Markov Processes Representing Traffic in Connecting Networks

By V. E. BENEŠ

(Manuscript received June 27, 1963)

A class of Markov stochastic processes x_t, suitable as models for random traffic in connecting networks with blocked calls cleared, is described and analyzed. These models take into account the structure of the connecting network, the set S of its permitted states, the random epochs at which new calls are attempted and calls in progress are ended, and the method used for routing calls.

The probability of blocking, or the fraction of blocked attempts, is defined in a rigorous way as the stochastic limit of a ratio of counter readings, and a formula for it is given in terms of the stationary probability vector p of x_t. This formula is

$$\frac{(p, \beta)}{(p, \alpha)}, \qquad \text{or} \qquad \frac{\sum_{x \in S} p_x \beta_x}{\sum_{x \in S} p_x \alpha_x},$$

where β_x is the number of blocked idle inlet-outlet pairs in state x, and α_x is the number of idle inlet-outlet pairs in state x. On the basis of this formula, it is shown that in some cases a simple algebraic relationship exists between the blocking probability b, the traffic parameter λ (the calling rate per idle inlet-outlet pair), the mean m of the load carried, and the variance σ^2 of the load carried. For a one-sided connecting network of T inlets (= outlets), this relation is

$$1 - b = \frac{1}{\lambda} \frac{2m}{(T - 2m)^2 - (T - 2m) + 4\sigma^2};$$

for a two-sided network with N inlets on one side and M outlets on the other, it is

$$1 - b = \frac{m}{(N - m)(M - m) + \sigma^2}.$$

The problem of calculating the vector p of stationary state probabilities is fully resolved in principle by three explicit formulas for the components of p: a determinant formula, a sum of products along paths on S, and an expansion in a power series around any point $\lambda \geqq 0$. The formulas all indicate how these state probabilities depend on the structure of the connecting network, the traffic parameter λ, and the method of routing.

CONTENTS

I. INTRODUCTION

A connecting system is a physical communication system consisting of (*i*) a set of terminals, (*ii*) a control unit which processes the information needed to set up calls, and (*iii*) a connecting network through which calls are switched between terminals.

Connecting systems have been described heuristically and at length in a previous paper.[1] Also, some of the algebraic and topological properties of connecting networks have been studied in another paper.[2] The models to be used here have been described (but not studied) in a third paper.[3] These papers are a source of background material for reading the present one; familiarity with them is desirable, but is not presupposed.

The principal problem treated here is the exact theoretical calculation of the grade of service (as measured by the probability of blocking) of a connecting network of given but arbitrary structure; the calculation is to be carried out in terms of a mathematical model for the operation of the network. The model used here is a Markov stochastic process x_t defined by some simple probabilistic and operational assumptions. The problem is first reduced to calculation of the stationary probability vector p of x_t from the "statistical equilibrium" equations. From the form of this reduction it follows that in many cases of practical interest the probability of blocking is uniquely determined by the mean and variance of the carried load, a fact heretofore known only for very simple systems.

In the past, the application of A. K. Erlang's very natural method of statistical equilibrium has been visited by a curse of dimensionality, that is, by the extremely large number of equations comprised in the equilibrium condition. This difficulty has not only put explicit solutions apparently out of the question; it has even made it effectively impossible to reach a reliable qualitative idea of the dependence of the blocking probability on the structure of the network, the method of routing, etc.

Three explicit formulas for the solution p of the equilibrium equations will be given. One is based on purely algebraic considerations, and the others largely on combinatory and probabilistic notions. Because of the generality of the model with respect to network structure, these formulas are of necessity rather complex. Except in simple cases, they cannot be regarded as giving a final (or even a working) solution to the problem of calculating equilibrium probabilities. Still, they expose the mathematical character of the problem, and provide a badly needed starting point for well grounded approximations. For only after one has studied and understood this character can he seriously consider ignoring some of it in approximations.

II. PRELIMINARY REMARKS AND DEFINITIONS

Various combinatory, algebraic, and topological features of the connecting network play important roles in the analysis of stochastic models for network operation. Some of these features will now be described, and terminology and notations for them introduced.

Let S be the set of permitted (i.e., physically meaningful) states of the connecting network under study. It has been pointed out in earlier work[1,2] that these states are partially ordered by inclusion $\leqq$, where

$$x \leqq y$$

means that state x can be obtained from state y by removing zero or more calls. Also, these states can be arranged (in fact, partitioned) in an intuitive manner in a *state-diagram*, the Hasse figure for the partial ordering $\leqq$. This figure is a graph constructed by partitioning the states in horizontal rows according to the number of calls in progress, the kth row consisting of all states with k calls in progress. The unique *zero*, or empty, state of the network, in which no calls are in progress, is placed at the bottom of the figure; above it comes the row consisting of states with exactly *one* call in progress, and so on. The figure is completed by drawing a graph with the states as nodes, and with adjacency matrix determined by the condition that states differing by exactly one call are adjacent. This means that in drawing the graph we place lines between states (in successive rows) that differ in point of one call. A *maximal* state is one that forms a summit of the state-diagram, i.e., has no states above it in the partial ordering.

For most systems the state-diagram has the following heuristic description: there is a unique "point" at the bottom corresponding to the zero state; there are usually many "points" at the top corresponding to the maximal states; and the diagram is very "fat" in the middle, because of the multitude of states with a moderate number of calls in progress.

We mention at this point that usually the number of states, i.e., the number of elements of S, is astronomically large. Indeed, this fact has been a principal obstacle to theoretical progress on problems of congestion in large connecting systems. For an illustration, in the network of No. 5 Crossbar type, illustrated in Fig. 1, made for 1000 lines out of square 10×10 switches, the number of *maximal states alone* is

$$(100 \times 10!)^4 = 1.734 \times 10^{34}.$$

The set of *inlets* of a connecting network is denoted by I, and the set

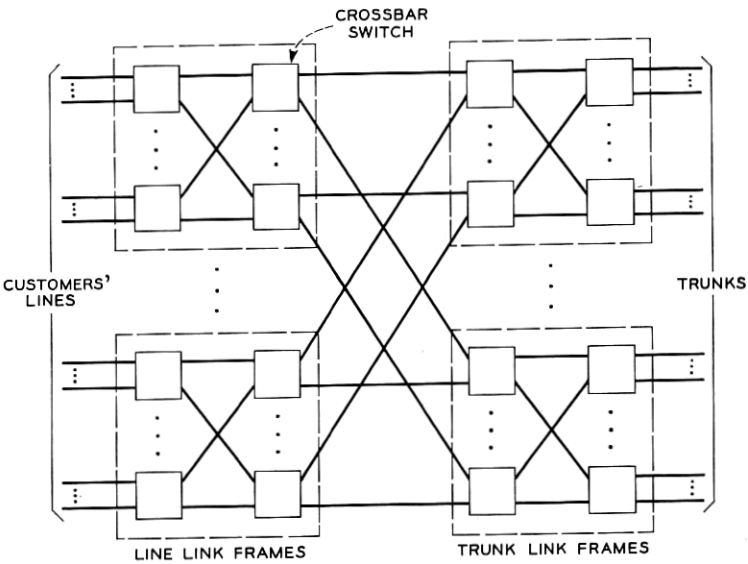

Fig. 1 — Structure of No. 5 Crossbar network.

of *outlets* by Ω. It is possible that $I \cap \Omega = \phi$, that $I \cap \Omega \neq \phi$, or even that $I = \Omega$, i.e., that all inlets are also outlets, depending on the "community of interest" aspects of the structure of the network. It is assumed that every call or connection is made only between an inlet and an outlet.

If x is a state, the notation $| x |$ (read "the norm of x") will denote the number of calls in progress in state x. If X is a set, then $| x |$ will denote the cardinality of X, i.e., the number of elements of X. We define the levels

$$L_k = \{x \in S: \quad | x | = k\}, \qquad k = 0, 1, \cdots, \max_{x \in S} | x |,$$

as the sets of states in which a specified number of calls is in progress. The $\{L_k\}$ form a partition of S,

$$\bigcup_k L_k = S$$

$$L_k \cap L_j = 0, \qquad k \neq j.$$

The "neighbors" of a state x are just those states which can be reached from x by adding or removing one call. These neighbors y of x can be divided into two sets according as $y > x$ or $y < x$; so we are led to define

$$\begin{aligned} A_x &= \text{set of neighbors } \textit{above } x \\ &= \text{set of states accessible from } x \text{ by } \textit{adding} \text{ one call} \\ B_x &= \text{set of neighbors } \textit{below } x \\ &= \text{set of states accessible from } x \text{ by } \textit{removing} \text{ one call.} \end{aligned}$$

III. SUMMARY

The basic probabilistic assumptions that define the randomness in the traffic models to be studied are given precise statement in Section IV. They are, briefly, (*i*) the hang-up rate per call in progress is unity, and (*ii*) the calling rate per idle inlet-outlet pair is a constant $\lambda > 0$. Various operational aspects, such as the disposition of lost calls, and the method of routing, are specified and discussed in Section V. It is assumed that lost calls are refused without a change in state, and that routes for calls are chosen in a way that depends both on the call being set up or processed and on the current state of the system. In Section VI these probabilistic and operational assumptions are summarized in a transition rate matrix, Q. In Section VII, a Markov stochastic process x_t (the mathematical model for the operating system) is defined, and the statistical equilibrium condition $Qp = 0$ for the stationary probability vector p of x_t is formulated.

In Section VIII, the probability b of blocking is defined as the (probability one) limit of a ratio of counter readings, and a formula for b is given in terms of the stationary vector p. From this formula it is shown, in Section IX, that a simple algebraic relationship often exists between the blocking probability b, the traffic parameter λ, the mean load carried, and the variance of the load carried.

The remainder of the paper is devoted to the study and calculation of the vector p of stationary probabilities. Two explicit solutions, one algebraic and one combinatory in character, are given in Section X. In Section XI it is shown that the combinatory solution is a special case of a general formula for the stationary measure of an ergodic Markov process. The dependence of $p = p(\lambda)$ on the network structure and the method of routing is analyzed in an elementary way in Section XII. It is first shown that $p(\cdot)/p_0(\cdot)$ has components that are analytic in a neighborhood of the nonnegative real axis, and so are expressible in the form

$$\frac{p_x(\mu + \epsilon)}{p_0(\mu + \epsilon)} = \sum_{m=0}^{\infty} \epsilon^m c_m(x,\mu).$$

For $\mu = 0$ and $\epsilon = \lambda$ sufficiently small, this gives an expansion of p in powers of λ. It is then shown that with $|x|$ the number of calls in progress in state x, p_x is of order $\lambda^{|x|}$ as $\lambda \to 0$. This result renders possible a recursive calculation (Sections XII and XIV) of the coefficients $c_m(x,0)$ from the partial ordering $\leqq$ of S and a matrix used to specify the method of routing. Once p is developed as a power series in λ, a similar expansion is readily given (Section XIII) for the probability b of blocking.

In Section XV, finally, we completely solve the problem of calculating the coefficients $c_m(x,\lambda)$ for arbitrary values of $\lambda > 0$, giving each such coefficient both a combinatory interpretation, and an explicit formula, viz., a sum of products along paths through S which are trajectories for x_t permitted by the routing rule.

IV. PROBABILITY

To construct a Markov process for representing the random trajectory of the operating network through the set S of states, we shall make two simple probabilistic assumptions. The traffic models to be studied embody what has come to be known as a "finite-source effect," that is, a dependence of the instantaneous total calling-rate on the number of idle inlets, and on that of idle outlets.

In an attempt to describe this dependence in a simple rational way, let us imagine a customer located at one of the inlets [outlets] of the connecting network, and seek to assign him a calling-rate, assuming that he is in an idle condition. We shall suppose that the traffic he offers is homogeneous in the sense that he calls every outlet [inlet] at the same rate, or with the same frequency. Indeed, we shall assume that all customers offer homogeneous traffic. Now on most occasions when he is making a call, a customer does not know whether the terminal he is calling is busy or idle. Thus, if he is on an inlet [outlet] it seems reasonable to suppose that there is a probability

$$\lambda h + o(h) \qquad \lambda > 0$$

that he attempts a call to a particular outlet [inlet] (distinct from his own) in the next interval of time of length h, as $h \to 0$, whether that outlet [inlet] is busy or not. The qualifying phrase "distinct from his own" is inserted to cover the case in which some inlets are also outlets, and in which it is reasonable to suppose that an idle terminal that is both an inlet and an outlet does not attempt to call itself.

We therefore make these two probabilistic assumptions:

(a) Holding-times of calls are mutually independent random variables, each with the negative exponential distribution of unit mean.

(b) If at time t the network is in a state x in which at least one member of the inlet-outlet pair $(u,v) \in I \times \Omega$ is idle (that is, one of u or v is not involved in a call in progress), the time elapsing from t until a call between u and v is attempted is a random variable having a negative exponential distribution with a mean $1/\lambda$, $\lambda > 0$. For different choices (u,v) and different occasions t, these times are all mutually independent and also independent of the call holding times.

These assumptions can be rendered in the informal terminology of "rates" as follows:

(*i*) The hang-up rate per call in progress is unity.

(*ii*) The calling-rate between an idle inlet u (outlet v) and an arbitrary outlet v (inlet u) with $u \neq v$ is $\lambda > 0$.

Assumptions (a) and (b) provide all the "randomness" needed to construct our models. The choice of a *unit* hang-up rate merely means that the mean holding-time is being used as the unit of time, so that only the one parameter λ need be specified.

V. OPERATION

To complete the description of the traffic models to be analyzed we must indicate how the network is operated. Since in the present work we are taking into account only the network configuration, and omitting consideration of the control unit, it suffices to describe how calls to busy terminals are handled, how blocked calls are treated, and how routes or paths through the network are chosen.

It will be assumed that attempted calls to busy terminals are rejected, and have no effect on the state of the system; similarly, blocked attempts to call an idle terminal are refused, with no change in the state of the system. All successful attempts to place a call are completed instantly, with some choice of route.

To describe how routes are assigned to calls, we introduce a *routing matrix* $R = (r_{xy})$, with the following properties: for each x let Π_x be the partition of A_x induced by the equivalence relation of "having the same calls up," or satisfying the same "assignment" (of inlets to outlets); then for each $Y \in \Pi_x$, r_{xy} for $y \in Y$ is a probability distribution over Y; in all other cases $r_{xy} = 0$.

The interpretation of the routing matrix R is this: any $Y \in \Pi_x$ represents all the ways in which a particular call c not blocked in x (between an inlet idle in x and an outlet idle in x) could be completed when the network is in state x; for $y \in Y$, r_{xy} is the chance that if this call c is attempted, it will be routed through the network so as to take the system to state y. That is, we assume that if c is attempted in x, then a state y is drawn at random from Y with probability r_{xy}, independently each time c is attempted in x; the state y so chosen indicates the route c is assigned. The distribution or probability $\{r_{xy}, y \in Y\}$ thus indicates how the calling-rate λ due to the call c is to be spread over the possible ways of putting up the call c. It is apparent that

$$\sum_{y \in A_x} r_{xy} = \text{number of calls which can actually be put up in state } x$$
$$= s(x), \qquad (\text{"successes" in } x),$$

the second equality defining $s(\cdot)$ on S. This account of the method of routing completes the description of the traffic models to be studied.

VI. TRANSITION RATES

For the purpose of defining a Markov stochastic process it is convenient and customary to collect the probabilistic and operational assumptions introduced above in a matrix $Q = (q_{xy})$ of *transition rates*, here given by

$$q_{xy} = \begin{cases} 1 & y \in B_x \\ \lambda r_{xy} & y \in A_x \\ -|x| - \lambda s(x) & y = x \\ 0 & \text{otherwise.} \end{cases}$$

The number q_{xy}, for $x \neq y$, has the usual interpretation that if the system is in state x, there is a chance

$$q_{xy}h + o(h)$$

that it will move to y in the next interval of time of length h, as $h \to 0$. Similarly

$$1 - q_{xx}h + o(h)$$

is the probability that the system will stay in x throughout the next interval of time of length h, as $h \to 0$

VII. MARKOV PROCESSES

In terms of the transition rate matrix Q it is possible to define a stationary Markov stochastic process $\{x_t, -\infty < t < +\infty\}$ taking values on the set S of states. The matrix $P(t)$ of transition probabilities

$$p_{xy}(t) = \Pr\{x_t = y \mid x_0 = x\}$$

of x_t satisfies the equations of Kolmogorov

$$\frac{d}{dt} P(t) = QP(t) = P(t)Q,$$

$$P(0) = I,$$

and is given formally by the formula

$$P(t) = \exp tQ.$$

Theorem 1: There exists a decomposition of the set S of states into a transient *set F and a single* ergodic *set $S - F$ containing the zero state; members of F have the property*

$$\lim_{t\to\infty} p_{xy}(t) = 0 \qquad y \in F, x \in S;$$

on $S - F$ there is a unique stationary *(or* equilibrium*) distribution $\{p_x, x \in S - F\}$ such that*

$$\lim_{t\to\infty} p_{xy}(t) = p_y > 0 \qquad y \in S - F, x \in S$$

$$\sum_{x \in S-F} p_x p_{xy}(t) = p_y \qquad y \in S - F, \quad \text{all } t$$

$$\sum_{x \in S-F} q_{xy} p_x = 0 \qquad y \in S - F.$$

Proof: The existence of the unique ergodic set $S - F$ follows from the fact that the zero state is accessible from every other state by hangups. The existence and character of the limit of $p_{xy}(t)$ as $t \to \infty$ is a consequence of exercise 19, p. 436 of Feller,[4] i.e., of the fact that the characteristic values r of Q satisfy $r = 0$ or $\mathrm{Re}(r) < 0$. (See also Bellman,[5] p. 294.)

To prove the uniqueness of p, suppose that q is a different probability vector on $S - F$ that also satisfies the "equilibrium" condition

$$\sum_{x \epsilon S-F} q_{xy} q_x = 0 \qquad y \epsilon S - F.$$

Then by Kolmogorov's equation

$$\frac{d}{dt} \sum_{x \epsilon S-F} g_x p_{xy}(t) = \sum_{x,z \epsilon S-F} q_x q_{xz} p_{zy}(t) = 0.$$

Integrating from 0 to t, and using $P(0) = I$, we find

$$\sum_{x \epsilon S-F} q_x p_{xy}(t) = q_y \qquad y \epsilon S - F.$$

Since $S - F$ is the only ergodic set, the left-hand side approaches p_y as $t \to \infty$. Hence $p = q$.

It is convenient to extend the dimension of p to $|S|$ by adding zero components for states in F, so that $p_{xy}(t) \to p_y \geqq 0$ for all $x,y \epsilon S$. The consideration of the transient set F is not just a mathematical fillip, since a "good" routing rule R may explicitly make certain "bad" states *unreachable* from the zero state, and thus place them in F to good purpose.

In the notation of Halmos,[6] p. 65, the stationary probability vector satisfies the equilibrium condition

$$Qp = 0.$$

This is the classical equation of state, or equation of statistical equilibrium, familiar in traffic theory. For our process x_t it takes on the rather simple form

$$[|x| + \lambda s(x)]p_x = \sum_{y \epsilon A_x} p_y + \lambda \sum_{y \epsilon B_x} p_y r_{yx}, \qquad x \epsilon S.$$

The left-hand side represents the average rate of exits from x, while the right-hand side is the average rate of entrances into x, in equilibrium. We define

$$p_k = \sum_{|x|=k} p_x = \Pr\{L_k\}.$$

Lemma 1: For $1 \leqq k \leqq w = \max_{x \epsilon S} |x|$

$$kp_k = \lambda \sum_{x \epsilon L_{k-1}} p_x s(x).$$

Proof: From the statistical equilibrium equation for $x = 0$ we obtain

$$\sum_{y \epsilon A_0} p_y = \lambda s(0) p_0 ,$$

which is Lemma 1 for $k = 1$. Assume that the lemma holds for a given $k \geqq 1$. Summing the statistical equilibrium equations over $x \epsilon L_k$ we find

$$kp_k + \lambda \sum_{x \epsilon L_k} s(x) p_x = \sum_{x \epsilon L_k} \sum_{y \epsilon A_x} p_y + \lambda \sum_{x \epsilon L_k} \sum_{y \epsilon B_x} p_y r_{yx} .$$

The second sum on the right is the same as

$$\lambda \sum_{y \epsilon L_{k-1}} p_y \sum_{x \epsilon A_y} r_{yx} ,$$

and by definition,

$$\sum_{x \epsilon A_y} r_{yx} = s(y).$$

Hence (induction hypothesis) the second sum equals kp_k . It is easy to see that in the first sum on the right each p_y is counted exactly $| y |$ times, i.e., $(k + 1)$ times, since for a given $y \epsilon L_{k+1}$ there are exactly $(k + 1)$ elements $x \epsilon L_k$ for which $y \epsilon A_x$. Thus the first sum is

$$(k + 1) \sum_{y \epsilon L_{k+1}} p_y = (k + 1) p_{k+1}$$

and the Lemma follows by induction. This result could also be obtained from the general observation that the statistical equilibrium equations are equivalent to the principle that for any set X of states the average rate of exits from X equals the average rate of entrances into X. (See Morris and Wolman.[7])

VIII. PROBABILITY OF BLOCKING

The fraction of calls that are refused because they are blocked, or the probability of blocking, is a quantity of particular interest to traffic engineers; they use it to assess the grade of service provided by an operating connecting network. The rigorous theoretical calculation of blocking probabilities has long been an outstanding problem of traffic theory. This problem is outstanding in both senses of the word: it is *conspicuous*, and it is *unsolved*. In fact, not even the definition (let alone the calculation) of the probability of blocking has received adequate treatment; for example, the otherwise monumental treatise of R. Syski[8] does not give a general account of blocking probability.

Since it is desirable to have a close connection between theoretical quantities and their physical meanings in terms of measurements, we shall approach the study of blocking probabilities by asking how these probabilities might be measured "in the field." The most natural method of measuring the fraction of blocked attempts seems to be this: to the control unit of the connecting system under consideration we attach two counters; the first will count up one unit every time an attempted call is blocked, and the second will register one unit every time a call is attempted; the ratio of the reading of the first counter to that of the second should, after a long time during which the system's parameters remain constant, be an approximate measure of the fraction of blocked attempts. For mathematical convenience, one can then *define* the probability of blocking to be the limit (as time increases without end) of this ratio of the counter readings. This mathematical definition was first proposed by S. P. Lloyd, although, of course, the ratio has been the practical definition for 50 years, being the "peg count and overflow ratio."

A precise mathematical version of this measurement procedure can be given as follows: on the same sample space as that of the process x_t that describes the operating network, we define two additional stochastic processes $\{b(t), t \geqq 0\}$ and $\{a(t), t \geqq 0\}$ by the (respective) conditions

$$b(t) = \text{number of blocked attempted calls in } (0,t],$$

$$a(t) = \text{number of attempted calls in } (0,t].$$

These stochastic processes are the mathematical analogs of the counter readings. It is reasonable to use the limit

$$\lim_{t\to\infty} \frac{b(t)}{a(t)}$$

of the ratio of $b(\cdot)$ to $a(\cdot)$ as a mathematical definition of the probability of blocking, provided that the limit exists in a suitable sense. We show that this limit exists and is constant with probability one, and we give a formula for it.

Theorem 2: The probability of blocking b, defined by

$$b = \lim_{t\to\infty} \frac{b(t)}{a(t)},$$

exists and is constant with probability one; its almost sure value is

$$b = \frac{(p,\beta)}{(p,\alpha)} = \frac{\sum_{x\epsilon S} p_x \beta_x}{\sum_{x\epsilon S} p_x \alpha_x},$$

where p is the vector of stationary probabilities, and

β_x = *number of idle inlet-outlet pairs that are blocked in state* x,

α_x = *number of idle inlet-outlet pairs in state* x.

Proof: It can be seen that $a(t)$ and $b(t)$ can be written as sums over S,

$$a(t) = \sum_{x \in S} a_x(t)$$

$$b(t) = \sum_{x \in S} b_x(t)$$

where

$a_x(t)$ = number of attempted calls made in $(0,t]$ with the system in state x,

$b_x(t)$ = number of blocked attempts made in $(0,t]$ with the system in state x.

Now a blocked attempt occuring at an epoch u such that $x_u = x$ does not change the state of the system. Such an epoch u is a regeneration point of the process x_t. A successful attempt occurring at an epoch u at which $x_u = x$ *does* change the state of the system. The time interval from u back to the last previous epoch v at which a successful attempt occurred in state x, however, is independent of the behavior of x_t for $t > u$; it depends only on the fact that the system left x by adding a new call, not on *what* new call it was, nor on where into A_x x_t went as this new call was completed. This can be seen as follows: we have

$$u - v = \tau - v + u - \tau$$

where τ is the epoch at which x was last entered prior to u. Now $\tau - v$ is independent of x_t for $t > \tau$ if $x_{\tau+0}$ is known to be x, because x_t is a Markov process.

Let U be an event measurable on $\{x_t, t > u\}$. Then

$$\Pr\{U \text{ and } u - \tau \leqq \mu \mid x_{\tau+0} = x\}$$

$$= \Pr\{u - \tau \leqq \mu \mid x_{\tau+0} = x\} \sum_{y \in A_x} \frac{r_{xy}}{s(x)} \Pr\{U \mid x_{u+0} = y\},$$

where

$$\Pr\{u - \tau \leqq \mu \mid x_{\tau+0} = x\} = \sum_{n=1}^{\infty} \left(\frac{|x|}{|x| + \lambda s(x)}\right)^{n-1} \frac{\lambda s(x)}{|x| + \lambda s(x)} \int_0^{\mu} \frac{[|x| + \lambda s(x)]^n \, t^{n-1} \, e^{-t[|x|+\lambda s(x)]}}{(n-1)!} \, dt.$$

Thus the time intervals $\beta_1, \beta_2, \cdots$ elapsing between successive blocked attempts in state x, and those $\alpha_1, \alpha_2, \cdots$ elapsing between successive attempts in state x, both form sequences of mutually independent, and except possibly for the first elements β_1 and α_1, identically distributed random variables. That is, the elements of each sequence are mutually independent, but the sequences are not independent since one consists of partial sums over blocks of the other.

Both these sequences can be studied, then, in terms of a sequence $x_1, x_2, \cdots$ of mutually independent random variables, all (except possibly x_1) identically distributed. We define for $t \geqq 0$ and $k \geqq 0$

$$S_0 = 0$$

$$S_n = \sum_{i=1}^{n} x_i, \qquad x_i = \alpha_i \quad \text{or} \quad \beta_i$$

$$n(t) = k \quad \text{if and only if} \quad S_k \leqq t < S_{k+1},$$

$$n(t) = a_x(t) \quad \text{or} \quad b_x(t).$$

It is now straightforward to show that $t^{-1}n(t)$ approaches a limit with probability one, and to find the limit. Let us put, for $t > S_1$

$$\frac{n(t)}{t} = \frac{n(t)}{S_{n(t)}} \cdot \frac{S_{n(t)}}{t}.$$

The first factor converges to $E^{-1}\{x_2\}$ with probability one, by the law of large numbers. The local suprema of

$$\frac{t - S_{n(t)}}{t}$$

for $t > S_1$ occur at the points

$$t = S_k \qquad k = 1, 2, \cdots,$$

and have the values

$$\frac{x_k}{S_k} = \frac{k}{S_k} \cdot \frac{x_k}{k} \qquad k = 1, 2, \cdots.$$

Again, the first factor converges to $E^{-1}\{x_2\}$ with probability one by the law of large numbers. Since $E\{x_2\} < \infty$, and $\{x_k, k \geqq 2\}$ are identically distributed,

$$\sum_{k=2}^{\infty} \Pr\{x_k > \epsilon k\} = \sum_{k=2}^{\infty} \Pr\{x_2 > \epsilon k\} < \infty$$

and it follows from the Borel-Cantelli lemma that for any $\epsilon > 0$

$$\Pr\{x_k > \epsilon k \text{ for infinitely many values of } k \geqq 2\} = 0.$$

Hence $x_k = o(k)$ as $k \to \infty$, with probability one, and with the same probability,

$$\frac{S_{n(t)}}{t} \to 1 \quad \text{as} \quad t \to \infty.$$

It follows that with probability one,

$$\lim_{t\to\infty} t^{-1}a_x(t) = E^{-1}\{\text{time interval between successive attempted calls in } x\},$$

$$\lim_{t\to\infty} t^{-1}b_x(t) = E^{-1}\{\text{time interval between successive blocked attempted calls in } x\}.$$

Furthermore (cf. Ref. 11, p. 247, equation (1.2) and p. 249)

$$\frac{d}{dt} E\{b_x(t) \mid x_0 = y\} = \lambda p_{yx}(t)\beta_x \to \lambda p_x\beta_x \quad \text{as} \quad t \to \infty.$$

However, by Feller's renewal theorem (cf. Ref. 11, p. 246), we know that

$$\lim_{t\to\infty} t^{-1}E\{b_x(t)\mid x_0 = y\} = E^{-1}\{\text{time interval between successive blocked attempted calls in } x\}.$$

Hence, with probability one,

$$t^{-1}b_x(t) \to \lambda p_x\beta_x \quad \text{as} \quad t \to \infty.$$

A similar argument shows that with probability one

$$t^{-1}a_x(t) \to \lambda p_x\alpha_x \quad \text{as} \quad t \to \infty$$

and completes the proof of Theorem 2.

IX. A BASIC FORMULA

Engineers have recognized (at least) four quantities as significant for the study and design of connecting networks carrying random traffic. These are the calling rate, the average load carried, the variance of the load carried, and the probability of blocking. In our model these quantities are given respectively by

λ = calling rate per idle inlet-outlet pair

$m = \sum_{x \epsilon S} |x| p_x$ = average number of calls in progress

$$\sigma^2 = \sum_{x \epsilon S} (|x| - m)^2 p_x$$

$$b = (p,\beta)/(p,\alpha).$$

It is natural to ask whether there exist any systematic relationships between these quantities, or between these and (possibly) other simple parameters of the network under study. Such relationships would be particularly useful and significant if they were largely independent of the structure or configuration of the connecting network, and were valid either for all networks or for large classes of them. We shall show that there often exists a simple *algebraic* relation among λ, m, σ^2, and b. Its exact form depends on which inlets are also outlets. First we prove

Theorem 3: The probability b of blocking can be written as

$$b = 1 - \frac{m}{\lambda \sum_{x \epsilon S} p_x \alpha_x} \tag{1}$$

or, in words, as

$$b = 1 - \frac{\textit{average load carried}}{(\textit{calling rate per idle}) \times (\textit{average number of idle pairs})}.$$

Proof: In equilibrium, the average rate of successful attempts must equal the average rate of hangups. Hence, intuitively,

$$\lambda \sum_{x \epsilon S} p_x s(x) = \sum_{x \epsilon S} |x| p_x = m. \tag{2}$$

Since $\beta_x = \alpha_x - s(x)$, the result follows from Theorem 1. The actual validity of the identity (2) can be inferred from Lemma 1, by summation on k.

Formula (1), rewritten in the form

$$1 - b = \frac{\text{average load carried}}{\text{average rate of attempts}},$$

should be viewed as a direct generalization of Erlang's classical loss formula for c trunks, blocked calls cleared, and calls arising in a Poisson process of intensity $a > 0$. In that case the probability of loss is

$$E_1(c,a) = \frac{\frac{a^c}{c!}}{\sum_{j=0}^{c} \frac{a^j}{j!}},$$

and it can be seen that

$$1 - E_1(c,a) = \frac{\sum_{j=0}^{c-1} \frac{a^j}{j!}}{\sum_{j=0}^{c} \frac{a^j}{j!}}$$

$$= \frac{\sum_{j=1}^{c} j \frac{a^j}{j!}}{a \sum_{j=0}^{c} \frac{a^j}{j!}}$$

$$= \frac{\text{average number of busy trunks}}{\text{total calling rate}}.$$

To exhibit useful special cases of the general formula (1) of Theorem 3, we introduce a partial classification of connecting networks. A network is called *one-sided* if $I = \Omega$, i.e., if all inlets are also outlets; a network is *two-sided* if $I \cap \Omega = \phi$, i.e., if no inlet is an outlet.

Corollary 1: For a one-sided network of T terminals

$$b = 1 - \frac{1}{\lambda} \frac{2m}{(T - 2m)^2 - (T - 2m) + 4\sigma^2}.$$

Proof: For the one-sided network in question, we have $I = \Omega$, $| I | = | \Omega | = T$, and so

$$\alpha_x = \binom{T - 2 | x |}{2}$$

$$\sum_{x \epsilon S} p_x \alpha_x = \tfrac{1}{2}\{T^2 - (2T - 1)2m - T + 4m^2 + 4\sigma^2\}.$$

Corollary 2: For a two-sided network with M terminals on one side and N on the other

$$b = 1 - \frac{1}{\lambda} \frac{m}{(M - m)\,(N - m) + \sigma^2}.$$

Proof: It is clear that in this case

$$\alpha_x = (M - | x |)(N - | x |)$$

so that

$$\sum_{x \epsilon S} p_x \alpha_x = (M - m)(N - m) + \sigma^2.$$

Each of the foregoing corollaries exhibits an explicit algebraic relationship between λ, m, σ, and b, based only on the one- or two-sidedness of the network.

The preceding corollaries can be used to show that in a large system, the numerical value of the constant λ will be small—indeed, of the order of the reciprocal of the number of inlets and outlets. This can be seen by the following heuristic argument, carried out for a one-sided network with T terminals: suppose that each terminal carries q $(0 \leqq q < 1)$ erlangs and that the blocking probability b is so small that we can ignore it and set

$$b = 1 - \frac{1}{\lambda} \frac{2m}{(T - 2m)^2 - (T - 2m) + 4\sigma^2} = 0.$$

Since the network is one-sided, any load carried by one terminal is also carried by some other terminal, and so

$$qT = 2m$$

whence

$$\lambda = \frac{1}{T} \frac{q}{(1 - q)^2 - \frac{(1 - q)}{T} + \frac{4\sigma^2}{T^2}}.$$

Because

$$\sigma^2 = \sum_{x \epsilon S} p_x(|x| - m)^2$$

and

$$(|x| - m)^2 \leqq \frac{T^2}{4},$$

we have $0 \leqq 4\sigma^2/T^2 \leqq 1$, and so

$$\lambda \approx \frac{\text{const.}}{T}$$

with

$$\frac{q}{1 + (1 - q)^2} \leqq \text{const.} \leqq \frac{q}{(1 - q)^2 - T^{-1}(1 - q)}.$$

X. SOLUTION OF THE EQUATIONS OF STATISTICAL EQUILIBRIUM

So far, we have shown that the theoretical determination of the blocking probability b reduces to that of the stationary vector p or, in many cases, to that of the mean m and variance σ^2 of the carried load. In either case, some knowledge of p is required. Most of the rest of this paper, therefore, is devoted to the calculation of p and to the study of its properties.

In the past, the application of A. K. Erlang's very natural method of "statistical equilibrium" to congestion in connecting networks has been visited by the curse of dimensionality, that is, by the extremely large number $|S|$ of equations comprised in the stationarity condition $Qp = 0$. This difficulty has not only put explicit solutions apparently out of the question; it has even made it effectively impossible to reach a reliable qualitative idea of the *dependence* of the state probabilities $\{p_x, x \in S\}$ on the structure of the network and on the method of routing.

To be sure, it has always been possible in principle to solve $Qp = 0$ by successive elimination of unknowns; however, when the dimension of p is of order 10^{40} or so, this remark is hardly helpful. Since successive elimination can be used to solve $Qp = 0$ for any "ergodic" transition rate matrix Q, it neither elucidates nor uses any of the special features of the matrices Q that arise in problems of congestion in networks. Thus, even were it algebraically feasible, the method of successive elimination treats our matrices Q as indistinguishable from other matrices possessing a zero characteristic value.

We shall give several explicit solutions of the equilibrium equations. One is based on purely algebraic considerations, and the others largely on combinatory and probabilistic notions. Because of the generality of our model with respect to network structure, the formulas appearing in the solutions are necessarily rather complex. Except in simple cases, they cannot be regarded as giving a final (or even a working) solution to the problem of calculating equilibrium probabilities. Nevertheless, they expose the mathematical structure of the problem and provide a badly needed starting point for well grounded approximations. For only after one has studied and understood this structure can he seriously think about throwing some of it away in approximations.

To describe the solutions in full detail, we need various preliminary definitions and conventions.

It will be shown in a later paper[9] that the minimum value of the blocking probability b is achieved by a routing matrix R consisting entirely of zeros and ones, i.e., by a deterministic rule. So it is assumed henceforth that R has only zeros or ones for entries.

A *path* on S of length $l \geqq 0$ is any ordered sequence $x_0, x_1, \cdots, x_l$ of $(l + 1)$ elements of S. A lower case pi, π, will be used as a symbol for a generic path on S, and we write

$$\pi = \{x_0, x_1, \cdots, x_l\}, \qquad l = l(\pi)$$

to indicate that π is a path of length $l(\pi)$ consisting of $x_0, x_1, \cdots, x_l$ in that order. Note that paths of length zero are countenanced. A path π is a *loop* if

$$x_0 = x_l,$$

and also

$$x_i \neq x_j$$

whenever $0 < i < j \leqq l(\pi)$. A loop of length zero is a path of length zero. If $\pi = \{x_0, x_1, \cdots, x_0\}$ is a loop, each element $x_0, x_1, \cdots$ etc. will be spoken of as being *on* π.

The elements x and y of S are called *adjacent* in the graph of $(S, \leqq)$, i.e., in the state diagram, if one of the following equivalent conditions holds:

(*i*) x covers y or y covers x
(*ii*) $y \in A_x$ or $x \in A_y$
(*iii*) x and y differ by exactly one call in progress.

A path on S is called *continuous* if successive elements of the path are adjacent.

In order that x_t have positive probability of following a path π, it is not enough that π be continuous. For evidently the action of the routing matrix R (assumed to consist solely of zeros and ones) is to prohibit certain paths on S as (parts of) possible realizations of the process x_t. Here "possible" of course means "having positive probability." There exists then a class of those paths that are permitted by R, definable in several ways. One such way is as follows: A path $\pi = \{x_0, x_1, \cdots x_l\}$ on S is *permitted by R* if for each i in the range $1 \leqq i \leqq l$,

$$x_i \in B_{x_{i-1}} \qquad \text{or} \qquad r_{x_{i-1}x_i} = 1.$$

The set of paths permitted by R is denoted by P.

With X a subset of S,

$$\text{perm}(X)$$

will denote the set of all *permutations* of X, i.e., one-to-one maps of X onto itself. We let

$$y_1, y_2, \cdots y_{|S|}$$

be an arbitrary simple ordering of S, and we define the *ordinal number* $\omega(x)$ of a state $x \in S$ by the condition

$$\omega(x) = n \quad \text{if and only if} \quad x = y_n, \qquad n = 1, 2, \cdots, |S|.$$

For each m,n in the region $1 \leqq m,n \leqq |S|$, we define a function $\zeta_{mn}(\cdot)$ on the domain $1 \leqq i \leqq |S|$ by the condition

$$\zeta_{mn}(i) = \begin{cases} i & 1 \leqq i < \min(m,n) \\ 0 & i = m \quad \text{or} \quad i = n \\ i - 1 & \min(m,n) < i < \max(m,n) \\ i - 2 & \max(m,n) < i \leqq |S|. \end{cases}$$

We observe that $\zeta_{mn}(\cdot)$ has an inverse for each m and n. Now let $\varphi(\cdot)$ be a permutation of the set of states with the mth and the nth removed; then

$$i \rightarrow \zeta_{mn}\{\omega[\varphi(y_{\zeta_{mn}^{-1}(i)})]\} \qquad i = 1, 2, \cdots, |S| - 2$$

defines the permutation $a_{mn}(\varphi)$ *associated with* φ. Also, sgn $a_{mn}(\varphi)$ is $+1$ or -1 according as the permutation $a_{mn}(\varphi)$ is even or odd.

The "hang-up" matrix $H = (h_{xy})$ is defined by the condition

$$h_{xy} = \begin{cases} 1 & \text{if} \quad y \in B_x \\ 0 & \text{otherwise.} \end{cases}$$

Let x and z be states, and suppose that $\pi = \{x_0, \cdots, x_l\}$ is a path in P beginning at z and ending at x, so that $x_0 = z$ and $x_l = x$. Suppose also that the trajectory represented by π contains m new calls, i.e., there are exactly m values of i in the range $1 \leqq i \leqq l$ such that

$$x_i \in A_{x_{i-1}}.$$

Since π starts at z, in which $|z|$ calls are in progress, and ends up at x, in which there are $|x|$ calls in progress, it is evident that

$$l(\pi) = 2m + |z| - |x|.$$

The set of paths which start at a state z, never return to z, and end up at $x \neq z$, is denoted by

$$K_{zx}.$$

Thus π belongs to K_{zx} if and only if $x_0 = z$, $x_i \neq z$ for $0 < i \leqq l(\pi)$, and $x_l = x$.

Let $\pi = \{x_0, x_1, \cdots, x_l\}$ be a path on S, and let $f(\pi, \cdot)$ be a function defined for $x_0, x_1, \cdots, x_l$. In terms of $f(\pi, \cdot)$ and π, we define a product along the path π by the expression

$$\prod_{\substack{i=1 \\ \pi=\{x_0,x_1,\cdots,x_l\}}}^{l(\pi)} f(\pi, x_i).$$

It is convenient to abbreviate this product by the readily understandable expression

$$\prod_{x \epsilon \pi} f(\pi,x).$$

In the special case that $f(\pi,\cdot)$ has the form

$$f(\pi,x_i) = h(x_{i-1}, x_i), \qquad i = 1, \cdots, l$$

we abbreviate the product by

$$\prod_{x \epsilon \pi} h(px,x).$$

The notation $p\cdot$ is supposed to represent the *predecessor* of an element in π.

The first and simplest solution of the equilibrium equation $Qp = 0$ to be given is based on an observation made by I. W. Sandberg, namely, that $\det(Q)$, and hence $\det(Q')$, are zero, so that $Q'\operatorname{adj}(Q') = 0$, and thus columns of the matrix of cofactors of Q should give solutions of $Qp = 0$. The author has not succeeded in elucidating the probabilistic significance of these simple algebraic facts. It will be seen later that the other solutions to be given are, on the other hand, natural, plausible, or even obvious from a probabilistic viewpoint, but are algebraically involved.

Theorem 4: Let m be an integer in the range $1 \leqq m \leqq |S|$. *An unnormalized nonnegative solution p of* $Qp = 0$ *is given by*

$$p_{y_n} = (-1)^{|S|-1+m+n} \sum_{\varphi \epsilon \text{ perm } (S-\{y_m,y_n\})} \operatorname{sgn} a_{mn}(\varphi)$$

$$\prod_{\varphi(z)=z} (-|z| - \lambda s(z)) \prod_{\varphi(z)\neq z} (h_{z\varphi(z)} + \lambda r_{z\varphi(z)}).$$

Proof: Since $\det(Q) = 0$, it follows that $\det(Q') = 0$, the prime indicating the transposed matrix. Hence (Birkhoff and MacLane,[10] p. 290) no matter what ordering of S is used.

$$Q'\operatorname{adj}(Q') = 0,$$

where 'adj' denotes the adjoint matrix, i.e., the transposed matrix of cofactors. Let $C = (c_{xy})$ be the matrix of cofactors of Q corresponding to the ordering $y_1, y_2, \cdots, y_{|S|}$ of S, and suppose that the entries of Q are also arranged according to this ordering. Then

$$C = \operatorname{adj}(Q')$$

and we find that

$$\sum_{z \epsilon S} q_{zx} c_{zy} = 0.$$

Thus any column of the matrix C of cofactors of Q gives a solution of the equilibrium equations. It follows from a result of W. Ledermann (Bellman,[5] p. 294, exercise 10) that

$$(-1)^{|S|-1} c_{xy} \geqq 0.$$

We see that all the cofactors c_{xy} have the same sign, and each column of the matrix C yields a nonnegative solution of $Qp = 0$. Hence all columns are proportional, because there is only one nonnegative solution, up to normalization. The theorem follows from the standard formula for a cofactor as a determinant.

Theorem 5: If $z \epsilon S$ is any state, then an unnormalized solution of the statistical equilibrium equations $Qp = 0$ is given by

$$p_z = 1$$

and for $x \neq z$,

$$p_x = \sum_{\pi \epsilon P \cap K_{zx}} \lambda^{\frac{l(\pi)+|x|-|z|}{2}} \prod_{y \epsilon \pi} \frac{1}{|y| + \lambda s(y)}.$$

Proof: The formula given can be verified by direct substitution in the equations

$$[\lambda s(x) + |x|] p_x = \sum_{y \epsilon A_x} p_y + \lambda \sum_{y \epsilon B_x} p_y r_{yx}, \qquad x \epsilon S.$$

Convergence of the infinite sum will follow from our Theorem 7 and exercise 19, p. 378 of Feller.[4]

XI. STATIONARY PROBABILITY MEASURES FOR ERGODIC MARKOV PROCESSES

In order to shed light on Theorem 5 (and also to prove it by a probabilistic argument) we shall consider in this section the general problem of calculating the stationary probability measure of an ergodic continuous parameter Markov process on a finite number of states. Our object is to give an explicit formula for the measure in terms of the transition rate matrix. Again, it is needless to mention that a formula of such generality must be fairly complex. Applied to familiar Markov processes whose stationary measures are well known, the formula to be given yields some unexpected combinatory identities, not pursued here.

We shall now use the notations x_t, S, Q, and $P(\cdot)$ to describe an *arbitrary* Markov stochastic process x_t in continuous time, taking values in a finite set S of states with transition rate matrix $Q = (q_{xy})$ and transition probability matrices $P(t) = (p_{xy}(t))$, t real. It is assumed that there is a single ergodic class of states. Such a general interpretation of notations already introduced (for specific processes describing traffic in connecting networks) is made to avoid defining new terminology; it is made in this section only, and should cause no confusion.

If z is a state, a *return to* z is defined to be an epoch of time at which x_t reaches z, i.e., u is a return to z if for some $\epsilon > 0$, $x_t \neq z$ for $u - \epsilon < t < u$ and $x_t = z$ for $u < t < u + \epsilon$. A *departure from* z is an epoch of time at which x_t leaves z, i.e., u is a departure from z if for some $\epsilon > 0$, $x_t = z$ for $u - \epsilon < t < u$ and $x_t \neq z$ for $u < t < u + \epsilon$. A *return time* to z is a period of time elapsing between a departure from z and the next return to z. We set, for $t \geqq 0$,

$$H_z(t) = E\{\text{number of returns to } z \text{ in } (0,t] \mid x_0 = z\}$$

$$\mu_z = E\{\text{return time to } z\}$$

$$q_z = -q_{zz} = E^{-1}\{\text{length of a stay in } z\}$$

The notation $H_z(\cdot)$ has been chosen because the defined quantity has an obvious resemblance to the classical renewal function. (See Smith.[11])

There is a simple relationship between the equilibrium probability of a state x, and the quantities μ_x and q_x; this is expressed in the next theorem which, though probably familiar, is included for completeness.

Theorem 6: For $x \in S$, $p_x = [1 + q_x\mu_x]^{-1}$.

Proof: The transition probability $p_{xx}(t)$ approaches p_x as $t \to \infty$, and is expressible as

$$p_{xx}(t) = e^{-q_x t} + \int_0^t e^{-q_x(t-u)}\, dH_x(u).$$

Since stays in x and returns to x are all mutually independent, the stays being identically distributed, and the returns also, the renewal theorem [Smith,[11] p. 247, formula (1.3)] implies that the right side approaches

$$\frac{\displaystyle\int_0^\infty e^{-q_x t}\, dt}{E\left\{\begin{matrix}\text{interval between successive} \\ \text{returns to } x\end{matrix}\right\}} = \frac{1}{1 + q_x\mu_x}.$$

Thus p_x can be calculated from μ_x where

$$\begin{aligned}\mu_x &= \int_0^t u\, d\Pr\{\text{return time to } x \leqq u\} \\ &= \int_0^\infty \Pr\{\text{return time to } x > u\}\, du.\end{aligned}$$

For our purposes it is convenient to approach the calculation of p_x in a slightly different way. Let z be any state, and let x be a state distinct from z, $z \neq x$. Define $q_{zx}(t)$ in $t \geqq 0$ to be the probability that if the stochastic process start at z at time zero, it be at x at time t *without having returned to* z. Thus

$$q_{zx}(t) = \Pr\{x_t = x \quad \text{and} \quad (\text{epoch of first return to } z) > t \mid x_0 = z\}.$$

For convenience, we set $q_{zz}(t) \equiv 0$ in $t \geqq 0$.

Lemma 2: For $z \neq x$, $t \geqq 0$,

$$p_{zx}(t) = \int_0^t q_{zx}(t - u)\, dH_z(u).$$

Proof: Let t_i, $i = 1, 2, \ldots$. be the epoch of the ith return to z in $t > 0$, and let $A_i(t)$ be the event

$$\{x_t = x \quad \text{and} \quad t_i \leqq t < t_{i+1}\}.$$

Then

$$\Pr\{A_i(t) \mid x_0 = z\} = \int_0^t q_{zx}(t - u)\, d\Pr\{t_i \leqq u \mid x_0 = z\}.$$

However (cf. Ref. 11, p. 251, formula (1.7)),

$$H_z(t) = \sum_{i=1}^{\infty} \Pr\{t_i \leqq t \mid x_0 = z\}$$

and

$$p_{zx}(t) = \sum_{i=1}^{\infty} \Pr\{A_i(t) \mid x_0 = z\}.$$

The integration and the summation can be interchanged by the monotone convergence theorem, and the lemma follows.

Lemma 3: For $z \neq x$,

$$p_x = \frac{q_z}{1 + q_z \mu_z} \int_0^{\infty} q_{zx}(u)\, du.$$

Proof: The integral on the right exists, since

$$\int_0^{\infty} q_{zx}(u)\, du = E\{\text{time spent in } x \text{ between successive returns to } z\} \leqq \mu_z.$$

The lemma follows from Lemma 2 and the renewal theorem.

The matrix A is defined by the condition $A = (a_{xy})$ with

$$a_{xy} = \begin{cases} \dfrac{q_{xy}}{q_x} & x \neq y \\ 0 & x = y. \end{cases}$$

It can be verified that A is a stochastic matrix, indeed, the one-step transition probability matrix of a Markov stochastic process $\{x_n, n \text{ an integer}\}$ taking values on S; x_n is a discrete-time analog of x_t obtained by ignoring the lengths of time spent in a state.

Lemma 4: For $z \neq x$,

$$\int_0^\infty q_{zx}(u)\, du = \frac{1}{q_x} E\left\{\begin{matrix}\text{number of arrivals at } x \text{ between} \\ \text{successive returns to } z\end{matrix}\right\}.$$

Proof: The integral is the expected time spent in x between successive returns to z. Each stay in x has mean length $1/q_x$, and the stays are independent of the rest of the trajectory followed.

Lemma 5: For $z \neq x$,

$$E\left\{\begin{matrix}\text{number of arrivals at } x \text{ between} \\ \text{successive returns to } z\end{matrix}\right\}$$

$$= \sum_{n=1}^{\infty} \Pr\{x_n = x \quad \text{and} \quad x_j \neq z \quad \text{for} \quad 1 \leqq j \leqq n \mid x_0 = z\}.$$

Proof: We remark that the expectation on the left is the same for both x_t and x_n. The lemma is then a special case of the theorem that if $\{A_i, i = 1, 2, \ldots\}$ are any events, then the expected number of A_i that occur is

$$\sum_{i=1}^{\infty} \Pr\{A_i\}.$$

Lemma 6:

$$\mu_z = \sum_x \int_0^\infty q_{zx}(u)\, du.$$

Proof: This is an immediate consequence of $q_{zz}(\cdot) \equiv 0$ and

$$\int_0^\infty q_{zx}(u) = E\{\text{time spent in } x \text{ between successive returns to } z\}.$$

for $z \neq x$.

Lemma 7: Let $x \neq z$. *Then*

$$\Pr\{x_n = x \quad \text{and} \quad x_j \neq z \quad \text{for} \quad 1 \leqq j \leqq n \mid x_0 = z\}$$

$$= \frac{q_x}{q_z} \sum_{\substack{\pi \in K_{zx} \\ l(\pi)=n}} \prod_{y \in \pi} \frac{q_{py,y}}{q_y}.$$

Proof: The event in question can occur in as many ways as there are paths of length n in K_{zx}. The probability that $\{x_j, 0 \leqq j \leqq n\}$ follow a path π from z to x is

$$\frac{q_x}{q_z} \prod_{y \in \pi} \frac{q_{py,y}}{q_y},$$

the quotient in front correcting for the end points.

Combining Theorem 6 with Lemmas 4, 5, 6, and 7, we obtain the following explicit formula for the stationary probabilities:

Theorem 7: If $x \neq z$, then

$$p_x = \frac{1}{1 + q_z\mu_z} \sum_{\pi \in K_{zx}} \prod_{y \in \pi} \frac{q_{py,y}}{q_y}$$

with

$$q_z\mu_z = \sum_{x \neq z} \sum_{\pi \in K_{zx}} \prod_{y \in \pi} \frac{q_{py,y}}{q_y},$$

and

$$p_z = \frac{1}{1 + q_z\mu_z}.$$

We remark that Theorem 5 follows from the above if we choose $z = 0 =$ zero state, omit normalization, and observe that only products along permitted paths ($\pi \in P$) are nonzero. Theorem 7 is an analog for continuous parameter processes of a theorem of Derman[12] for Markov chains.

XII. EXPANSION OF THE STATIONARY VECTOR p IN POWERS OF λ

We now turn to examining, in an elementary way, the analytical dependence of the state probabilities $\{p_x, x \in S\}$ on the calling rate λ, on the structure of the network, and on the routing matrix R. It will be shown that the partial ordering $\leqq$ of the set S of states can be used to calculate the elements of p by expanding the ratios

$$\frac{p_x}{p_0} \qquad x > 0$$

in powers of the traffic parameter λ in a neighborhood of $\lambda = 0$, and then determining the coefficients of this expansion from the structure of the network and the routing matrix by a recursive procedure. The solution so obtained is later (Section XV) extended to arbitrary real positive values of λ by analytic continuation, and the coefficients are calculated.

Our approach to studying the stationary probability vector p will be guided by these intuitive remarks: it is known that in various simple models (of connecting systems carrying random traffic with blocked calls refused) the probability that k calls be in existence is proportional to the kth power of a constant associated with the calling rate divided by k factorial. For example, in Erlang's model for c trunks with blocked calls cleared, the chance that k calls are in progress is proportional to

$$\frac{a^k}{k!}, \qquad 0 \leqq k \leqq c$$

where a is the calling rate. Note that the exponent of a is the number of calls in progress, i.e., the current difference between the cumulative number of new calls and that of hang-ups, assuming that the system started in the zero state. The factorial in the denominator is the number of orders in which the k calls in progress could all hang up, or alternatively, could all have arisen.

The situation in our model is very similar. Each call still in progress required an event occurring at the rate λ to put it in existence; for each state x, there are exactly $|x|!$ orders in which the $|x|$ calls in progress in x could arise, or terminate. These circumstances suggest that for $x > 0$, p_x might be of order $\lambda^{|x|}$ as $\lambda \rightarrow 0$, and that the coefficient of $\lambda^{|x|}$ in p_x might involve $|x|!$ in the denominator. These conjectures are true, and are the first step in the systematic calculation of p_x by expansion in powers of λ, to be carried out in this section.

We first record some analytical properties of p as preliminary results. Some of these results could be obtained as consequences of the basic solutions given in Section X. Most of the proofs to be given, however, are independent of Section X, and proceed by simple arguments from the equilibrium equation.

When we need to view p_x as a function of the parameter λ, we write $p_x = p_x(\lambda)$, $x \in S$, or in vector form, $p = p(\lambda)$.

Lemma 8: $\lim_{\lambda \rightarrow 0} p_x(\lambda) = \delta_{x0}$.

Proof: Let x be a maximal state in the partial ordering $\leqq$ of the set S of all states. Then $s(x) = 0$, and

$$|x| p_x(\lambda) = \lambda \sum_{y \in B_x} p_y(\lambda) r_{yx} .$$

Since $0 \leqq p_y(\lambda) \leqq 1$ for all $\lambda > 0$ and all $y \in S$, the lemma is true for maximal states. Assume, as a hypothesis of induction, that the lemma is true for all y with $|y| \geqq k + 1$. Then for $x \in L_k$, $k > 0$,

$$[|x| + \lambda s(x)] p_x(\lambda) = \sum_{y \in A_x} p_y(\lambda) + \lambda \sum_{y \in B_x} p_y(\lambda) r_{yx} ,$$

and so $p_x(\lambda) \to 0$ as $\lambda \to 0$. The proof is completed by observing that for each $\lambda > 0$,

$$p_0(\lambda) = 1 - \sum_{x>0} p_x(\lambda).$$

Lemma 9: For each $x \in S$, p_x is the restriction to real positive argument of a rational function $p_x(\cdot)$ of a complex variable μ. The function $p_x(\cdot)$ has no poles in a neighborhood of the half-line $\operatorname{Re}(\mu) \geqq 0$, $\operatorname{Im}(\mu) = 0$, *and an expansion*

$$\frac{p_x(\mu + \epsilon)}{p_0(\mu + \epsilon)} = \sum_{m=0}^{\infty} \epsilon^m c_m(x,\mu),$$

with real coefficients $c_m(x,\lambda)$ is valid for $\operatorname{Re}(\mu) \geqq 0$, $\operatorname{Im}(\mu) = 0$, *and $|\epsilon|$ small enough.*

Proof: The equation $Qp = 0$ can be solved for a normalized (i.e., probability) vector $p(\lambda)$ by successive elimination or by use of Theorem 4. Either procedure gives rise to an algebraic expression for $p_x(\lambda)$, $x \in S$. Let $p_x(\mu)$ be that rational function of a complex variable μ defined by substituting μ for λ in this algebraic expression. Since $0 \leqq p_x(\lambda) \leqq 1$, $p_x(\mu)$ has no poles in a neighborhood of the nonnegative real axis. To justify the expansion we show that $p_x(\cdot)/p_0(\cdot)$ is also analytic in that neighborhood. But this is immediate because by Lemma 8,

$$p_x(\lambda) \to \delta_{0x} \quad \text{as} \quad \lambda \to 0,$$

and by Theorem 1, $p_0(\lambda) > 0$ for $\lambda \geqq 0$ because the zero state belongs to the ergodic class $S - F$.

Setting $\mu = 0$ and $\epsilon = \lambda$ in Lemma 9, we obtain an expansion of p_x/p_0 in powers of the traffic parameter λ,

$$\frac{p_x(\lambda)}{p_0(\lambda)} = \sum_{m=0}^{\infty} \lambda^m c_m(x,0),$$

valid for λ small enough.

Theorem 8: For $k \geqq 0$ and $x \in L_k$,

$$p_k = \sum_{x \in L_k} p_x = O(\lambda^k) \quad \text{as} \quad \lambda \to 0,$$

and

$$p_x = O(\lambda^k) \quad \text{as} \quad \lambda \to 0.$$

Proof: We prove both results simultaneously by induction. By Lemma 8, the result is true for $k = 0$. Assume that it is true up through $k - 1 \geqq 0$. From Lemma 1 and the induction hypothesis, we find

$$p_k = \frac{\lambda}{k} \sum_{y \epsilon L_{k-1}} p_y s(y)$$
$$= O(\lambda^k) \quad \text{as} \quad \lambda \to 0.$$

By Lemma 9, for λ small enough p_x can be written as a power series around $\lambda = 0$

$$p_x = \sum_{m=0}^{\infty} \lambda^m c_m(x,0).$$

Thus

$$\sum_{m=0}^{\infty} \lambda^m \sum_{x \epsilon L_k} c_m(x,0) = p_k = O(\lambda^k) \quad \text{as} \quad \lambda \to 0.$$

The first nonvanishing coefficient in the expansion on the left must be positive, else $p_k < 0$ for $\lambda > 0$ small enough, which is impossible since $p_x(\lambda) \geqq 0$ for $\lambda > 0$. Hence

$$\sum_{x \epsilon L_k} c_0(x,0) \geqq 0.$$

However, the first nonvanishing coefficient in the expansion of p_x must also be positive, for the same reason as above, namely, that $p_x(\lambda) \geqq 0$ for $\lambda \geqq 0$. Thus $c_0(x,0) \geqq 0$. Hence $p_k = O(\lambda^k)$ as $\lambda \to 0$ implies $c_0(x,0) = 0$. We apply the same argument successively to show that for $x \epsilon L_k$, the coefficients $c_1(x,0), \ldots, c_{k-1}(x,0)$ are all zero, and the theorem is proven.

Theorem 9: For $x > 0$

$$p_x = p_0 \frac{\lambda^{|x|}}{|x|!} r_x + o(\lambda^{|x|}) \quad \text{as} \quad \lambda \to 0$$

where

$r_x = (R^{|x|})_{0x}$
= *the* 0,x *entry of the* $|x|$th *power of the routing matrix* R
= *the number of permitted strictly ascending paths from* 0 *to* x.

Proof: The equation of statistical equilibrium that defines p is

$$[\,|x| + \lambda s(x)]p_x = \sum_{y \epsilon A_x} p_y + \lambda \sum_{y \epsilon B_x} p_y r_{yx}, \qquad x \epsilon S.$$

For convenience, suppose that $|x| = k$. We divide the equation by p_0, use Lemma 9 to expand the components of p/p_0 in powers of λ, and equate the coefficients of λ^k on each side of the equation. This gives

$$kc_k(x,0) + s(x)c_{k-1}(x,0) = \sum_{y \epsilon A_x} c_k(y,0) + \sum_{y \epsilon B_x} c_{k-1}(y,0) r_{yx}.$$

By Theorem 8, $c_{k-1}(x,0) = 0$ and $c_k(y,0) = 0$ for $y \in A_x$. Therefore

$$kc_k(x,0) = \sum_{y \in B_x} c_{k-1}(y,0) r_{yx},$$

or in general, with γ the vector with components $c_{|y|}(y,0)$,

$$\begin{aligned} c_{|x|}(x,0) &= \frac{1}{|x|} \sum_{y \in B_x} c_{|y|}(y,0) r_{yx} \\ &= \frac{1}{|x|} c_{|y|}(y,0) r_{yx} \\ &= \frac{1}{|x|} (R\gamma)_x . \end{aligned}$$

Iterating this relation $|x|$ times, we find

$$c_{|x|}(x,0) = \frac{1}{|x|!} (R^{|x|} \gamma)_x .$$

Now it is easily seen that the y,x entry of R^k is zero unless $k = |x| - |y|$, and in particular, if $k = |x|$, this entry is zero unless $y = 0$. Thus

$$c_{|x|}(x,0) = \frac{1}{|x|!} (R^{|x|})_{0x} c_0(0,0),$$

and it is obvious from the definition of the $c_m(y,0)$ that $c_0(0,0) = 1$.

Theorem 10: Let the sequences $\{c_m(x,0),\ m \geqq 0,\ x \in S\}$ be defined recursively by

$$\begin{aligned} c_m(0,0) &= \delta_{m0} \\ c_m(x,0) &= 0 \quad \textit{for} \quad 0 \leqq m < |x| \\ &\qquad \textit{and} \quad x > 0 \end{aligned}$$

$$c_{|x|}(x,0) = \frac{r_x}{|x|!}$$

$$|x| c_m(x,0) + s(x) c_{m-1}(x,0) = \sum_{y \in A_x} c_m(y,0) + \sum_{y \in B_x} c_{m-1}(y,0) t_{yx}$$

$$\textit{for} \quad m > |x| \quad \textit{and} \quad x > 0.$$

If for $x > 0$

$$\lambda < (\limsup_{m \to \infty} | c_m(x,0) |^{1/m})^{-1}$$

then the component p_x of p is given by

$$p_x = p_0 \sum_{m=0}^{\infty} \lambda^m c_m(x,0).$$

If

$$\lambda < \min_{x>0} \left(\limsup_{m\to\infty} | c_m(x,0) |^{1/m} \right)^{-1},$$

then the probability p_0 *of the zero state is determined by the normalization condition* $\sum_{x \epsilon S} p_x = 1$ *as*

$$p_0 = \frac{1}{1 + \sum_{x>0} \sum_{m=0}^{\infty} \lambda^m c_m(x,0)} .$$

Proof: This result follows immediately from Lemma 9 and Theorem 9, using the standard formula for the radius of convergence of a power series.

XIII. EXPANSION OF THE PROBABILITY OF BLOCKING IN POWERS OF λ

With a method of calculating equilibrium state probabilities for small λ at hand (in principle, at least) we now show how the probability b of blocking can be calculated, to any desired degree of accuracy, by an expansion in powers of the traffic parameter λ, assumed sufficiently small. In most connecting networks of practical interest, none of the states near the bottom of the state-diagram has any blocked calls, so that it is necessary for a state x to have certain minimum number of calls in progress before it can have any blocked idle pairs. To take advantage of this situation in our calculation, we let

$$n = \text{least } k \text{ such that some call is blocked in a state of } L_k .$$

Theorem 11: The probability b of blocking can be expanded in a power series in λ *in a neighborhood of* $\lambda = 0$; *only terms of order higher than or equal to* λ^n *appear.*

Proof: From Theorem 2 we have, since $\beta_x = 0$ for $|x| < n$, and $c_k(x,\lambda) = 0$ for $k < |x|$,

$$\begin{aligned} b &= \frac{\sum_{n \leqq |x|} p_x \beta_x}{\sum_{x \epsilon S} p_x \alpha_x} \\ &= \frac{\lambda^n \sum_{j=0}^{\infty} \lambda^j \sum_{n \leqq |x| \leqq n+j} c_{n+j}(x,0)\beta_x}{\sum_{j=0}^{\infty} \lambda^j \sum_{j \leqq |x|} c_j(x,0)\alpha_x} = \frac{\lambda^n B(\lambda)}{A(\lambda)} . \end{aligned} \tag{3}$$

Since the denominator is not zero in a neighborhood of $\lambda = 0$, $b = b(\lambda)$ is analytic there and can be expanded in powers of λ. Up to terms of order λ^{n+2} this expansion is

$$\begin{aligned} b = \lambda^n \frac{B(0)}{A(0)} &+ \lambda^{n+1} \frac{A(0)B'(0) - A'(0)B(0)}{[A(0)]^2} \\ &+ \lambda^{n+2}\left(\frac{A'(0)B''(0) - A''(0)B(0)}{2[A(0)]^2} \right. \\ &\quad \left. - \frac{A'(0)A(0)B'(0) - A'^2(0)B(0)}{[A(0)]^3}\right) \\ &+ o(\lambda^{n+2}). \end{aligned}$$

The coefficients in the first two terms can be obtained by the following calculations:

$$\begin{aligned} A(0) &= \sum_{|x| \geqq 0} c_0(x,0)\alpha_x \\ &= c_0(0,0)\alpha_0 \\ &= \alpha_0 , \\ B(0) &= \sum_{|x| \geqq n} c_n(x,0)\beta_x \\ &= \sum_{|x|=n} c_n(x,0)\beta_x \\ &= \frac{1}{n!} \sum_{x \epsilon L_n} r_x\beta_x , \\ A'(0) &= \sum_{|x| \geqq 1} c_1(x,0)\alpha_x \\ &= \sum_{|x|=1} c_1(x,0)\alpha_x \\ &= \sum_{x \epsilon L_1} r_{0x}\alpha_x , \\ B'(0) &= \sum_{|x| \geqq n+1} c_{n+1}(x,0)\beta_x \\ &= \sum_{x \epsilon L_{n+1}} c_{n+1}(x,0)\beta_x + \sum_{x \epsilon L_n} c_{n+1}(x,0)\beta_x \\ &= \frac{1}{(n+1)!} \sum_{x \epsilon L_{n+1}} r_x\beta_x + \sum_{x \epsilon L_n} c_{n+1}(x,0)\beta_x . \end{aligned}$$

The constants $\{c_{k+1}(x,0), |x| = k\}$ can be determined by the following recurrence, obtained from Theorem 10:

$$c_1(0,0) = 0,$$

$$c_{n+1}(x,0) = \frac{1}{n}\left\{\frac{1}{(n+1)!}\sum_{y\epsilon A_x} r_x + \sum_{y\epsilon B_x} c_n(y,0)r_{yx} - \frac{s(x)r_x}{n!}\right\}.$$

Our results can be put in a slightly more explicit form by expanding $\log b$ rather than b, and using the fact that $A(\cdot)$ and $B(\cdot)$, as defined by (3), are *generating functions*. We have

$$\log b = n \log \lambda + \log B(\lambda) - \log A(\lambda).$$

Except for the systematic absence of factorials, the coefficients in the expansion of $\log B(\lambda)$ are related to those in the expansion of $B(\lambda)$ as cumulants are to moments. Set

$$b_j = \sum_{n \leqq |x| \leqq n+j} c_{n+j}(x,0)\beta_x\,, \qquad j = 0, 1, \ldots,$$

$$a_j = \sum_{j \leqq |x|} c_j(x,0)\alpha_x\,, \qquad j = 0, 1, \ldots,$$

so that

$$B(\lambda) = \sum_{j=0}^{\infty} \lambda^j b_j\,,$$

$$A(\lambda) = \sum_{j=0}^{\infty} \lambda^j a_j\,.$$

Then, by a standard formula (Riordan,[13] p. 37),

$$\log B(\lambda) = \sum_{j=0}^{\infty} \lambda^j z_j(b),$$

$$\log A(\lambda) = \sum_{j=0}^{\infty} \lambda^j k_j(a),$$

where for $u = a$ or b (sequences)

$$\kappa_n(u) = \sum (u_1)^{k_1} \cdots (u_n)^{k_n} \frac{(-1)^{k-1}(k-1)!}{(k_1)! \cdots (k_n)!},$$

with $k = k_1 + k_2 + \cdots + k_n$, and the sum over all partitions of n, i.e., all solutions in nonnegative integers of $k_1 + 2k_2 + \cdots + nk_n = n$.

XIV. COMBINATORY INTERPRETATION AND CALCULATION OF THE CONSTANTS $\{c_m(x,0),\ x \epsilon S,\ m \geqq 0\}$

We shall now evaluate the coefficients in the power series expansion of p around $\lambda = 0$ explicitly as sums of products on paths in S. Additional combinatory notions that enter this calculation are discussed first.

A path is said to *contain a loop* if it returns (one or more times) to a place where it has been previously. Thus $\pi = \{x_0, \cdots, x_l\}$ has a loop if there are integers i and j, $0 \leqq i < j \leqq l$, such that

$$x_i = x_j .$$

A *circuit* is a path that ends where it begins. Thus the generic path $\pi = \{x_0, \ldots, x_l\}$ is a circuit if $x_0 = x_l$. A circuit is a *loop* if it contains no subcircuits, i.e., π is a loop if $x_0 = x_l$ and $i \neq j$ implies

$$x_i \neq x_j$$

for any $0 \leqq i \leqq l$ and $0 < j < l$.

A circuit of length two is of necessity a loop. Furthermore, the circuits of length two on S can be partitioned into two classes according to the direction in which they are traversed. Each such loop can be thought of as obtained either by first adding a new call and then removing that same call, or by removing a call and then replacing that same call. Thus if $\pi = \{x_0, x_1, x_2\}$ is a loop of length two on S, then

$$x_0 = x_2$$

and either

$$x_2 = x_0 \in B_{x_1}$$

or

$$x_2 = x_0 \in A_{x_1}$$

and, of course, not both. In the first instance we say that π is of the *first kind*. Thus a loop of the first kind is a path π of the form $\{x,y,x\}$ with $y \in A_x$, i.e., it is a trajectory in S obtained by starting at a state x, adding a new call to go to a state y, and then removing that very same call to return to x.

For a path $\pi = \{x_0, \ldots, x_l\}$ and a state x we say that x is *on* π if x is one of $x_0, x_1, \ldots, x_l$. If x is on π, we say that a loop of the first kind on π ends at x if for some i in the range $2 \leqq i \leqq l$,

$$x_i = x_{i-2} = x$$

and

$$x_{i-1} \in A_x ,$$

that is, the subpath $\{x_{i-2}, x_{i-1}, x_i\}$ is a loop of the first kind, and x_i is x.

With π and x as above we define

$$g(\pi,x) = \begin{cases} |x|, & \text{if } x \text{ is on } \pi \text{ and a loop of the first kind on } \pi \text{ ends at } x \\ 1, & \text{otherwise,} \end{cases}$$

and we denote by $\nu(\pi)$ the total number of loops of the first kind contained in π.

With these combinatory preliminaries behind us, we are ready to prove

Theorem 12: The coefficients $\{c_m(x,0),\, x \in S,\, m \geqq 0\}$ *are given by*

$$c_m(0,0) = \delta_{m0}\,,$$

and for $x > 0$,

$$c_m(x,0) = \sum_{\substack{\pi \in P \cap K_{0x} \\ l(\pi)=2m-|x|}} (-1)^{\nu(\pi)} \prod_{y \in \pi} \frac{g(\pi,y)}{|y|}. \tag{4}$$

Before proving the theorem it is probably helpful to state it in words, thus: To calculate $c_m(x,0)$, consider all the paths π that are permitted by R, start at 0, never return to zero, end up at x, and have length $2m - |x|$ (i.e., consist of m new calls and $m - |x|$ hangups); along each such path π take the product of the reciprocals of the numbers of calls in progress in the states traversed by π, omitting states at which a loop of the first kind ends; weight the product positive or negative according as π has an even or an odd number of loops of the first kind; add up all the weighted products.

Proof of Theorem 12: We already know that $c_m(x,0) = \delta_{m0}$ and that $c_m(x,0) = 0$ for $x > 0$ and $m < |x|$. The latter result is consistent with Theorem 12 because no path from 0 can reach x in fewer than $|x|$ steps. Consider then the case $m - |x| \geqq 0$. Any path π from 0 to x of length $2m - |x| = m$ consists entirely of new calls, and for such a π

$$\prod_{y \in \pi} \frac{g(\pi,y)}{|y|} = \frac{1}{|x|!}.$$

The number of such paths, summed over in (4), is easily seen to be r_x, the number of permitted strictly ascending paths from 0 to x. Thus for $m = |x|$, (4) states that

$$c_m(x,0) = \frac{r_x}{|x|!},$$

as was proven in Theorem 9.

The remainder of the proof is by upward induction on m and finite downward induction on x, using the recurrence formula

$$| x | c_m(x,0) + s(x)c_{m-1}(x,0) = \sum_{y \epsilon A_x} c_m(y,0) + \sum_{y \epsilon B_x} c_{m-1}(y,0)r_{yx} ,$$

given by Theorem 10. The starting step of the induction is the fact that formula (4) holds for $x \epsilon S$ and m such that

$$0 < | x | \leqq m \leqq 1.$$

This is a consequence of the fact, already proven, that (4) holds for $| x | = m$.

Let us assume, as a hypothesis of induction, that the theorem holds for all $x \epsilon S$ and m such that

$$0 < | x | \leqq m \leqq k. \tag{5}$$

We shall prove from this, by downward induction on x, that it also holds for $x \epsilon S$ and m such that

$$0 < | x | \leqq m = k + 1.$$

This last condition exactly describes the new cases covered in extending (5) to $k + 1$.

Where π is a path, we use the natural notation πx to denote the path obtained from π by adding x to π as a new ultimate element, assuming that x is adjacent to the last element of π. We now observe that if πx is a path that does not end in a loop of the first kind, then $g(\pi x,x) = 1$, and so

$$\prod_{z \epsilon \pi x} \frac{g(\pi x,z)}{| z |} = \frac{1}{| x |} \prod_{z \epsilon \pi} \frac{g(\pi,z)}{| z |}, \qquad \nu(\pi x) = \nu(\pi.) \tag{6}$$

A state x is *maximal* in the partial ordering $\leqq$ if no new calls can be put up in x, for whatever reason. If x is maximal, then A_x is empty and $s(x) = 0$. A state x is *maximal in (a set)* $X \subseteq S$ if $x \geqq y$ for every $y \epsilon X$, and $x \epsilon X$.

Let $0 \leqq | x | \leqq m = k + 1$ and suppose first that x is maximal. Then

$$c_m(x,0) = \frac{1}{| x |} \sum_{y \epsilon B_x} c_{m-1}(y,0)r_{yx} .$$

No path ending at a maximal state x can end in a loop of the first kind, and any such path must have a $y \epsilon B_x$ with $r_{yx} = 1$ as a penultimate element if it is to be a permitted path. Then clearly

$$\{\pi \in P \cap K_{0x}: l(\pi) = 2m - |x|\}$$
$$= \bigcup_{y \in B_x} \{\pi x: \pi \in P \cap K_{0y} \text{ and } l(\pi) = 2m - |x| - 1 \text{ and } r_{yx} = 1\}.$$

Let now $y \in B_x$, $\pi \in P \cap K_{0y}$, $l(\pi) = 2m - |x| - 1$, and $r_{yx} = 1$. Then πx satisfies $g(\pi x, x) = 1$ and also formula (6). Thus formula (4) holds for maximal x and $0 \leqq |x| \leqq m = k + 1$, by the hypothesis of induction.

Next, consider states x that are maximal in the set

$$\{y \in S: 0 \leqq |x| \leqq k + 1\}.$$

These are just the elements of L_{k+1}, i.e., the states x with $|x| = k + 1$. Since we are assuming $m = k + 1$, the result (4) holds for these x by Theorem 10.

Finally, assume as a hypothesis of downward induction (on $|x|$) that the result is true for $y \in S$ and m such that

$$1 < j + 1 \leqq |y| \leqq m = k + 1,$$

and suppose that $|x| = j$. Then

$$c_m(x,0) = \frac{1}{|x|} \sum_{y \in A_x} [c_m(y,0) - r_{xy} c_{m-1}(x,0)] + \frac{1}{|x|} \sum_{y \in B_x} c_{m-1}(y,0) r_{yx}.$$

If π is a path on S the notation $\text{pe}(\pi)$ denotes the *penultimate element* of π, i.e., $\text{pe}(\pi) = z_{l-1}$ for $\pi = \{z_0, z_1, \dots, z_l\}$. The notation $\text{ape}(\pi)$ denotes the *antepenultimate element* of π, i.e., for $\pi = \{z_0, z_1, \dots, z_l\}$, $\text{ape}(\pi) = z_{l-2}$.

A path of length $2m - |x|$ belonging to $P \cap K_{0x}$ reaches x either via A_x or via B_x. In the latter case the path cannot end in a loop of the first kind. By the hypothesis of induction,

$$\frac{1}{|x|} \sum_{y \in B_x} c_{m-1}(y,0) r_{yx} = \sum_{y \in B_x} r_{yx} \sum_{\substack{\pi \in P \cap K_{0y} \\ l(\pi) = 2(m-1) - |y|}} \frac{(-1)^{\nu(\pi)}}{|x|} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|}$$
$$= \sum_{\substack{\pi \in P \cap K_{0x} \\ l(\pi) = 2m - |x| \\ \text{pe}(\pi) \in B_x}} (-1)^{\nu(\pi)} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|}, \tag{7}$$

the second equality following from (6).

Now consider a path πx of length $2m - |x|$, belonging to $P \cap K_{0x}$, reaching x via A_x, and not ending in a loop of the first kind, i.e., with $\text{pe}(\pi) \neq x$. Using the hypothesis of downward induction, we find

$$\frac{1}{|x|} \sum_{y \in A_x} \sum_{\substack{\pi \in P \cap K_{0y} \\ l(\pi)=2m-|x|-1 \\ \text{pe}(\pi) \neq x}} (-1)^{\nu(\pi)} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|} = \sum_{\substack{\pi \in P \cap K_{0x} \\ \text{pe}(\pi) \in A_x \\ \text{ape}(\pi) \neq x}} (-1)^{\nu(\pi)} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|}, \tag{8}$$

the second equality again following from (6).

Finally, we consider those paths of length $2m - |x|$ belonging to $P \cap K_{0x}$ which *do* end in a loop of the first kind. Such a path is of the form

$$\pi y x$$

with $\pi \in P \cap K_{0x}$, $l(\pi) = 2m - |x| - 2$, and $r_{xy} = 1$, for some $y \in A_x$. We observe that in this case $|y| = |x| + 1$ and that

$$\begin{aligned} \frac{(-1)^{\nu(\pi)}}{|x|} \left\{ \prod_{z \in \pi y} \frac{g(\pi y,z)}{|z|} - \prod_{z \in \pi} \frac{g(\pi,z)}{|z|} \right\} &= \frac{(-1)^{\nu(\pi)}}{|x|} \left\{ \frac{1}{|y|} - 1 \right\} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|} \\ &= -\frac{(-1)^{\nu(\pi)}}{|y|} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|} \\ &= (-1)^{\nu(\pi)+1} \frac{g(\pi y x,y)}{|y|} \\ &\quad \cdot \frac{g(\pi y x,x)}{|x|} \cdot \prod_{z \in \pi} \frac{g(\pi,z)}{|z|} \\ &= (-1)^{\nu(\pi)+1} \prod_{z \in \pi y x} \frac{g(\pi y z,z)}{|z|}. \end{aligned} \tag{9}$$

We note that $\nu(\pi y x) = \nu(\pi) + 1$, and that $\nu(\pi y) = \nu(\pi)$. By summing formula (9) over paths of length $2m = |x| - 2$ belonging to $P \cap K_{0x}$ *and* over $y \in A_x$ such that $r_{xy} = 1$, we obtain

$$\begin{aligned} \frac{1}{|x|} \sum_{\substack{y \in A_x \\ r_{xy}=1}} \sum_{\substack{\pi \in P \cap K_{0x} \\ l(\pi)=2m-|x|-2}} &\left\{ (-1)^{\nu(\pi t)} \prod_{z \in \pi y} \frac{g(\pi y,z)}{|z|} \right. \\ &\left. - (-1)^{\nu(\pi)} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|} \right\} = \sum_{\substack{\pi \in P \cap K_{0x} \\ l(\pi)=2m-|x| \\ \text{pe}(\pi) \in A_y \\ \text{ape}(\pi)=x}} (-1)^{\nu(\pi)} \prod_{z \in \pi} \frac{g(\pi,z)}{|z|}. \end{aligned} \tag{10}$$

By the hypothesis of induction

$$\frac{1}{|x|}\sum_{\substack{y\epsilon A_x \\ r_{xy}=1}}\sum_{\substack{\pi\epsilon P\cap K_{0x} \\ l(\pi)=2m-2-|x|}}(-1)^{\nu(\pi)}\prod_{z\epsilon\pi}\frac{g(\pi,z)}{|z|}=\frac{1}{|x|}\sum_{y\epsilon A_x} r_{xy}c_{m-1}(x,0). \quad (11)$$

Also, it can be seen that for $y \epsilon A_x$

$$\begin{aligned}&\{\pi \epsilon P\cap K_{0y}: l(\pi) = 2m - |x| - 1 \quad \text{and} \quad \mathrm{pe}(\pi) \neq x\}\\ &\cup \{\pi y: \pi \epsilon P\cap K_{0x}, r_{xy} = 1, \quad \text{and} \quad l(\pi) = 2m - |x| - 2\} \quad (12)\\ &= \{\pi \epsilon P\cap K_{0y}: l(\pi) = 2m - |y|\}.\end{aligned}$$

Combining (8) and (10), using (11) and (12), and applying the hypothesis of downward induction, we find

$$\begin{aligned}&\sum_{\substack{\pi\epsilon P\cap K_{0x} \\ l(\pi)=2m-|x| \\ \mathrm{pe}(\pi)\epsilon A_x}}(-1)^{\nu(\pi)}\prod_{z\epsilon\pi}\frac{g(\pi,z)}{|z|}\\ &= \frac{1}{|x|}\sum_{y\epsilon A_x}\sum_{\substack{\pi\epsilon P\cap K_{0y} \\ l(\pi)=2m-|y|}}(-1)^{\nu(\pi)}\prod_{z\epsilon\pi}\frac{g(\pi,z)}{|z|} - \frac{1}{|x|}\sum_{y\epsilon A_x} r_{xy}c_{m-1}(x,0) \quad (13)\\ &= \frac{1}{|x|}\sum_{y\epsilon A_x}\{c_m(y,0) - r_{xy}c_m(x,0)\}.\end{aligned}$$

Together, (7) and (13) complete the inductive step.

XV. CALCULATION OF $c_m(x,\lambda)$

In order to give an explicit formula for $c_m(x,\lambda)$ for $\lambda > 0$ we suppose that, for each path π on S, all the upward (in $\leqq$) transitions on π are of two kinds, denoted by the symbols λ and ϵ. (The calculation we present is more easily understood if the new calls labeled ϵ are thought of as due to the increment ϵ in calling rate, while those labeled λ are due to the original calling rate λ.) In other words, we consider the set of all paths π on S as (partially) labeled by assigning either λ or ϵ to each upward transition. Formally, we define a *labeling* $\lambda(\cdot)$ of a path $\pi = \{x_0, x_1, \ldots, x_l\}$ to be any function defined for $x_1, x_2, \ldots, x_l$ with the property

$$\lambda(x_i) = \begin{cases}\epsilon \quad \text{or} \quad \lambda & \text{if} \quad x_i \text{ covers } x_{i-1},\\ 0 & \text{otherwise.}\end{cases}$$

The set of *all possible labelings* of a path π is denoted by $\Lambda(\pi)$, and membership therein by the notation $\lambda(\cdot) \epsilon \Lambda(\pi)$. The *functions* $\lambda(\cdot)$

should not be confused with the *constant* λ. A path π together with a labeling $\lambda(\cdot)$ of π will be called a *labeled path* and denoted by $(\pi,\lambda(\cdot))$. The *index* $\epsilon(\pi,\lambda(\cdot))$ of a labeled path is the number of times $\lambda(x)$ assumes the value ϵ for $x \in \pi$.

Let $\lambda(\cdot) \in \Lambda(\pi)$ be a labeling of $\pi = \{x_0, x_1, \ldots, x_l\}$. The function $h(\pi,\lambda(\cdot),\cdot)$ is defined by the condition

$$h(\pi,\lambda(\cdot),x) = |x| + \lambda s(px)$$

if x is on π and a loop of the first kind on π ends at x, with the first (i.e., upward going) leg of the loop labeled ϵ by $\lambda(\cdot)$, i.e.,

$$\lambda(px) = \epsilon,$$

and by

$$h(\pi,\lambda(\cdot),x) = 1$$

in all other cases.

For a path π and a labeling $\lambda(\cdot) \in \Lambda(\pi)$, the function $\zeta(\pi,\lambda(\cdot))$ is defined by

$$\zeta(\pi,\lambda(\cdot)) = \begin{cases}\text{the number of loops of the first kind on } \pi\\ \text{labeled } \epsilon \text{ on the upward leg by } \lambda(\cdot).\end{cases}$$

Theorem 13: For $\lambda > 0$ and $m \geqq 0$

$$c_0(0,\lambda) = \delta_{0m}$$

and for $x > 0$,

$$c_m(x,\lambda) = \lambda^{-m} \sum_{\substack{\pi \in P \cap K_{0x}\\ \lambda(\cdot) \in \Lambda(\pi)\\ \epsilon(\pi,\lambda(\cdot))=m}} \lambda^{(l(\pi)+|x|)/2}(-1)^{\zeta(\pi,\lambda(\cdot))} \prod_{y \in \pi} \frac{h(\pi,\lambda(\cdot),y)}{|y| + \lambda s(y)}. \quad (14)$$

Note that by Theorem 5, with $z = 0$, formula (14) for $m = 0$ reduces, as it should, to

$$c_0(x,\lambda) = \frac{p_x(\lambda)}{p_0(\lambda)}.$$

Also, formula (14) agrees with formula (4), Theorem 12, as λ is allowed to approach zero, since in this limit only $(\pi,\lambda(\cdot))$ which are full of ϵ's contribute, with $m = \frac{1}{2}(l(\pi) + |x|)$.

Proof of Theorem 13: The values of $c_0(x,\lambda)$ and $c_m(0,\lambda)$ are consequences of the definition

$$\sum_{m=0}^{\infty} \epsilon^m c_m(x,\lambda) = \frac{p_x(\lambda + \epsilon)}{p_0(\lambda + \epsilon)}. \quad (15)$$

The equilibrium condition $Qp = 0$ for traffic parameter $\lambda + \epsilon$ comprises exactly the equations

$$[\,|x| + (\lambda + \epsilon)s(x)]p_x(\lambda + \epsilon) = \sum_{y \epsilon A_x} p_y(\lambda + \epsilon) + (\lambda + \epsilon) \sum_{y \epsilon B_x} p_y(\lambda + \epsilon) r_{yx}, \qquad x \epsilon S.$$

Dividing by $p_0(\lambda + \epsilon) > 0$, substituting the expansion (15), and collecting coefficients of like powers of ϵ, we find that $c_m(x,\lambda)$ for $x \epsilon S$ satisfies the equation

$$[\,|x| + \lambda s(x)]c_m(x,\lambda) = \sum_{y \epsilon A_x} [c_m(y,\lambda) - r_{xy}c_{m-1}(x,\lambda)] + \sum_{y \epsilon B_x} r_{yx}[c_{m-1}(y,\lambda) + \lambda c_m(y,\lambda)]. \tag{16}$$

It can be verified, using the fact that for any labeling $\lambda(\cdot)$ of a path π the number of times $\lambda(x)$ has the value λ for $x \epsilon \pi$ is

$$\frac{l(\pi) + |x|}{2} - \epsilon(\pi,\lambda(\cdot)),$$

that formula (14) gives a formal solution of equations (16).

To prove the theorem it suffices to show that the infinite sum over $\pi \epsilon P \cap K_{0x}$ in formula (14) is absolutely convergent, and that the left-hand side of formula (15), with the $c_m(x,\lambda)$ as given by the theorem is absolutely convergent for ϵ small enough.

We first observe that for each π and $\lambda(\cdot)$ summed over in formula the factors $h(\pi,\lambda(\cdot),y)$ in

$$\prod_{y \epsilon \pi} \frac{h(\pi,\lambda(\cdot),y)}{|y| + \lambda s(y)}$$

are uniformly bounded, and that at most $\min(m,\nu(\pi))$ of them are greater than unity. Also, the number of upward transitions (new calls) along a path $\pi \epsilon P \cap K_{0x}$ of length $l(\pi)$ is just

$$\frac{l(\pi) + |x|}{2}.$$

Thus there are exactly

$$\binom{\frac{1}{2}l(\pi) + \frac{1}{2}|x|}{m}$$

ways of labeling the upward transitions on a path $\pi \epsilon P \cap K_{0x}$ with length $l(\pi)$ and index m.

Hence for some constant $a > 0$

$$| c_m(x,\lambda) | \leqq a^m \sum_{\pi \in P \cap K_{0x}} \lambda^{(l(\pi)+|x|)/2} \binom{\frac{l(\pi) + | x |}{2}}{m} \prod_{y \in \pi} \frac{1}{| y | + \lambda s(y)}.$$

By Lemma 7, with $\{x_i, i \text{ an integer}\}$ of the lemma defined in terms of the matrix Q appropriate to our congestion problem,

$$\sum_{x>0} \frac{| x | + \lambda s(x)}{\lambda s(0)} \sum_{\substack{\pi \in P \cap K_{0x} \\ l(\pi)=k}} \lambda^{(k+|x|)/2} \prod_{y \in \pi} \frac{1}{| y | + \lambda s(y)} = \Pr\{x_i \neq 0 \quad \text{for} \quad 1 \leqq i \leqq k \mid x_0 = 0\}.$$

By ex. 19, p. 378 of Feller,[4] there exists $0 < q < 1$ such that the probability on the right is at most q^k for all $k \geqq |S|$. Hence

$$\sum_{x>0} | c_m(x,\lambda) | \leqq \text{const. } a^m \sum_{k=m} \binom{k}{m} q^k.$$

This proves that (14) converges absolutely, and that the left side of (15) converges absolutely for $|\epsilon|$ small enough.

REFERENCES

1. Beneš, V. E., Heuristic Remarks and Mathematical Problems Regarding the Theory of Connecting Systems, B.S.T.J., **41,** July, 1962, pp. 1201–1247.
2. Beneš, V. E., Algebraic and Topological Properties of Connecting Networks, B.S.T.J., **41,** July, 1962, pp. 1249–1274.
3. Beneš, V. E., A "Thermodynamic" Theory of Traffic in Connecting Networks, B.S.T.J., **42,** May, 1963, pp. 567–607.
4. Feller, W., *An Introduction to Probability Theory and its Applications*, **1,** 2nd Ed., John Wiley & Sons, New York, 1957.
5. Bellman, R., *Introduction to Matrix Analysis*, McGraw-Hill Book Co., Inc., New York, 1960.
6. Halmos, P. R., *Finite Dimensional Vector Spaces*, 2nd Ed., D. Van Nostrand Company, Inc., Princeton, 1958.
7. Morris, R., and Wolman, E., A Note on 'Statistical Equilibrium,' Operations Research, **9,** 1961, pp. 751–753.
8. Syski, R., *Introduction to Congestion Theory in Telephone Systems*, Oliver and Boyd, London, 1960.
9. Beneš, V. E., Theoretical Problems of Optimal Routing in Communications Networks, to appear.
10. Birkhoff, G., and MacLane, S., *A Survey of Modern Algebra*, The Macmillan Company, New York, 1950.
11. Smith, W. L., Renewal Theory and its Ramifications, J. Roy. Stat. Soc., Ser. B., **20,** 1958, pp. 243–302.
12. Derman, C., A Solution to a Set of Fundamental Equations in Markov Chains, Proc. Amer. Math. Soc., **5,** 1954, pp. 332–334.
13. Riordan, J., *An Introduction to Combinatorial Analysis*, John Wiley & Sons, New York, 1958.