cent cobalt. An extended two phase region from 76.5 to 88.5 weight per cent cobalt was established at 600 degrees Cent. (1110 degrees Fahr.).

An order-disorder transition occurs in the alpha phase in the region of 50 weight per cent cobalt. The critical temperature of order is in the neighborhood of 700 degrees Cent. (1290 degrees Fahr.) depending upon the composition. The ordered arrangement has the cesium chloride structure.

The lattice constants of the alpha phase deviate widely from a linear function of the cobalt content. The first additions of cobalt increase the cell size to a maximum at approximately 20 per cent cobalt. Further additions result in a contraction in the cell size to the limit of the alpha phase. Compositions in the region of 50 per cent cobalt exhibit an increase in cell size on ordering.

Determination of Microphone Performance.[3] F. L. Hopper and F. F. Romanow. Methods of determining the performance characteristics of microphones by acoustic measurements are described. Work factors involving the accuracy of the methods are discussed. The correlation between a microphone's performance as determined by acoustic measurement and by listening tests is reported. Application of both types of test to a studio type of cardioid microphone is given as an example.

[2] *Trans. Amer. Soc. for Metals*, June 1941.
[3] *Jour. Soc. of Motion Picture Engineers*, April 1941.

Room Noise Spectra at Subscribers' Telephone Locations.[4] DANIEL F. HOTH. That room noise can be a distinct handicap to conversation by masking the speech sounds in the ear of the listener and thus impairing the ease and accuracy of reception is of considerable concern to the telephone engineer. Room noise not only complicates the problems involved in the design and engineering of telephone systems capable of affording satisfactory service, but it is also one of the factors which affect the costs of the telephone plant. The effects of noise on telephone conversation depend, of course, upon the characteristics of the noises which occur at the places where telephones are being used. The arrangements and practices necessary for reducing the effects of noise depend upon a knowledge of these characteristics. As a result numerous measurements of room noise have been made from time to time over a period of many years by Bell System engineers. For the most part such measurements have involved the determination of a single figure to represent the noise measured, as in the recent survey of sound levels described by Mr. D. F. Seacord in the July 1940 issue of *The Journal of the Acoustical Society of America.* While such measurements are invaluable in providing information on the frequency of occurrence of different noise levels at telephone locations, their value is enhanced by additional measurements of the distribution of the noise energy throughout the frequency band involved in the reception of speech. The present paper describes such measurements and shows the effects of a number of contributing factors on the spectrum of the noise. It is shown that the spectrum of room noise has a characteristic shape.

Film Scanner for Use in Television Transmission Tests.[5] AXEL G. JENSEN. This paper describes the design and construction of a television film scanner primarily intended for use as a testing tool in designing circuits suitable for television program transmission.

The equipment employs electronic scanning and the image dissector is used as the electronic pickup device. The image dissector has a high degree of linearity between light input and signal output and the picture signal is not accompanied by any spurious shading signals. Furthermore, the direct-current component of the television signal is directly available at the output of the tube. The lower sensitivity of the dissector tube is not important in this case since a highly efficient optical projection system makes it possible to override noise to a high degree.

In film scanners for entertainment purposes it is desirable to use ordinary 24-frame motion pictures and such film scanners therefore include a me-

[4] *Jour. Acous. Soc. Amer.*, April 1941.
[5] *Proc. I.R.E.*, May 1941.

chanical or optical translating mechanism for translating the 24-frame film picture into a 30-frame interlaced television picture. In the present equipment it was found more expedient to simplify the construction by allowing the use of specially printed film. Ordinary 24-frame film is "stretched" by printing every other frame twice and the remaining frames three times in succession, thereby producing a film with a total of 60 frames instead of the original 24. Vertical scanning is then obtained by the continuous motion of this film at the rate of 60 frames per second and horizontal scanning by a simple electronic line sweep in the dissector tube.

Acoustic Design Features of Studio Stages, Monitor Rooms, and Review Rooms.[6] D. P. Loye. A survey was made of studio experience, and measurements were made of stages, review rooms, and other units. These data were correlated and used as a valuable guide in the determination of the optimum characteristics and dimensions recommended for major studio scoring stages, monitor rooms, dubbing rooms, review rooms, and studio theaters.

Information regarding Hollywood preview theaters is included in an Appendix.

A New Microphone Providing Uniform Directivity over an Extended Frequency Range.[7] R. N. Marshall and W. R. Harry. A new microphone is described which consists of a moving coil pressure element combined with an improved ribbon pressure gradient element to give a cardioid directional characteristic. The theory of operation is reviewed, and consideration is then given to variations in directivity caused by diffraction, separation of the elements, and disparities in their phase and response characteristics. It is then shown how these variations are largely eliminated by equalization in the electrical circuit so that the resulting directivity is practically independent of frequency throughout the range from 70 to 8000 cycles. The use of a moving coil pressure element makes high efficiency possible, while the design of an unusually rugged ribbon element provides a marked reduction in noise due to air currents. Several useful directional patterns in addition to the cardioid pattern are provided in the new microphone, and the theory and merits of these patterns are presented. Finally some of the results which were obtained in field trials of the new microphone are discussed.

The Magnetostriction, Young's Modulus and Damping of 68 Permalloy as Dependent on Magnetization and Heat Treatment.[8] H. J. Williams, R. M.

[6] *Jour. Soc. of Motion Picture Engineers*, June 1941.
[7] *Jour. Acous. Soc. Amer.*, April 1941.
[8] *Phys. Rev.*, June 15, 1941.

Bozorth and H. Christensen. This paper describes measurements of the changes in certain physical properties of 68 Permalloy that result from different thermal and mechanical treatments and considers them in relation to the domain theory. The magnetostriction varied with heat treatment from 2.5×10^{-6} to 22×10^{-6}. The change in Young's modulus with magnetization to saturation varied from 0.09 to 10.5 per cent. The damping of mechanical vibrations was also measured as dependent on magnetization and heat treatment. Young's modulus and the damping constant were determined by measuring the natural frequency of vibration and the width of the resonance curve of a hollow rectangle magnetized parallel to its sides so that the magnetic circuit was complete without air gaps or end effects.

Contributors to this Issue

Karl K. Darrow, B.S., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., University of Chicago, 1917. Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. As Research Physicist, Dr. Darrow has been engaged largely in writing on various fields of physics and the allied sciences.

Thornton C. Fry, A.B., Findlay College, 1912; A.M., University of Wisconsin, 1913; Ph.D., 1920. Mathematician, Western Electric Company, 1916–24; Bell Telephone Laboratories, since 1924. Dr. Fry is now Mathematical Research Director in the Laboratories.

C. M. Hebbert, B.S., Otterbein College, 1911; M.S., University of Illinois, 1914; Ph.D., University of Illinois, 1917; Research Assistant for National Advisory Committee on Aeronautics, 1919; Instructor in Mathematics, University of Illinois, 1919–20. American Telephone and Telegraph Company, 1920–29; Bell Telephone Laboratories, 1929–. Dr. Hebbert is engaged in the study of general transmission engineering problems.

F. R. Stansel, B.S. in E.E., Union College, 1926; M.E.E., Polytechnic Institute of Brooklyn, 1934; D.E.E., Polytechnic Institute of Brooklyn, 1941. Bell Telephone Laboratories, 1926–. Dr. Stansel has been engaged in the design of special laboratory measuring and testing apparatus.

VOLUME XX OCTOBER, 1941 NUMBER 4

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

50c per copy *$1.50 per Year*

THE BELL SYSTEM TECHNICAL JOURNAL

Published quarterly by the

American Telephone and Telegraph Company

195 Broadway, New York, N. Y.

SUBSCRIPTIONS

Subscriptions are accepted at $1.50 per year. Single copies are 50 cents each. The foreign postage is 35 cents per year or 9 cents per copy.

The Bell System Technical Journal

Vol. XX *October, 1941* *No. 4*

The Reliability of Holding Time Measurements

By ROGER I. WILKINSON

I—The Problem to be Solved

ONE of the fundamental quantities in traffic engineering is the average duration of subscribers' calls. This figure in seconds multiplied by the average number of calls expected over a given route in an hour, and divided by 3600, gives the traffic load submitted in average simultaneous calls—or "the average" as it is commonly called. Tables and curves are widely available which may then be consulted to find the number of paths to be provided so that no more than a desired small percentage of the calls presented will find all paths busy.

The direct measurement of call lengths with a stop watch occurs to one as being the simplest means for obtaining a sample of holding times. It is seldom used, however, due to the relative slowness with which a large number of observations are accumulated coupled with the not inconsiderable expense of the small army of observers required, each looking at one call at a time.

A second direct method of obtaining holding time measurements is by recording mechanically or electrically the length of each call passing over a group of switches or trunks during a certain interval of time. Various holding time recorders or "cabinets" following this principle have been used more or less extensively in the Bell System. Their chief disadvantage has lain in requiring considerable time and labor for summarizing the results. Problems of the perfect maintenance of the measuring equipment have also been present.

To make possible the rapid accumulation of holding time data on a considerable number of calls at relatively slight expense, the method of switch or plug counts has been introduced. This consists in scanning mechanically, electrically, photographically or by eye the group of paths at regular intervals, and recording each time the number found busy.[1] Such data give estimates immediately of the average load being carried, and by a

[1] This number will be highly variable, and even on properly engineered groups great concern need not be felt should few, or even no, cases of "all paths busy" appear since such peaks are of short duration and might easily be missed except in a very long series of counts.

relatively simple analysis a measure of the reliability of such an estimate can be obtained. If in addition for the same period a record is kept on a call- or peg-count meter of the number of calls passing over the group, it is possible also to obtain estimates of the average call holding time and the reliability of such an average.

Direct measurement of holding times or switch counts should naturally be made on groups during periods which are presumably typical of those toward which the engineering is ultimately directed. Usually, although not always, this will be the busy or busiest hours of the day during the busy season of the year. In order to decide intelligently how long a period needs to be studied in any given case some knowledge of the persistence of the same holding time universe is necessary. This might be obtained through relatively small holding time samples made in the hours of interest every day for several weeks in the busy season. If spottiness or "lack of control" is not apparent, the problem will be comparatively simplified. If trends are present, however, it will be necessary to investigate their nature (such as whether some one day of the week shows high holding times) and apportion the main sampling procedure in a fashion to give these peculiarities their proper weighting.

It will be of interest to examine in this respect certain limited data at hand taken by the pen register method some years ago on an inter-office trunk group in Newark, New Jersey. The kind of examination made here will serve to indicate the procedure which may be found suitable in some degree for application to other groups whose characteristics are relatively little known.

II—Preliminary Study of Newark Data

It has long been known that local subscriber call holding times, t, follow remarkably closely the simple exponential frequency distribution,

$$f(t)dt = ke^{-kt}dt, \tag{1}$$

where $\frac{1}{k}$ = the average holding time.[2] This was found to be substantially true of the data collected on the inter-office trunk group in Newark as shown in Fig. 1 for 7385 calls observed in 19 hours having loads in the range 15.0–16.0 average simultaneous calls. The fit of the exponential curve having an average equal to the observed average of 2.380 minutes is seen to be quite good. It may be further noted that in the exponential distribution the standard deviation, σ, equals the mean $\bar{t}$. In practice σ

[2] A. K. Erlang apparently was the first to notice this holding time distribution, "Nyt Tidsskrift for Matematik" (Denmark), 1909.

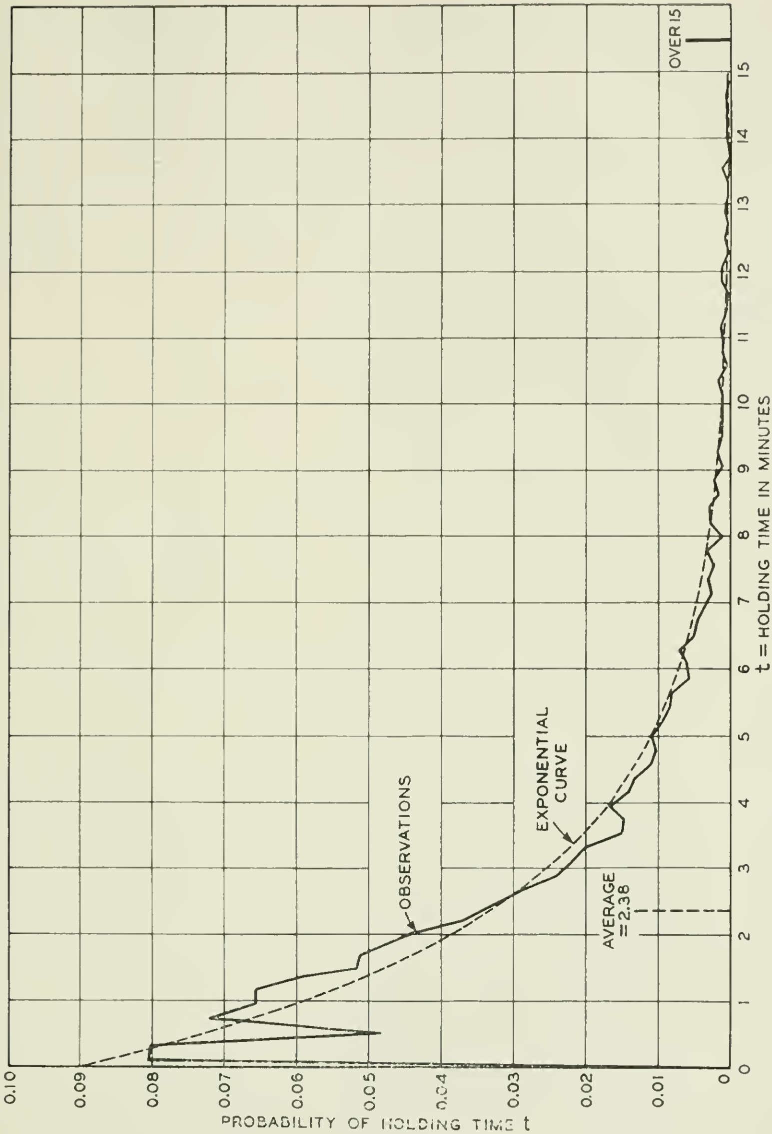


Fig. 1—Distribution of 7385 local holding times, Newark

is usually found to be slightly larger than $\bar{l}$ although not markedly so. We may use this information to test the homogeneity of the holding time universe should all hours of the days be grouped indiscriminately.

TABLE I
HOURLY HOLDING TIME DATA, NEWARK
(Figures in body of table are average holding times in seconds)

Day	Date	Hour of day							
		9-10 am	10-11 am	11-12 am	12-1 pm	1-2 pm	2-3 pm	3-4 pm	4-5 pm
Monday	7- 8-18	131.0	139.0	140.5					
Tuesday	7- 9-18	151.1	151.0	159.0					
Wednesday	7-10-18	146.3	161.4	140.0					
Thursday	7-11-18	138.5	133.7	151.4					
Friday	7-12-18	—	146.0	138.1					
Saturday	7-13-18	123.7	139.5	130.5					
M.	7-15-18	135.1	152.6	147.0					
T.	7-16-18	134.5	138.4	—					
W.	7-17-18	138.4	151.3	159.8					
Th.	7-18-18	148.2	148.0	—					
F.	7-19-18	147.1	136.4	—					
S.	7-20-18	146.9	131.8	—					
M.	7-22-18	145.2	146.7	148.7					
T.	7-23-18	154.5	145.3	143.7					
W.	7-24-18	132.5	137.7	157.5					
Th.	7-25-18	—	149.0	—					
F.	7-26-18	138.0	157.9	174.4					
S.	7-27-18	128.5	142.0	150.1					
M.	7-29-18	132.3	141.5	—		166.7			
T.	7-30-18	151.3	143.4	139.5					
W.	7-31-18	142.1	129.4	144.1					
Th.	8- 1-18	141.1	134.4	154.1					
F.	8- 2-18	161.6	150.5	150.3					
S.	8- 3-18	148.7	147.7	134.5					
M.	8- 5-18	139.4	131.0	142.9					
T.	8- 6-18	158.0	141.4	158.5					
T.	8-13-18	162.8	141.6	—					
W.	8-14-18	136.0	150.8	—					
Th.	8-15-18	153.0	141.8	139.0					
F.	8-16-18	141.0	160.1	151.6					
Th.	9- 5-18	144.7	158.9	—					
F.	9- 6-18	—	139.5	—					
W.	9-25-18	—	—	144.3				157.6	
Th.	9-26-18	152.1	143.1	134.2					
F.	9-27-18	139.7	160.5	149.9					
M.	9-30-18	—	132.8	128.9		138.7			152.6
T.	10- 1-18	129.7	137.5	150.0		158.3			166.0
W.	10- 2-18	138.5	135.2	132.5			161.0		
Th.	10- 3-18	142.0	143.0	152.0	174.3	153.0	165.0		
F.	10- 4-18	128.4	136.9	145.7		150.5			
S.	10- 5-18	137.0	138.4	138.5					
M.	10- 7-18	—	—	—		150.0	139.2	137.7	
T.	10- 8-18	131.0	136.4	145.1			150.0	145.1	152.2
W.	10- 9-18	138.3	144.4	142.0		174.8		145.4	151.6
Th.	10-11-18	135.4	149.8	—					
Summary:									
No. Hours		39	43	33	1	7	4	4	4
Average		141.63	143.67	146.01	174.3	156.3	153.8	146.4	155.6

In Table I and on Fig. 2 are shown the holding time averages for 135 hours observed at various times of the day over a period of 3 months. At first glance these appear to fall in two rather distinct groups, those before noon and those after noon. If the 115 hours before noon be considered as defining a homogeneous group, could those holding time averages found in the afternoon be reasonably considered as coming from the same universe? We first find the average holding time of the 115 forenoon hours to be 143.5 seconds. Since these hours averaged about $n = 390$ calls each, the standard deviation of the means $\sigma_{\bar{t}}$ should, by theory, be closely

$$\sigma_{\bar{t}} = \frac{\sigma}{\sqrt{n}} = \frac{\bar{t}}{\sqrt{390}} = 7.29.$$

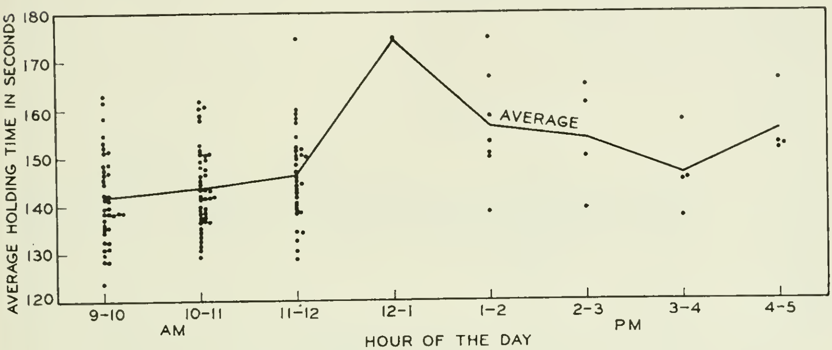


Fig. 2—Day to day holding time averages by hours of the day, 135 hours, Newark

The standard deviation observed is 9.26, some 27% higher, which, however, agrees with the observation made in the previous paragraph. On the hypothesis that the universe of 115 early hours has the parameters of $\bar{t} = 143.5$ and $\sigma = 7.29$, we see that the observations for each of these three clock hours could readily have occurred. The deviations of their averages from 143.5 are 1.9, .17 and 2.5 seconds, respectively, and according to theory the corresponding standard errors in these averages are 1.168 $\left(= \frac{7.29}{\sqrt{39}}\right)$, 1.111 $\left(= \frac{7.29}{\sqrt{43}}\right)$, and 1.270 $\left(= \frac{7.29}{\sqrt{33}}\right)$. All the deviations are well within two times the standard error of the assumed mean of the holding time universe. The remaining 20 observations from noon on, however, average 154.6 seconds, and if they could reasonably have come from the hypothesized universe, this figure should not differ from 143.5 by more than, say, three times the standard error 1.630 $\left(= \frac{7.29}{\sqrt{20}}\right)$. Actually the difference is more than six times the standard error, strongly indicating a significant difference between the forenoon and afternoon holding times. We conclude that between 9 a.m. and noon the holding times are satisfactorily controlled but that we should not attempt to include observations on afternoon hours with them. Since the heaviest loads here occurred generally in the morning we should confine our direct measurements or switch counts to these hours for determining the engineering holding time.

It may occasionally be well to investigate the possibility that certain days of the week have, on the average, longer holding times than other days. If the Newark 9–12 a.m. data are plotted by days of the week as in Fig. 3 we see that the averages for each day fluctuate considerably as shown by the heavy dots. In testing these points the simple average of the σ's for

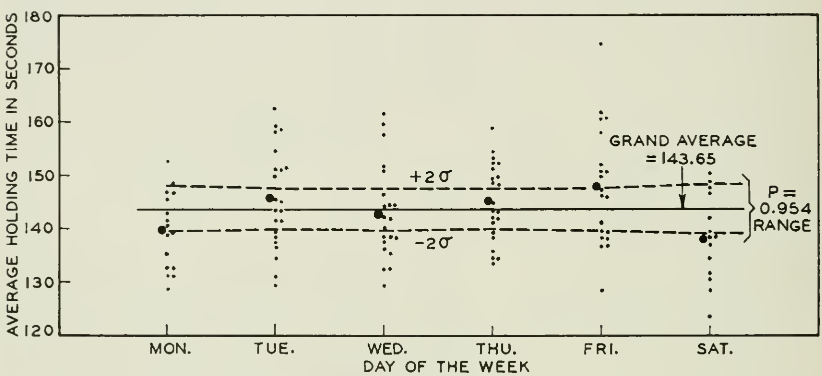


Fig. 3—Variations in average holding times, 115 9–12 a.m. hours, by days of the week, Newark

each day's hours is taken as an estimate of the standard deviation of the homogeneous universe from which all the hours are presumed to be drawn. Then with the weighted arithmetic means of the daily averages as the best estimate available for the mean of the universe, each day's average is tested to see whether it could reasonably have arisen from it. The $\pm 2\sigma$ lines which should include some 95% of the day-averages are shown on Fig. 3. It is seen that two of the six points fall slightly outside these limits indicating a moderately significant difference in the holding time conditions for Friday and Saturday. The sampling procedure to follow in such cases of non-controlled populations is not rigorously definable. However, it is clear that the samples should be drawn from the various groups of controlled elements which probably go to make up the universe, and roughly in proportion to the importance to be assigned to each such group. In our example here we would probably want to draw samples of about equal size from the calls of each week day in the week.

Finally there may be some question as to the busy season, its length and stability. Plotting of the same data for the 9–12 a.m. hours as in Fig. 4A and 4B will help to decide these points. The morning hours' holding time averages for each day of the week are plotted for several weeks during the suspected busy season. Wide changes in the load through the passage

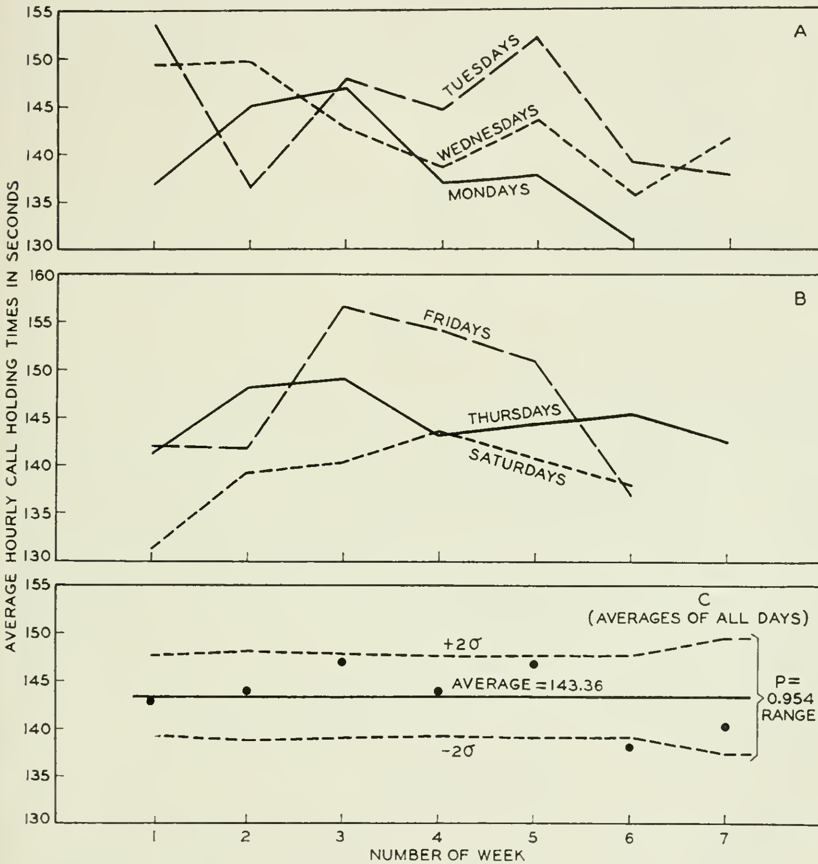


Fig. 4—Seasonal trend in holding time averages, 105 9–12 a.m. hours, Newark

of weeks can be noted by eye. In the case illustrated there appears no consistency of movement. Applying the identical test used in the previous paragraph for the day-of-the-week changes, we find in Fig. 4C that in the first five-week period no movements of significance took place. However, the sixth week, which followed the first five after an interval of about six weeks, showed a significant drop suggesting the approach of a lower level of traffic. The traffic engineer would probably decide to schedule his holding time observations during the weeks numbered one to five inclusive.

Having determined something as to the character of the holding time trends, if any, with hours of the day, with days of the week, and seasons of the year the traffic engineer is in a better position to lay out a program for sampling. He will especially want to apportion the total sample between the hours or days which show significant differences among themselves roughly in proportion to the relative traffics flowing at those levels. The less specific the information on the traffic flow characteristics the more important it will be to spread the observations over a variety of hours, days or weeks.

III—A Satisfactory Sample of Directly Measured Holding Times

If the standard deviation σ of individual call lengths is known, we can estimate the standard error of the average of n measurements as

$$\sigma_{\text{avg}} = \frac{\sigma}{\sqrt{n}}. \tag{2}$$

Since n will usually be several hundred we can obtain a good figure for σ by calculating the standard deviation, S, of the n observations. As noted before, for exponential calls this will be not far from the average holding time $\bar{t}$ which may be substituted for σ if great accuracy is not required. In fact if the sampling is representatively made from a universe not strictly homogeneous, the better figure for σ may be the average $\bar{t}$, instead of the standard deviation found in the sample since in so-called Poisson Sampling of stable but nonhomogeneous universes the standard error of the average may be somewhat reduced from $S/\sqrt{n}$.

We may now make the statement that for n large the probability is P that the true average holding time does not differ from that observed by more than $\pm z \dfrac{\bar{t}}{\sqrt{n}}$ seconds, where P and z are given in the table below.

TABLE II

P	z	P	z
.50	.6745	.95	1.960
.85	1.440	.99	2.576
.90	1.645	.999	3.291

For example if we have measured the individual lengths of 900 calls which show an average of 150.3 seconds, we are then 99% sure that the true holding time average for the sampled universe lies closely within the range

$$150.3 \pm 2.576 \frac{150.3}{\sqrt{900}},$$

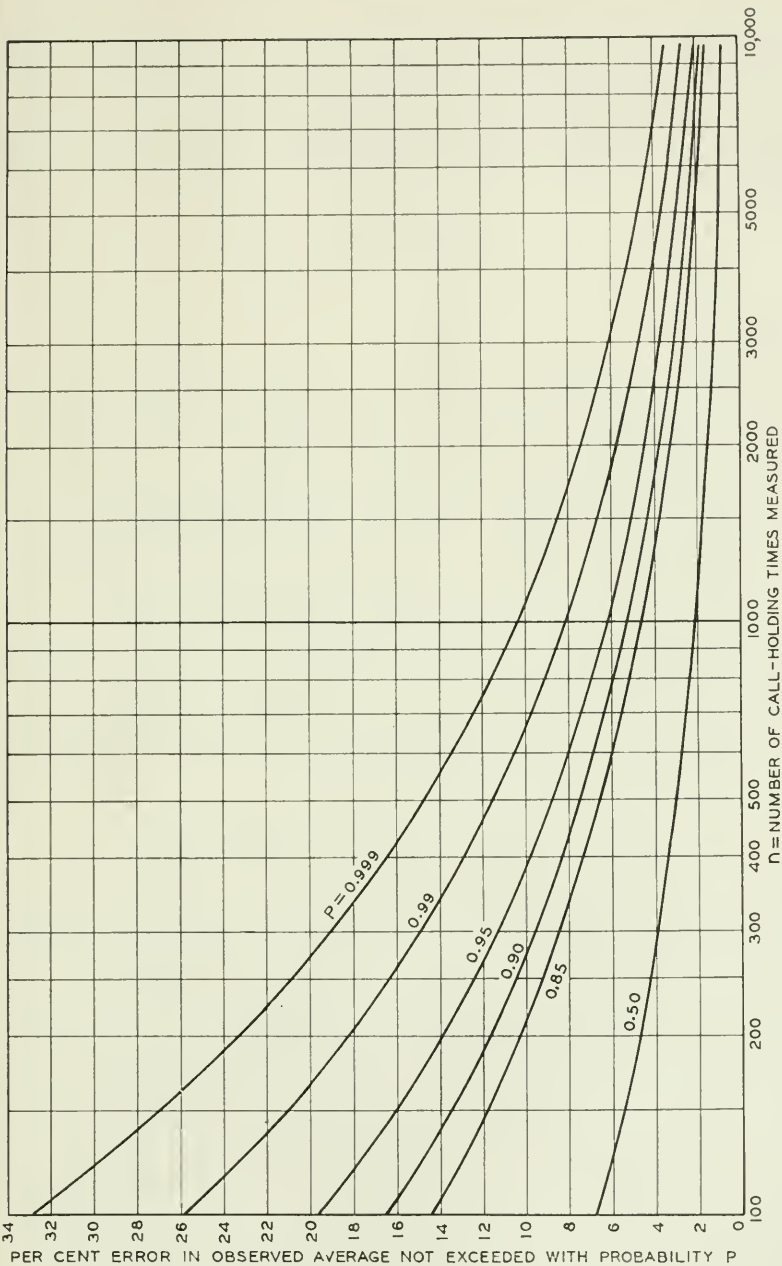


Fig. 5—Error in exponential-holding-time average determined by stop watch method

that is between 137.4 and 163.2 seconds. (The best single estimate, of course, is the observed average of 150.3 seconds.) Or conversely, if one should desire to determine the true average holding time within 5 seconds with a surety of $P = .90$, he may use an approximate value of the standard deviation (or the average holding time) based on past experience and substitute in

$$n = \left(\frac{1.645\bar{t}}{5}\right)^2,$$

to obtain the number of calls to be measured. If $\bar{t} = 150$ seconds here, $n = 2421$.

The same information is contained in Fig. 5 which gives the *per cent* error, $\pm 100 z_p/\sqrt{n}$, in the observed average not exceeded with probability P.

All this is based on the assumption that each of the n call lengths is accurately enough measured so that no appreciable error is introduced from this source. Obviously there is no point in expending much effort in carefully "proportioning" a sample so as to be representative of the vagaries of the universe if each of the calls so chosen is not pretty accurately measured. This would be quite as futile as measuring very accurately the holding times of a number of calls chosen during some short time period which might turn out to be wholly untypical of certain of those important periods coming earlier or later. For these direct measurement cases it will probably be quite satisfactory if each call is measured with a maximum error of not over one-tenth of $\bar{t}/\sqrt{n}$. In our example of 900 calls this would be .501 seconds, that is measurement of each call to the nearest second.

IV—Holding Times by Switch Count Methods

If each call's holding time is not measured with considerable accuracy it is immediately clear that additional calls must be observed in order to compensate therefor. This is the situation in the method of switch counts which is in effect a means for noting at regular intervals i whether a particular call does or does not exist. Thus none of the calls are at all closely measured for their individual lengths. Other errors will also have to be considered since at the beginning of the period some switch counts are inevitably included on a number of calls from the preceding hour and at the end of the period some of the calls registered on the peg count meter will end beyond the period with the loss of part of their proper switch counts. As a result there are in this method three distinct sources of holding time error whose magnitudes we shall proceed to investigate in turn:

a. Errors at the start of the observation period;
b. Errors at the end of the observation period;
c. Errors at the beginning and end of each call.

The theoretical conclusions will be compared at various points with certain data available.

a. *Errors at the Start of the Observation Period*

If the period of observation T be divided into r equal intervals of length i, and switch counts are made at the beginning (and end) of each interval, we shall have a total of $r + 1$ observations. The #1 count will give us immediately the number of calls extending into the period from the preceding hours. We have no means, however, of segregating their contributions to subsequent switch counts, so a theoretical estimate of this amount is required.

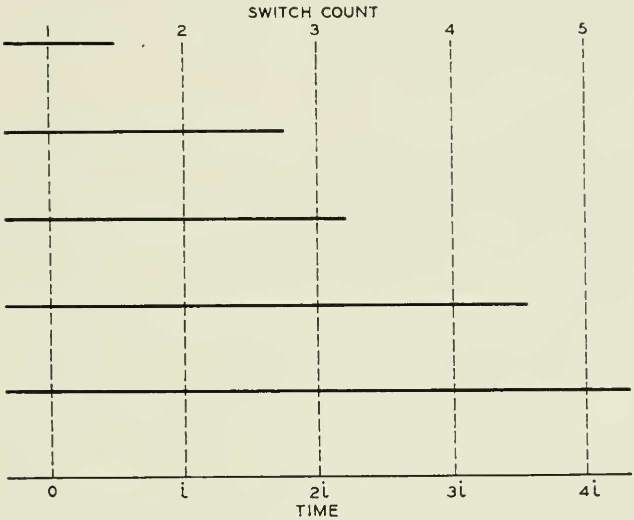


Fig. 6—Diagram of switch counts at beginning of the period

If the average holding time of the calls is $\bar{t}$, and they follow closely the exponential law of distribution, we may reason as follows.[3] Consider the case of a single call passing time 0 at the start of the observation period as in Fig. 6. Then the probability that it will be included only in switch count number 1, that is that it ends between time 0 and time i, is

$$p_1 = P(<i) = 1 - e^{-\frac{1}{\bar{i}}}.$$

[3] Of course we do not know $\bar{t}$ exactly since that is the ultimate object of our study; however, for the present purpose great accuracy will not be required, and $\bar{t}$ can usually be taken as the first estimate of holding time obtained by the switch count method without corrections.

Similarly the probability that exactly two switch counts will be contributed by such a call is

$$p_2 = P(>i) - P(>2i) = e^{-\frac{i}{\bar{t}}} - e^{-\frac{2i}{\bar{t}}}.$$

Likewise,

$$p_3 = P(>2i) - P(>3i) = e^{-\frac{2i}{\bar{t}}} - e^{-\frac{3i}{\bar{t}}},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$p_u = P[>(u-1)i] - P(>ui) = e^{-\frac{(u-1)i}{\bar{t}}} - e^{-\frac{ui}{\bar{t}}}.$$

If now there have been m such calls observed on switch count number 1, we shall need to add m variables of the type

$$f(u) = e^{-\frac{(u-1)i}{\bar{t}}} - e^{-\frac{ui}{\bar{t}}} = e^{-\frac{ui}{\bar{t}}}(e^{\frac{i}{\bar{t}}} - 1) = ce^{-\frac{ui}{\bar{t}}},$$

where u may take all values from 1 to $r + 1$. The exact addition of these variables when m is more than a small number, say 3 or 4, becomes quite complex. However, in such cases (which may be the rule) we revert to the method of combining their individual moments to obtain the moments (and parameters) of the resultant distribution. We find for a single variable,

$$\bar{u} = c\sum_{u=1}^{r+1} ue^{-u\frac{i}{\bar{t}}} = \frac{1}{1 - e^{-\frac{i}{\bar{t}}}}\left[1 - e^{-(r+1)\frac{i}{\bar{t}}}\right], \tag{3}$$

and

$$\sigma_u = \frac{e^{-\frac{i}{2\bar{t}}}}{1 - e^{-\frac{i}{\bar{t}}}}\left[1 - 2(r+1)\left(1 - e^{-\frac{i}{\bar{t}}}\right)e^{-r\frac{i}{\bar{t}}} - e^{-(2r+1)\frac{i}{\bar{t}}}\right]^{\frac{1}{2}}. \tag{4}$$

The factors shown in brackets in equations (3) and (4) will approach unity very closely in any practical applications of the present type; they will therefore be omitted in the subsequent analysis.

The mean and standard deviation of the sum of m such variables are readily determined, of course, as[4]

[4] It is interesting to note that if the first switch count had been omitted so that only $\bar{m}$ could have been estimated from the average of the switch counts from #2 onward, we might have assumed a Poisson distribution for m, that is $\rho_m = \frac{\bar{m}^m e^{-\bar{m}}}{|\underline{m}}$, and thereby have obtained an estimate of the switch counts contributed by calls from the preceding period as follows,

$$\bar{s}'_m \approx \frac{\bar{m}}{1 - e^{-\frac{i}{\bar{t}}}}, \tag{5'}$$

$$\sigma'_m \approx \frac{\sqrt{\bar{m}(e^{-\frac{i}{\bar{t}}} + 1)}}{1 - e^{-\frac{i}{\bar{t}}}}. \tag{6'}$$

$$\bar{s}_m = m\bar{u} = \frac{m}{1 - e}, \tag{5}$$

$$\sigma_m = \sqrt{m}\,\sigma_u = \sqrt{m}\,\frac{e^{-\frac{i}{2t}}}{1 - e^{-\frac{i}{t}}}. \tag{6}$$

In Fig. 7 is shown a comparison between this theory and the actual numbers of 1-minute switch counts contributed by these carry-over calls

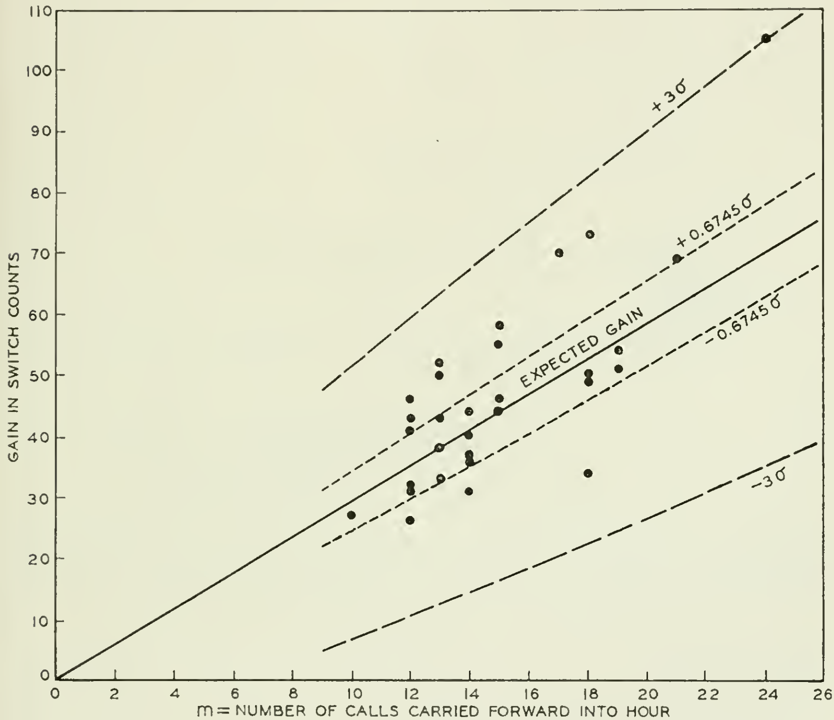


Fig. 7—Gains over true switch counts at the start of the observation period

in 30 hourly periods of observation on the Newark interoffice trunk group previously described. Since $\bar{m}$, the average number carried forward is about 15, it was thought that the sum-distribution of so many variables would rather closely approach normality. On this assumption the ±50% and ±99.73% "control" lines have been drawn on Fig. 7. One point falls on the latter limit lines, while 16 of the 30 fall within the 50% lines, thus providing a gratifying corroboration of the theory.

b. Errors at the End of the Observation Period

If switch counts have been made at regular intervals i so that the $r+1$st count occurs at the exact end of the period for which the number of originating calls has been registered, then the situation closely resembles that at the start of the period. The $r+1$st observation tells us immediately how many calls are continuing into the next hour. A particular one of these calls may extend to the areas 0 to i, i to $2i$, $2i$ to $3i$, etc., measured beyond the end of the period as in Fig. 8. A call ending in the interval $2i$ to $3i$,

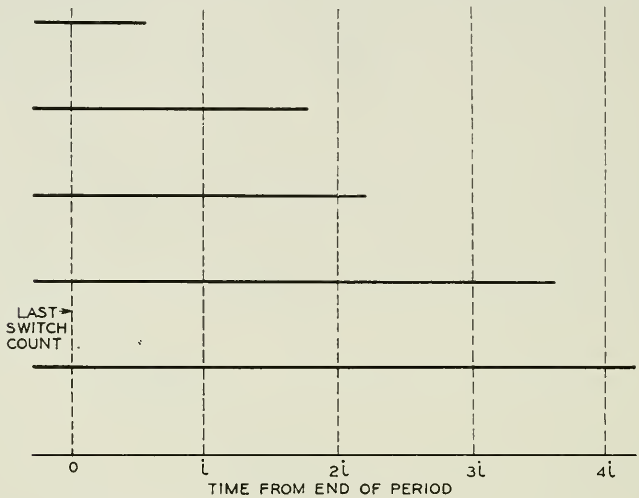


Fig. 8—Diagram of switch counts at the end of the period

for instance, would fail to have two switch counts marked up if the counting stopped at 0. Then the probability of losing exactly zero switch counts is

$$p_0 = P(<i) = P(>0) - P(>1i) = 1 - e^{-\frac{i}{\bar{t}}},$$

and in general

$$p_v = P(>vi) - P[>(v+1)i] = e^{-\frac{vi}{\bar{t}}} - e^{-\frac{(v+1)i}{\bar{t}}} = e^{-\frac{vi}{\bar{t}}}\left(1 - e^{-\frac{i}{\bar{t}}}\right)$$

$$= c'e^{-v\frac{i}{\bar{t}}} \tag{7}$$

where v varies from o to ∞. The average and standard deviation of a single variable will be[5]

$$\bar{v} = \sum_{v=0}^{\infty} vp_v = \frac{e^{-\frac{i}{\bar{t}}}}{1 - e^{-\frac{i}{\bar{t}}}}, \tag{8}$$

[5] The value $\bar{v}$ is one less than $\bar{n}$ shown in equation (3) after neglecting the minute correction factor, since each call there received a switch count at 0 time which is omitted here. The standard deviation is identical with equation (4) without the correction factor, since a constant deduction has simply been made on each call.

and

$$\sigma_v = \frac{e^{\frac{i}{2\bar{i}}}}{1 - e^{-\frac{i}{\bar{i}}}}. \tag{9}$$

The frequency distribution of the sum of w such discrete variables v is readily found to be

$$f(v_w) = \left(1 - e^{-\frac{i}{\bar{i}}}\right)^w \frac{\underline{/w + v_w - 1}}{\underline{/w - 1}\ \underline{/v_w}}\ e^{-v_w \frac{i}{\bar{i}}}.$$

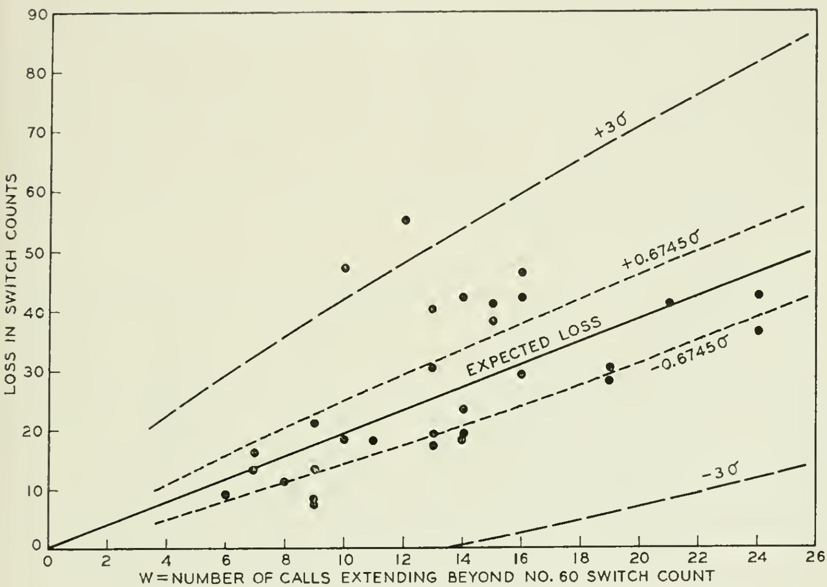


Fig. 9—Loss of true switch counts at the end of the observation period

The parameters of this sum distribution are

$$\bar{s}_w = w\bar{v} = \frac{we^{-\frac{i}{\bar{i}}}}{1 - e^{\frac{i}{\bar{i}}}}, \tag{10}$$

$$\sigma_w = \sqrt{w}\,\sigma_v = \sqrt{w}\,\frac{e^{-\frac{i}{2\bar{i}}}}{1 - e^{-\frac{i}{\bar{i}}}}. \tag{11}$$

A check on these last formulas is shown in Fig. 9 where the numbers of switch counts lost on w calls carried beyond the last (or r + 1st) switch count are recorded for 30 busy hour periods. Two of the hours show results outside the theoretical $\pm 3\ \sigma$ normal curve limits but the falling of almost exactly half of the points within the $\pm 50\%$ lines reassures us that the estimated parameters are probably quite good. (If instead we had calculated, say, $\pm 99\%$ limit lines based on the skew distribution $f(v_w)$ just derived, it seems likely that even the two "unusual" points of Fig. 9 might have fallen in the "reasonable" range.)

At this point it will be well to point out that in making switch counts considerable care needs to be exercised in two directions. First, switch counts $\#1$ and $\#r + 1$ should coincide very closely with the beginning and end, respectively, of the observation period, the intermediate counts of course being uniformly spaced. Second, each count should be taken as quickly as possible so that a substantially instantaneous reading of calls in progress is obtained. These two desiderata can usually be attained readily in schemes using mechanical, electrical or photographic means for recording the switch counts. Counts made by observers, however, may be subject to highly variable errors since in some cases a substantial portion of the interval i may be required to complete a count. Such uncertainties naturally increase the end-effect errors, and, consequently, the overall error in an average-holding-time determination.

To estimate the magnitude of the increased errors resulting from failure to meet the above switch count specifications would require a special study for each kind and type of failure; these would probably differ in every application. An idea of the sensitiveness of switch counting to such irregularities and the likely order of magnitude of the increased errors may be gained by examining certain of the Newark data. Here the last switch count in many of the hours, although taken instantly, failed to coincide well with the end of the observation period T. The last count fell at points varying from a little after the period closed to nearly an interval i ahead of this instant as shown in Fig. 10. In most hours this permitted a small number of calls to mark up on the peg count register after the last switch count was taken.

As a result, if corrections were not made, the estimate of switch counts lost on calls extending beyond the end of the hour would have omitted about a third of those properly included, with a consequent lowering of the average holding time estimate by approximately one per cent.

If the time j which has elapsed between the last switch count and the end of the period T is known, certain corrections for this particular irregularity can be attempted. We shall indicate the formulas required since they will be useful in an analysis of the Newark data. If an average of $\alpha = \frac{j}{i} a$ calls are assumed to originate in the interval j, and they follow a Poisson distribution $\frac{\alpha^x e^{-\alpha}}{\underline{/x}}$, then the expected number of switch counts lost is found to be

$$E = \alpha \frac{\frac{\bar{t}}{j}\left(e^{\frac{i}{\bar{t}}} - 1\right)\left(1 - e^{-\frac{j}{\bar{t}}}\right)e^{\frac{j-2i}{\bar{t}}}}{\left(1 - e^{-\frac{i}{\bar{t}}}\right)^2}, \tag{12}$$

and the standard deviation is

$$\sigma' = \sqrt{\alpha}\left[\frac{\bar{t}}{j}\left(e^{\frac{i}{\bar{t}}} - 1\right)\left(1 - e^{-\frac{j}{\bar{t}}}\right)e^{\frac{j-2i}{\bar{t}}} \frac{1 + e^{-\frac{i}{\bar{t}}}}{\left(1 - e^{-\frac{i}{\bar{t}}}\right)^3}\right]^{\frac{1}{2}}. \tag{13}$$

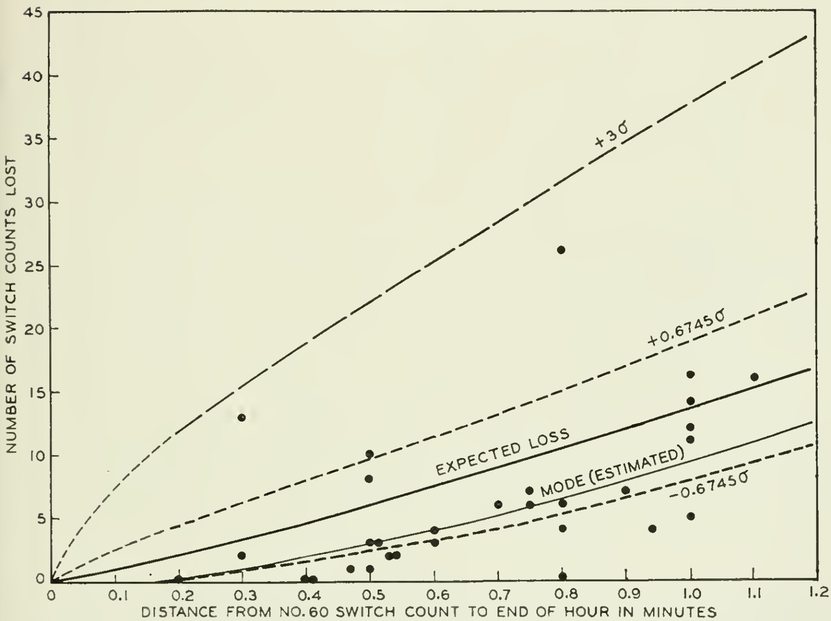


Fig. 10—Switch counts lost on calls which originated after the last switch count

If j is small, α will be correspondingly small making the average switch counts lost, E, small. The distribution of lost switch counts will then be very skew since the case of zero lost will be prominent. For larger j's the distribution should assume a unimodal form, gradually becoming less skew. A number of Newark busy hours were studied in this fashion, and the results are compared with theory on Fig. 10. The agreement is seen to be fair and a modal line estimated by eye falls substantially below the expected mean corroborating the decided skewness just predicted. The wide fluctuations in lost switch counts, each one of which if incorrectly estimated results in a considerable error (in the present example about .2 second) in the estimate of the average holding time, will serve to indicate the desirability where possible of eliminating altogether this and other supplementary errors by seeing that individual counts are taken very quickly, and that the first and last switch counts coincide closely with the ends of the observation period.

c. Errors in Measuring Each Call

Due to the method of counting the switches at finite intervals, an exact measurement of the length t of any one call will seldom if ever be made. We shall attempt in what follows to determine the magnitude and char-

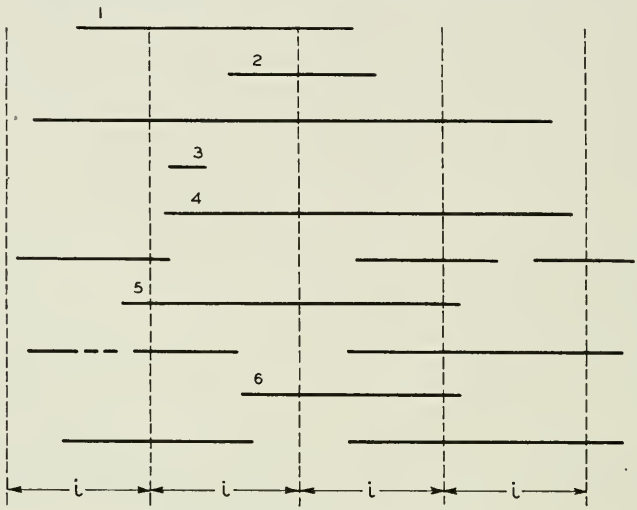


Fig. 11—Typical field of switch counts

acteristics of these errors in measuring individual calls. In any field of switch counts, such as in Fig. 11, there will be calls of types #1 and #2 which receive about one switch count for each i call seconds of length. There will likewise be many others such as #3 and #4 which will be substantially undercounted, and about as many others, #5 and #6, which will be overcounted. We shall proceed with certain special cases, and then obtain the general result desired.

Case 1. t lies between 0 and i

If the holding times t follow some law of fluctuation $f(t)$, a certain proportion of them will have lengths lying in the range $t = 0$ to $t = i$. Such a call will either cross one of the switch count points, or it will not. Upon the assumption of a random instant of origination the probability of its crossing will be simply $\frac{t}{i}$. That is, if the origination point falls within t seconds to the left of a switch count, the call will be marked up; if not it will not be counted. If it is marked up there will occur for that call a plus estimation error of $x = i - t$ seconds since we shall eventually assume that every switch count infers i call seconds of use on the trunk group. Likewise there is a probability of $\frac{i - t}{i}$ that the call will not be counted with the resultant negative error of $x = -t$ seconds. In summary the total probability of a positive error of $x = i - t$ is

$$f(t)\ dt\ \frac{t}{i};$$

and for a negative error of $x = -t$,

$$f(t)\ dt\ \frac{i - t}{i}.$$

Case 2. t lies between i and 2i

Such a call may be included either once or twice in the switch counts. The probability that it is counted twice is $\frac{t - i}{i}$ with a resultant plus error of $x = 2i - t$. The probability that it is counted but once is $1 - \frac{t - i}{i} = \frac{2i - t}{i}$, with the corresponding negative error of $x = i - t$. The overall probabilities of course will be formed by weighting these as in the first case with the probability, $f(t)dt$, that the holding time of length t to $t + dt$ actually occurs.

General Case. t lies between qi and (q + 1)i

It will readily be seen that by extending the reasoning of the two cases above to the case of t lying between qi and $(q + 1)i$ we shall have a plus error (due to the call being marked up $q + 1$ times) of $x = (q + 1)i - t$, with a probability of occurrence of $\frac{t - qi}{i}$, and a negative error (the call marked up q times) of $x = qi - t$ with a probability of $\frac{(q + 1)i - t}{i}$.

Summarizing the above cases, a *negative* error of size x can occur in a great number of ways, due to t taking the values $-x$, $i - x$, $2i - x$, $\cdots$ $qi - x$, $\cdots$ with the corresponding probabilities of occurrence of the call lengths, $f(-x)$, $f(i - x)$, $f(2i - x)$, $\cdots$ $f(qi - x)$ $\cdots$, respectively. In addition each time such a call length does occur we must introduce the contingent probability $\frac{i+x}{i}$ that a negative and not a positive error will occur. The total probability of making an error of x, where $x \overline{<} 0$, on any call is then,

$$p_{x \overline{<} 0}(x)\,dx = \frac{i+x}{i}\,[f(-x) + f(i-x) + f(2i-x) + \cdots]\,dx. \quad (14)$$

Similarly we find the total probability of making a positive error of magnitude x, on any call, as

$$p_{x \overline{>} 0}(x)\,dx = \frac{i-x}{i}\,[f(i-x) + f(2i-x) + f(3i-x) + \cdots]\,dx. \quad (15)$$

It will now be of interest to apply equations (14) and (15) to some particular types of holding time distributions.

(*a*). *Constant Holding Times, $t = h$*[6]

If t is constant and equal to h, it will necessarily fall within some one of the special cases enumerated above. Suppose h lies between qi and $(q+1)i$. There will be but one value of the error x_1 possible in the negative range and it will equal $qi - h$, with a corresponding single value x_2 in the positive range equal to $(q+1)i - h$. It will be seen that equations (14) and (15) reduce simply to

$$p_{x_1 \leq 0}(x_1) = p(qi - h) = \frac{i+x_1}{i}\,f(qi - x) = \frac{i+x_1}{i}\,f(h) = \frac{i+x_1}{i}, \quad (16)$$

and

$$\begin{aligned} p_{x_2 \geq 0}(x_2) &= p[(q+1)i - h] \\ &= \frac{i-x_2}{i}\,f[(q+1)i - x_2] = \frac{i-x_2}{i}\,f(h) = \frac{i-x_2}{i}. \end{aligned} \quad (17)$$

The mean and standard deviation of this two-valued variable are found to be

$$\bar{x} = 0, \quad (18)$$

$$\sigma_x = \sqrt{-x_1(i+x_1)} = \sqrt{(i-x_2)x_2} = \sqrt{-x_1x_2}. \quad (19)$$

It may be noted that σ_x attains a maximum of $i/2$ when $x_2 = i/2$, and approaches 0 for $x = 0$. This is of importance when one has to choose an observation interval i for switch counting constant or relatively constant holding times.

[6] The particular error distributions for cases *a* and *c* were obtained by G. W. Kenrick in 1923.

Example: An hour with 372 calls having a constant holding time per call of $h = 131.8$ seconds was subjected to a 60 second switch count study, records being kept of the errors in measurement on individual calls. As shown in Fig. 12, 284 or 76.3% gave counts of "2" with an error of $120 - 131.8 = -11.8$ seconds. The remaining 88 calls, or 23.7% received counts

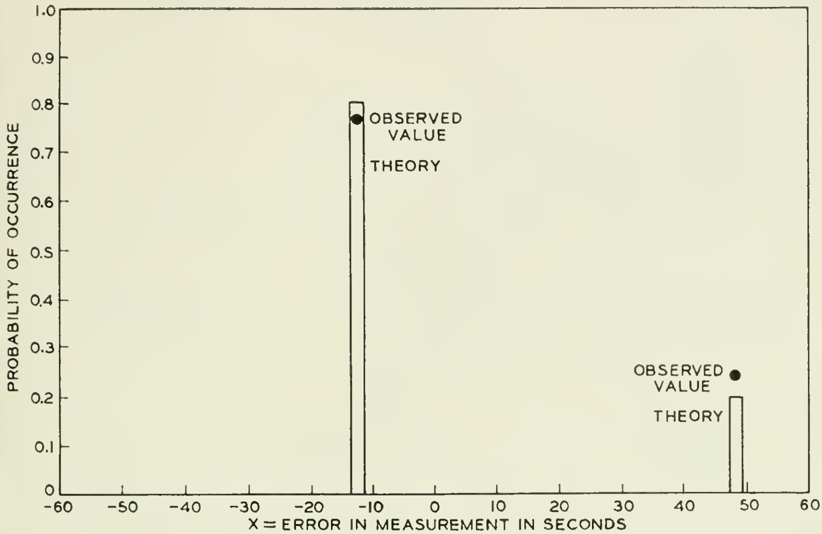


Fig. 12—Error distribution for measurements on individual calls with constant holding times

of "3", with errors of $180 - 131.8 = 48.2$ seconds. Applying the theory just developed to this case gives,

$$p(2i - h) = p(-11.8) = \frac{60 - 11.8}{60} = .803,$$

$$p(3i - h) = p(48.2) = \frac{60 - 48.2}{60} = .197.$$

As indicated on Fig. 12, these theoretical values check very satisfactorily with the observations. The observed average holding time = 134.2 seconds as against the true value of 131.8 seconds; the error of 2.4 seconds is quite compatible with $\sigma_x = \sqrt{11.8(48.2)} = 23.85$ seconds and the $n = 372$ calls observed.

(*b*). *Equally Likely Distribution of Holding Times between Adjacent Multiples of i*

Imagine a holding time distribution of any general form but with a constant probability of occurrence between adjacent pairs of multiples of i, such as in Fig. 13. Such a distribution would probably occur but rarely, if ever, in practice. However, if the intervals i are short compared to the average holding time $\bar{t}$, such an assumption may not introduce any serious discrepancy in whatever form is simulated.[7]

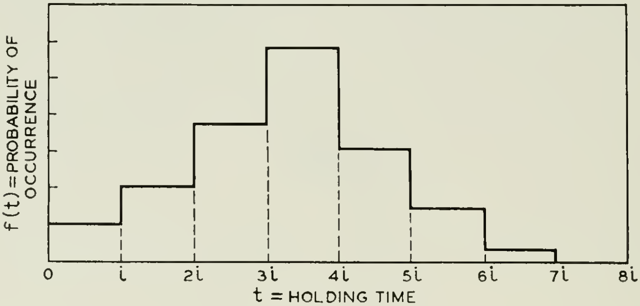


Fig. 13—A varying holding time with a number of equally likely ranges

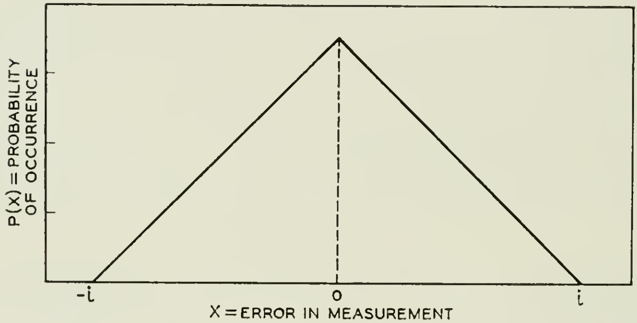


Fig. 14—Distribution of call measurement errors for "equally likely" holding time distributions

In this case it is obvious upon inspection that the sum of the terms in the brackets in equation (14) is a constant for all values of x, and likewise in equation (15), and that they equal each other. Hence

$$p_{x \lesseqgtr 0}(x) = K \frac{i + x}{i}, \tag{20}$$

and

$$p_{x \gtreqless 0}(x) = K \frac{i - x}{i}, \tag{21}$$

[7] The analytics of the allied case in which the errors at the ends of a call were assumed to fall equally likely between $\pm i/2$, were discussed by E. C. Molina in an unpublished memorandum dated September 7, 1920.

give the isosceles triangular distribution of errors on individual calls shown in Fig. 14. In this the average error is

$$\bar{x} = 0, \tag{22}$$

and the standard deviation is

$$\sigma_x = .408i. \tag{23}$$

(c). Holding Times Exponentially Distributed, $f(t) = ke^{-kt}$, *where* $k = \frac{1}{\bar{t}}$

With holding times of the exponential type the sum of the terms in the brackets of equations (14) and (15) will depend on the particular magnitudes of the errors x assumed. If in equation (14), we substitute exponential expressions for the f-functions, we have

$$\begin{aligned} p_{x \leq 0}(x)\,dx &= \frac{i+x}{i}\left(ke^{-k(-x)} + ke^{-k(i-x)} + ke^{-k(2i-x)} + \cdots\right)dx \\ &= \frac{i+x}{i}\,ke^{kx}\left(1 + e^{-ki} + e^{-2ki} + \cdots\right)dx \\ &= \frac{i+x}{i}\,ke^{kx}\,\frac{1}{1-e^{-ki}}\,dx \\ &= k'\,\frac{i+x}{i}\,e^{\frac{x}{\bar{t}}}\,dx, \end{aligned} \tag{24}$$

where

$$k' = \frac{1}{\bar{t}\left(1 - e^{-\frac{i}{\bar{t}}}\right)}.$$

Similarly we find

$$p_{x \geq 0}(x)\,dx = k'\,\frac{i-x}{i}\,e^{-\frac{i-x}{\bar{t}}}\,dx. \tag{25}$$

The mean and standard deviation of this unusual-shaped distribution of x are found to be

$$\bar{x} = 0, \tag{26}$$

$$\sigma_x = \sqrt{\bar{t}}\,\sqrt{\frac{(2\bar{t}+i)e^{-\frac{i}{\bar{t}}} - 2\bar{t} + i}{1 - e^{-\frac{i}{\bar{t}}}}}. \tag{27}$$

In Fig. 15 is shown the distribution of the individual errors found by 60-second switch counts on 746 varying holding time calls (2 hours on the

Newark group). Their true average holding time was 131.45 seconds. The mean error was found to be +1.84 seconds and the standard deviation 25.55 seconds. The corresponding theoretical distribution is found to have a standard deviation of 24.56 seconds with a mean, of course, of zero. The theoretical distribution is superposed on the data of Fig. 15 and is seen to give quite a good fit.

It is interesting that the theoretical average error for each of these three widely dissimilar holding time distributions should be zero, while their standard deviations and analytical forms assume quite different characteris-

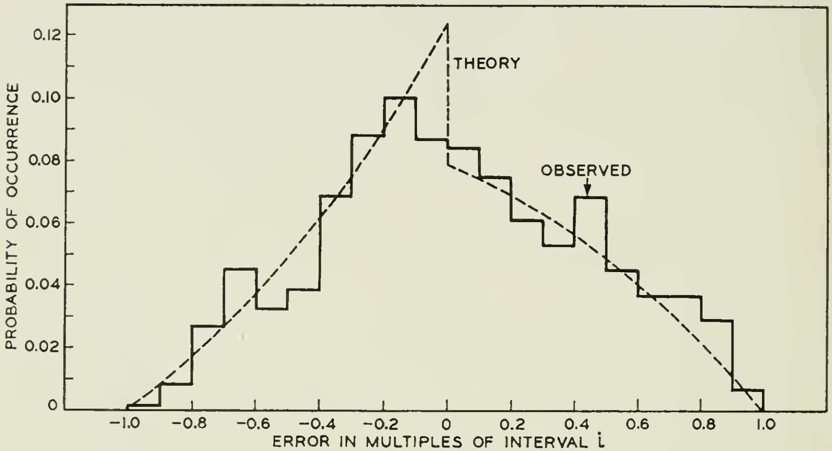


Fig. 15—Distribution of call measurement errors for exponential holding times

tics.[8] A comparison of the σ's obtained from equations (19), (23) and (27) for a typical choice of values, $\bar{l} = 145$ seconds, $i = 60$ seconds, gives

$$\sigma_x \text{ constant h.t.} = 29.58 \text{ seconds,}$$

$$\sigma_x \text{ ``Equally likely'' h.t.} = 24.48 \text{ seconds,}$$

$$\sigma_x \text{ Exponential h.t.} = 24.03 \text{ seconds.}$$

The σ_x constant h.t. is largely a function of whether $\bar{l}$ is closely a whole multiple of i; comparing the values for the other two σ_x's indicates it is slightly advantageous that most of the variable holding time calls to be switch counted in practice are of a roughly exponential form.

The Total Error on n Calls

We shall now attempt to combine the errors from the three sources just discussed and formulate some general conclusions for making the most

[8] It may readily be shown that the average error will be zero for any assumed holding time distribution, by noting that each length of call therein may momentarily be segregated and considered under paragraph "a" as a constant holding time.

of the switch count method of estimating average holding times. Suppose we have completed a succession of $r + 1$ switch counts spaced uniformly by the interval i. The period covered will then be ri units long, and we shall assume that the call count has covered exactly the same total interval so as to eliminate certain of the correction difficulties described under IV-b above. Suppose the first switch count (at the beginning of the period) showed m calls up, the last switch count (at the exact end of the period) showed w calls up, and the sum of all of the $r + 1$ counts, *including* the first, totalled the number s.

As we found in equations (5) and (6) the average correction in the number s to be made on account of calls from the previous period being switch counted is

$$\bar{s}_m = \frac{m}{1 - e^{-\frac{i}{\bar{t}}}},$$

and the standard deviation in the correction here is

$$\sigma_m = \sqrt{m}\,\frac{e^{-\frac{i}{2\bar{t}}}}{1 - e^{-\frac{i}{\bar{t}}}}.$$

At the end of the hour the corresponding correction was found in equations (10) and (11) to have an average of

$$\bar{s}_w = \frac{we^{-\frac{i}{\bar{t}}}}{1 - e^{-\frac{i}{\bar{t}}}},$$

and a standard deviation of

$$\sigma_w = \sqrt{w}\,\frac{e^{-\frac{i}{2\bar{t}}}}{1 - e^{-\frac{i}{\bar{t}}}}.$$

These are quite independent corrections if the observation period is several times the average holding time, the usual case. Then our best estimate of the number of switch counts we would have obtained if all (and only) those associated with the r calls originated in the period had been counted is,

$$s' = s - \bar{s}_m + \bar{s}_w\,, \tag{28}$$

and the standard deviation of s' is

$$\sigma_{s'} = \sqrt{\sigma_m^2 + \sigma_w^2}. \tag{29}$$

We now obtain immediately a preliminary figure for the average holding time of the n calls as

$$\bar{t}_1' = \frac{s'i}{n}, \tag{30}$$

and a standard deviation of this average by

$$\sigma_{\bar{t}_1'} = \frac{\sigma_{s'}i}{n}. \tag{31}$$

This last standard deviation comprehends the uncertainty in the holding time average caused by our inability to measure exactly the number of switch counts which properly should be associated with the n calls. We must now modify this measure of dispersion by the added fluctuation inherent in the method of switch counting at finite intervals. These variations were found to cause no change in the "expected" or most likely value of the average holding time (as shown by $\bar{x} = 0$ in equations (18), (22) and (26)), but the σ_x's of equations (19), (23) and (27) showed sizeable uncertainties which must be included. Since σ_x is for errors in the measurement of individual calls we obtain the standard error of the average as

$$\sigma_{\text{avg}} = \frac{\sigma_x}{\sqrt{n}}. \tag{32}$$

This error is built up on each call throughout the period and hence is practically independent of those errors arising at the ends of the hour. Their joint effects are additive so we obtain the estimates of the final parameters of the average holding time as

$$\bar{t}' = \bar{t}_1' + 0 = \frac{(s - \bar{s}_m + \bar{s}_w)i}{n}, \tag{33}$$

$$\sigma_{\bar{t}'} = \sqrt{\sigma_{\bar{t}_1'}^2 + \sigma_{\text{avg}}^2} = \frac{1}{n}\sqrt{(\sigma_m^2 + \sigma_w^2)i^2 + n\sigma_x^2}. \tag{34}$$

If the holding times are exponential these equations may be written as

$$\bar{t}' = \frac{i}{n}\left[s + \frac{we^{-\frac{i}{\bar{t}}} - m}{1 - e^{-\frac{i}{\bar{t}}}}\right], \tag{35}$$

$$\sigma_{\bar{t}'} = \frac{1}{n\left(1 - e^{-\frac{i}{\bar{t}}}\right)} \times \sqrt{\left((m + w)i^2 + n\bar{t}(2\bar{t} + i)\left(1 - e^{-\frac{i}{\bar{t}}}\right)\right)e^{-\frac{i}{\bar{t}}} - n\bar{t}(2\bar{t} - i)\left(1 - e^{-\frac{i}{\bar{t}}}\right)}. \tag{36}$$

It may be noted that the unknown average holding time $\bar{t}$ enters into both equations (35) and (36). They are not very sensitive to this value, however, and a first approximation obtained from $\bar{t} = \frac{(s - m)i}{n}$ will usually suffice. Further refinement may be obtained by recalculating using the new value $\bar{t}'$ found from equation (35). The form of the distribution represented by the parameters of equations (35) and (36) is not known of

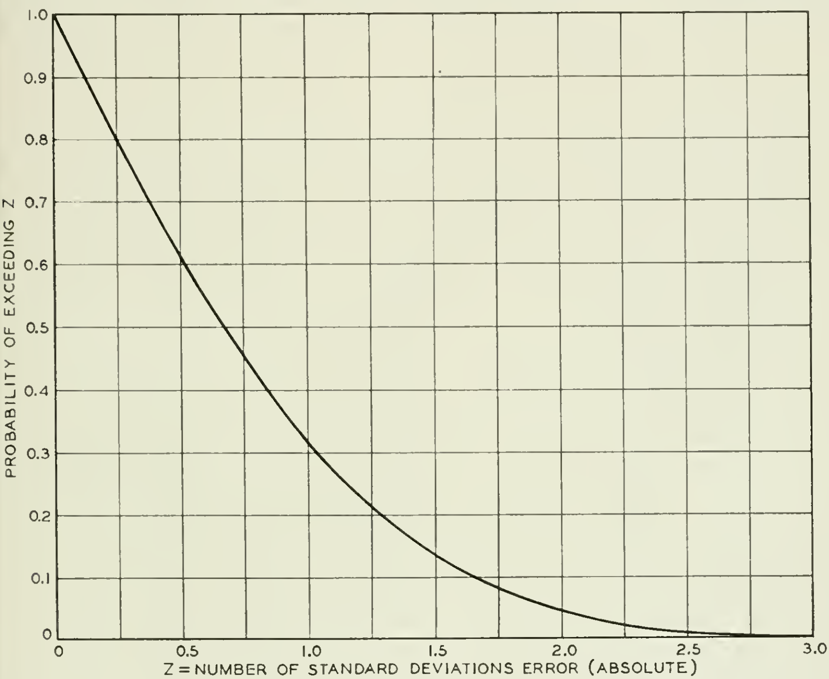


Fig. 16—Cumulative distribution of overall errors in average holding time, regardless of sign

course, but since the errors are essentially the sum of three primary error distributions which are inclined to be unimodal themselves (except for very small trunk groups), we shall probably be not far wrong to assume the normal form for $\bar{t}'$. If the magnitude only of the errors in the estimate $\bar{t}'$ of the unknown true holding time $\bar{t}$ of the n calls under observation is desired, we can readily construct a distribution of these discrepancies Figure 16 is the theoretical cumulative half-normal frequency curve, and when σ is equated to $\sigma_{\bar{t}'}$ found from equation (36), the probability of exceeding any given error in the holding time estimate may be read off directly.

Example: Suppose we have made 60-second switch counts on a group of

T

SWITCH COUNT D

Hour No.	At Beginning of the Hour			At End of the Hour						
	No. Calls Carried Forward m	Expected Switch Count Gain** $m \frac{1}{1-e^{-c}}$	St. Dev. of Total Gain** $\sqrt{m} \frac{e^{-\frac{c}{2}}}{1-e^{-c}}$	No. Calls Beyond #60 Count w	Distance between #60 and end of Hour (min.) j	Due to calls which Pass #60 Sw. Count: Expected Loss** $w \frac{e^{-c}}{1-e^{-c}}$	Due to calls which Pass #60 Sw. Count: St. Dev. of Total Loss** $\sqrt{w} \frac{e^{-\frac{c}{2}}}{1-e^{-c}}$	Due to Calls Originating bet #60 S.C. and End of the H: Avg. No. Expected $\alpha = \frac{j}{t} a$; $(a = 13.44)$	Expected S.C. Loss on these calls (Equation 12)	St. of S.C. (Eq 1
1	2	3	4	5	6	7	8	9	10	
1*	12	35.13	8.23	9	.4	17.34	7.13	2.24	4.71	4.
2	19	55.61	10.35	24	.5	46.25	11.64	2.80	5.99	5.
3	12	35.13	8.23	8	1.1	15.42	6.72	6.16	15.05	8.
4	15	43.91	9.20	13	.9	25.05	8.56	5.04	11.78	7.
5	10	29.27	7.51	13	1.0	25.05	8.56	5.60	13.40	8.
6	18	52.69	10.08	9	1.0	17.34	7.13	5.60	13.40	8.
7	14	40.98	8.89	10	1.0	19.27	7.51	5.60	13.40	8.
8	13	38.05	8.56	7	.8	13.49	6.28	4.48	10.23	7.
9	18	52.69	10.08	15	.7	28.91	9.20	3.92	8.78	6.
10	15	43.91	9.20	7	.6	13.49	6.28	3.36	7.36	5.
11†*	19	55.61	10.35	24	.2	46.25	11.64	1.12	2.24	3.
12	24	70.26	11.64	10	.5	19.27	7.51	2.80	5.99	5.
13	13	38.05	8.56	19	1.0	36.61	10.35	5.60	13.40	8.
14‡	19	55.61	10.35	10	.3	19.27	7.51	1.68	3.45	4.
15	13	38.05	8.56	14	.5	26.98	8.89	2.80	5.99	5.
16	15	43.91	9.20	12	1.1	23.13	8.23	6.16	15.05	8.
17	21	61.47	10.88	11	.8	21.20	7.88	4.48	10.23	7.
18	12	35.13	8.23	14	.75	26.98	8.89	4.20	9.50	6.
19	13	38.05	8.56	8	.6	15.42	6.72	3.36	7.36	5.
20	12	35.13	8.23	21	.5	40.47	10.88	2.80	5.99	5.
21	14	40.98	8.89	16	.5	30.83	9.50	2.80	5.99	5.
22	17	49.76	9.79	14	.94	26.98	8.89	5.26	12.43	7.
23	12	35.13	8.23	13	.53	25.05	8.56	2.97	6.41	5.
24‡	14	40.98	8.89	14	.8	26.98	8.89	4.48	10.23	7.
25‡	18	52.69	10.08	9	.47	17.34	7.13	2.63	5.61	5.
26*	18	52.69	10.08	6	.8	11.56	5.82	4.48	10.23	7.
27	13	38.05	8.56	13	.53	25.05	8.56	2.97	6.41	5.
28	14	40.98	8.89	16	.3	30.83	9.50	1.68	3.45	4.
29	14	40.98	8.89	19	1.0	36.61	10.35	5.60	13.40	8.
30	12	35.13	8.23	15	.4	28.91	9.20	2.24	4.71	4.
31	15	43.91	9.20	16	.75	30.83	9.50	4.20	9.50	6.

Notes: * In these hours additional errors are present because #1 switch count did not coincide well time 0.

† Hour No. 11 was only 59 minutes long.

‡ In each of these hours 1 pen registered call lasting for over 50 minutes was excluded as a tro condition.

** c in the formulas at column headings stands for $\frac{i}{t}$, and was estimated as .418055 from $i = 60.2$ $t = 144$.

III

…31 Hours, Newark

…witch Counts in the Hour											
…o. …itch …unts …bs'd. s	Corr. No. s' $= s -$ (3) + (7) + (10)	St. Dev. of Corr. Total S.C. $[(4)^2 + (8)^2 + (11)^2]^{1/2}$	Length of S.C. Interval $i = 60\,j\,\frac{60-j}{59}$ (seconds)	No. Calls Origi-nated in the Hour n	Estimated Average Holding Time $= \frac{s' \cdot i}{n}$ (seconds)	Std. Dev. of Avg. Holding Time due to Variation in No. Sw. Counts $\frac{(14) \cdot i}{n}$	Std. Dev. of Avg. Holding Time due to Errors in Measuring Each Call $= \frac{\sigma_x}{\sqrt{n}}$ (Equation 27)	Total Std. Dev. of Avg. Hold. Time $= [(18)^2 + (19)^2]^{1/2}$	Actual Measured Average Holding Time in the Hour (seconds)	Error in Estimated Holding Times in Multiples of the Theoretical Std. Dev. $\frac{(21)-(17)}{(20)}$	Percent Error $\frac{(21)-(17)}{(21)}$
12	13	14	15	16	17	18	19	20	21	22	23
69	856	11.89	60.61	355	146.15	2.03	1.31	2.40	142.0	1.73	2.92
37	834	16.48	60.51	359	140.57	2.78	1.30	3.05	139.9	.22	.48
50	845	13.63	59.90	371	136.80	2.21	1.28	2.56	133.0	1.49	2.86
48	841	14.67	60.10	328	154.10	2.69	1.36	3.02	147.2	2.28	4.69
37	846	13.96	60.00	311	163.22	2.69	1.40	3.04	159.9	1.09	2.08
68	846	14.74	60.00	372	136.45	2.38	1.28	2.71	131.8	1.72	3.53
74	766	14.16	60.00	315	145.90	2.70	1.39	3.04	149.0	−1.02	2.08
42	828	12.75	60.20	332	150.14	2.31	1.35	2.68	143.7	2.40	4.48
93	778	15.13	60.30	350	134.04	2.60	1.32	2.92	138.0	−1.36	2.86
41	818	12.64	60.41	308	160.44	2.48	1.40	2.84	157.8	.93	1.67
65	858	15.92	60.81	343	152.11	2.82	1.33	3.09	149.0	1.01	2.08
01	856	14.86	60.51	374	138.49	2.40	1.28	2.73	131.1	2.71	5.64
11	823	15.67	60.00	341	144.81	2.76	1.34	3.07	144.0	.26	.56
01	766	13.43	60.71	318	146.24	2.56	1.38	2.88	144.5	.60	1.21
09	804	13.47	60.51	351	138.60	2.32	1.32	2.66	141.7	−1.16	2.19
02	796	15.01	59.90	305	156.33	2.95	1.41	3.28	158.7	− .72	1.49
39	809	15.17	60.20	337	144.52	2.71	1.34	3.02	141.8	.90	1.92
92	793	13.88	60.25	332	143.91	2.52	1.35	2.86	141.7	.77	1.56
19	804	12.42	60.41	313	155.17	2.40	1.39	2.76	151.8	1.22	2.22
80	791	14.67	60.51	311	153.90	2.85	1.40	3.16	157.5	−1.14	2.28
05	801	14.08	60.51	380	127.55	2.24	1.26	2.56	128.9	− .53	1.05
63	853	15.34	60.06	303	169.08	3.04	1.42	3.36	160.5	2.55	5.35
34	730	13.12	60.48	290	152.24	2.74	1.45	3.08	158.0	−1.87	3.65
71	667	14.41	60.20	287	139.91	3.02	1.46	3.35	136.5	1.02	2.50
45	815	13.40	60.54	321	153.71	2.53	1.38	2.87	153.5	− .07	.14
48	817	13.60	60.20	338	145.51	2.42	1.34	2.77	145.8	− .10	.20
05	798	13.33	60.48	324	148.96	2.49	1.37	2.83	150.5	− .55	1.02
66	759	13.64	60.71	316	145.82	2.62	1.39	2.94	150.0	−1.42	2.78
55	764	15.85	60.00	340	134.82	2.80	1.34	3.11	137.6	− .90	2.02
83	781	13.23	60.61	337	140.46	2.38	1.34	2.72	1 5.2	−1.74	3.26
75	771	14.86	60.25	327	142.06	2.74	1.36	3.06	145.5	−1.13	2.36

20 trunks for a period of one hour, finding $m = 7$ and $w = 13$ as the number of calls up at the beginning and the end of the hour, respectively. Suppose also we have a total of $s = 680$ switch counts, which includes the first count of 7 at time zero. If the register recording number of calls originated in the hour reads 282, what is the best estimate of the average holding time of the n calls, and what is the probability that the true holding time is within 3 seconds of this estimate?

We find our initial estimate for $\bar{t}$ from

$$\frac{(s - m)i}{n} = \frac{(673)60}{282} = 143.19 \text{ seconds.}$$

Substituting in (35) and (36) we find

$$\overline{t'} = 145.65 \text{ seconds,}$$

$$\sigma_{\bar{t}'} = 2.685 \text{ seconds.}$$

Then from Fig. 16 we read that the probability that this estimate of $\bar{t}$ is more than 3 seconds, that is $\frac{3}{\sigma_{\bar{t}'}} = 1.12$ standard deviations, in error is .263. Likewise we may read that the probability is .94 that the error in the average is *not* over 5 seconds.

As something of a final and overall comparison of theory and observation, the actual errors in holding times when estimated by the switch count method for the 31 busy hours in Newark have been tabulated in Table III. The analysis of these pen register records was complicated by the fact that the intervals i varied somewhat from switch count to switch count, and even more from hour to hour, so that the last switch count often came near the midpoint of the 60th minute. To some extent these irregularities of counting correspond more closely to the timing variations in manual switch counts than if they had been taken by machine at perfectly uniform intervals. Such corrections as could be managed by the application of equations (12) and (13) were made to the individual hours. In spite of these precautions the actual errors were somewhat larger than those which could be explained by theory although all large discrepancies were run down and accounted for. The absolute errors are shown plotted in Fig. 17a in terms of the theoretical standard deviations for each hour and in 17b in per cent of the observed holding times. This case will again serve to show that the switch count method is quite sensitive to variations from a perfect application of the rules, and that very considerable care is required to remain within the error limits indicated by the theory.

It is interesting to see what portion of the error is contributed by the two end effects and what by the errors made through "measuring" the calls by switch counts at finite intervals. Hence on Fig. 18 is shown the overall error distribution for the illustrative example given above with the distributions of its two elements as well. The method of combining standard deviations (equation (34)) explains why the two error distributions of Fig.

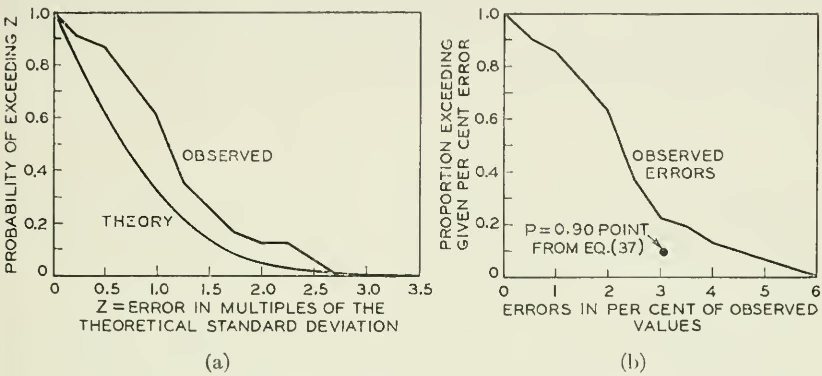


Fig. 17—Comparison of observed and theoretical overall errors

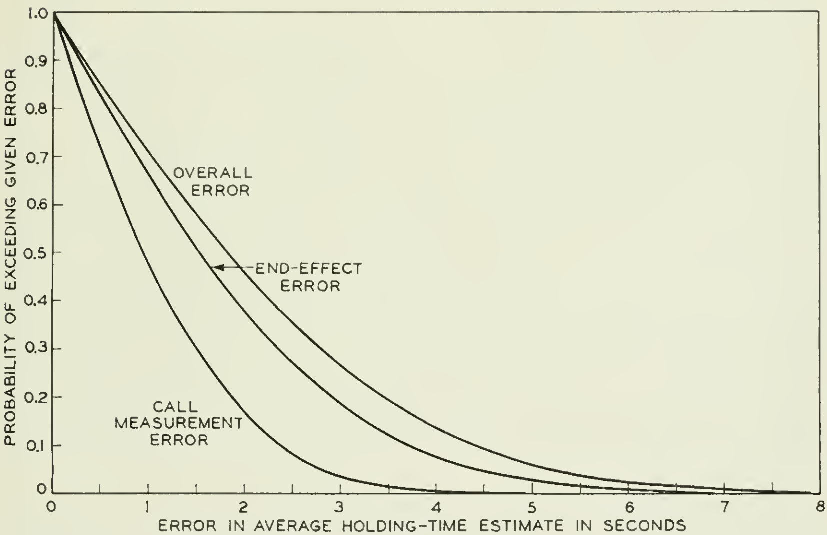


Fig. 18—Typical overall distribution of errors in estimating average holding times

18 do not add directly to give the overall error curve. We may immediately draw the important conclusion from Fig. 18 that the errors due to the uncertainties at the ends of the observation period may in certain cases be considerably greater than those due to the lack of precise measurement of the length of each individual call.

The relative prominence of these two errors will not change very rapidly with different sizes of observation periods (that is, the number of calls, n), since the end effects' standard deviation (equation 31) will vary inversely as n, while the standard deviation due to the error in measuring individual calls (equation 32) will vary inversely as the $\sqrt{n}$. Doubling the length of the observation period would then decrease the first σ to one-half, and the second σ to .707 of its former value.

Equation (36) shows that

$$\sigma_{\bar{t}'} = f(m, w, n, \bar{t}, i).$$

We will not know m and w before making the switch counts but we can probably substitute the average value of their sum which is $2n\bar{t}/T$, without seriously disturbing the average value of $\sigma_{\bar{t}'}$. This gives

$$\sigma_{\bar{t}'} \approx \frac{\sqrt{\frac{\bar{t}}{n}}}{1 - e^{-\frac{i}{\bar{t}}}} \sqrt{\frac{2\bar{t}^2 e^{-\frac{i}{\bar{t}}}}{T} + \left[(2\bar{t} + i)e^{-\frac{i}{\bar{t}}} - 2\bar{t} + i\right]\left(1 - e^{-\frac{i}{\bar{t}}}\right)}, \quad (37)$$

where now $\sigma_{\bar{t}'}$ is a function of only four variables, n, $\bar{t}$, i and T.

It is of interest to compare the errors predicted by equation (37) with those found by the theory formulated for the particular m and w observations found in each hour's observations at Newark. An average of n = 332 calls per hour was observed with an average holding time in the order of $\bar{t}$ = 145 seconds. The switch count interval, i, was approximately 60 seconds, and the observation period was T = 3600 seconds. If we take P = .90, the per cent error corresponding will be $\pm 100\ (1.645)\ \sigma_{\bar{t}'}/\bar{t}$. Using the above estimate of $\sigma_{\bar{t}}$, we find an error of about 3.05 per cent, or $\pm .0305$ (145) = ± 4.42 seconds. This point has been plotted on Fig. 17b and bears about the same relationship to the observed errors as does the theoretical curve in Fig. 17a in which the comparison takes into account the actual calls carried beyond the start and end of each hour. The discrepancy in Fig. 17b is largely accounted for by the same discussion given heretofore for Fig. 17a.

The Overall Error in Estimating the Average Holding Time

The engineer who has the problem of devising a switch count schedule will want to be able to estimate at least roughly the order of accuracy he will actually obtain in the average holding time found from the data in a number of observation periods. Up to this point in this section we have concerned ourselves with discovering only the errors inherent in measuring the average length of a *particular n calls* of the exponential type in an observation period of length T. As we saw in section III, even when such an average length $\bar{l}$ is accurately known for the n calls, it may not exactly or even closely coincide with the true average of the universe of calls of which the n are presumed to form a random sample. The errors we have just studied and those described in section III must now be combined to give us a measure of the overall accuracy of the switch count method.

Equation (2), when applied to the exponential holding times we are here concerned with, gives us for the standard error of the average in a sample

$$\sigma_{\text{sampling}} = \frac{\sigma}{\sqrt{n}} \approx \frac{\bar{l}}{\sqrt{n}}. \tag{38}$$

This error is independent of that represented by $\sigma_{\bar{l}'}$, so we may determine the overall (oa) standard error by

$$\sigma_{oa} = \sqrt{\sigma_{\text{sampling}}^2 + \sigma_{\bar{l}'}^2}. \tag{39}$$

We shall be particularly interested in knowing how much the value of $\sigma_{\bar{l}'}$ contributes to the overall standard deviation, σ_{oa}. This may be conveniently expressed by writing the ratio

$$q = \frac{\sigma_{oa}}{\sigma_{\text{sampling}}} = \sqrt{1 + \frac{\sigma_{\bar{l}'}^2}{\sigma_{\text{sampling}}^2}}$$

$$= \sqrt{1 + \frac{\frac{2i}{\bar{l}}\frac{i}{T}e^{-\frac{i}{\bar{l}}} + \left[\left(2 + \frac{i}{\bar{l}}\right)e^{-\frac{i}{\bar{l}}} - 2 + \frac{i}{\bar{l}}\right]\left(1 - e^{-\frac{i}{\bar{l}}}\right)}{\left(1 - e^{-\frac{i}{\bar{l}}}\right)^2}}. \tag{40}$$

Now it is readily seen from (40) that q depends on $\bar{l}$, i and T. Hence if T is held constant we may plot curves between q, i and $\bar{l}$ as shown on Fig. 19. What is more, if T is varied, say increased by a factor k, equation (40) shows that if i and $\bar{l}$ are also increased by the same factor the values of q resulting may still be read directly from Fig. 19.

For example: If approximately 100-second exponential calls are to be switch counted for an hour with observation intervals of 120 seconds, we read on Fig. 19 that the *overall* standard error (or P = .50, .90, .99, etc. error) in estimating the true average holding time is q = 1.134 times the basic standard error resulting from taking a random sample of n calls from a very large universe of calls. That is to say, the residual sampling error present in a stop watch measurement of the n calls is increased by 13.4 per cent due to our resort to switch count methods.

If a continuous period of T = 2 hours (i.e. k = 2) is switch counted in just the same manner and under the same conditions, we should now read on the $\bar{l}$ = 50 seconds curve opposite i = 60 seconds, giving us an increase over the basic sampling error of 12.4 per cent. This meets one's common

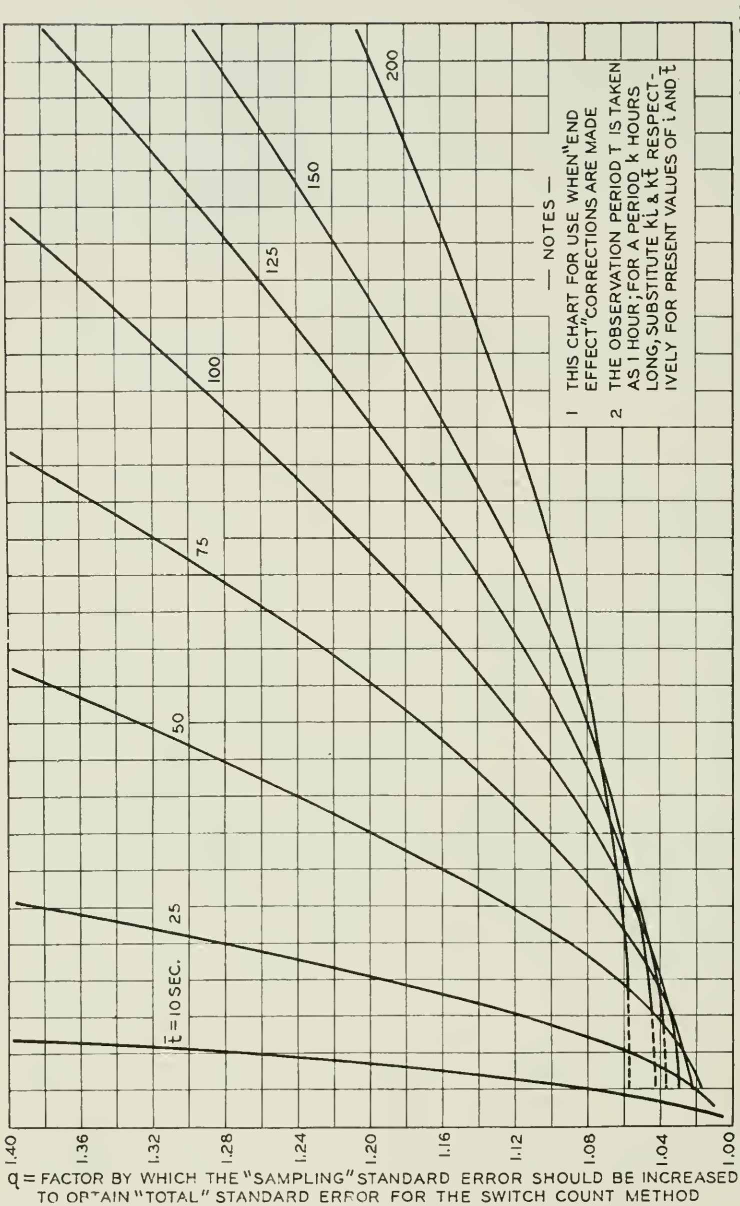


Fig. 19—Increase in "overall error" over "sampling error" in the estimate of holding time averages by switch counts (with "end effect" corrections)

sense conclusions since, as previously suggested, when the observation period is increased by a factor k,

a. The "sampling" error decreases as $\frac{1}{\sqrt{k}}$, since we are looking at k times as many calls,

b. The error of measurement on the individual calls decreases also as $1/\sqrt{k}$, and for the same reason, and

c. The "end effect" errors are unchanged in actual magnitude but are now prorated over k times as many calls; hence, the effect of this element decreases as $1/k$.

The overall effect then as k increases is for σ_{oa} to decrease faster than σ_{sampling} as we have just seen in the example. Thus not only is the sampling error decreased by lengthening the observation period, but the overall error decreases even faster. It is clear then that an observation period as long as is consistent with a "controlled" universe of holding times is to be preferred.

Further important deductions may be drawn from a study of Fig. 19. If we wish to minimize the effect of errors introduced by a use of the switch count method we shall need to select our observation interval i so that it will be relatively small compared with the average holding time. Apparently we should do well to keep $\frac{\bar{l}}{i} \gtreqqless 1.5$; the higher this ratio the better, although the improvement beyond, say, 2.0, is slow. Roughly, for commonly observed local subscriber holding times if the holding time is at least twice the observation interval the increase in error occasioned by the switch count method over the stop watch method need not exceed 7 per cent.

Fig. 19 has been constructed for use when $\bar{l}$ is estimated as $\bar{l}'$ from equation (35),

$$\bar{l}' = \frac{(s - s_m + s_w)i}{n} = \frac{i}{n}\left[s + \frac{w e^{-\frac{i}{l}} - m}{1 - e^{-\frac{i}{l}}}\right], \qquad (35)$$

in which s is the sum of $r + 1$ switch counts (which includes counts at both the extreme ends of the period), m is the count at the beginning, and w the count at the end of the period.

It will be of interest to estimate something of the enlarged error when $\bar{l}'$ is found merely by taking

$$\bar{l}' = \frac{(s - m)i}{n} \qquad (41)$$

which entirely neglects the special information contained in the first and last switch counts. $\left(\text{If only } r\left(= \frac{T}{i}\right) \text{counts are made they should be at the}\right.$

end of each i-interval, the last one coming at the exact end of the whole period.) This is the common case in which we merely sum all the switch counts, multiply by the counting interval and divide by the number of calls shown on the peg count meter as originating in the period T. The standard error for each end effect will then be approximately that given by σ_m' in equation (6′) where no attention is paid to the end switch count values of m and w. Substituting σ_m' for σ_m and σ_w in (29) and following the same analysis as before gives for q' (instead of q),

$$q' = \sqrt{1 + \frac{\frac{2i}{\bar{t}}\frac{i}{T}\left(e^{-\frac{i}{\bar{t}}} + 1\right) + \left[\left(2 + \frac{i}{\bar{t}}\right)e^{-\frac{i}{\bar{t}}} - 2 + \frac{i}{\bar{t}}\right]\left(1 - e^{-\frac{i}{\bar{t}}}\right)}{\left(1 - e^{-\frac{i}{\bar{t}}}\right)^2}}. \quad (42)$$

A plot of this last expression is given in Fig. 20. By comparing points on Fig. 20 with corresponding ones on Fig. 19, one obtains an idea of the increase of error due to failure to correct the switch counts for the end effects as indicated in equation (35). For example, with 100-second calls switch counted at 120-second intervals we find $q = 1.134$ while $q' = 1.203$, indicating quite a marked increase in the overall error. The particular errors resulting in any given circumstance coupled with the cost of making the end effect corrections will determine the practical desirability of which method to adopt, that is whether the factor for increasing the basic sampling standard error shall be read from Fig. 19 or from Fig. 20.

Finally, a chart has been drawn up as Fig. 21, by which a measure of the overall error in estimating the unknown true holding time may readily be determined. The right hand section of the chart is a redrawing of Fig. 5 given in section III for the sampling errors of individually measured calls. Scale A is carried across and reproduced at C permitting the small nomograph B C D to give easily the product of the sampling error and the q (or q') factor at D. The left hand chart is based simply on the fact that the overall error decreases inversely as the square root of the number of periods switch counted. From it the number of periods required to obtain any desired accuracy can be read.

The estimate of the average holding time will be found from a simple average of the estimates made for individual observation periods,

$$\bar{\bar{t}}' = \frac{\bar{t}_1' + \bar{t}_2' + \cdots + \bar{t}_g'}{g}. \quad (43)$$

If a certain per cent error in the estimated average holding time is obtained for a single period the improvement for the combination of g periods

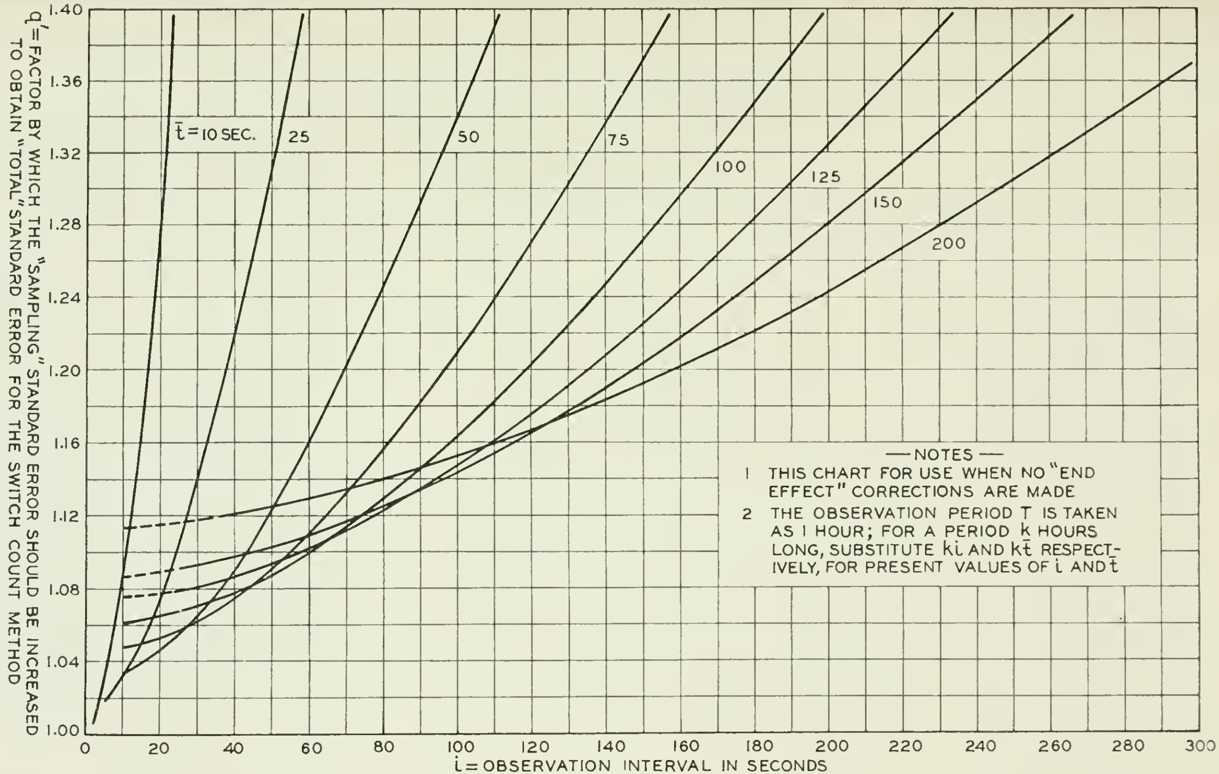


Fig. 20—Increase in "overall error" over "sampling error" in the estimate of holding time averages by switch counts (no "end effect" corrections)

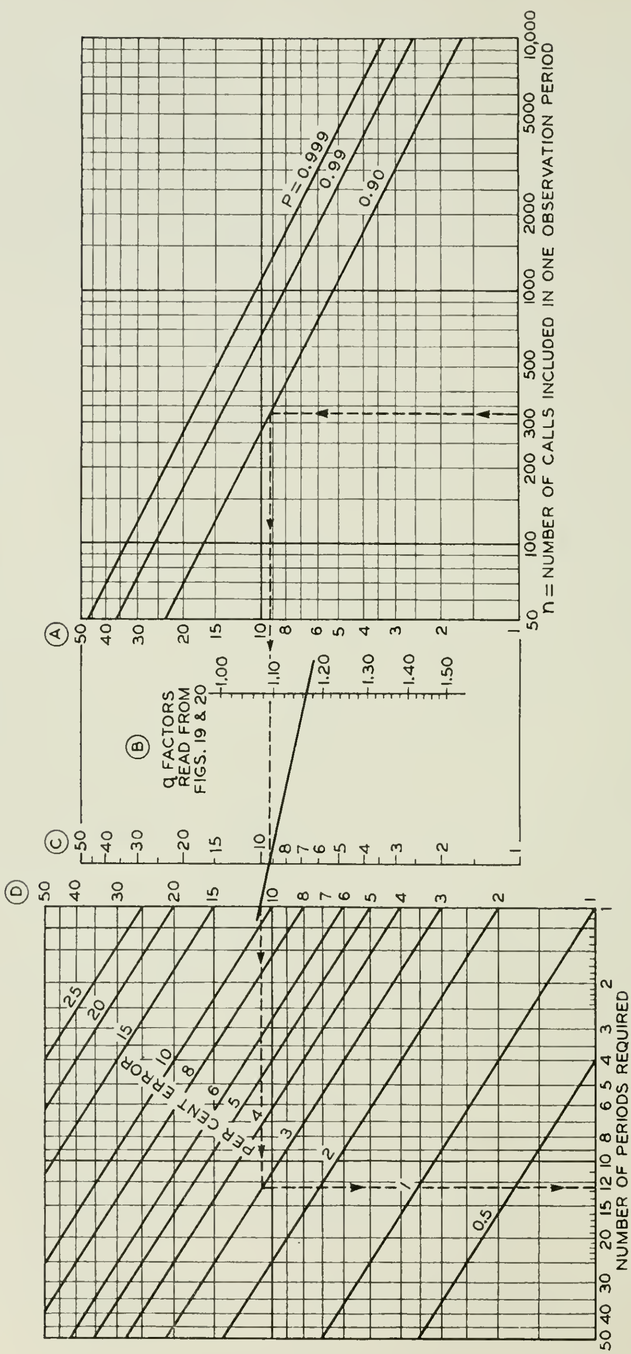


Fig. 21—Determination of number of observation periods required to produce any desired accuracy in estimating holding times by switch counts

is then,

$$\frac{1}{\sqrt{g}} \text{ (single period error).}$$

Illustrative Example. If calls of approximately 135 seconds holding time are to be switch counted at 3 minute intervals by 1 hour periods, how many such periods will be required to give us an assurance of $P = .90$ that the resulting holding time estimate does not differ from the true by more than 3 per cent? Assume an average of 325 calls per hour over the group, and that end effect corrections (a) will be made, and (b) will not be made.

Solution (a) We first read on Fig. 19 that $q \approx 1.165$. Turning to Fig. 21, we find that opposite $n = 325$ and $P = .90$ we have an error of 9.2 per cent, which we carry over to the C scale of the nomograph. Laying a straight-edge across this point and $q = 1.165$ determines a point on D somewhat above the 10 per cent line. Projecting this point across to the desired 3 per cent error line and dropping to the lower edge of the chart we find that 12.7, say 13, such one-hour observation periods will be necessary to meet the accuracy specifications of the problem. This is the effort required when the end effect corrections are made.

(b) If the end effect corrections are to be ignored, we determine our factor from Fig. 20 to be $q' \approx 1.268$. Proceeding on Fig. 21 exactly as before we find that the number of one-hour observation periods required is increased to 15.1, say 15. Thus a failure to make the end effect corrections causes us to increase our switch counts by about 20 per cent. It is by such comparisons as these that one decides the practical desirability of making the end effect corrections.

Summary of Switch Count Procedures for Exponential Holding Times

The choice of the size of unit observation period should rest primarily on the considerations discussed in Section II, that is the homogeneity of the holding time data from hour to hour, day to day, etc. Other things being equal, we should select the longest period consistent with the view that enough periods must be included so that representative sampling of all known or suspected major variations in holding time character is accomplished. The length of the switch count interval will likely already have been decided by the equipment at hand or by other considerations. If a choice is available, however, a short interval will produce more reliable results than a long one, by the amounts indicated on Figs. 19 and 20. With these matters decided, Fig. 21 is consulted to see how many periods must be switch counted in order to obtain the desired accuracy.

Having actually made the switch counts, exercising the cautions we have mentioned, the average holding time $\bar{t}'$ for each period is obtained from equation (35) if the correction of the counts for the end effects in each period is made. If no such corrections are to be made, equation (41) is used instead of (35). The arithmetic mean, equation (43), of the values obtained in the various periods will then be the best available estimate of the unknown true average holding time. The reliability of this figure should be substantially that which the schedule was designed to produce.

Switch Count Errors for Non-Exponential Holding Times

If switch counts are made on calls with other than exponential holding times the resultant errors may be greater or less than those shown by Figs. 19, 20 and 21. The comparison of typical σ_x's calculated from equations (23) and (27) would suggest that for varying holding times the error in the measurement of individual calls is not greatly dependent on the form of the holding time distribution *as long as the average call length covers several intervals i.* In such a case the charts developed for exponential holding times can probably be used with little allowance for the discrepancy present.

On the other hand for calls with an unusual or extreme fluctuation about an average $\bar{l}$ less than i, the errors due to assuming the situation to be equivalent to the exponential case may be no longer negligible. The only procedure then would appear to be either to work out the errors actually present, reverting to the basic error equations (14) and (15), and approximating the new end effect corrections, or to revise the switch counting program to materially shorten the interval i.

For relatively constant holding times the value of σ_x can be reduced to a small figure by choosing the switch count interval i so that it is contained in the average holding time $\bar{l}$ closely a whole number of times. Then by equation (19) the error in individual call measurements must, of necessity, be small in nearly every instance. It will be noted that the above specification permits choosing $i = \bar{l}$; moreover, it may readily be seen that in this case the end effect corrections will tend to disappear, giving a highly accurate measurement with relatively few observations. Just how many, of course, will depend on how constant is the quantity measured, and how closely the switch count interval i approaches the true average $\bar{l}$.

V—General Summary

The general problem of determining the average holding times of subscribers' or other calls by sampling methods has been discussed. The need for a proper apportionment of the sample is emphasized and examples are given from telephone experience to illustrate typical analysis procedures. Methods for estimating the reliability of these sampling results for both directly measured holding times and for switch count studies are given along with various curves and charts calculated to assist the traffic engineer in devising a working schedule for the sampling of holding times, particularly those of an exponential character.

Electrical and Mechanical Analogies

By W. P. MASON

Introduction

During the past few years, apparatus which transfers electrical into acoustical or mechanical energy has received wide application. This came about through the popular use of radios, phonographs, public address systems, and sound motion pictures. While the fundamental principles of such electro-mechanical or electro-acoustic transducers have been known for decades, it is safe to say that the rapid progress and excellent design obtained have been due in a large part to the knowledge derived from the related subject of electrical network theory.

Two examples may be cited to show the nature and extent of the improvement. Barton in his "Theory of Sound" (1914) cites measurements on the efficiency of acoustic foghorns operated from an electrical source of power and finds that the efficiency of conversion from electrical into acoustic energy is less than one per cent. Today large loud speakers have been developed which can be used for similar purposes and these have efficiencies of conversion greater than 50 per cent. Another and more striking example is the mechanical phonograph. From the days of its invention by Edison, mechanical phonograph reproducers had been constructed from such mechanical units as needles, diaphragms, horns, and their connecting mechanical elements. As late as 1925 the best of such units was capable of reproducing only three octaves. About this time, another mechanical phonograph[1] was constructed from the same sort of elements, but with their dimensions and relationships designed according to relations developed in electrical network theory, and the resulting structure was able to reproduce a frequency band corresponding to five octaves with greater uniformity and an increase in the efficiency of conversion.

The type of electrical network which has received the greatest application in the design of mechanical and acoustical systems is the electrical wave filter invented by Dr. G. A. Campbell. This may seen surprising at first sight, since the filter is usually regarded as a device for attenuating unwanted frequency bands while passing other frequency bands which it is desired to receive. The filter has two properties which make it of interest in electro-acoustic transducer systems. These are: first, the filter is able to coordinate

[1] Maxfield and Harrison, *Bell Sys. Tech. Jour.*, Vol. 5, No. 3, p. 493, 1926.

the action of several resonant elements to produce a device with a uniform transmission over a wide frequency range; and second, the dissipationless filter, with matched impedance terminations, is a device which delivers to its output all of the energy impressed upon it over the widest possible frequency range consistent with the elements composing it. These properties of the filter have been made use of in purely electrical networks to determine the largest band width a vacuum tube with known characteristics can have and still deliver a specified gain at a specified impedance level. Applied to electro-mechanical transducer systems, the filter theory shows how to combine resonant mechanical or electro-mechanical elements to produce a uniform conversion of electrical to mechanical energy, or vice versa, over a wide frequency range. Also, it is able to determine the greatest band width that can be obtained without loss of efficiency for any type of conversion element.

This transfer of knowledge from one branch of science, electrical network theory, to another branch of science dealing with mechanical and electro-mechanical structures is one example of a long line of such interchanges that have been going on for over a hundred years. These interchanges are made possible by the fundamental analogies which exist between electrical and mechanical systems and which rest finally on the fact that electrical motions and mechanical motions satisfy the same type of differential equations. Since such analogies have been very productive in the past and are likely to continue to be so in the future, it seems worthwhile to examine their foundation and development.

Early Borrowings of Electrical from Mechanical Theory

The equations of motion of mechanical bodies and mechanical media were developed and studied long before the equations for electrical wave propagation were derived. Under these circumstances it is natural that attempts should have been made to explain electrical wave propagation as a mechanical phenomenon. The view that electrical actions are ultimately dynamic was one whose development in the hands of Maxwell led to notable advances in the science, and it was the view toward which most of the early authorities leaned. In support of this point of view Maxwell showed that the forces on any system of charged bodies could be attributed to a system of stresses in the medium in which they are embedded. Since magnetic energy is associated with the presence of charge in motion while electro-static energy is present for charges at rest, an identification was made between kinetic and magnetic energy and between electro-static and potential energy. Applied to a concentrated system, this point of view indicates that an inductance is the analogue of a mass, while a capacitance is the analogue of a spring.

A case for which this point of view bore useful results was the case of anomalous dispersion in optics. It was found experimentally that when light was sent through certain substances the velocity of propagation depended markedly on the frequency in the neighborhood of a certain critical frequency. Below this critical frequency the velocity decreased as the frequency approached it, going rapidly toward zero as the critical frequency was approached. Above this critical frequency the phase velocity was greater than the velocity of light in the material and gradually approached that value for high frequencies. At the critical frequency a large absorption of light occurred. This was first explained by Sellmeier as being due to some element in the medium having a resonant frequency at the critical frequency In obtaining his equations he used a mechanical model in which the resonant elements were spaced at equal intervals and excited by waves propagated by virtue of the mass and elasticity of the substance.

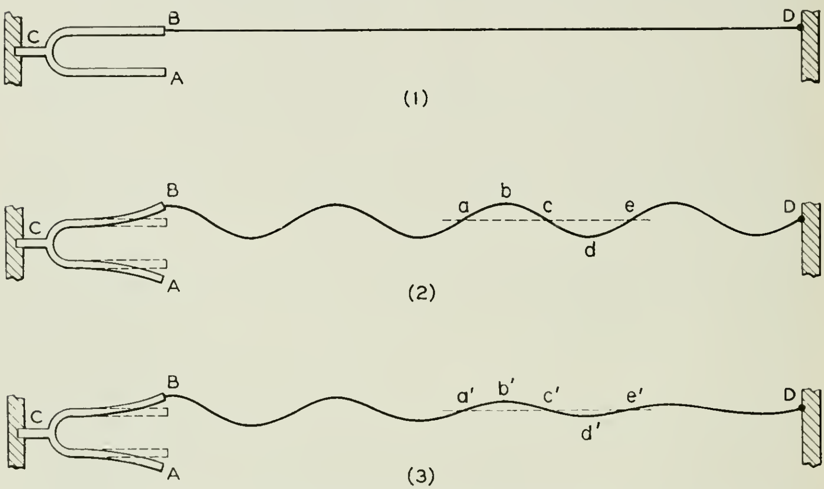
The case of greatest interest from the communication viewpoint is the influence of mechanical theory on the theory of the loaded transmission line. Wave propagation in a mechanical bar or stretched string has similar characteristics to that of a dissipationless electrical line, but when the effect of series resistances and shunt leakances were taken account of, effects appeared for the electric line which had not previously been studied in mechanical systems. These were high attenuation, which cut down the amount of power delivered to the output, and distortion, which caused the shape of the signal received at the end of the line to be different from that sent into the line. Heaviside showed that the distortion could be removed by having a certain relationship between the inductance, capacitance, resistance and leakance, and moreover that a smaller attenuation and a lower distortion would result, if an inductance were uniformly distributed along the line.

It was not a practical matter, however, to put in extra inductance at every point of the line so Heaviside suggested and tested out the effect of placing inductances at discrete points along the line, and found no beneficial results. It was not until Campbell and Pupin independently showed that the inductances had to have discrete values and be placed at definite separations that any progress was made in approaching the desired conditions.

Pupin's method of arriving at the solution is well illustrated by the following extract from his paper.[2] "The main features of the theory are extremely simple and can be explained by a simple mechanical illustration. Consider the arrangement of Fig. 1. A tuning fork has its handle C rigidly fixed. To one of its prongs is attached a flexible inextensible cord BD. One terminal of the cord is fixed at D. Let the fork vibrate steadily, the vibration being maintained electromagnetically or otherwise. The motion of the cord will be a wave motion If the frictional resistances opposing the motion

[2] "Wave Transmission Over Non-uniform Cables and Long Distance Air Lines," M. I. Pupin, *Trans. A. I. E. E.*, Vol. XVII, May 19, 1900.

of the cord are negligibly small the wave motion will be approximately that of stationary waves as in Fig. 2. The direct waves coming from the tuning fork and the reflected waves coming from the fixed point D will have nearly equal amplitudes and by their interference form approximately stationary waves. If, however, the frictional resistances are not negligibly small, then there will be dissipation of the propagated wave energy. Hence, the direct and the reflected waves will not result in stationary waves. The attenuation of the wave is represented graphically in Fig. 3. Experiment will show that, other things being equal, increased density of the string will diminish attenuation, because a larger wave requires a smaller velocity in order to store up a given quantity of kinetic energy and a smaller velocity brings



Figs. 1 to 3—Standing waves and damped waves on a mechanical transmission line (taken from Pupin's paper)

with it a smaller frictional loss. This is a striking mechanical illustration of a wave conductor of high inductance. It should be observed here that an increase of the density will shorten the wave-length.

"Suppose now that we attach a weight, say a ball of beeswax, at the middle point of the string, in order to increase the vibrating mass. This weight will become a source of reflection and less wave energy will reach the point D than before. The efficiency of transmission will be smaller now than before the weight was attached. Subdivide now the beeswax into three equal parts and place them at three equidistant points along the cord. The efficiency of wave transmission will be better now than it was when all the wax was concentrated at a single point. By subdividing still further the efficiency will be still more improved; but a point is soon reached when further subdivision produces an inappreciable improvement only. This point is reached when the cord thus loaded vibrates very nearly like a uniform cord of the same mass, tension, and frictional resistance."

Campbell, who first arrived at the design formulae for the coil loaded line, was guided by the solution, given by Lagrange over 100 years earlier, of the propagation of a wave along a string loaded with masses at discrete intervals, and a generalization of it made by Charles Godfrey,[3] for he states[4] "For the method of treatment which I first employed I am indebted to an interesting article by Mr. C. Godfrey on the 'Discontinuities of Wave Motion Along a Periodically Loaded String.' " The spacing of the coils arrived at is the same as the spacing of the massive beads along a string, namely, that π coils or beads occur per actual wave-length of the highest frequency to be transmitted. The added result not given in the mechanical case was that the addition of such coils reduced the attenuation and decreased the distortion.

The different points of view of the two inventors are well illustrated by the quotations. Pupin was attempting to obtain a system which approached a uniform line while Campbell was investigating the propagation characteristics of the structure without particular regard as to whether the transmission was the same as that which would be provided by an equal amount of inductance uniformly distributed along the line. It was the broader point of view of Campbell which proved of wide significance and which in particular led to the invention of the electrical wave filter.

Development of Electrical Network Theory

The first structures capable of transmitting bands of frequencies and attenuating all other frequencies were mechanical structures, although this was not generally realized at the time or made use of. The first structure of this sort was the string loaded with massive beads, which was first studied by Lagrange. Introducing the approximation that the mass of the string could be neglected, Lagrange showed that all of the natural frequencies of the device came below a certain critical frequency f_c. Routh[5] after discussing Lagrange's solution, points out that there may be a period of excitation of the string which is "so short that no motion of the nature of a wave is transmitted along the string."[6] An acoustic forerunner of the wave filter was the combination of two tubes of different lengths first proposed by

[3] *Phil. Mag.* 45, 456 (1898).

[4] See "Collected Papers of George A. Campbell," page 16.

[5] "Advanced Rigid Dynamics," p. 260, Paragraph 411 (1892).

[6] It should be noted that all mechanical filter theory of this time had to do with the natural resonances of unterminated filters or the transmission through misterminated sections. The idea of matched impedance terminations to introduce power into and absorb power from a filter is a development of electrical network theory.

Herschel. This structure passes low frequencies but attenuates strongly frequencies for which the difference in path length of the two tubes is an odd number of half wave-lengths.

It is interesting to note that all of the fore-runners of the filter were of the continuously distributed type which had their elements distributed uniformly along the length of the device. All such devices have an infinite number of pass bands, usually harmonically related. This is true also of the dissipationless loaded line considered as a filter structure. It was only by such abstractions as neglecting the weight of the loaded string that single pass bands were obtainable.

The low-pass electrical filter grew out of a network to simulate the operation of a long cable. By using series coils to simulate the loading and the distributed inductance of the line, condensers to simulate the distributed capacitance of the line, series resistances to simulate the ohmic resistance of the line and the resistance of the coils, and shunt resistances to simulate the leakance of the line, Campbell was able to obtain in a small space, a device which had the same propagation characteristics up to the cut-off frequency as a long section of loaded cable. Furthermore, by making the resistance small, he was able to obtain a frequency range from zero frequency up to a cut-off frequency f_c with small attenuation, and a high attenuation at higher frequencies, and thus obtained the first true low-pass filter. He also put his filter to practical use for he says[7] "I have made use of these results by employing artificial loaded lines for cutting out harmonics in generator circuits. The harmonics may all be cut down as far as desired by the use of a sufficient number of sections, while the attenuation of the fundamental can be reduced at pleasure by decreasing the resistance."

Having developed the fundamental idea of a filter as a device for transmitting without loss one frequency range and attenuating all other frequencies, he went on to extend the idea to other types of filters in which different frequency ranges were passed. The band-pass filter was already realized in 1903 for Campbell says[7] "Combining condensers and inductances, we may make a system which will not only cut out higher frequencies, but also all frequencies below a certain limit." The high-pass and band elimination filters followed shortly after. With the invention of the electrical wave filter, electrical network theory can be considered as well started.

The science of electrical networks did not progress much farther for a number of years. This was due primarily to the lack of application for any of the structures developed. However, with the development of carrier current transmission on telephone lines, the necessary stimulus was received. Carrier current systems were an adaptation of radio communication systems

[7] "On Loaded Lines in Telephonic Transmission," *Philosophical Magazine*, Ser. 6, Vol. 5, pp. 313–330, March 1903.

to wire lines, with the line taking the place of the ether as the wave transmitting medium. Previously, radio systems had been developed which would transmit messages in definite frequency ranges. These transmitting ranges were selected from other frequency ranges by means of electrical tuned circuits, which were themselves a borrowing from the acoustic resonators of Helmholtz devised many years earlier. Tuned circuits are not advantageous for selecting out channels in a carrier system, because with them it is not possible to regulate the band width received or to get the necessary discrimination between the pass band and the attenuated region. It was found, however, that filters could meet these requirements and consequently they were applied in separating the channels of the first carrier systems.

This use stimulated the further development of electrical network theory. Filters with sharper discriminating properties, composite filters containing sections of like image impedances but different attenuation characteristics, transforming filters, impedance corrected filters for reducing reflections, filters using mutual inductances, attenuation and phase correcting networks are among the later developments. These investigations were carried out by a large number of individuals among whom may be mentioned Bartlett, Bode, Carson, Cauer, Foster, Fry, Guillemin, Johnson, Norton, Shea, Wagner, and Zobel. Electrical network theory has progressed to such an extent that it is now possible to select substantially any desired frequency range, with very little of the frequency range wasted in obtaining the desired selectivity, and to control the amplitude and phase of the currents received over long distance lines so that a high degree of fidelity of the received signal can be maintained.

Borrowings of Mechanical Theory from Electrical Network Theory

While this development of electrical theory was progressing, very little development of a parallel nature was being carried out for mechanical theory, due probably to the lack of a corresponding stimulus. With the advent of the vacuum tube, public address system, and radio broadcasting, however, a demand developed for loud speakers and related equipment. It was shortly realized that the parallel developments of electrical network theory provided a base for the design of such equipment. One of the first to recognize this possibility was Professor A. G. Webster,[8] who pointed out the usefulness of the concept of impedance in mechanical systems. He applied the concept to the phonograph and developed the first theory of the action of acoustic horns. After this occurred the widespread application of the

[8] A. G. Webster, "Acoustic Impedance, and The Theory of Horns and of the Phonograph," *Nat'l. Acad. of Science*, Vol. 5, p. 275, 1919.

electrical network theory to the design of electro-mechanical systems mentioned in the introduction.

Aside from this electro-mechanical field special applications have been made in acoustic and mechanical apparatus where problems occur similar to those solved by electrical means. In all of these applications it is the filter type structure that is applied.

One of the first of these applications was the acoustic filter. In ventilating ducts, automobile, and other types of engines, and for many other uses, it is desirable to pass a steady or slowly varying stream of air, and attenuate the more rapid vibrations which constitute the undesired noise. Furthermore, it is desirable to pass the low-frequency variations with little or no loss, since such loss increases the back pressure on the engine or blower and greatly decreases their efficiency. For this purpose the low-pass filter type structure is well suited since it passes a low-frequency band with little or no attenuation and strongly suppresses higher frequency components.

The rudimentary idea of the acoustic filter probably dates back to Herschel (1833) who suggested the use of combinations of tubes capable of suppressing certain frequencies. Following the development of electrical wave filters, Professor G. W. Stewart[9] showed that combinations of tubes and resonators could be devised which would give transmission characteristics at low frequencies similar to electrical filters. This theory worked well as long as the structure was small or the frequency low, but broke down for large structures and high frequencies due to the essentially distributed nature of the elements. A theory of acoustic filters was given by the writer in 1927,[10] which took account of wave motion in the elements, and this theory could account for the properties of the filters to much higher frequencies. Since then, Lindsay[11] and his collaborators have discussed a number of acoustic type filters with various types of side branches and obstructions.

Mufflers existed long before the theory of acoustic filters was worked out but they were designed as a series of baffles, which introduced considerable back pressure on the engine. Most recently designed mufflers have a straight conducting path with side-branches in conformance with acoustic filter theory and are proportioned to attenuate most of the frequencies above 100 cycles. As a result they are considerably more effective than

[9] *Phys. Rev.* 20, 528 (1922); 23, 520 (1924); 25, 590 (1925). See also Stewart and Lindsay "Acoustics," Chap. VII. D. Van Nostrand.

[10] A Study of the Regular Combination of Acoustic Elements, with Applications to Recurrent Acoustic Filters, Tapered Acoustic Filters, and Horns. *B.S.T.J.* Vol. VI, pp. 258–294, April 1927.

[11] An excellent review and resumé of the literature on gaseous and solid acoustic filters is given by Lindsay, "The Filtration of Sound I," *Jour. App. Phys.* 9, 612 (1938); "The Filtration of Sound II," *Jour. App. Phys.* 10, 620 (1939).

early mufflers and introduce considerably less back pressure on the engine or blower.

Other uses to which mechanical filters have been put are in obtaining shockproof mountings and vibration damping devices, in obtaining vibration and noiseproof walls and floors, and in obtaining constant-speed motors in which the effects of gear irregularities are removed by the use of a low-pass mechanical filter.

Mechanical and Electro-Mechanical Counterparts of Electrical Filters

Although combinations of electrical elements were first studied and applied in wave filters and other structures, it does not follow that they have any inherent advantages over analogous combinations of mechanical or electro-mechanical elements which can be used as filters. In fact, elements which depend on mechanical motion have the great advantage that they have very little energy dissipation associated with their motion and, hence, the equivalent mechanical elements have a higher ratio of reactance to resistance, or "Q," than do their electrical counterparts. The result is that considerably more selective filters can be made from mechanical or electro-mechanical elements than can be obtained by employing electrical coils and condensers.

The first attempts[12,13] along this line were made in substituting masses for coils and springs for condensers in standard electrical filter configurations. This work resulted in usable filters up to several thousand cycles in frequency, which have been used for certain special purposes.

More recently, electro-mechanical elements have been used to take the place of some or all of the electrical elements of a filter and this work has resulted in filters with markedly superior characteristics to those obtained with filters using only electrical elements. The type of electro-mechanical element which has been used most extensively in selective filters is the piezo-electric crystal and particularly the quartz crystal. This element has the advantage of an electro-mechanical converting system in the piezo-electric effect and a very high mechanical Q. Moreover, a quartz crystal is very stable mechanically and can be cut so that its frequency changes very little over a wide temperature range. For these reasons, quartz crystals have been applied extensively when it is desirable to obtain a narrow band filter or a very selective filter.

[12] This work was carried out principally by Messers. Hartley, Lane and Wegel.

[13] The use of a mechanical filter in visually studying the properties of a wave filter is described in a paper "A Mechanical Demonstration of the Properties of Wave Filters," C. E. Lane, *S.M.P.E. Jour.* Vol. 24, pp. 206–220, March 1935.

Such filters have received a wide variety of uses. Very narrow band filters have been used in carrier systems as pilot channel filters for separating out the pilot or control frequency from the other frequencies present; in radio systems for separating the carrier frequency from the sideband frequencies; and in heterodyne sound analyzing devices for analyzing the frequencies present in industrial noises, speech, and music. Wider band filters employing coils as well as crystals have provided very selective devices which are able to separate one band of speech frequencies from another band different by only a small percentage frequency range from the desired band. This property makes it possible to space channels close together with only a small frequency separation up to a high frequency, and such filters have had a wide use in the high frequency carrier systems and in the coaxial system which transmits up to 480 conversations over one pair of conductors. In radio systems such filters have been used extensively in separating one sideband from the other in single sideband systems.

The History of Electrical Resonance

By JULIAN BLANCHARD

OUR earliest knowledge of electricity was of the static kind; later came the voltaic cell and the direct current. But not until the discovery of alternating or oscillating currents of electricity could the phenomenon of electrical resonance make its appearance. Today, as we turn the dials of our radio receivers and "tune in" on the station we want it is recognized how widespread its application has become. Nevertheless, it seems that few have given thought to how this important principle came to light and how and when it got into common use.

The Oscillatory Nature of the Leyden Jar Discharge

The Leyden jar, discovered in 1746, was for many years one of the most important instruments in the meager equipment of electrical experimenters. When the jar was charged by an electrical machine and the discharging knobs brought close enough together a spark would jump between them. The savants of those days reasoned that this doubly coated jar was a storer of electricity, a condenser; that before the spark passed there was an accumulation of positive charge on one coating and of negative on the other; and when the spark passed these charges neutralized each other and the jar was discharged. But they did not know or suspect that this discharge was oscillatory, that first one side and then the other became positively charged, until the motion gradually came to rest.

The view that such was the case seems first to have been put forward in 1826 by Felix Savary, in France. It had been observed by him, and very likely by others as well, that a steel needle magnetized by the discharge of a Leyden jar did not in all circumstances have the same polarity. In the following words he suggested the idea that the results were due to the oscillatory discharge of the jar:

"An electric discharge is a phenomenon of movement. Is this movement a continuous translation of matter in a determined direction? Then the opposite polarity of magnetism observed at different distances from a straight conductor, or in a helix with gradually increasing discharges, would be due entirely to the mutual reactions of the magnetic particles in the steel needles. The manner in which the action of a wire changes with its length appears to me to exclude this supposition.

"Is the electric movement during the discharge, on the other hand, a series of oscillations transmitted from the wire to the surrounding medium and soon attenuated by resistances which increase rapidly with the absolute velocity of the moving particles?

"All the phenomena lead to this hypothesis which makes not only the intensity but the polarity of the magnetism depend on the laws in accordance with which the small movements die out in the wire, in the surrounding medium, and in the substance which receives and conserves the magnetism."[1]

Some fifteen years later Joseph Henry in America was experimenting with the Leyden jar and studying the currents induced in adjoining conductors by the discharge through another conductor. To determine the direction of the induced current he observed the polarity of a small steel needle magnetized by the current. In describing his experiments at a session of the American Philosophical Society he made reference to the work of Savary and stated that he had undertaken to repeat this investigator's experiments before attempting any new advances. He observed the same effect, the occasional reversal of the polarity of the needle after a discharge, and arrived at the same explanation:

"This anomaly which has remained so long unexplained, and which at first sight appears at variance with all our theoretical ideas of the connection of electricity and magnetism, was after considerable study satisfactorily referred by the author to an action of the discharge of the Leyden jar which had never before been recognized. The discharge, whatever may be its nature, is not correctly represented (employing for simplicity the theory of Franklin) by the single transfer of an imponderable fluid from one side of the jar to the other; the phenomena require us to admit *the existence of a principal discharge in one direction, and then several reflex actions backward and forward, each more feeble than the preceding, until the equilibrium is obtained.* All the facts are shown to be in accordance with this hypothesis, and a ready explanation is afforded by it of a number of phenomena which are to be found in the older works on electricity, but which have until this time remained unexplained."[2]

The published account of Henry's observations is not precisely in his own words but apparently in those of the reporter or secretary of the Society before which he spoke. It would seem from the above quotation, if it correctly represented the author, that Henry had overlooked the conclusions drawn by Savary, for they appear to be the same as his own.

This was in 1842. At a meeting of the Physical Society of Berlin in 1847 Helmholtz read his celebrated paper "On the Conservation of Force" (*Über die Erhaltung der Kraft*). Among the many illustrations of the conservation of energy principle in various branches of physics he discussed the case of the Leyden jar discharge, and incidentally noted another bit of evidence in favor of its oscillatory nature, an experiment by Wollaston in electrolysis. Commenting on the energy relations found to hold in this case he said:

"It is easy to explain this law if we assume that the discharge of a battery is not a simple motion of the electricity in one direction, but a backward and forward motion between the coatings, in oscillations which become continually smaller until the entire *vis viva* is destroyed by the sum of the resistances. The notion that the current of discharge consists of alternately opposed currents is favored by the alternately opposed magnetic actions of the same; and secondly by the phenomena observed by Wollaston while attempting to decompose water by electric shocks, that both descriptions of gases are exhibited at both electrodes."[3]

It may be interesting to note in passing that this now famous memoir by Helmholtz on the conservation of energy was considered so advanced and speculative as to be refused publication in the leading German scientific journal of the time. In it there was set forth, with far more thoroughness and generality than had been done before (by Mayer and Joule, for instance), the theorem that in any closed system the sum total of the energy is constant; a principle that at once denies the possibility of perpetual motion. It was privately published in pamphlet form in 1847. Its author, later to be recognized as the greatest German physicist of the century, was then an obscure young army surgeon, just twenty-six years of age.

By this time, it can be assumed from the foregoing, it was generally accepted by the learned in electrical science that the spark of a Leyden jar discharge was an oscillatory motion of electricity. This conclusion was arrived at as a logical and reasonable deduction from the results of various experiments, although the mode of action was still obscure. It was time now for a more analytical examination of the subject, and this was soon to appear.

In 1853 the British physicist Sir William Thomson, afterwards Lord Kelvin, published a paper with the title "On Transient Electric Currents,"[4] which, like that of Helmholtz, became in time a classic. In this paper the generalized problem of the discharge of a condenser through a conductor was treated mathematically. In addition to resistance and capacity he recognized the effect of inductance (called by him the "electrodynamic capacity") upon the discharge, and established an equation expressing the fact that the energy of the charged condenser at any instant during discharge is partly being dissipated as heat and partly conserved as current energy in the circuit; his equation being, in present day terminology,

$$-\frac{d}{dt}\left(\frac{1}{2}\frac{q^2}{C}\right) = \frac{d}{dt}\left(\frac{1}{2}Li^2\right) + Ri^2,$$

or

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = 0,$$

an equation that is easily solved. He analyzed the various solutions, which depend upon the relative values of the constants, or their ratios, and showed that under certain conditions the discharge is unidirectional and under others it is oscillatory, but damped. This beautiful bit of mathematical analysis, exact and thorough as it was, and clarifying the entire phenomenon, passed almost unnoticed at the time; but it came into its own with the arrival of wireless telegraphy.

There followed a few years later a direct experimental verification of the theory of the oscillatory nature of the Leyden jar discharge. In 1858 Feddersen[5] examined the spark by means of a revolving mirror, and extended his researches during the following year by the use of photography. There was thus obtained visual evidence of the reversal of direction of the discharge, and it was even possible to determine the frequency; photographs of these oscillatory sparks were sent to Thomson, who had suggested in his paper this very possibility of proof. Other experimenters followed with variations of this method of investigation, and in 1890 Boys[6] improved upon it by photographing the spark by means of a series of rapidly revolving lenses. Shortly before this a very important discovery had been made in connection with the spark discharge, Hertz's discovery of electric waves, and as a consequence more physicists were turning to a study of its characteristics, chiefly with the aid of photography.

The Effect of Capacity in an Alternating Current Circuit

It will be noted that the foregoing account is concerned with the very rapid, and transient, motion of electricity in open circuits. While knowledge of this sort of electric current was being advanced, Faraday's (and Henry's) discoveries in electromagnetic induction had made possible the invention of the dynamo and the production of a sustained alternating current—ordinarily of much slower motion. This generator, at first in the form of the feeble magneto, was for a long time not much more than a toy, and experience continued to be limited largely to the direct "galvanic" current. When eventually alternating currents began to be employed to an appreciable extent, in experiment and in industry, there were some new phenomena encountered, and those less theoretically grounded were slow to realize the peculiar effect of a condenser in the circuit, although the choking effect of an inductance alone was easily apparent.

The name of the great genius Maxwell now comes into our history. As we have seen, Lord Kelvin was the first to give a mathematical treatment of the oscillations of a Leyden jar discharge. So Clerk Maxwell was the first to publish an analysis of the effect of capacity in a circuit containing inductance and resistance and an impressed alternating electromotive force, and to show the conditions for resonance. The way in which he came to solve this problem makes an interesting story, and it was told in a characteristically interesting manner by the late Professor Pupin in the course of a discussion at a meeting of the A.I.E.E. Some preliminaries to this story may first be related.

For long the Ruhmkorff induction coil and the magneto-electric machine had been familiar objects in physical laboratories. In 1866 there appeared a description of Henry Wilde's striking experiments in which he virtually reinvented and introduced the separately excited dynamo, passing the small (commutated) current from a magneto-electric generator through the field magnet windings of another machine, from the armature of which a very much larger current was obtained. Sir William Grove, reading of these experiments, had the idea that a magneto might also be used advantageously to operate a Ruhmkorff coil with an alternating current. Induction coils had always been excited by means of a battery with a self-acting circuit breaker to interrupt the primary current, and in order to prevent sparking at the contacts and to stop the current more abruptly a condenser was connected across the contact terminals. Grove screwed up and kept closed the contact breaker, thus short-circuiting the condenser, and applied an ordinary medical magneto-electric machine to the primary terminals of his induction coil. To his surprise he found that he could get no secondary discharge at all; but by holding open the contact breaker, and so putting the condenser permanently in series with the primary coil and the armature of the magneto-electric machine, he obtained sparks nearly a third of an inch in length between the ends of the secondary. He saw that the effect was dependent upon the presence of the condenser in the circuit; "But why there should be no effect, or an appreciable one, when the primary circuit is completed, the current being alternated by the rotation of the coils of the magneto-electric machine, I cannot satisfactorily explain," he said.[7]

And now Professor Pupin's story:

". . . Maxwell, I think, was the first to show the effect of introducing a condenser into an alternating current circuit, and it is very interesting to observe this circumstance. Maxwell was spending an evening with Sir William Grove who was then engaged in experiments on vacuum tube discharges. He used an induction coil for this purpose, and found that if he put a condenser in parallel [it was in series, rather] with the primary circuit of his induction coil he could get very much larger sparks, which meant, of course, that he got a very much larger current through his primary coil, an alternating current generator being used to feed the primary. He could not see why. Maxwell, at that time, was a young man. That was about 1865, if I do not err. [It was 1868.] Grove knew that Maxwell was a splendid mathematician, and that he also had mastered the science of electricity as very few men had, especially the theoretical part of it, and so he thought he would ask this young man how it was possible to obtain such powerful currents in the primary circuit by adding a condenser. Maxwell, who had not had very much experience in experimental electricity at that time, was at a loss. But he spent that night in working over his problem, and the next morning he wrote a letter to Sir William Grove explaining the whole theory of the condenser in multiple [series] connection with a coil. It is wonderful what a genius can do in one night! He pointed out the exact relations between the condenser, the self-induction and the frequency which would give the largest current, and he was the first to do this, so far as I know . . ."[8]

Maxwell's letter, which began with the sentence, "Since our conversation yesterday on your experiment on magneto-electric induction, I have considered it mathematically, and now send you the result," was dated March 27, 1868; it was sent by Grove to the Philosophical Magazine, where it was published in May. Preliminary to the mathematical treatment Maxwell gave in this letter an unusually clear exposition of the analogy existing between certain electrical and mechanical effects, and from the standpoint of pedagogy as well as physics it will be interesting to see the language he used. He expressed himself thus:

"The machine produces in the primary wire an alternating electromotive force, which we may compare to a mechanical force alternately pushing and pulling at a body.

"The resistance of the primary wire we may compare to the effect of a viscous fluid in which the body is made to move backwards and forwards.

"The electromagnetic coil, on account of its self-induction, resists the starting and stopping of the current, just as the mass of a large boat resists the efforts of a man trying to move it backwards and forwards.

"The condenser resists the accumulation of electricity on its surface, just as a railway buffer resists the motion of a carriage towards a fixed obstacle."[9]

Using such concepts as these he gave a simple and lucid explanation of the problem without resort to mathematics; and then in a postscript, or appendix, he gave the mathematical theory of the experiment, employing a schematic diagram of the apparatus. Using different, but equivalent, symbols, he derived and solved the now familiar expression for the current i in such a circuit,

$$E \sin \omega t = L \frac{di}{dt} + Ri + \frac{1}{C} \int i dt$$

This is recognizable as similar to that set up by Lord Kelvin for the discharge of a condenser $\left(E \text{ being zero in that case, and } i \text{ being equal to } \frac{dq}{dt} \right)$. The solution of this equation is

$$i = \frac{E}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}},$$

from which Maxwell pointed out that the current would be a maximum when $\omega L = \frac{1}{\omega C}$ (ω being proportional to the frequency and L and C being the inductance and capacity, all in their proper units). Effecting this condition would of course be a case of "electrical resonance", brought about by electrical "tuning", though Maxwell himself did not specifically make use of these terms.

It is apparent, therefore, that all the knowledge necessary for the realization of electrical tuning was available in 1868 when this little communication of Maxwell's was published. Yet it was some years before electrical resonance was encountered to any extent in practice, or any practical use deliberately made of it. It was something known about by advanced physicists, perhaps, but outside the ken of most of those experimenting with electricity at that period; it must be remembered that there was no profession of electrical engineering as early as this. Not until 1884 do we find any other published discussion of the effect of a condenser in an alternating current circuit. This was in a technical paper by John Hopkinson, Cambridge trained physicist and later to become one of the foremost electrical engineers of his day. The paper had to do with alternating current theory and the operation of alternating current machines, and in it occurs the following:

> "Some time ago Dr. Muirhead told me that the effect of an alternating-current machine could be increased by connecting it to a condenser. This is not difficult to explain: it is a case of resonance analogous to those which are so familiar in the theory of sound and in many other branches of physics.
>
> "Take the simplest case, though some others are as easy to treat. Imagine an alternating-current machine with its terminals connected to a condenser; it is required to find the amplitude of oscillation of potential between the two sides of the condenser. . ."[10]

By setting up an equation similar to that used by Maxwell the required expression was found; and by assuming certain reasonable values of the frequency, resistance, inductance and capacity, he calculated that the amplitude of the potential difference across the condenser might be many times the voltage of the generator. It is apparent from the language he used that he had a perfectly clear understanding of electrical resonance.

The First Electrical Resonance Curve

Following Maxwell, there was another brilliant young physicist destined to become famous who showed a thorough acquaintance with electrical resonance and who made good use of it in his celebrated researches. This was Heinrich Hertz, in Germany, who applied it in the detection of electric waves produced by a spark discharge, the oscillatory nature of which had already been well investigated, as we have seen. The simple device he used for exploring the field in the vicinity of the discharge was a rectangle or circle of wire containing a minute spark gap, the loop being of such dimensions as to be in resonance with his high frequency oscillator. It was by this careful exploration that Hertz demonstrated, for the first time, the existence of electromagnetic waves in space. In the first of his series of papers describing these experiments, "On Very Rapid Electric Oscillations," published in 1887, he devotes one section to a discussion of "Resonance Phenomena." An extract from this will show how he was thinking:

"But it seemed to me that the existence of such oscillations might be proved by showing if possible symphonic relations between the mutually reacting circuits. According to the principle of resonance, a regularly alternating current must (other things being similar) act with much stronger inductive effect upon a circuit having the same period of oscillation than upon one of only slightly different period. If, therefore, we allow two circuits, which may be assumed to have approximately the same period of vibration, to react on one another, and if we vary continuously the capacity or coefficient of self-induction of one of them, the resonance should show that for certain values of these quantities the induction is perceptibly stronger than for neighbouring values on either side."[11]

A series of experiments along these lines demonstrated the effect conclusively. In the secondary circuit the length of spark that could be obtained across the adjustable gap increased to a maximum when the two circuits were in tune. The first electrical resonance curve ever published is given in the above mentioned paper,[12] a relation between the length of wire in the detecting loop and the greatest length of spark obtainable for each length of wire, all other conditions remaining unchanged. The curve shows the familiar sharp peak at the point of resonance. In all his succeeding researches on electric waves Hertz used this simple tuned circuit as a detector. It was the forerunner of the resonance type of wave-meter to be used later in the yet unborn art of radio.

Among the prominent British physicists Oliver Lodge at this time was also experimenting with electrical resonance and writing and lecturing about it. He had from the first taken a keen interest in the work of Hertz and in fact came close to anticipating Hertz in the discovery of electric waves through his notable work on lightning conductors. In a brief article published in 1890 he described a method, which he had used a year before in a lecture, of displaying the spark producing power of electric radiation by tuning the circuit of one Leyden jar to that of another containing a spark gap and excited in the usual way.[13] When the secondary circuit was in resonance with the first its Leyden jar would "overflow." But Lodge objected to the use of the term "resonance" and preferred the term "syntony"; "the name 'resonance' is too suggestive of some acoustic reverberation phenomenon to be very expressive," he maintained.[14] Although he and some of the other English writers continued to say "syntony" and "syntonic", this terminology did not permanently stick.

During the next decade, as electrical engineering developed somewhat, especially in alternating currents, we find more attention being paid to this subject, both by physicists and engineers. Among those interesting them-

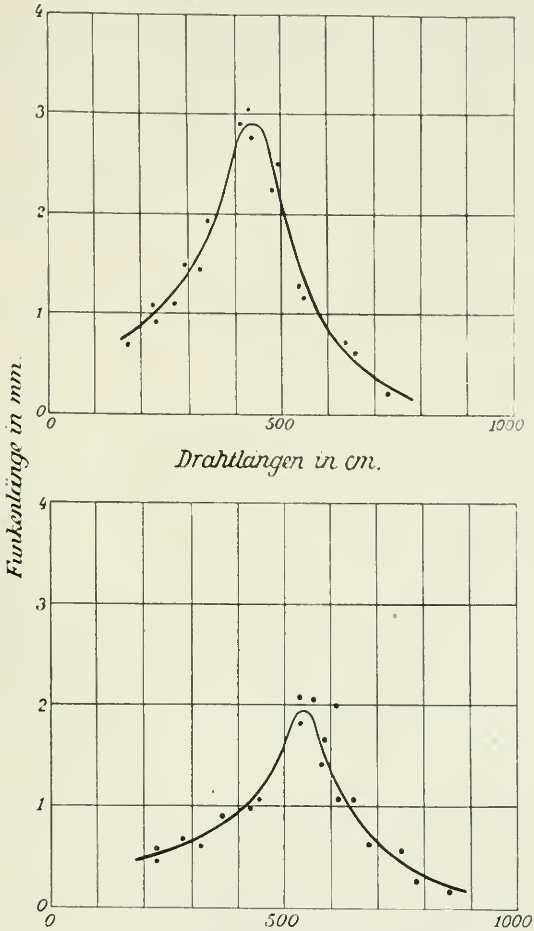


Fig. 1—The first electrical resonance curves published, by Hertz, 1887; showing the greatest length of spark obtainable in his detecting loop for various lengths of wire in the loop.

selves in the matter may be mentioned such men as T. H. Blakesley, Gisbert Kapp, J. A. Fleming, R. T. Glazebrook, James Swinburne, Maurice Hutin and Maurice Leblanc, Frederick Bedell and A. C. Crehore, Nikola Tesla, M. I. Pupin, and John Stone Stone; some of these being concerned with high frequencies, some with low. As indicating the general state of knowledge at this time concerning alternating current theory, a statement in a textbook by Blakesley, with preface dated May, 1889, is illuminating. This author says:

"It is often taken for granted that the simple form of Ohm's Law, total E.M.F. ÷ total resistance = total current, is true for alternating currents. That is to say, the E.M.F. employed in the formula is taken to be the sum of the impressed E.M.F.'s alone. That there are causes which modify the value of the current as deduced from this simple equation, such as mutual or self-induction, or the action of condensers, is often acknowledged in textbooks, and the values and laws of variation of the current are correctly stated for certain cases of instantaneous contact and breaking of circuit. But the effect of an alternating E.M.F. upon a circuit affected by self-induction, mutual induction, and condensing action, has not been, so far as I know, put into a tangible working form."[15]

Somewhat similar observations were expressed by Kapp in an article in the *Electrician* a year or so later. Referring to the paper by Hopkinson mentioned above, he commented as follows:

". . . he showed that with a certain capacity, periodic time, self-induction and resistance in circuit, the potential difference between the plates of the condenser may be 80 times the E.M.F. of the alternator. Startling as such a result must naturally appear, it failed at the time to attract much attention from practical engineers who, no doubt, preoccupied with the problems relating to continuous-current work, were content to let such an intricate and apparently abstruse problem lie at rest until such time as its consideration should be forced upon them. This time has now come, and what in 1884 was merely an interesting laboratory experiment, having no further application than perhaps the breaking down of a condenser, is at present an interesting practical problem, which the electrical engineer has to face. Phenomena arising from the effects of capacity in alternate-current circuits are forcing those who have to do with such circuits to give attention to the problems connected with the phenomena."[16]

This quotation gives a fair picture of the situation with respect to electrical engineering around 1890. Familiarity with capacity and resonant effects, it appears, was beginning to grow with the enlargement of professional experience.

Resonance in Electric Communication

While resonance, or an approximation thereto, is occasionally encountered in ordinary power engineering and electric lighting, here it is generally a case of something to be avoided, evidence of something gone wrong. An unintentional resonant condition in a power circuit could result in considerable damage due to excessive current flow. In electric communication, on the other hand, where frequencies are higher, and where *frequency* itself is one of the fundamental elements, and currents comparatively small, resonance is of prime importance and may be of great practical value.

As we consider the use of resonance in electric communication there may occur to some readers a recollection of a very early system of multiplex signaling known as the "harmonic telegraph", representing the attempts of Elisha Gray, Alexander Graham Bell, E. Mercadier and others to transmit simultaneously a number of telegraph messages over the same line; experiments which, in the case of Bell, led to the invention of the telephone. These various schemes, however, were all based on the principle of mechanical resonance; electromagnetically driven tuned reeds at the receiving end were set to vibrating by signaling currents generated by corresponding reeds at the transmitting end. The principle of electrical resonance was not involved in such methods.

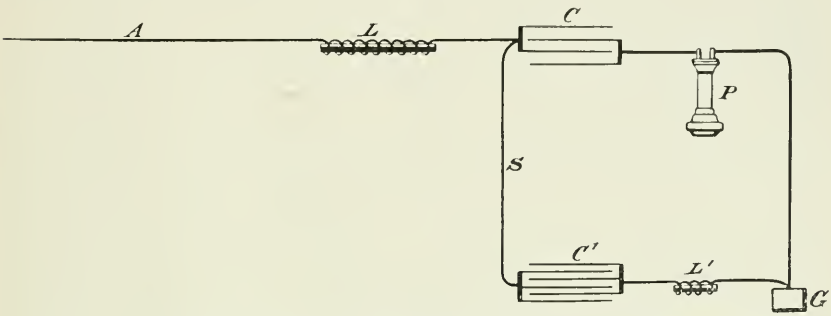


Fig. 2—Combination of series and shunt resonant elements to lessen interference of power source low frequencies with higher telephone frequencies; from U. S. patent of Stanley and Kelly, 1891.

Suggestions for the practical application of electrical resonance began to appear in the early 1890's. By this time, as our history shows, the phenomenon was generally understood by the technically trained and the well informed; it was one of the facts of science open to all. Henceforth, progress in the putting to use of it was largely in the hands of inventors and its history is to be found in the study of patents.

In telephony, one of the earliest proposals is illustrated by a United States patent issued to Stanley and Kelly in 1891,[17] showing methods for preventing interference with telephone currents by lower frequency currents induced in the line by power sources. One of the methods described was the insertion in series with the receiver of a capacity making a combination resonant to the mean speech frequency, supplemented by a shunt combination of capacity and inductance resonant to the interfering frequency. It need hardly be said that such an arrangement, favoring only a narrow band of the speech frequencies, would greatly promote distortion and would find little favor with telephone engineers.

Another application, for a different purpose, appeared in a French patent issued to Hutin and Leblanc during the same year.[18] These engineers were pioneers in the attempt at multiplex telephony by means of high-frequency carrier currents, a method now so greatly extended. In May, 1891, several months before their patent papers were filed, they had reviewed in a French journal the theory of resonance in an inductively coupled circuit in the course of a general article on alternating currents, and had briefly suggested therein its application to multiplex signaling.[19] The scheme disclosed in their patent comprised the transmission over the line of a number of super-audible frequencies, now called carrier currents, the modulation of each by a

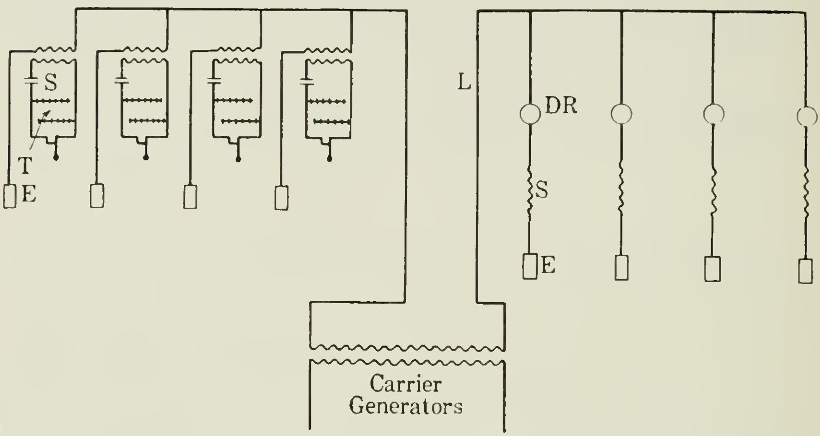


Fig. 3—Multiplex carrier telephone circuit of Hutin and Leblanc, as in their French patent of 1891. At each terminal are shown four branches, each of these branch circuits being tuned to one of the carrier frequencies.

separate telephone transmitter, and means for separating and detecting the individual messages. At both the transmitting and receiving ends, in their plan, there were branches from the connecting line, each of these branch circuits being tuned by means of a capacity which balanced the inductance in the circuit, so that each responded by resonance to its own carrier frequency to the exclusion of the others.* Here was a substitution of electrical resonance for the mechanical resonance of the older harmonic telegraph.

At about this same time Professor Pupin in his early research work at Columbia had developed a method of analyzing a complex current by picking out its components in an inductively coupled resonant circuit—a

* While their patent drawing fails to show tuning condensers in the receiving branches, without which the scheme would be inoperative, and the description on this point is vague, in the interference cases that later developed in the U. S. Patent Office (reference 24) it was claimed that this omission was an oversight. In a later French patent, No. 234,785, granted March 5, 1894, as in their U. S. patent, No. 838,545, this defect was corrected.

tool, or technique, that is now quite familiar in electrical laboratories. He made use of this in the study of the harmonics generated in a circuit by the magnetic reactions of an iron core upon the magnetizing current, an effect that had been observed and for the first time correctly explained by Rowland at Johns Hopkins, and a description of his method was published in 1893.[20] The following year, yielding to the suggestions of his scientific friends, according to the account he has written ("I often regretted it, because it involved me in a most expensive and otherwise annoying legal contest"[21]), he made application for a patent on "Multiple Telegraphy", applying this idea of selecting by resonance to the problem of separating the signals.[22] Very soon afterwards another inventor, John Stone Stone, appeared upon the scene with practically the same idea,[23] and interference cases thereupon resulted in the U. S. Patent Office and the courts involving these two and the French inventors Hutin and Leblanc, who had also filed in the United States.[24] Upon the claims of the contestants and the differences that characterized their schemes for multiplex signaling we need not dwell; suffice it to say that in the matter of priority Pupin was adjudged the winner.* It appears that this distinguished scientist was not unimpressed with what he considered the originality of his ideas about the practical use of resonance. In the inimitable story of his life, "From Immigrant to Inventor," he refers to this as "my invention of electrical tuning,"[25] and says again, "I called it electrical tuning, a term which has been generally adopted in wireless telegraphy."[25] In another place, and on another occasion, he said, "It was badly needed and I had it developed several years before Marconi had made his invention. . ."[26]

Before passing to other applications in the field of electric communication, chiefly in the radio art, it might be said that in these early proposals for multiplex operation the separation of the carrier frequencies could not be successfully achieved by so simple a means as an ordinary resonant circuit. For one thing, the distortion introduced would be prohibitive, unless the carrier channels were placed so far apart as to be uneconomic. It remained for the Campbell band filter, invented about twenty years later, to enable the frequencies to be squeezed close together and distortion and other difficulties to be overcome. Furthermore, the whole art had to wait for the invention of the vacuum tube as the perfect generator of the kind of currents required, as well as modulator, amplifier and demodulator of these currents.

Later developments in the intricate and complex technique of wire

* An examination of the report of the interference hearings (reference 24) shows that Pupin claimed to have conceived the idea of using electrical resonance in multiplex telegraphy in the summer of 1890, following a careful study he had made of the investigations of Hertz, and to have begun experimental work on it in October of that year, thus antedating Hutin and Leblanc. Upon the adjudication of this contest in favor of Pupin on the main issues, patents on some of their claims were also allowed Stone and the French inventors.

communication have brought forth more useful applications of the principle of electrical resonance than the examples cited above. But let us now turn to radio. Here the application of resonance is elemental and fundamental. But not so in the beginning, however. When Marconi brought his embryonic outfit to England in 1896 and demonstrated his best accomplishments over the next three or four years, the problem of selectivity was non-existent. Further, the type of detector then available, the Branly

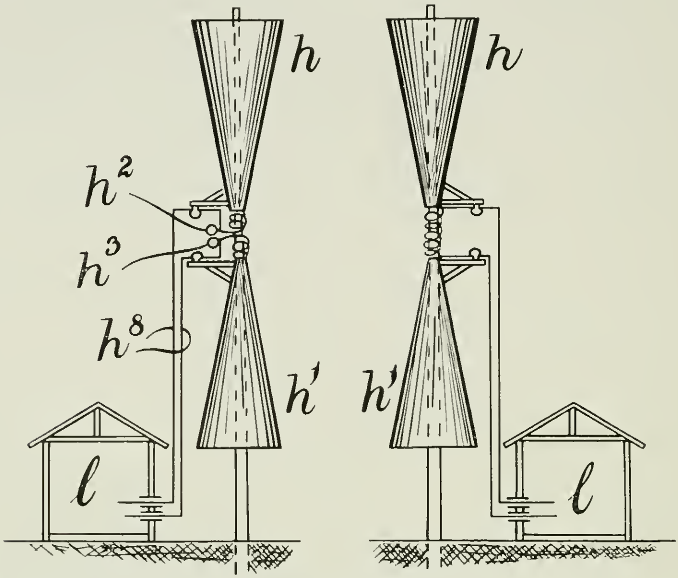


Fig. 4—First tuned radio transmitting and receiving antennas, as proposed in Lodge's British patent of 1897; tuning accomplished by inductance coils between the capacity areas *h* and *h'*.

metal filings coherer or modifications thereof, was responsive to electric waves varying considerably in frequency, so the need of tuning to obtain sensitivity was at first not actually imperative. This detector was connected directly in the untuned antenna circuit. Then the ambition to increase the distance of reception led to a search for greater sensitivity, and as a first step (1898) Marconi introduced into the receiving circuit his "jigger", or oscillation transformer.[27] The primary, of few turns, was in the antenna circuit; the tuned secondary, wound with an eye to the reduction of capacity, stepped up the voltage and applied it to the coherer. Here no adjustable tuning was provided, but instead there were different jiggers wound to suit the transmitted wave-lengths employed and thus secure the maximum effect.

It was foreseen in the early days of radio that if it were ever to become a commercial practicability it would be necessary to provide means for receiving one wave-length to the exclusion of others—to provide selectivity. Crookes in his prophetic Fortnightly Review article of 1892 had clearly envisaged this.[28] As a solution of this problem Lodge in 1897 applied for a British patent on "Improvements in Syntonized Telegraphy without Line Wires,"[29] the stated object of his invention being "to enable an operator by means of what is known as Hertzian wave telegraphy to transmit messages

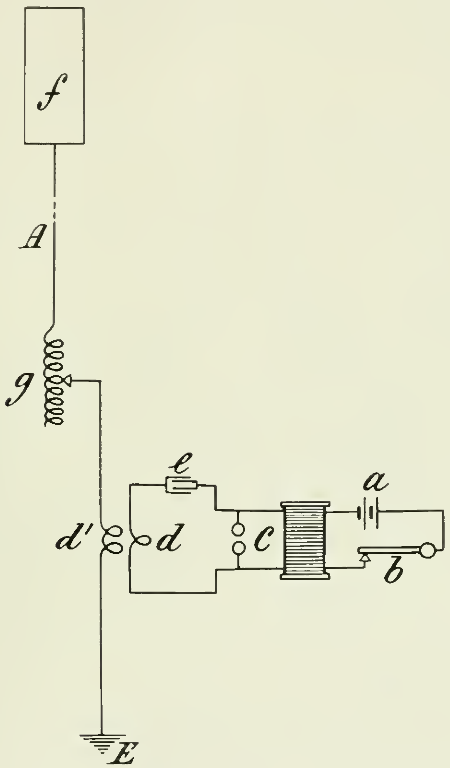


Fig. 5—The tuned inductively coupled two-circuit radio transmitter adopted by Marconi in 1900.

across space to any selected one or more of a number of different individuals in various localities each of whom is provided with a suitably arranged receiver." His radiator, modeled after Hertz, was a pair of "capacity areas", or triangular shaped metal plates (one of them preferably grounded), separated by a spark gap and having interposed an inductance coil of a few turns, for the purpose of tuning. This coil was not continuously adjustable but was to be replaced by others for changes of wave-length. The receiving station was provided with a similar arrangement except that in place of the spark gap there was connected a Branly type coherer as a wave detector.

While the particular forms of apparatus shown were never adopted in practice, nevertheless Lodge's tuning patent was upheld as valid in a legal contest for priority later on (1911) and it was purchased by the Marconi

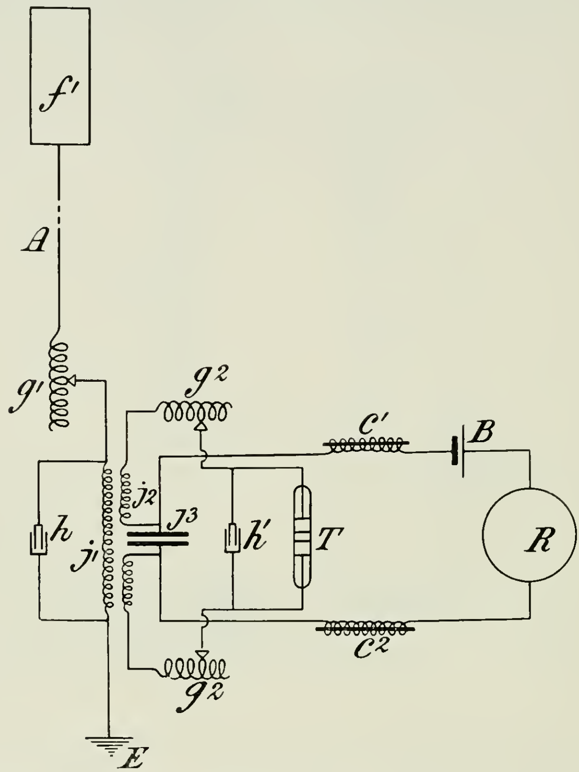


Fig. 6—Marconi's tuned radio receiving antenna and circuit of 1900.

Company as a result of the litigation.[30] Lodge, in his autobiography "Past Years", speaking with the usual modesty of the inventor rather than the scientist that he was, says, "Real selective tuning became possible through my patent of 1897. . ."[31]

In Marconi's first patent on wireless telegraphy, applied for in 1896 (British patent No. 12,039, the very first patent in radio), there was no reference whatever to the matter of tuning. But his well-famed patent No. 7777 of 1900[32] was primarily concerned with this object and went much ahead of his jigger ideas of 1898. Here there were four tuned circuits. At the sending station the spark gap circuit was inductively coupled to the transmitting aerial (an improvement credited to Ferdinand Braun of Germany), and by means of a variable condenser in the former and a variable inductance in the latter these two circuits could be tuned and brought into resonance with each other. This accomplished the production of a much more persistent train of oscillations in the aerial and a more efficient radiation of energy. At the receiving end the aerial was tuned to the incoming waves by means of a variable inductance, and the inductively coupled detector circuit was in turn tuned to resonance, likewise by means of a variable inductance. It was partly through such steps in the realization of greater sensitivity as well as selectivity that Marconi eventually succeeded with transoceanic telegraphy.*

In this patent the inventor gave the specifications for nine different *tunes*, as he called the different frequencies intended for different stations, or for different distances; that is to say, the details of design of the aerials, transformers, inductances and capacities of the transmitting and receiving circuits for each tune. Thus interference between one station and another might be avoided by using different frequencies. It may be observed here, however, that the matter of selectivity was not so easy at that time when the rather broad-spectrum spark transmitter was the only kind available. Very sharp tuning had to wait upon the advent of continuous waves, supplied first by the Poulsen arc or the high-frequency alternator and then by the vacuum tube. But many other improvements, and new wonders besides, were waiting on the vacuum tube.

It hardly seems necessary to pursue our subject further than this point, considering how it so quickly thereafter became a commonplace item in our electrical storehouse. Our interest was chiefly in how it got started. We have seen how it had its roots in certain experiments with the Leyden jar; how the results of experiments were clarified by mathematical analysis and a correct theory formulated; and then, as the need and opportunity arose, how the principle was applied and made use of by inventive minds: the wilderness first entered, then surveyed, and at last inhabited. So it is, we find, with most new ideas in the scientific world.

References

1. Felix Savary, "Memoire sur l'Aimantation" ("Memoir on Magnetization"), *Annales de Chimie et de Physique*, vol. 34, pp. 5–57, 1827. (Read before the Academy of Sciences, Paris, July 31, 1826.)

* Tesla's brilliant experiments with resonance and high-frequency currents during this period and his knowledge and handling of tuned coupled circuits should be noted here. Although his work for a time was concerned largely with the conversion of ordinary power source currents into currents of very high frequency and voltage (his "Tesla coil" of 1891 is still well known) for a proposed system of electric lighting by vacuum tube discharges, much of it was applicable to wireless telegraphy. Particularly, his synchronous discharger with adjustable electrodes and provision for tuning the low-frequency circuit to resonance, patented in 1896, could very readily have been incorporated into a wireless transmitter.

2. Joseph Henry, "On Induction from Ordinary Electricity; and on the Oscillatory Discharge," *Proc. Am. Phil. Soc.*, vol. 2, pp. 193–196; June 17, 1842. "Scientific Writings of Joseph Henry," vol. 1, pp. 200–203, published by the Smithsonian Institution, Washington, D. C., 1886.
3. Hermann von Helmholtz, "Über die Erhaltung der Kraft," published in pamphlet form by G. Reimer, Berlin, 1847. Reprinted in volume 1 of "Wissenschaftlichen Abhandlungen von Hermann Helmholtz," by J. A. Barth, Leipzig, 1882. Reprinted in Ostwald's "Klassiker der exacten Wissenschaften," No. 1, by Wilhelm Engelmann, Leipzig, 1889. Translated into English by John Tyndall, in "Scientific Memoirs, Natural Philosophy," Article IV, pp. 114–162; edited by John Tyndall and William Francis, printed by Taylor and Francis, London, 1853.
4. William Thomson, "On Transient Electric Currents," *Phil. Mag.*, vol. 5, pp. 393–405; June, 1853. "Mathematical and Physical Papers," vol. 1, pp. 540–553.
5. W. Feddersen, "Beiträge zur Kenntniss des elektrischen Funkens," *Ann. der Phys.*, vol. 103, pp. 69–88, 1858; also vol. 108, pp. 497–501, 1859; vol. 112, pp. 452–459, 1861; vol. 113, pp. 437–467, 1861; vol. 115, pp. 336–343, 1862; vol. 116, pp. 132–170, 1862.
6. C. V. Boys, "Notes on Photographs of Rapidly Moving Objects and on the Oscillating Electric Spark," *Proc. Roy. Soc. Lond.*, vol. 11, pp. 1–15; Nov., 1890.
7. W. R. Grove, "An Experiment in Magneto-Electric Induction," *Phil. Mag.*, vol. 35, pp. 184–185; March, 1868.
8. M. I. Pupin, *Trans. A.I.E.E.*, vol. 10, p. 223, 1893. (Discussion.)
9. James Clerk Maxwell, "On Mr. Grove's 'Experiment in Magneto-Electric Induction'. In a Letter to W. R. Grove, F.R.S.," *Phil. Mag.*, vol. 35, pp. 360–363; May, 1868. "Scientific Papers," vol. 2, pp. 121–124.
10. John Hopkinson, "On the Theory of Alternating Currents, Particularly in Reference to Two Alternate-Current Machines Connected to the Same Circuit," *Jour. Soc. Tel. Engrs.*, vol. 13, pp. 496–515; Nov. 13, 1884.
11. Heinrich Hertz, "Electric Waves," p. 42. Translated by D. E. Jones. The Macmillan Company, London and New York, 1893 and 1900.
12. Reference 11, page 43.
13. Oliver Lodge, "Easy Lecture Experiment in Electric Resonance," *Nature*, p. 368, Feb. 20, 1890.
14. Oliver Lodge, "Modern Views of Electricity," p. 339, 2nd ed., 1892.
15. T. H. Blakesley, "Papers on Alternating Currents of Electricity, for the Use of Students and Engineers," p. 1, 3rd ed., 1891. Whittaker & Co., London.
16. Gisbert Kapp, "Capacity and Self-Induction in Alternate Current Working. I.," *The Electrician*, vol. 26, pp. 197–199; Dec. 19, 1890.
17. William Stanley, Jr., and John F. Kelly, "Means for Preventing the Interference with Speech on Telephone Circuits by Induced or Other Currents"; *U. S. Patent* No. 464,529, filed May 22, 1891, issued Dec. 8, 1891.
18. Maurice Hutin and Maurice Leblanc, *French Patent* No. 215,901, dated Sept. 3, 1891, delivered Dec. 15, 1891. See also *U. S. Patent* No. 838,545, filed May 9, 1894, issued Dec. 18, 1906.
19. Maurice Hutin and Maurice Leblanc, "Étude sur les Courants Alternatifs et Leur Application au Transport de la Force," *La Lumière Électrique*, vol. 40, p. 259; May 9, 1891.
20. M. I. Pupin, "On Electrical Oscillations of Low Frequency and Their Resonance," *Am. Jour. of Sci.*, vol. 44, pp. 325–334, April; pp. 420–429, May; pp. 503–520, June, 1893. "Practical Aspects of Low-Frequency Electrical Resonance," *Trans. A.I.E.E.*, vol. 10, pp. 370–394; May, 1893.
21. M. I. Pupin, "From Immigrant to Inventor," p. 300. Charles Scribner's Sons, New York, London, 1923.
22. M. I. Pupin, "Multiple Telegraphy," *U. S. Patent* No. 707,007; orig. filed Feb. 23, 1894, issued Aug. 12, 1902.
23. John Stone Stone, "Method of Electrical Distribution and Selective Distribution," *U. S. Patent* No. 726,368; orig. filed Apr. 4, 1894, issued Apr. 28, 1903; also 726,476, 1903; 729,103, 1903; 729,104, 1903; 638,152, 1899.
24. Interference in the U. S. Patent Office No. 17196 and No. 17197.
25. Reference 21, p. 19. See also pp. 294–302.
26. M. I. Pupin, "Fifty Years' Progress in Electrical Communications," *Jour. A.I.E.E.*, vol. 46, pp. 59–61, Jan.; pp. 171–174, Feb., 1927. (Presidential Address, Dec. 27, 1926.)

27. G. Marconi, "Improvements in Apparatus Employed in Wireless Telegraphy," *British Patent* No. 12,326; filed June 1, 1898, accepted July 1, 1899.
28. William Crookes, "Some Possibilities of Electricity," *The Fortnightly Review*, London, pp. 173–181; Feb., 1892.
29. Oliver Lodge, "Improvements in Syntonized Telegraphy without Line Wires," *British Patent* No. 11,575; filed May 10, 1897, accepted Aug. 10, 1898.
30. Oliver Lodge, "Past Years, An Autobiography," p. 236. Charles Scribner's Sons, New York, 1932.
31. Reference 30, p. 233.
32. G. Marconi, "Improvements in Apparatus for Wireless Telegraphy," *British Patent* No. 7777; filed April 26, 1900, accepted April 13, 1901.

A Rapid Visual Test for the Quantitative Determination of Small Concentrations of Calcium in Lead*

By EARLE E. SCHUMACHER and G. M. BOUTON

A method is described for estimating calcium in lead which consists in casting a test ingot in a prescribed manner and comparing its surface appearance with the surface appearance of standards. The calcium content can be determined by inspection.

IN THE manufacture of lead-calcium sheath it is desirable to control the calcium content to $0.028 \pm .005$ per cent in order to obtain the most desirable combination of properties. Since calcium is a very active element chemically, special manufacturing procedures were developed to minimize the contact of molten lead-calcium alloy with the air. Despite the improved techniques, some calcium is always lost and must be replaced. Before this can be accomplished satisfactorily, obviously, the calcium content of the alloy must be determined. Conventional chemical procedures are accurate but not entirely satisfactory for plant control use because they are time consuming and too costly. The best of the chemical methods introduces a lag of at least two hours in melting kettle control. Quantitative spectrographic analysis methods were carefully tested, and while they showed some advantage over conventional methods, they were still unduly time consuming.

With the ever increasing interest in lead-calcium alloys for cable sheath, storage-battery grids, and other applications, it became desirable that a rapid, reliable and not too costly method be developed for determining their calcium content. In approaching the problem, several methods of attack involving physical, chemical, or electrical properties suggested themselves. A few of the methods investigated were:

1. Observations of the rate of oxidation or tarnish of freshly cut surfaces using a variety of atmospheres and temperatures.
2. Thermal EMF measurements against pure lead.
3. Measurements of hardness or strength of samples after various heat treatments.
4. Measurements of electrode potentials in various solutions.
5. Use of various metallographic techniques.
6. Observation of recrystallization tendencies after the samples had been deformed.

* This article is being published in *Metals and Alloys*.

Unfortunately, none of these methods proved adequate. Either the properties involved were insufficiently sensitive to changes in calcium content, or other factors masked the effect of calcium.

Early in the study of lead-calcium it was noticed that the molten alloys quickly filmed over with oxide. Careful observation of the characteristics of these molten alloys revealed no phenomena that varied sufficiently with calcium content to serve as a clue to the composition. However, when these alloys were chill cast with as little agitation as possible, the surface of the ingots became progressively duller with increasing calcium content to a certain value. Further increase in calcium content resulted in the fissuring of the surface oxide leaving bright metallic areas exposed. This fissuring phenomenon, which is illustrated in Fig. 1, is the type of composition-sensitive indicator desired. When samples of lead-calcium are melted and cast under controlled conditions the surface markings are reproduced with considerable fidelity in respect to areas of dull and bright surface. The ratio of these areas is dependent on the calcium content. For clearness of illustration the samples were photographed under lighting conditions that made the bright highly reflecting surfaces appear dark in the photograph.

The success of the method is dependent to a large extent on the details of procedure that are given below. Since calcium is readily removed from lead by oxidation, a melting and casting procedure for the test specimens was adopted that resulted in a minimum loss of calcium. Fluxes and inert atmospheres, which normally provide adequate protection against oxidation, could not be used here since they seriously interfere with the fissuring phenomenon that is the basis of the method. The means finally adopted consists in melting a strip of the cable sheath to be tested in a hemispherical sheet iron crucible about two inches in diameter. A Bunsen burner flame of sufficient intensity to melt a 100-gram sample in about two minutes is applied to the bottom of the crucible. The bottom edge of the sample melts first and the balance of the sample slides smoothly into the pool of metal first formed with a minimum of rupture of the surface. The broad round shape of the crucible permits it to be tilted until the lip is but a fraction of an inch from the surface of the mold before the metal starts to pour, thus subjecting the stream of molten alloy to only a brief exposure to the atmosphere during pouring. The molten alloy should never be stirred nor should the crucible be shaken unnecessarily during the casting operation. Under the melting conditions described, the casting temperature of the melt is controlled sufficiently if the crucible is removed from the flame three or four seconds after the last portion of the sample has melted. By slight modifications in technique, samples for analysis may be taken directly from the commercial melting kettles. The mold used is probably

Fig. 1—Surface appearance of test ingots of chemical lead-calcium alloys cast in atmosphere containing 0.02% carbon dioxide and having 50% relative humidity. The dark areas are very bright when viewed with directly reflected light. Approximately full size

not critical in shape or size. The one used in this development is an iron plate $\frac{3}{4}$ in. x 4 in. x 8 in. with a tapered depression milled in its surface. The test ingot is about 4 in. long, $\frac{5}{8}$ in. wide, $\frac{1}{32}$ in. thick at the casting end and $\frac{3}{8}$ in. thick at the other end. It weighs 80 to 100 grams. The effect of mold temperature has been studied and is not critical in the range from room temperature to that reached by the mold as a result of casting into it at intervals of a few minutes.

Control of the atmosphere over the surface has been found necessary to insure reproducible results. For this reason the cellophane enclosed casting chamber shown in Fig. 2 was devised. There is a door on the right hand side for insertion of the crucible and a rubber inlet tube on the left for entry of the gas mixture. It has been found by experiment that both moisture and the carbon dioxide content of the air influence the results obtained. Therefore, both are removed chemically and then re-introduced

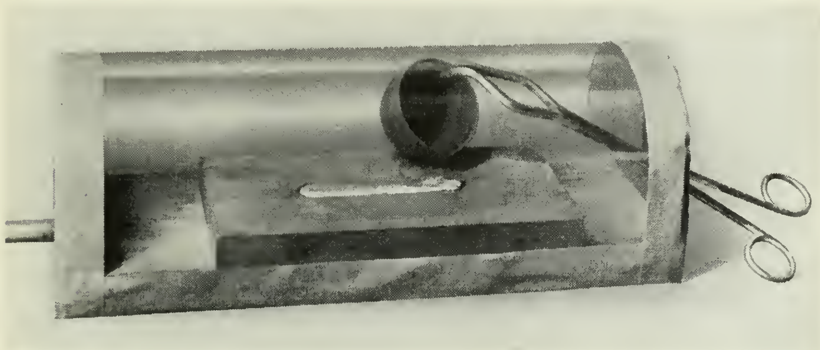


Fig. 2—Apparatus used in casting lead-calcium ingots under controlled atmospheres

in metered quantities to the air passing into the casting chamber. By means of appropriate flow meters, dry air could be mixed in definite proportions with air saturated with moisture to produce the desired humidity. Another and perhaps simpler system is to pass the dry air over certain salt solutions of known vapor pressure. Carbon dioxide is conveniently made available by placing solid carbon dioxide (dry ice) in a Dewar flask having a stopper with two exit tubes. By means of an escape valve on one tube any desired CO_2 pressure can be built up in the flask to force the gas through a flow meter and into the air line leading to the casting chamber. One satisfactory arrangement of apparatus for controlling the composition of the atmosphere is shown in Fig. 3.

The surface appearances shown in Fig. 1 were obtained by casting in an atmosphere having 50 per cent relative humidity and 0.02 per cent carbon dioxide. Increasing the carbon dioxide or decreasing the moisture content causes the surfaces to become brighter and vice versa. This provides a few thousandths of a per cent latitude in adjusting the sensitive range of the method to the median calcium content desired. Alloys having calcium contents outside the range of the method may be estimated by admixture with known amounts of lead-calcium alloys of known calcium content. Pure lead, in general, cannot be used for dilution because the oxide and possibly traces of impurities present in it cause the loss of some calcium from the mixture.

To date most experience with the use of the visual test for calcium has been on alloys made with chemical lead. This grade of lead is substantially free from As, Sn, Bi, Fe, Sb and Zn, and contains about 0.004 per cent Ni, 0.06 per cent Cu and 0.007 per cent Ag. The variation in concentration of these elements in the commercial supply has not been found to be great enough to cause serious interference with the indications given by the casting test. However, the method is influenced by certain variations in impurity content which are in excess of those normally encountered in the

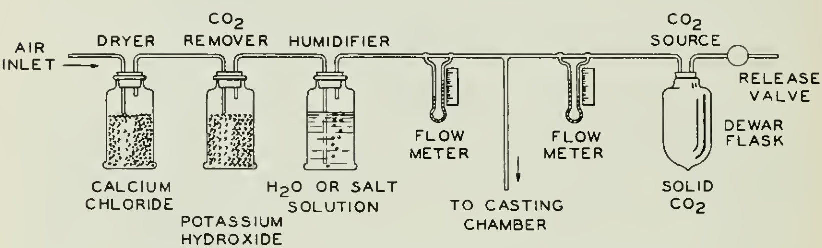


Fig. 3—Apparatus used for regulating atmosphere composition

usual supply. The presence of tin in the order of a few thousandths of one per cent causes low indications of calcium content. The use of high purity lead in place of "chemical" lead has a similar effect. Bismuth additions up to 0.1 per cent are inconsequential. Arsenic and antimony form compounds with calcium which dross off in the mixing kettle so that these elements would not be found in the finished sheath. The effects of the interfering impurities referred to above can be corrected for, when their presence is known, by varying the atmosphere in the casting chamber or by admixture with a known proportion of alloy having a pre-determined higher calcium content. By these procedures fissuring can be made to occur in alloys that otherwise do not give this manifestation. In practical operation the brand of lead being used will be known and the necessary adjustments can be made in the conditions of the test.

Sufficient analytical data have been collected to establish the fact that the method here presented is rapid, reliable, and extremely helpful for plant control application.

Abstracts of Technical Articles by Bell System Authors

Electron Microscopes and their Uses.[1] JOSEPH A. BECKER and ARTHUR J. AHEARN. Three and a half centuries have passed since Zacharias Janssen, a spectacles maker of Middleburg, Holland, put two lenses in a six-feet-long tube and thereby made the first known compound microscope. In the years since then, the microscope, now grown into a powerful and intricate instrument, has played an important role in the discovery of much of man's knowledge of the physical world. There is, however, much that the microscope has been unable to reveal because of its limited range of useful magnification. Today a new type of magnifying instrument, the electron microscope, is extending the range of useful magnification far beyond its old limits and promises to supplement the traditional microscope in many fields of scientific research. In this article the authors describe types of electron microscopes, tell how they function, and outline how they are being used in physics, chemistry, metallurgy and the biological sciences. A number of pictures are shown to illustrate these uses.

Recent Developments in Protective Metallic Coatings.[2] R. M. BURNS. The prevention of corrosion is accomplished by two general methods: (1) the provision of a non-corrosive environment, and (2) the interposition of a protective film to exclude the corrosive environment from the metal. As an example of the first method one may cite the de-aeration of boiler feed waters and air conditioning in which moisture is controlled and dust, sulphur gases, etc. eliminated.

Referring to the second method, corrosion protective films may be divided into two main classes: the first consisting of those films formed naturally through the production of corrosion products on the surface of the metal to give a thin protective coating; and second, comprising films of paints, varnish, ceramic products or metals which themselves develop protective films. One natural type of protective film is the chemical conversion coating produced by various treatments, such as phosphate or chromate dipping or anodic oxidation.

Zinc is the most important of metallic coatings, 45% of the metal consumed in the United States being used in this manner. Hot galvanized coatings on steel have been improved by suitable pre-treatment of the

[1] *The Scientific Monthly*, October 1941.
[2] *The Monthly Review of the American Electroplaters' Society*, September 1941.

steel surface, such as results from an alternate oxidation and reduction and by the addition of small amounts of aluminum to the zinc bath.

Electroplated zinc deposits have the advantage of being applicable in greater thicknesses than hot-dipped coatings. Electroplating methods have made considerable progress, particularly in the wire field, with the speeding up of plating rates as much as twenty-five fold.

Bright zinc coatings have been developed in response to the demand for improved appearance and this finish is gradually replacing the older dull type.

The protective value of zinc depends directly upon the thickness of the coating. Experiments have listed environments in the order of increasing attack as follows: rural, tropic marine, temperate marine, suburban, urban and highly industrial. The resistance of zinc coatings to corrosion under water depends largely upon the degree of circulation of the water and its oxygen content. When a submerged zinc-coated armored cable is lapped with jute, thereby stagnating the water, the capacity of zinc to resist corrosion is increased.

Cadmium plate has good color and is very satisfactory for indoor use. It does not possess corrosion resistance equal to zinc under conditions of outdoor exposure. Bright nickel coatings or semi-bright coatings requiring mild buffing have largely replaced the older type of nickel coatings.

A very promising process of protecting steel, known as "Corronizing" consists in the application of a layer of nickel plate followed by either zinc or tin. The duplex coating is heated to 700–1000°F. yielding alloys practically free from pores which show high resistance to the salt spray test.

Seventy per cent of the production of tin plate is used in cans. The hot-dipped process is old and well established but is being challenged by continuous rolling processes involving electroplating methods of application.

Recent progress in the protection of metals by coatings of other metals is largely in the direction of electroplating and continuous processes.

Measurements of the Delay and Direction of Arrival of Echoes from Near-By Short-Wave Transmitters.[3] C. F. Edwards and Karl G. Jansky. Observations on pulses radiated by a high-power beam transmitter operating in the short-wave range show that when the receiver is located within the skip zone, echoes are observed having delays of from 1 to 50 milliseconds. These echoes are the result of scattering and three different types may be recognized, each arising from a different source.

Echoes of the multiple type were found to occur the most frequently and to have many of the characteristics of signals transmitted over long

[3] *Proc. I.R.E.*, June 1941.

distances. Components were observed from regions up to 4000 miles distant. Direction-of-arrival measurements using steerable arrays operating on the musa principle indicate that these multiple echoes are scattered from regions along the transmitted beam. Vertical angle-of-arrival measurements using a musa receiving system indicate that the surface of the earth may be the source of scattering.

Similarities between multiple echoes and southerly deviated waves from European transmitters have been found which indicate that the same phenomena may be responsible for both.

Evolution by Design.[4] REGINALD L. JONES. The evolution of the telephone plant is characterized by planning and invention. This article describes the Bell Telephone Laboratories' method of developing new and improved telephone apparatus by integrating the creative efforts of technicians in various fields of research and engineering. For telephone readers the recent development histories of telephone drop wire, relay, and station receiver are chosen to illustrate the problems encountered. Other apparatus, some of which is shown by the figures, follows a similar development pattern. New materials, improved processes of manufacture, economy in maintenance, and a better understanding of convenient use—all play continuing parts in the forward march of telephone design.

Television Transmission.[5] M. E. STRIEBY and C. L. WEIS. Experiments in the transmission of television signals over wire lines have been made from time to time as the television art has developed. The present paper discusses experiments made during the summer of 1940 with 441-line, 30-frame interlaced signals transmitted over coaxial cable and other telephone facilities. Some of the general problems of wire transmission have been included. In particular, the results of transmission studies on a system linking New York and Philadelphia are reported.

[4] *Bell Telephone Magazine*, August 1941.
[5] *Proc. I.R.E.*, July 1941.

Contributors to this Issue

Julian Blanchard, A.B., Trinity College (now Duke University), 1905; A.M., Columbia University, 1909; Ph.D., 1917. Professor of Engineering, Trinity College, 1909–1912; Research Assistant in Physics, Columbia University, 1912–1915. Physicist, Research Laboratory, Eastman Kodak Company, 1915–1917; Engineering Department, Western Electric Company, 1917–1925; Bell Telephone Laboratories, 1925–. Dr. Blanchard's work has been concerned primarily with special studies in connection with the development of vacuum tubes and radio. More recently he has been attached to the Technical Consulting Staff of the Laboratories.

G. M. Bouton, Ch.E., Brooklyn Polytechnic Institute, 1926. Bell Telephone Laboratories, 1926–. Mr. Bouton is in charge of a group concerned primarily with the metallurgy of lead cable sheath alloys and solders and with general microscopic examination of metallic materials.

W. P. Mason, B.S. in Electrical Engineering, University of Kansas, 1921; M.A., Columbia University, 1924; Ph.D., 1928. Bell Telephone Laboratories, 1921–. Dr. Mason has been engaged in investigations on carrier systems and in work on wave transmission networks both electrical and mechanical. He is now head of the department investigating piezoelectric crystals.

Earle E. Schumacher, B.S., University of Michigan; Research Assistant in Chemistry, 1916–18. Engineering Department, Western Electric Company, 1918–25; Bell Telephone Laboratories, 1925–. As Associate Research Metallurgist, Mr. Schumacher is in charge of a group whose work relates largely to research studies on metals and alloys.

Roger I. Wilkinson, B.S. in E.E., Iowa State College, 1924. Northwestern Bell Telephone Company, 1920–1921; American Telephone and Telegraph Company, 1924–1934; Bell Telephone Laboratories, Inc., 1934–. Mr. Wilkinson has studied principally the application to telephone problems of the mathematical theory of probability.