BASIC TELEPHONE TRAFFIC THEORY

(This E.I. supersedes and cancels E.I. TELEPHONE Exchanges Automatic T 0080 and T 0100.)

1. INTRODUCTION.

1.1 A most important aspect of the Traffic Engineer's work is to determine the quantities of automatic switching equipment required to carry, adequately and economically, the traffic at all stages in a telephone network.

1.2 This E.I. provides an understanding of the nature of telephone traffic and an introduction to the traffic theory which has been developed as a basis for Engineering Traffic Tables. These tables are used in practice to determine equipment quantities from known traffic data.

2. AVERAGE TELEPHONE TRAFFIC.

2.1 Definition. The telephone traffic carried by a group of circuits, during any specified time, is equal to the average number of simultaneous connections taking place in the group during the period concerned. If in a group of N circuits, the average number simultaneously engaged during an observation period is A, then a traffic of A Erlang is said to be flowing during this period. From this it follows that the total occupancy of a circuit group during a time T is equal to AT.

2.2 Units. The internationally accepted unit of telephone traffic is the Erlang (E). Other units of traffic measurement in use are :-

Unit	Explantion	Value in Erlangs
T.U.	Traffic Unit	1.0
C.C.S.	100 Call Seconds	0.0277
Speech Period	Three minutes of conversation.	0.05

3. OCCUPANCY OF A CIRCUIT GROUP.

3.1 From the definition above :-

$$A = \frac{\text{Total occupancy during T}}{T}$$

$$= \frac{t_1 + t_2 + t_3 + t_4 + \dots + t_x + \dots t_n}{T} \qquad (1)$$

where t_1, t_2 ... t_n are the holding times of the 1st, 2nd ... and Nth calls carried by the circuit group, and T is the time considered. The holding-times of the individual calls must be measured in the same units as the total time, T.

4. AVERAGE HOLDING TIME.

4.1 If n calls of various lengths are carried by a number of circuits during time T, the average duration of all calls is given by -

$$h = \frac{t_1 + t_2 + \dots\dots\dots\dots + t_n}{n} \qquad (2)$$

This average holding-time may be expressed as a fraction of the total time T, i.e.

$$H = \frac{h}{T} = \frac{t_1 + t_2 + \dots\dots\dots\dots + t_n}{nT} \qquad (3)$$

From equations (1) and (3) we obtain :-

$$H = \frac{A}{n} \qquad \text{or} \qquad A = nH \qquad (4)$$

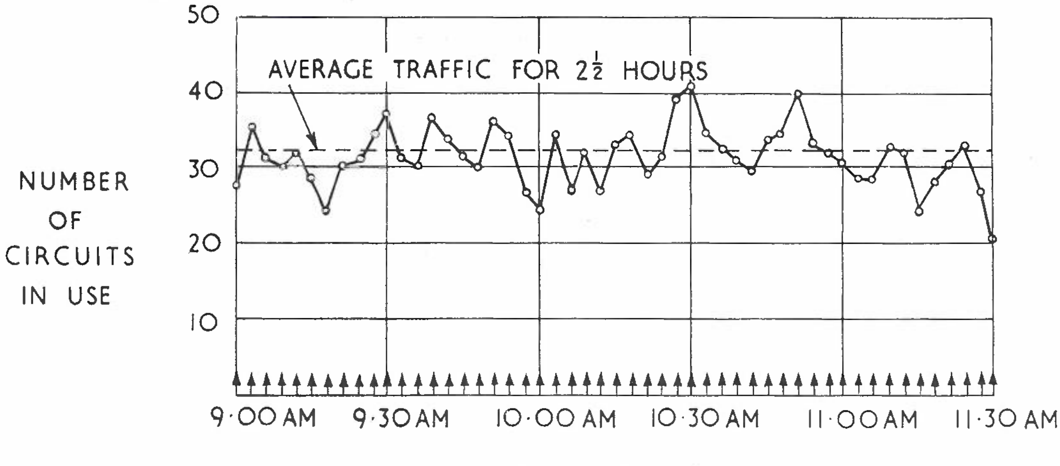
Equation (4) is an important relationship which states that the average traffic in a given time is equal to the product of the number of calls occurring in that time and the average holding-time of these calls expressed as a fraction of that time.

Usually T is taken as one hour, in which case, H is expressed in hours and n is the number of calls per hour.

Another important relationship is that the average traffic is the number of calls occurring in one holding time.

5. MEASUREMENT OF TELEPHONE TRAFFIC.

5.1 The normal method of telephone traffic measurement is to observe the number of occupied circuits at regular intervals during the period of measurement. Fig. 1 is the graph of a typical result obtained by this method, in which samples were taken at 3 minute intervals.



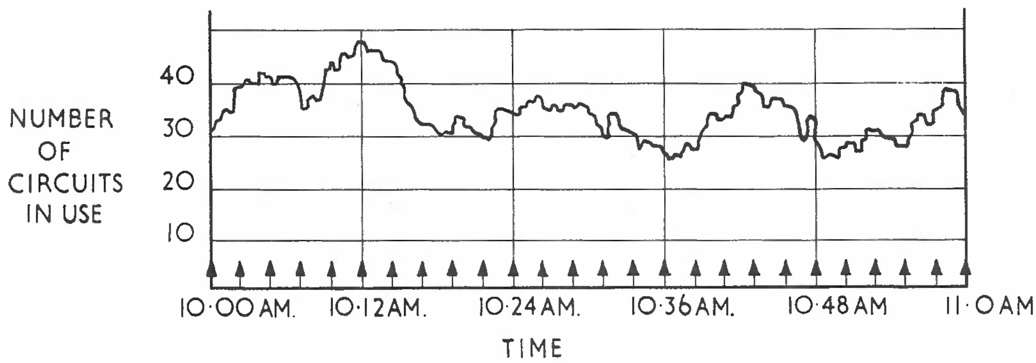
CIRCUITS IN USE AT 3 MINUTE INTERVALS.

FIG. 1.

The graph obtained from this sampling process is a simplified picture of the traffic flow. The average of the same readings is a good estimate of the traffic if the intervals between the successive observations are sufficiently small.

6. NATURE OF TELEPHONE TRAFFIC.

6.1 Short Term Fluctuations. Fig. 2 is a reproduction of a typical continuous pen recording of the traffic flow over a group of circuits during one hour.

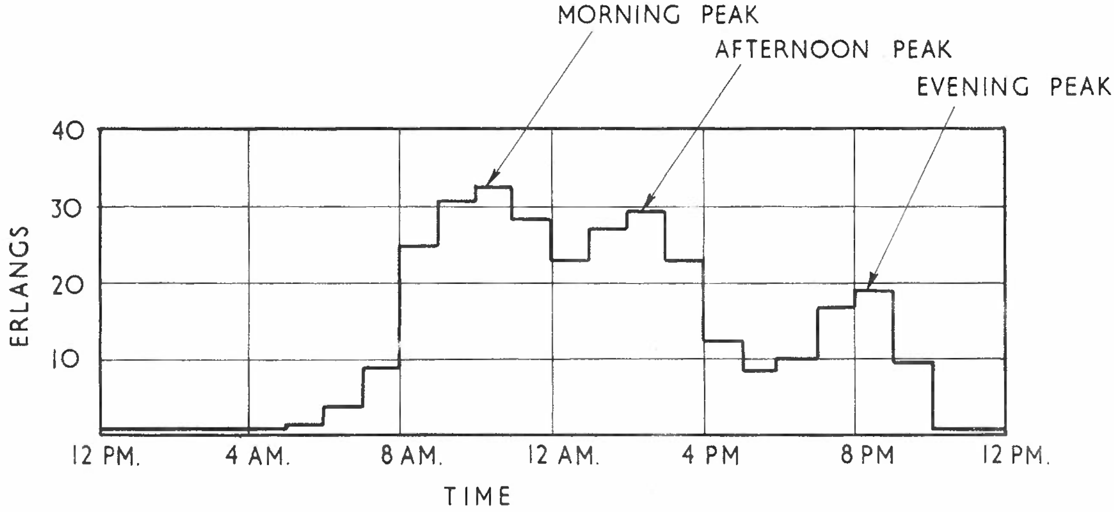


TRAFFIC FLOW FOR ONE HOUR.

FIG. 2.

6.2 <u>Hourly Variation</u>. Fig. 3 is a graph of the average hourly traffic passing a group of circuits for each hour of a 24 hour period.

Traffic measurements show that in large exchanges the busiest hour usually occurs during the same period of each day. This graph is typical of an industrial area whereas in the case of a residential area the evening peak is more marked.

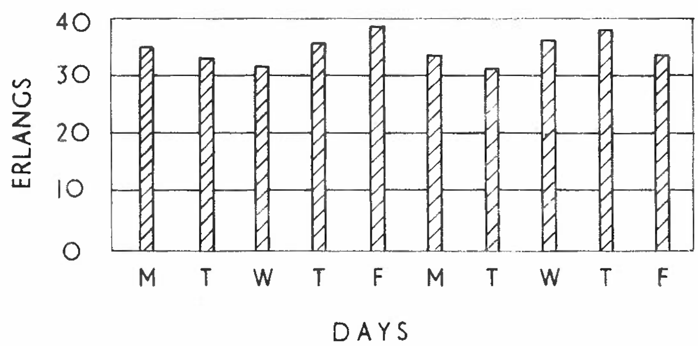


AVERAGE HOURLY TRAFFIC FLOW FOR PERIOD OF 24 HOURS.

FIG. 3.

6.3 <u>Daily Variations</u>. The daily busy hour traffic flowing over a particular circuit group usually varies from day to day. Fig. 4 is a typical case for 5 working days of two weeks.

Traffic measurements are made over a period, usually the five working days of the week, in order to obtain a representative estimate of the average traffic.



DAILY BUSY HOUR TRAFFIC FOR PERIOD OF 2 WEEKS.

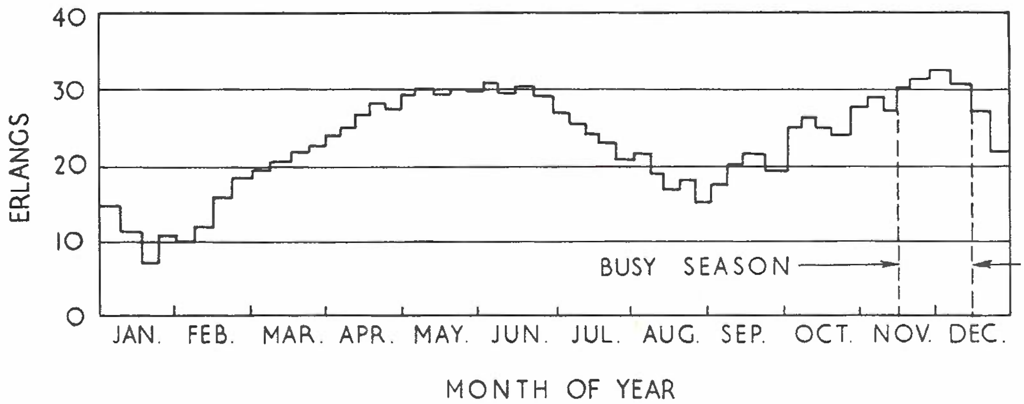
FIG. 4.

6.4 Busy Hour. The busy hour of the traffic carried by a group of circuits on a number of measuring days is obtained as follows:-

(i) Determine from the traffic readings the average traffic for each half hour of the measuring period of each day.
(ii) The average traffic carried during a particular half hour on each measuring day, e.g. 9.00 a.m. to 9.30 a.m., is then averaged for all the measuring days.
(iii) Select busy-hour by choosing the two consecutive half hours in which the sum of the two averages obtained from (ii) above is largest.
This method of busy hour selection is often referred to as Time-Consistent.

6.5 Seasonal Variations in Telephone Traffic. Fig. 5 is a typical graph of the "busy hour" traffic for each week of the year.

In this example there is a busy season of four weeks in November-December. This type of seasonal behaviour is common in telephone traffic. Busy season demands are met by designing equipment quantities to carry the average "busy hour" traffic during the busiest four consecutive weeks of the year.



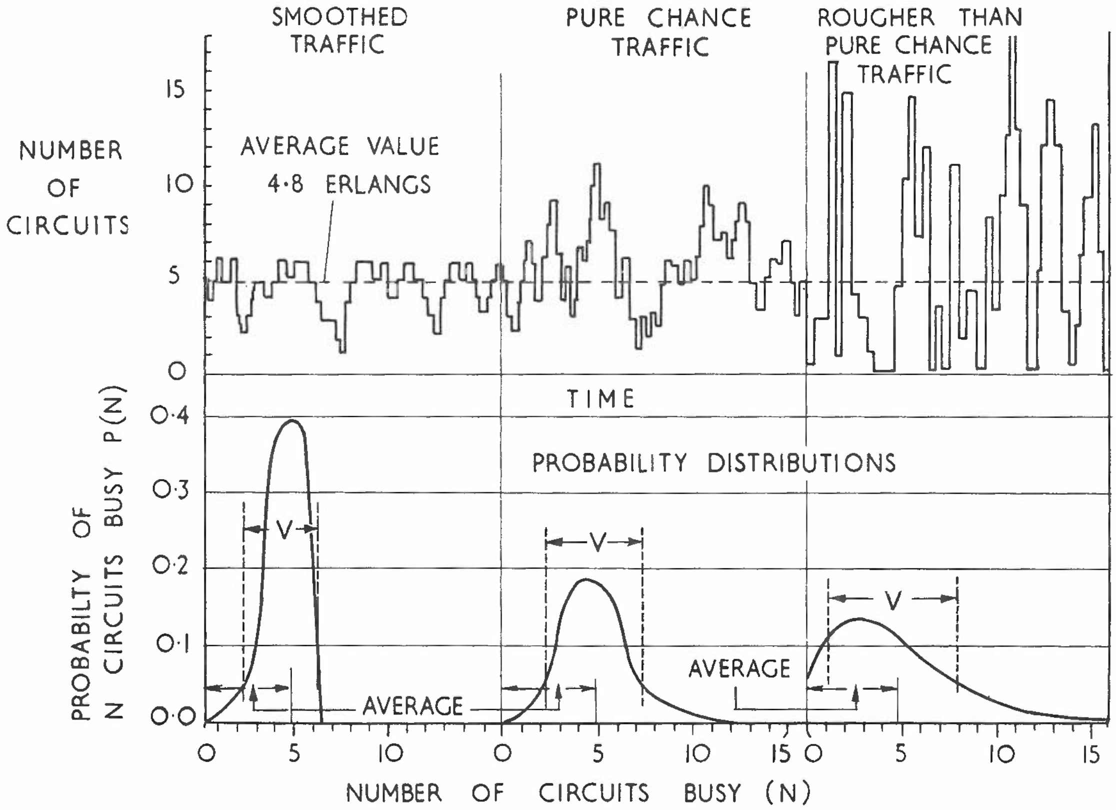
BUSY HOUR TRAFFIC FOR EACH WEEK OF YEAR.

FIG. 5.

6.6 Character of Telephone Traffic. Random traffic fluctuations may be represented by means of a probability diagram. This diagram indicates the probability of finding one, two, three or N circuits in the group engaged. These probabilites are equal numerically to the proportion of the time that one, two, three or N of the circuits are engaged. The number of circuits corresponding to the mean value of the probability distribution is equal to the magnitude of the average traffic flowing. The amount of spread of the probability curve about this mean value is a measure of the degree of roughness of the traffic.

The statistical term Variance is used as a measure of the magnitude of this spread. (Reference may be made to any text book on statistics.) The Variance of a distribution $V = \sigma^2$ where σ = the standard deviation for the distribution. If the traffic is pure chance, the variance is equal numerically to the mean (or average traffic value). If the traffic is smoother than pure chance the variance is less than the mean and if the traffic is rougher than pure chance the variance is greater than the mean.

Figure 6 indicates the instantaneous fluctuations and the corresponding probability diagram for these three classes of traffic.



THREE TYPES OF TELEPHONE TRAFFIC.

FIG. 6.

Pure Chance Traffic. Traffic which is approximately Pure Chance is generated by a large number of independent subscribers. In order to obtain this type of traffic two conditions must be satisfied:-

(i) The probability of a further call arising should not be affected by the number of calls already in progress. This implies that a very large number of subscribers generate the traffic.

(ii) The individual subscribers make their calls independently of the calls of the other subscribers. That is, a subscriber is as likely to make a call at any one time as at any other.

In practice groups of subscribers larger than 200 generate traffic of this sort.

Traffic Smoother than Pure Chance. Traffic of this type occurs whenever the probability of a further call being handled falls as the number of calls in progress increases. This type of traffic occurs on the early choices of a grading and on the direct or first choice routes where alternate routing is used.

Traffic Rougher than Pure Chance. Rough traffic occurs whenever the probability of a further call being handled is increased as the number of calls in progress increases. This type of traffic occurs on the late choice outlets of a grading or as the traffic overflowing from direct routes to an alternate route.

7. BASIS OF CIRCUIT PROVISION.

7.1 The general basis of circuit provision is a compromise between the standard of service to the subscribers and the cost of providing circuits. It is uneconomic to provide sufficient circuits to meet subscribers demands for service at all times. On the other hand, the provision of too few circuits will produce serious subscriber inconvenience due to the inability to obtain service on demand. A solution between these conflicting requirements is achieved by presenting a probability that a subscriber will fail to obtain a free circuit when attempting to make a call. In the case of a busy signal system this probability is termed "probability of loss" or "grade of service". In the delay system it is termed the "probability of delay".

8. TYPES OF TELEPHONE SYSTEMS.

8.1 If a subscriber attempts to make a call when all the circuits are busy, one of two things may happen, depending on which type of telephone system is in use at the particular switching stage:-

Busy Signal Systems. In this system, busy-tone is connected to the calling subscriber's line if all circuits are busy and the attempted call is said to be lost; the subscriber must call again in order to establish the connection. Typical examples are uniselectors and group selectors.

Delay Systems. This type of system is one in which those calls which find all circuits busy are stored until a circuit becomes available. Typical examples of Delay Systems are linefinders and call queues.

Combined Systems. In some networks both Busy Signal and Delay Systems are used at different switching stages, such as linefinders (delay) trunk to first selectors (busy-signal).

9. BUSY SIGNAL SYSTEMS.

9.1 Grade of Service. In Busy Signal Systems, the grade of service which is sometimes also known as the Call Congestion is the proportion of call attempts which are lost, that is $\frac{\text{Calls Lost}}{\text{Total Calls Offered}}$.

Typical Busy Hour grades of service which are used in local networks during the busy season of the year are 0.01 (1 in 100), 0.005 (1 in 200) and 0.002 (1 in 500).

The following terms are sometimes used instead of Grade of Service:-

(i) Time Congestion, equals $\frac{\text{Time all circuits are busy}}{\text{Total time}}$.

(ii) Traffic Congestion, equals $\frac{\text{Traffic Overflowing}}{\text{Traffic Offered}}$.

For smooth traffic such as that which flows from a well balanced and mixed 1st selector grading, less circuits are required to give the same grade of service than if the traffic were pure chance.

For rough traffic, such as that which is offered to a backbone route from a desired route, more circuits are required.

9.2 Traffic Theories. Theories have been developed to determine the relationship between the Grade of Service (B), the number of circuits (N) and the average traffic (A) for a particular type of trunking.

In these relationships, A is the average traffic offered to the group of circuits, not the average traffic carried by the group. However, for good grades of service, the traffic offered and the traffic carried differ only slightly. Several relationships have been derived for Grade of Service which use different assumptions regarding the way in which calls arise and are dealt with by the switching system.

Full Availability Trunking. This is a method of trunking wherein all the calling subscribers have access to all of the circuits provided.

Erlang's Formula. The most commonly used relationship for grade of service with full availability trunking is that due to A.K. Erlang (Copenhagen Telephone Co., 1907-29). This is given by the following expression -

$$B = \frac{\frac{A^N}{N!}}{1 + A + \frac{A^2}{2!} + \dots\dots\dots\dots + \frac{A^N}{N!}}, \qquad (5)$$

Where A is the average traffic offered in erlang,

N is the number of circuits provided, and

$$N! = N.(N-1). \quad (N-2) \dots\dots\dots\dots 3.2.1.$$

Erlang showed also that the probability of exactly X out of N circuits being engaged was -

$$P(X,N) = \frac{\frac{A^X}{X!}}{1 + A + \dots\dots\dots\dots \frac{A^N}{N!}}. \qquad (6)$$

For the special case of X = N,

$P(X,N) = P(N,N) = B$

= Probability of all circuits being busy

= Probability of loss as shown in formula 5.

Assumptions. This theory applies if the following assumptions are valid :-

(a) The individual subscribers are each as likely to orginate a call at one instant as any other during the busy hour.

(b) The probability of a further call arising is independent of the number of calls already in progress.

(c) Lost calls are cleared from the system. This means that subscribers who seek to make a call when all the circuits are busy, abandon the call; that is, they do not attempt to gain service by making repeated attempts until a circuit becomes available.

(d) The traffic "A" is the true average traffic offered to the group of circuits.

(e) Conditions of statistical equilibrium obtain, that is, the average traffic is not changing with time.

Validity of the Assumptions. In practice these assumptions are good approximations to reality but several of them break down under special conditions.

Assumption (a) is true unless the individual subscribers commence to make bursts of calls in rapid succession. This only happens when the Grade of Service is poor and the bursts of calls are abortive attempts to obtain service.

Assumption (b) is true if the number of calls in progress is always very much smaller than the number of subscribers producing the traffic. Theoretically this assumption is only true if the number of subscribers is infinite, but in practice any number in excess of 200 is satisfactory.

Assumption (c) is frequently questioned since, in practice, a subscriber who encounters congestion makes a repeat attempt fairly promptly. This mode of behaviour, whilst common, does not lead to serious errors unless the grade of service is bad. Even in this case the input traffic is very nearly Pure Chance although its value is inflated by the additional call attempts. If the subscribers concerned wait between repeat attempts, the input traffic will not be greatly inflated and the repeats can be regarded as pure chance calls. When this assumption breaks down completely, an alternative theory of "Lost Calls Held" is used.

Assumption (d) is correct if traffic measurements have been made for a sufficient number of busy hours. In most cases readings made for five busy hours are sufficiently accurate for practical purposes.

Assumption (e) is correct because the average traffic does not change signifioantly during the busy hour.

Average Duration of Congestion. Congestion occurs when all the N circuits available to a calling source are busy. After a period H, where H is the average holding time of all calls, these N calls will have been replaced, on the average, by A new calls, where A is the average offered traffic. Hence the average rate at which the congestion decreases is N - A calls per holding time, therefore, the average time taken for one circuit to become free and thus eliminate the congestion is, $\frac{H}{N-A}$. This, then is the average duration of congestion.

The Poisson Distribution. As the Grade of Service on a group of circuits improves, the ratio of overflow traffic to offered traffic becomes progressively smaller. Referring to equation (5), $\frac{A^N}{N!}$ becomes very small in relation to the sum -

$$\left[1 + A + \frac{A^2}{2!} + \dots\dots + \frac{A^N}{N!}\right] \text{which can be written} \sum_{\ell=0}^{N} \frac{A^\ell}{\ell!}$$

Also the terms $\left[\frac{A^\ell}{\ell}\right]$ between $\ell = N$ and $\ell = \infty$ are very small compared with the terms between $\ell = 0$ and $\ell = N$ and the sum to N terms approaches the sum to infinity. That is,

$$\sum_{\ell=0}^{\ell=N} \frac{A^\ell}{\ell} \longrightarrow \sum_{\ell=0}^{\ell=\infty} \frac{A^\ell}{\ell!}$$

However, $$\sum_{\ell=0}^{\ell=\infty} \frac{A^{\ell}}{\ell!} = e^{A}$$

where e is the base of natural logarithms, and thus for good grade of service the Erlang probability

$$P(X,N) = \frac{\frac{A^X}{X!}}{1 + A + \dots \frac{A^N}{N!}}$$

approaches $P(X) = \frac{A^X}{X!} e^{-A.}$ (6A)

Equation 6A is known as the Poisson distribution. This is particularly valuable since the Poisson distribution has many other applications in statistics and has been extensively tabulated for values of A and X which are beyond the range of the available tabulations of equations 5 and 6.

Pure Chance Traffic - Table A. The Pure Chance Traffic capacity table used for full availability circuit groups is derived from Erlang's solution given by Equation 5.

9.3 Traffic Carried by a Group of Circuits. The traffic carried by a group of circuits can be obtained from Erlang's equation if the traffic offered is known.

Traffic carried = Traffic Offered - Traffic Overflowing.

$$= A - AB$$

$$= A\ (1-B) \qquad (7)$$

Where B is obtained from Equation 5.

This relationship holds good irrespective of the method of search over the group.

Figs. 7 and 8 (attached) give the traffic passed from groups of full availability circuits when offered a known Pure Chance traffic in Erlang.

These curves may be used for the determination of grades of service since the traffic passed from the group equals AB. They may also be used to calculate the traffic capacities of grading.

Example.

A group of 8 circuits is offered a traffic of 7 erlang, how much traffic is carried by the group?

Fig. 7 shows that the overflow traffic equals 1.25 erlang. Hence the group carries 7 - 1.25 = 5.75 erlang.

9.4 Traffic Carried by Individual Circuits. The traffic which, on the average, will be carried by a particular circuit can be calculated if the method of search over the group to which it belongs is known. Two methods of search are used with Australian telephone equipment.

(i) Non Homing. With this method the search for the first free circuit in the group, that is, outlet, starts from the point in the bank where the selector was last positioned. In these groups each individual circuit carries the same traffic (a) on the average, where -

$$a = \frac{A\ (1-B)}{N} \tag{8}$$

(ii) Homing. With this method the search for the first free outlet always starts from the same point. The traffic carried by the first circuit, that is, outlet, can be obtained as follows:-

When A erlangs are offered to the first circuit in a group the probability of loss or overflow is given by using Equation (7) in the form for N = 1, thus:-

Probality of loss $= B = \frac{A}{1 + A}$.

Hence the traffic passed from the first outlet -

$$a = \left[\frac{A}{1 + A}\right] A$$

Hence, if the traffic carried by the first outlet is a_1, -

$$a_1 = A - \frac{A^2}{1 + A}$$

$$= \frac{A}{1 + A}.$$

The traffic carried by the second outlet may be obtained similarly as -

$$a_2 = \frac{A^2}{1 + A} - \left[\frac{\frac{A^2}{2!}}{1 + A + \frac{A^2}{2!}}\right] \times A.$$

The same method of calculation can be used to determine the traffic carried by each subsequent outlet in a group offered traffic by homing selectors.

(iii) Graph of Traffic Carried by Each Choice. Figs. 9 and 10 (attached) have been produced by the methods described in the previous paragraphs. These curves give the traffic carried by each choice of a full availability group when offered a known total traffic. For example, if 5 erlang are offered to a full availability group then by referring to Fig. 9 it will be seen that:-

The first	outlet	carries	0.83E
The second	"	"	0.79E
The third	"	"	0.73E
The fourth	"	"	0.66E

and so on.

(iv) <u>The Traffic Carried by the Last Choice in a Full Availability</u> Group may be obtained by the method given above but the following approximate formula is also useful.

Last Trunk Traffic = B (N-A)

where B is the Grade of Service,

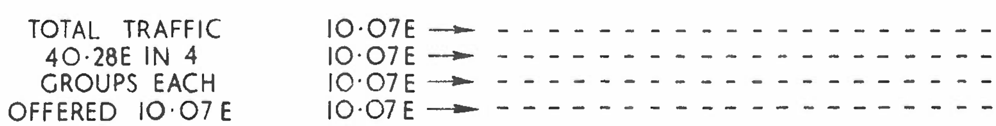
N is the number of trunks,

and A is the offered traffic in erlang.

<u>Note:</u> This formula applies only to groups with sequential hunting.

9.5 <u>Limited availability of Trunking</u>. The single full availability group is the most efficient form of trunking since congestion can only occur when every circuit on a route is busy. However, full availability trunking schemes are usually uneconomic for large traffics because selectors designed to hunt over large numbers of outlets are prohibitively costly.

Selectors are usually designed to hunt over 10 or 20 outlets per level and a form of limited availability is employed. The simplest form of limited availability trunking for both homing and non homing selectors consists of a number of separate full availability groups, for example, Fig. 11.

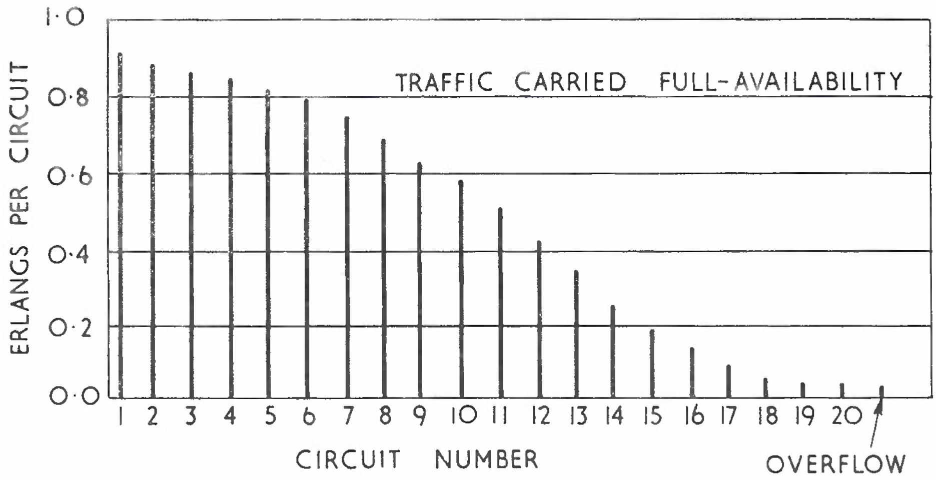


<u>LIMITED AVAILABILITY TRUNKING.</u>

<u>FIG. 11.</u>

This particular arrangement contains 80 trunks and will carry 40.28 erlang at a grade of service of 0.002. The efficiency or average traffic carried per trunk, is 0.504. This is low since a traffic peak in any one group cannot be absorbed by free circuits in the other groups. A single full availability group of 58 circuits would carry this traffic at the same grade of service with an efficiency of 0.695E per trunk.

Fig. 12 shows the average traffic by each circuit in each of the four groups of Fig. 11 when homing selectors are used. It can be seen that the later choices carry very little traffic.



AVERAGE TRAFFIC BY EACH CIRCUIT IN ANY ONE OF 4 GROUPS.

FIG. 12.

Interconnecting or Trunk Sharing Schemes. A number of special trunking schemes have been devised from time to time in an attempt to retain the economic advantages of limited selector availability and yet avoid the inherently limited efficiency of arrangements of the type shown in Fig. 11. These interconnecting schemes all have a basic characteristic. They all represent practical attempts to minimise the number of circuits which can be free in a group of trunks when one of the feeds is congested.

All interconnecting schemes are started by trunking a number of groups or splits away from the switching stage as is done when separate full availability groups are to be used. The first basic difference is that at least twice as many groups are formed. The second is that the trunks are distributed in the form of multiple appearances between the groups or splits to form a pattern of interconnections such that each individual circuit appears once in the availability of each of two or three or more, splits. Thus each split or group has access to a number of trunks on the route equal to its availability and no two splits have access to exactly the same group of trunks although they may have access to several in common.

All interconnecting schemes have the above characteristics and differ only in the exact form of the connection patterns which are used. One of the most frequently used methods of interconnecting is the grading which is only used for homing selectors. Very different patterns are used for non homing selectors. For these, see the special section on Interconnecting Schemes for Crossbar Exchanges.

Grading of Outlets. Fig. 13 shows a typical grading scheme. This type of arrangement gives a considerable improvement in efficiency by combining suitable numbers of the later choices between the groups without changing the grade of service. The commoning of the later choices which is characteristic of the grading patterns produces an improvement in efficiency in two ways which may be illustrated by reference to Fig. 13.

(i) The commoning between the grading groups increases the total traffic offered to that particular outlet, for example, in the case of the first pair it is doubled. A study of Figs. 9 and 10 will show that the traffic carried by an outlet increases with the traffic offered to it.

For example -

Traffic offered to a 10th Trunk	1.38	2.05	2.70	3.50
Traffic carried by the 10th Trunk	0.42	0.51	0.58	0.65

Thus the last eleven outlets from the grading of Fig. 13 each carry more traffic than each of the later outlets of the separate full availability groups of Fig. 11.

(ii) Not only do the later outlets each carry more traffic, but there are fewer of them. This factor is responsible for much of the gain in efficiency.

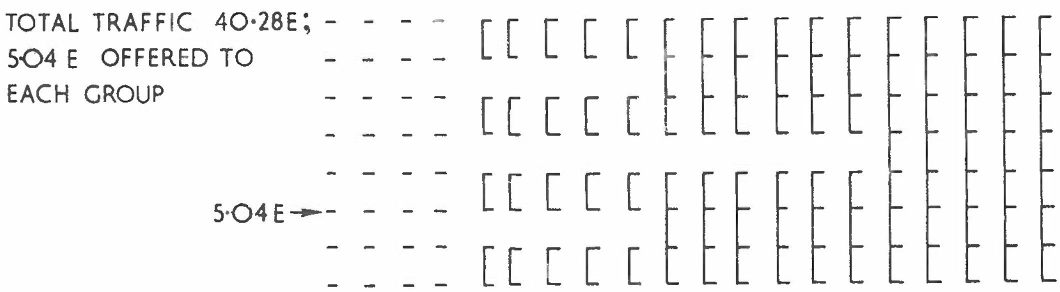


FIG. 13. TYPICAL GRADING SCHEME.

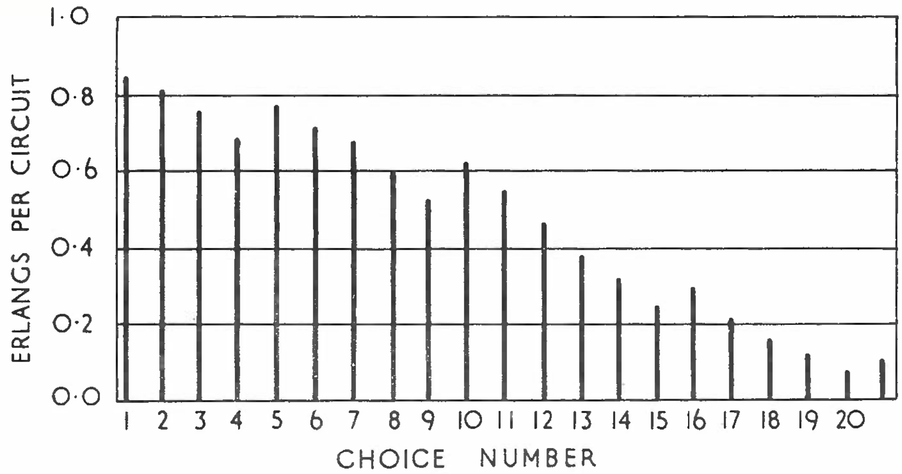
As an alternative explanation, commoning the later outlets between a number of previously separate full availability groups requires that the traffic offered to each group be reduced if the original grade of service is to be preserved. Hence, if a grading is to be constructed to handle the 40.28 erlang offered to the 4 full availability groups shown in Fig. 11 at the same grade of service it must have more than 4 grading groups. Usually, the required number of grading groups is at least twice the number of separate full availability groups required to handle the same traffic, except where the number of necessary trunks is not much greater than the availability, when the increase in grading groups may be less.

However, it is always found that, in the transition from separate full availability groups to a grading of the same traffic capacity the average number of trunks per grading group can be reduced proportionally faster than the rate at which the number of groups must be increased. Thus the total number of circuits is reduced resulting in a gain in efficiency.

It can be shown that for a given number of grading groups and grade of service the efficiency rises when the number of trunks from the grading is increased and reaches a peak after which it begins to drop, since the grading is then tending to revert to separate full availability groups. Generally, this peak in efficiency occurs when the number of trunks is between $\frac{1}{4}$ and $\frac{3}{4}$ the total number possible, i.e. $\frac{1}{4}$ to $\frac{3}{4}$ the product of the availability and the number of grading groups.

In addition, it has been found that when maximum efficiency has been reached in a grading with a fixed number of groups, the efficiency could be further raised with the same number of trunks by slightly increasing the number of grading groups. Therefore, given the availability, grade of service, and the number of trunks (the usual starting point in grading design, a particular number of grading groups can be found beyond which no further gain in efficiency will be obtained. In gradings with more than 6 groups the maximum efficiency is reached when the number of trunks is approximately one third the product of availability and the number of grading groups.

The traffic carried by each choice of the grading given in Fig. 13 is shown in Fig. 14.



TRAFFIC CARRIED BY EACH CIRCUIT IN THE GRADING IN FIG. 13.

FIG. 14.

The efficiency or average traffic carried per circuit of this arrangement is -

$$\frac{40.28}{69} = 0.584$$

which is appreciably better than the value of 0.503 obtained for the first arrangment of separate full availability groups shown in Fig. 11. Grading is an efficient system for limited availability trunking and is the usual A.P.O. method for interconnecting homing selectors.

9.6 Grading Theory. The grading theory underlying the British and Australian Post Office Traffic Tables for graded groups is due to G. F. O'Dell. ("The Influence of Traffic of Automatic Exchange Design", G. F. O'Dell, B.Sc., M.I.E.E., I.P.O.E.E. Paper No. 85, 1920.) This theory is developed from a simple approximate relationship due to A. K. Erlang for large interconnecting schemes of the pattern which Erlang believed to be "ideal", and good grades of service. Erlang derived his relationship in the following way:-

If N circuits are offered a pure chance traffic of A Erlang and the grade of service is so good that the traffic overflowing is small, the average carried per circuit will be $\frac{A}{N}$ approximately.

The probability that a hunting selector will find a particular circuit busy is numerically equal to the traffic which the circuit carries since this is equal to the proportion of time during which it is busy. Hence, on the average, the probability that a hunting selector will find a particular x circuits busy is $\frac{A^x}{N}$. However, if x equals the availability K, then the call will be lost. Thus,

$$B = \left[\frac{A}{N}\right]^K \tag{8}$$

Where A equals the traffic offered,
N equals the number of circuits,
K equals the availability of the hunting selector.

From this, he deduced that the traffic capacity of large gradings was given by,

$$A = N \left[B\right]^{1/K} \qquad (9)$$

O'Dell Modification. O'Dell modified this theory to take account of its inaccuracies when applied to the small gradings. He assumed that the first K circuits had a traffic capacity equal to that of a full availability group of K circuits, at the specified grade of service. He also assumed that each additional circuit added a traffic capacity of $\left[B\right]^{1/K}$ erlang to the total. Hence, O'Dell arrived at the relationship

$$A = A_{FA/K} + \left[B\right]^{1/K} (N-K) \qquad (10)$$

Where $A_{FA/K}$ is the traffic capacity of a full availability group of K circuits at the specified grade of service and the other symbols are defined as before.

Empirical Modification. Practical measurements of the traffic capacities of gradings with Pure Chance traffic inputs indicated that the actual increase in efficiency with each individual circuit was only 53% of the expected increase, with respect to full availability efficiency. However, these experiments did establish that the gain in traffic capacity for each additional circuit was very nearly constant and independent of the number of circuits already present. This led to the following empirical relationship for the traffic capacity of gradings offered Pure Chance traffic.

$$A = A_{FA/K} + \left[0.53 \left[B\right]^{1/K} + 0.47 \frac{A_{FA/K}}{K}\right] (N - K)$$

$$= \left[0.53 \left[B\right]^{1/K} + 0.47 \frac{A_{FA/K}}{K}\right] N - \left[0.53 \left[B\right]^{1/K} K - 0.53 A_{FA/K}\right] \qquad (11)$$

where B = Grade of Service,
A = Traffic Offered,
N = Number of Trunks,
K = Availability,

$A_{FA/K}$ = Traffic offered to a full availability group of K circuits for a grade of service equal to B.

A slightly more detailed account of this theory is given in "Traffic and Trunking Principles in Automatic Telephony" by G. S. Berkley whilst the original paper on the subject is "An outline of the Trunking Aspects in Automatic Telephony" by G. F. O'Dell, P.O.E.E.J., Vol. 65, No. 362, February, 1925.

The equations are the basis for the following traffic tables.

9.7 Smoothed Traffic Table - Table B. Equation (10) is the basis of the "B" or Smoothed Traffic Table as experiments have shown that it is a reasonable approximation for traffic which has been partially smoothed through one or more preceding gradings. The following equations give the traffic capacity A for a group of N circuits when offered Smooth Traffic if the grade of service B is 0.002.

A = 0.5375 (N) - 1.942.	Availability 10
A = 0.7329 (N) - 4.588.	" 20
A = 0.7632 (N) - 5.290.	" 23

Note: These equations only apply for numbers of circuits greater than 70 and make no provision that the grade of service shall not fall below 0.01 if an overload of 10% occurs. The standard (B = 0.002) table does make this provision.

9.8 Pure Chance Traffic Table - Table C. Equation (11) is the basis of the "C" or Pure Chance Traffic Table for gradings. In practice, the table slightly underestimates the traffic capacity as a safeguard against defects in grading construction and unbalanced traffic inputs. The following equations give the traffic capacity A for a group of N circuits when offered Pure Chance Traffic if the grade of service B is 0.002.

A = 0.4459 (N) - 1.029.	Availability 10
A = 0.6251 (N) - 2.432.	" 20
A = 0.6551 (N) - 2.804.	" 23

Note: These equations make no allowance to prevent the grade of service falling below 0.01 for a 10% traffic overload. The standard (B = 0.002) table does include this allowance.

9.9 Traffic Table BC. Some years after O'Dell produced the B and C tables, C. Palm published the following congestion formula -

$$B = \frac{E_{1,\,N}\,(A)}{E_{1,\,N-K}\,(A)} \qquad (12)$$

where

B = Grade of Service, K = Availability

$E_{1,\,N}\,(A)$ = Erlang loss probability for N circuits offered a traffic of A Erlangs, and

$E_{1,\,N-K}\,(A)$ is similarly defined for N - K circuits offered the same traffic.

The formula was derived from Homogeneous Interconnecting Schemes with very large number of splits. In practice, a well constructed grading with a large number of groups will achieve efficiencies near to those predicted by the formula if the gradings between the preceding switching stages are also well constructed and good allocation schemes are used. The formula has been used as the basis of traffic table BC. It applies for good grades of service - 0.002 or better.

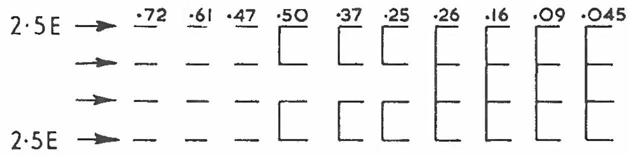
9.10 Traffic Carried by Each Trunk in a Grading. Erlang's theory was used to calculate the traffic carried by the individual trunks in a full availability group. This theory can be extended to gradings by the following assumption regarding the nature of the overflow traffic.

If the traffic overflows from two or more grading groups to a common trunk, the character of the traffic offered to the common depends only on the number of the choice from which it overflowed, and not on the number of groups which feed the common.

Hence, any common trunk behaves as if it were the same numbered choice in a full availability group with a new input such that the same amount of traffic is passed to the common.

An example is given below:-

Consider the four group arrangement shown in Fig. 15, that is, 22 circuits graded on an availability of 10. It is assumed that 10 erlang are offered to the grading with 2.5 erlang offered to each group.



22 CIRCUITS GRADED ON AN AVAILABILITY OF 10.

FIG. 15.

The traffic carried by the first, second and third choice trunks can be read from Fig. 9 directly. The traffic offered to the fourth choice is twice that overflowing from the third choice, that is, 2 x 0.706 = 1.412 erlang, which, from Fig. 7, equals the traffic offered to the fourth choice of a full availability group with a pure chance input traffic of 3.51 erlang. Fig. 9 can now be used to obtain the traffic carried by the fourth, fifth and sixth choices by reading the traffics carried for an input of 3.51 erlang. Fig. 7 gives the overflow traffic for the sixth choice of a full availability group offered 3.51 erlang as 0.29 erlang. Hence, 0.580 erlang are offered to the seventh choice. This traffic would be offered to the seventh choice of a single full availability group with an input traffic of 4.27 erlang.

Fig. 9 can now be used with this input traffic to obtain the traffic carried by the seventh, eighth, ninth and tenth choices.

The traffic carried by all the trunks in the grading has now been determined.

Fig. 7 may be used for an equivalent input of 4.27 erlang to obtain the traffic overflowing from the tenth choice as 0.033 erlang.

The significance of this figure is that this lost traffic may be divided by the total input traffic to obtain the grade of service. In this case, the grade of service is $0.033 \div 10 = 0.0033$.

Experiments with traffic machines have confirmed the practical accuracy of this method by offering artificial traffic to simulated gradings.

10. DELAY SYSTEMS.

10.1 Introduction. In delay systems those calls which arise when all the circuits are busy are stored until a circuit or operator becomes available and the call can proceed. If the calls are handled in the order in which they arrive, this is known as "Queueing". If the calls are handled in groups and these groups comprise those calls which arrive during a particular period of time, this is known as "Gating" or "Corralling".

Delay working may be used to effect economies in the provision of groups of expensive circuits where costs dictate that the traffic carrying efficiency of the circuits must be increased at the expense of some subscriber inconvenience since the usual method for achieving a high efficiency with a small group of circuits is to work with a high probability of finding all circuits busy.

Where a group of circuits has a high time congestion, the subscriber finds a delay system preferable to a busy signal system since, once he has entered his call into the system no further call attempt is needed. From the individual subscribers point of view, call queueing is preferable since the probability of a very long delay which would occur if he were passed over, that is, missed his turn several times, is greatly reduced.

If queueing cannot be arranged, gating is still preferable to unqueued operation since the calls in a particular group or "corral" must be handled before those in subsequent groups. This method ensures that the high probabilities of very long or very short delays inherent to unqueued operation may be considerably reduced and may in some cases nearly achieve the values for queued operation.

Delay systems in which certain particular subscribers or lines always have preference over all others are to be avoided or if used must operate with a very low average delay on calls delayed. A good example is a bimotional linefinder system without a slipped bank multiple where the lower numbered subscribers always take precedence over the higher numbered subscribers.

10.2 Performance Standards. The service provided to the calling subscriber by a delay system may be measured in a number of ways. The service standard normally employed is the proportion of all calls which are delayed by more than some specified time. The simplest of these standards is the probability of delay, which equals the proportion of time during which all the circuits provided are busy under delay working conditions. It does not equal the Erlang probability of loss.

Alternative standards of performance for delay systems are:-

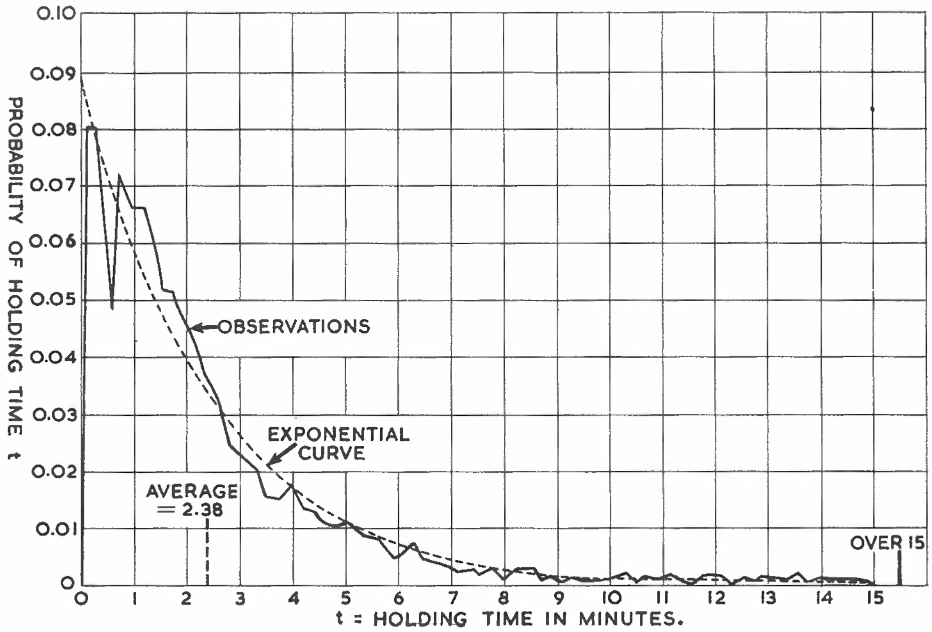
(i) The average delay on those calls which are delayed, and

(ii) The average delay on all calls. The use of this standard enables the total delay to be computed by addition for successive stages.

10.3 Distribution of Holding Times. The service standards mentioned above are affected by the holding time characteristic of the calls handled. Generally, either of two model holding time distributions is taken for the purposes of analysis. In practice, these two models cover those cases likely to be encountered sufficiently accurately. The models are:-

(i) Constant Holding Time. In this model, all calls have a duration equal to the average holding time, that is $P(t)$ is equal to unity for $t = H$ and zero for all other values of t.

(ii) Exponential Holding Times. The holding time distribution of untimed calls controlled by subscribers has been found to be a close approximation to this model. The dotted exponential curve shown in Fig. 16 may be seen to be in good agreement with the actual distribution found for a sample of 7385 local calls.



DISTRIBUTION OF 7385 LOCAL HOLDING TIMES.

FIG. 16.

The exponential curve depicts the probability that a call will have a holding time t; its equation is :-

$$P(t) = \frac{1}{H} e^{-t/H}$$

where e = base of natural logarithms,

H = the average holding time.

The great utility of this particular holding time distribution derives from the fact that it applies if the probability that a call which has already persisted for a time t will end during a further time dt, is independent of t and depends only on the average holding time H.

10.4 Delay Formulae. Theoretical formulae for the performance of delay systems with Pure Chance Traffic inputs have been developed and fall into one of four classes. In each case, the name of the persons who derived them are given.

Distribution of Holding Times	Order of Removal from Store	
	Order of Arrival (Queued)	Random (Unqueued)
Exponential	Erlang	Vaulot Riordan
Constant	Erlang Molina Crommellin Pollaczek	None known

10.5 Probability of Delay. Of the three performance standards used for delay systems, the expression for the "Probability of Delay" denoted by symbol D is common for all the holding time and order of service assumptions. It is given by -

$$D = \frac{\frac{A^N}{N!} \left(\frac{N}{N - A}\right)}{1 + A + \frac{A^2}{2!} + \dots\dots\dots\dots + \frac{A^{N-1}}{(N-1)!} + \frac{A^N}{N!} \left(\frac{N}{N - A}\right)} \qquad (13)$$

$$= \frac{N.E_{1,N}(A)}{N - A + A.E_{1,N}(A)} \; ,$$

where N = Number of Circuits,

A = Average Traffic Erlangs,

D = Probability of Delay,

$E_{1,N}(A)$ = Erlang Probability of loss when A erlangs offered to N trunks.

The second form of the formula is included for convenience in evaluation since $E_{1,N}(A)$ has been extensively tabulated by Conny Palm in his "Table of the Erlang Loss Formula" published by Kungl, Telestyrelsen, Stockholm, 1954.

10.6 Waiting Times. The probability that a call once delayed will have to wait for time "t" depends on the distribution of the holding time of the calls carried by the system and the order in which waiting calls are served.

The solution for queued operation, and the exponential distribution of holding times is due to A. K. Erlang. The following expression gives the probability that a call will be delayed for a time t or longer.

$$W(t) = D.e^{-(N-A)\,t/H} \qquad (14)$$

where e = Base of the natural logarithms,

N and A are defined as above,

H = the average holding time of all calls,

D = the probability of delay (see Equation 13).

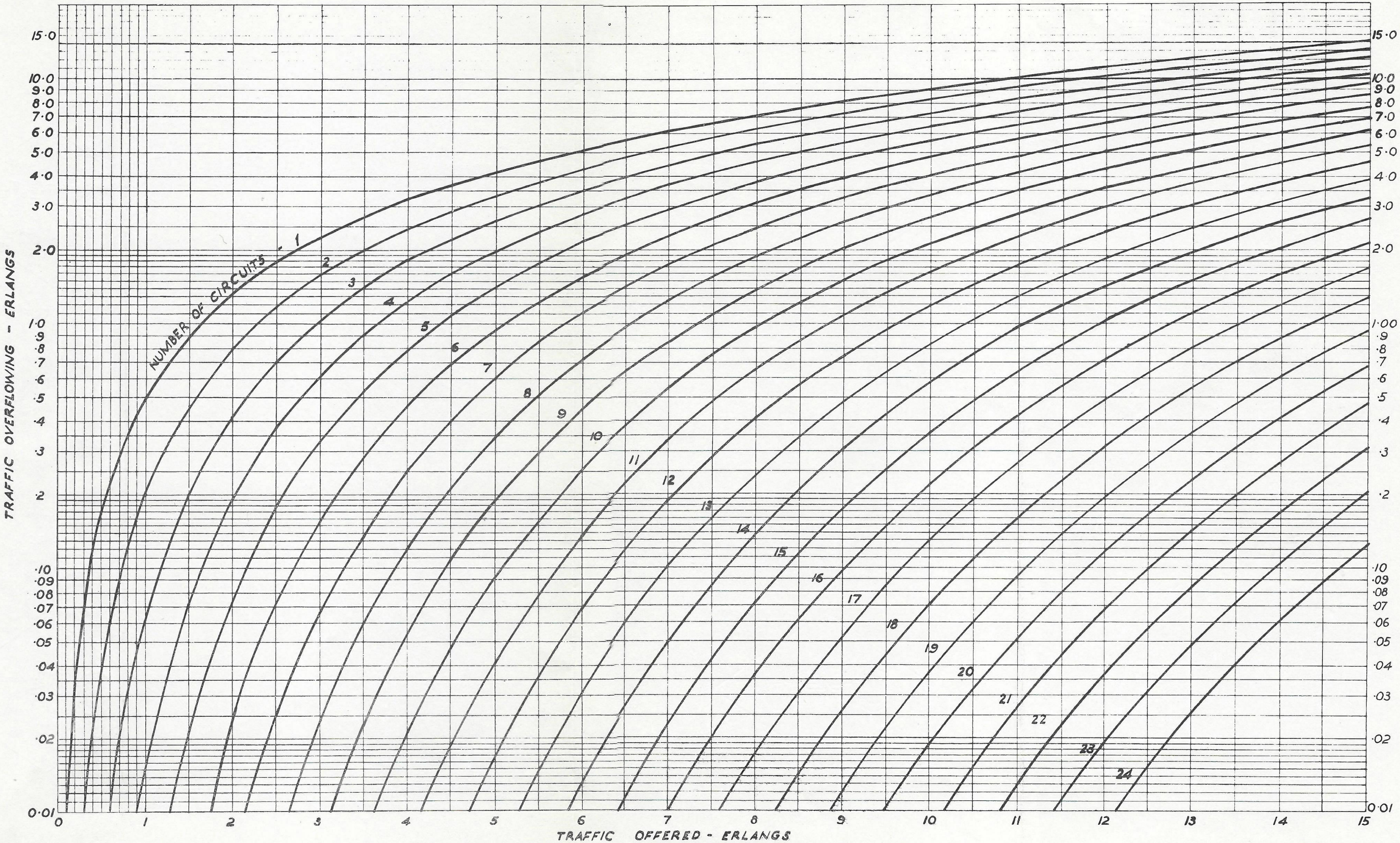


FIG. 7. TRAFFIC OVERFLOWING FROM GROUPS OF TRUNKS (FROM DRAWING CP.200).

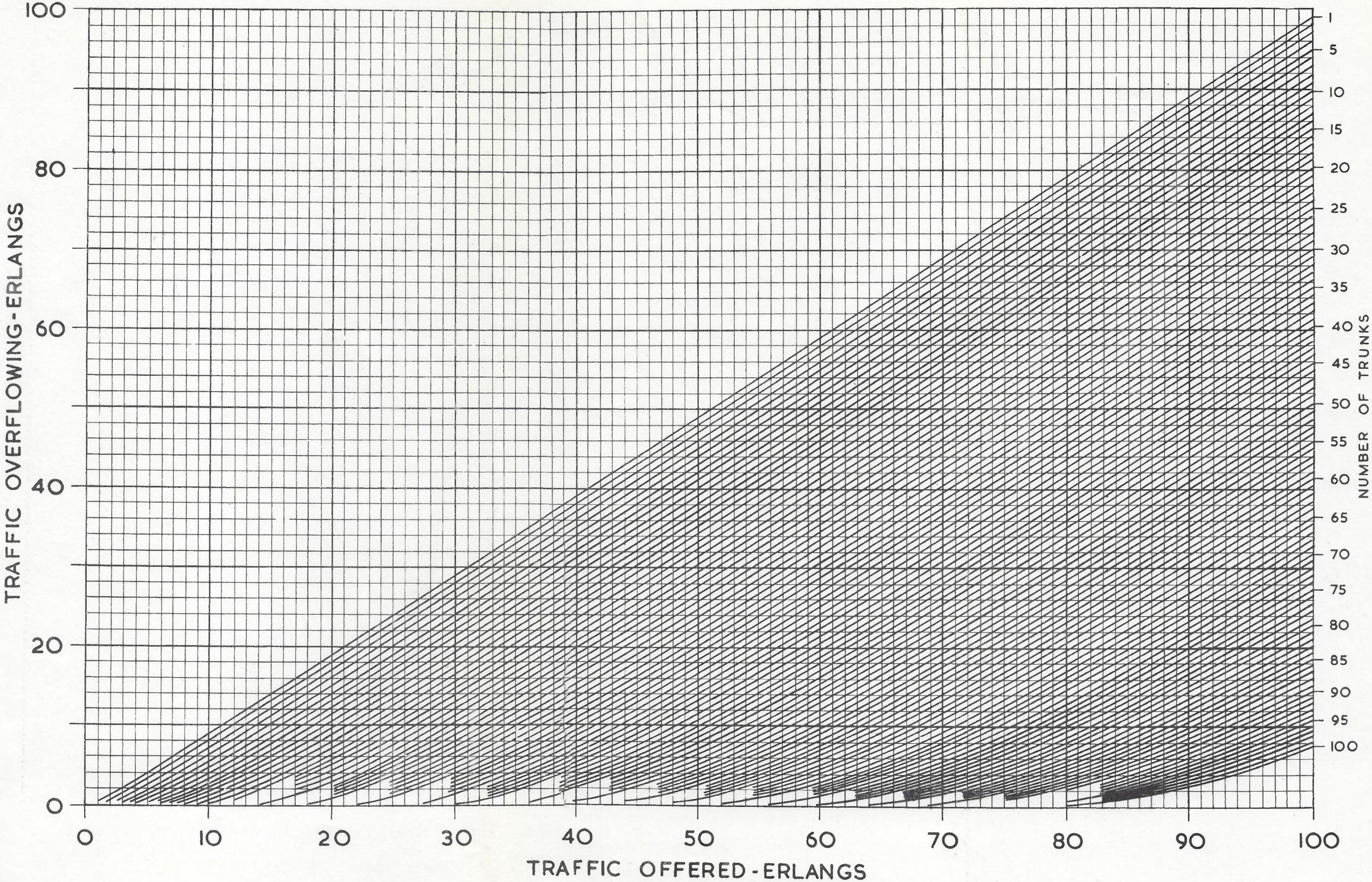


FIG. 8. TRAFFIC PASSED FROM GROUPS OF CIRCUITS (FROM DRAWING CP.200).

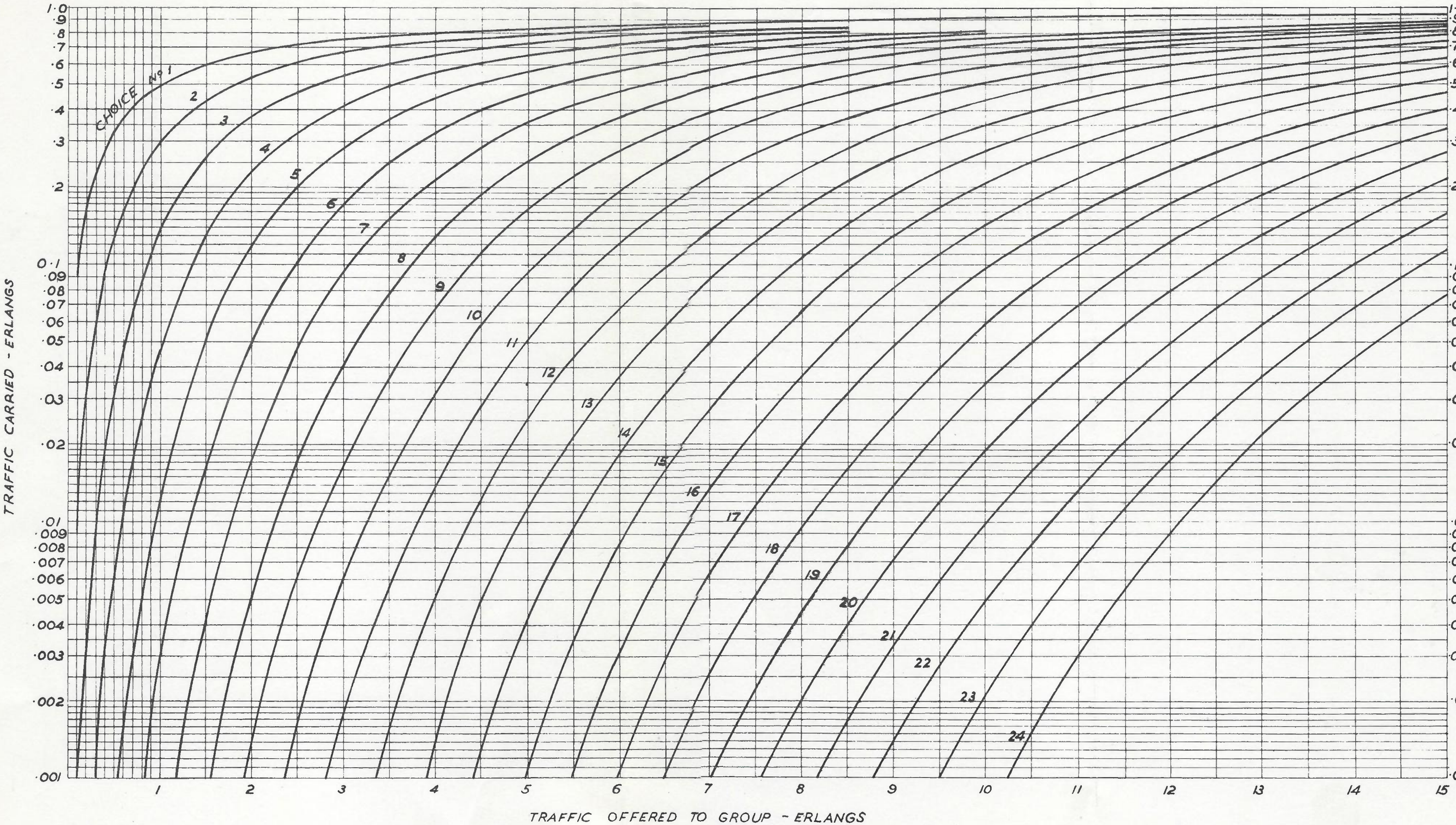


FIG. 9. TRAFFIC CARRIED BY INDIVIDUAL TRUNKS (FROM DRAWING CP.201).

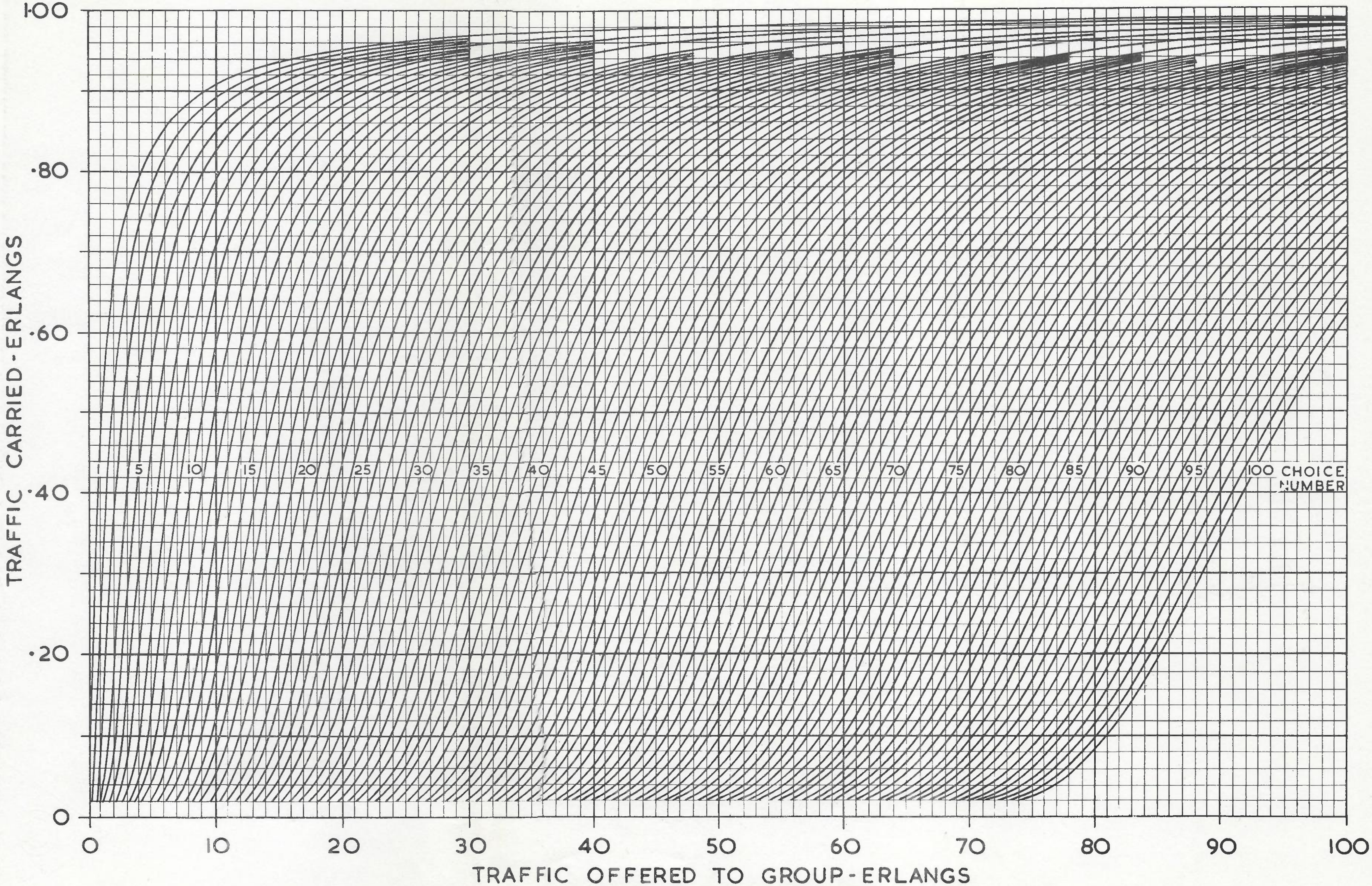


FIG. 10. TRAFFIC CARRIED BY EACH CHOICE (FROM DRAWING CP.201).

E N D.