https://ntrs.nasa.gov/search.jsp?R=19710024075 2018-07-24T20:10:41+00:00Z

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

NGL 21-002-008

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND

FACILITY FORM 602

N71-33551
(ACCESSION NUMBER)
(THRU)
41
(PAGES)
G3
(CODE)
CR 121429
(NASA CR OR TMX OR AD NUMBER)
08
(CATEGORY)

JUL 1971
RECEIVED
NASA STI FACILITY
INPUT BRANCH

Technical Report TR-161 June 1971
AFOSR 71-1982

A CRITICAL STATISTIC FOR CHANNELS WITH MEMORY

by

J-P. A. Adoul, B. D. Fritchman and L. N. Kanal

A CRITICAL STATISTIC FOR CHANNELS WITH MEMORY+

by

J-P. A. Adoul*, B. D. Fritchman** and L. N. Kanal***

Accepted for publication by IEEE Transactions on Information Theory.

+ Research supported by the Mathematical and Information Sciences Directorate, Air Force Office of Scientific Research, Air Force Systems Command, Arlington, Va., under grants Af-AFOSR 68-1390B to Lehigh University and AFOSR 71-1982 to the University of Maryland.

* Department of Electrical Engineering, University of Sherbrooke, Sherbrooke, Province of Quebec, Canada.

** Department of Electrical Engineering, Lehigh University, Bethlehem, Pa. 18015.

*** Computer Science Center, University of Maryland, College Park, Md. 20742.

ABSTRACT

We present a new descriptive statistic for channels with memory and show its utility (a) in evaluating and comparing existing models for such channels, and (b) as a theoretical tool in defining the error-gap distribution characteristics of real channels. We demonstrate that certain kinds of real channel behavior cannot be adequately described by previously proposed models and offer an example of a better model which includes many of the earlier models as special cases.

I. Introduction

In recent years many models have been proposed to characterize the error sequences encountered in real digital communication links [1-10]. For the most part these models have been developed to represent certain expermentally measured statistics, though inferences are often made about their applicability to more general situations. Unfortunately it is not always easy to ferret out all of the implicit assumptions about a model in order to determine its applicability to a specific channel. Furthermore, different models which apparently represent the same channel can lead to different conclusions about the behavior of communication processes. Thre reason is that the models do not in reality adequately represent some statistic which is critical in the analysis.

To circumvent such difficu'ties in this paper, we examine a number of basic properties which a model must satisfy if it is to adequately represent a real channel or class of real channels. A consequence of our analysis is the demonstration that certain kinds of real channel behavior cannot be properly described by previously developed models. We then offer an alternative model as an example. This model is a function of a slowly spreading Markov chain and coincidently includes many of the previously proposed models as special cases.

Though a number of important statistical properties of real channels are uncovered in the following analysis, it is by no means exhaustive. It should, however, go a long way toward the improvement of the process of selecting models to represent various classes of real communication channels. It should also lend insight into the varied behavior of real channels.

Consider the channel model shown in Figure 1. It consists of an input and output alphabet whose symbols are the q elements of the Galois field GF(q). Between the input and output the channel introduces a random discrepancy which is represented mathematically as the addition of a noise symbol, i.e.,

$$y_i = x_i + n_i, \tag{1}$$

where y_i and x_i are respectively the output and input symbols, n_i is the noise symbol which is also an element of GF(q) and "+" is addition over the same field.

The noise sequence $\{n_i\}$ can be thought to originate from a hypothetical random generator called the noise source. Throughout it will be assumed that the input sequence $\{x_i\}$ and noise sequence $\{n_i\}$ are statistically independent of each other, implying that the statistical properties of the channel are exhibited in the statistical properties of the noise source.

Whenever n_i is different from zero, an error is said to occur. In this way the error source is distinguished from the noise source. The channel error source generates the error sequence $\{e_i\}$ which is a mapping of the noise sequence $\{n_i\}$ onto $\{0,1\}$, i. e.,

$$e_i = e(n_i) = \begin{cases} 0 & \text{if } n_i=0 \\ 1 & \text{otherwise.} \end{cases} \tag{2}$$

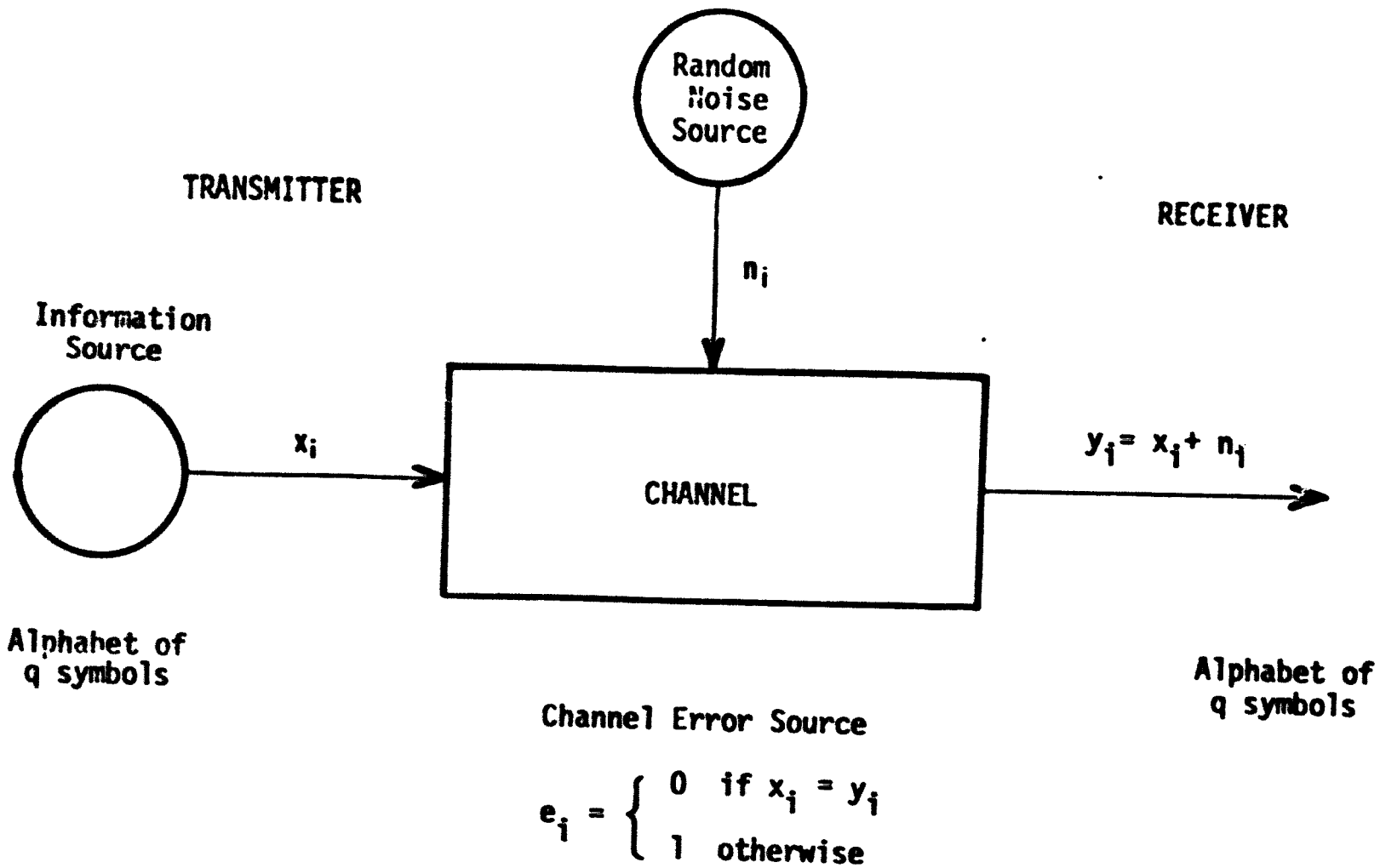


FIG. 1 DIGITAL COMMUNICATION CHANNEL

In the common case of binary transmission the noise and error sequences are equivalent. Both the error source and noise source are discrete-time, stochastic processes.

II. The Error-Gap Process.

The state space of the error process is composed of two states: "Error" or "1" and "Error Free" or "0". A positive integer exponent is used to indicate the number of consecutive symbols of the same type. In this notation the sequence 10000000010001100 is written as $10^8 10^3 1^2 0^2$.

In probability expressions we use the notation,

$$\Pr\{e_o=0,\ e_1=0,\ldots,e_{n-1}=0,\ e_n=1\} = P(0^{n-1}1) \quad (3)$$

A conditional probability is also written in terms of sequences say Pr{sequenceB|sequenceA}. If no other indication is given in the conditional part, it is understood sequence B directly follows sequence A.

Sequences of zeros between two errors ("1" states) are called error gaps (also error-free runs). The length of a gap is defined as one plus the total number of zeros in the sequence between two 1's. By defining the gap length in this way the sum of all gap lengths equals the total length of the error sequence.

Corresponding to the error process, the gap process $\{G_n\}$ can be introduced by treating the binary, error process as a succession of gaps of length G_n. The state space of this new process is the denumerable set of positive integers.

The probability

$$Pr\{G_n=m\} = P(X_{n+1}= 0, \ldots X_{n+m-1}= 0, X_{n+m} = 1|X_n= 1) \quad (4)$$

for all positive integers m is called the error gap probability mass function (EGPMF).

Assuming stationarity we have

$$Pr\{G_n=m\} = P(0^{m-1}1|1). \quad (5)$$

Let

$$F(m+1) \triangleq Pr\{G_n \geq m+1\} = P(0^m|1), \quad (6)$$

where

$$P(0^m|1) = \sum_{k=m}^{\infty} P(0^k1|1). \quad (7)$$

$P(0^m|1)$ is called the error-gap or error-free run distribution (EGD). From the definition it is observed that $P(0^m|1)$ is a monotonically decreasing function of m. Moreover $P(0^m1|1) = P(0^m|1) - P(0^{m+1}|1)$ and the expectation

$$\begin{aligned} E(G_n) &= \sum_{m=0}^{\infty} (m+1)\ P(0^m1|1) \\ &= \sum_{m=0}^{\infty} (m+1)[P(0^m|1) - P(0^{m+1}|1)] \quad (8) \\ &= \sum_{m=0}^{\infty} P(0^m|1) = \sum_{m=0}^{\infty} \frac{P(10^m)}{P(1)}. \end{aligned}$$

From the stationarity assumption it follows that $P(10^i) = P(0^i1)$. As $P(0^i1)$ is the probability that beginning with any symbol in the error sequence the first error will not be encountered for i symbols, the events i = 0, 1, 2, . . . are mutually

exclusive and exhaustive, Thus if and only if $P(0^{\infty}1) = 0$,

$$\sum_{m=0}^{\infty} P(10^m) = 1 . \tag{9}$$

Therefore

$$E(G_n) = \frac{\sum_{m=0}^{\infty} P(10^m)}{P(1)} = \frac{1}{P(1)} , \tag{10}$$

where $P(1)$ is the probability of error and is equal to $E(e_n)$. For real channels the probability of an error is greater than zero, which implies that

$$E(G_n) = \sum_{m=0}^{\infty} P(0^m|1) < \infty. \tag{11}$$

$E(G_n)$ is the expected number of symbols between two errors, i. e., the average number of symbols which will be transmitted before the recurrence of an error. Consequently, when $E(G_n)$ exists and is independent of n, it will be called the recurrence time, R_t. On the average, the number of non-errors associated with a single event G_n is R_t, which implies the number of events in the corresponding gap process is reduced by a factor R_t over the number of events in the binary error process. For good communication channels the probability of error, $P(1)$, is smaller than 10^{-3}, which means $R_t \geq 10^3$. For such channels the number of events dealt with in the corresponding gap process is reduced by a factor of at least 10^3 over the events in the error process.

III. The Descriptive Statistic.

We now introduce a new statistic, defined as the slope of the error-gap distribution plotted in logarithmic coordinates. We show how it can be used (a) in evaluating and comparing existing models both of the generative and descriptive type, and (b) as a theoretical tool in defining the error-gap distribution characteristics of real channels.

The error-gap distribution is often determined during the measurement of channel statistics. Consider a monotonically decreasing continuous function $\tilde{P}(0^m|1)$ which for all non-negative integers m is equal to $P(0^m|1)$; a sketch of $\tilde{P}(0^m|1)$ as a function of m is shown in Figure 2. Suppose this function is plotted in logarithmic coordinates as follows:

$$y = \log_e \tilde{P}(0^m|1), \tag{12}$$

$$x = \log_e m. \tag{13}$$

Now define the function

$$\alpha(x) \triangleq -\frac{dy(x)}{dx}. \tag{14}$$

Note that in this x-y coordinate system, the terms of the harmonic series, i.e., $P(0^m|1) = \frac{1}{m}$, fall on a straight line of slope equal to -1.

Now it is an elementary fact that if $|a_m| \le c_m$ for $m \ge m^*$, where m^* is some fixed integer, and if $\sum c_m$ converges, then $\sum a_m$ converges. Also, for non-negative a_m, if $a_m \ge d_m \ge 0$ for $m \ge m^*$, and if $\sum d_m$ diverges, then $\sum a_m$ diverges. But, as just demonstrated in (11) for real channels $E(G_n) < \infty$.

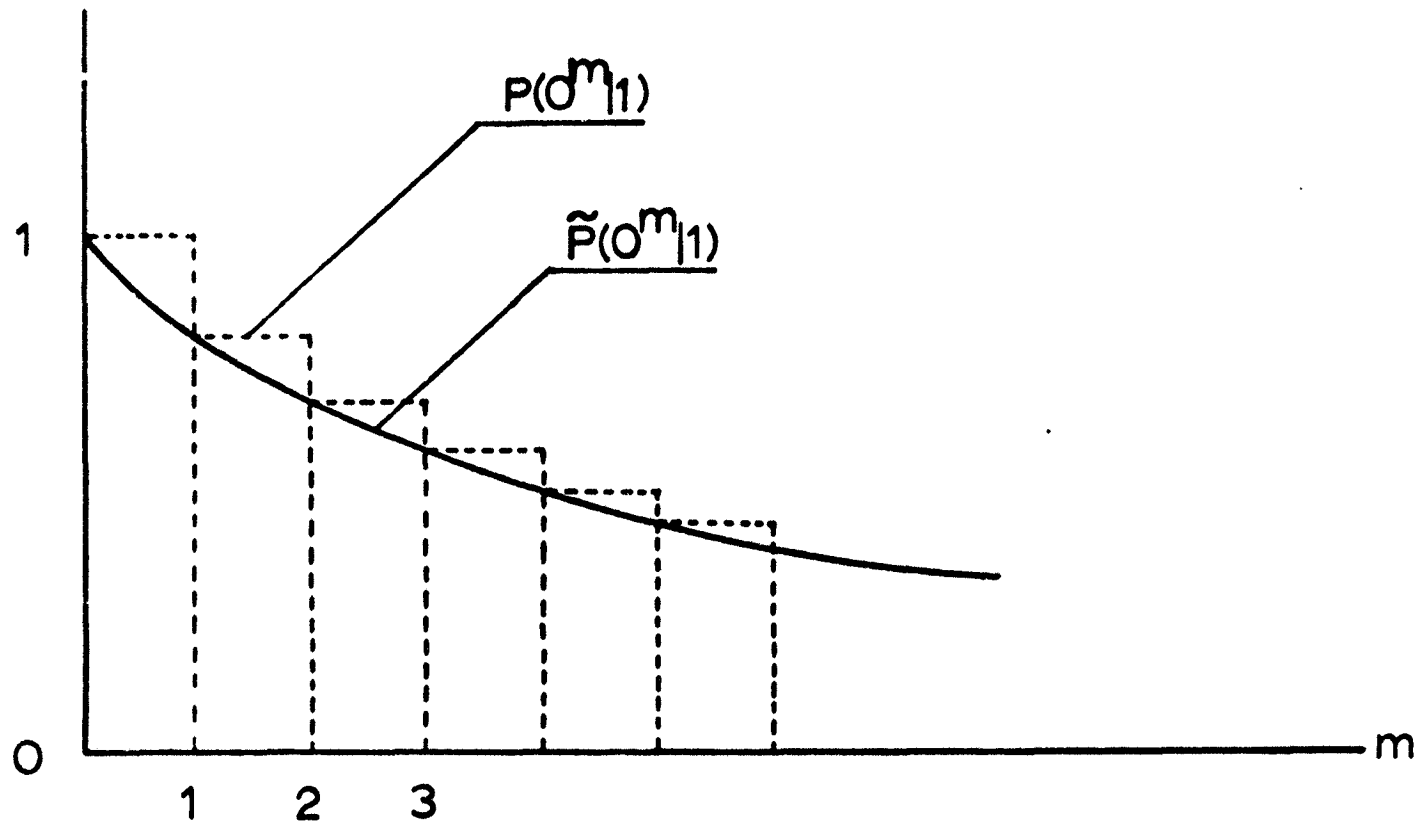


FIG. 2: FUNCTION EQUAL TO $P(0^m|1)$ AT INTEGER m.

Therefore any model having an EGD $P(0^m|1)<1/m^{(1+\epsilon)}$, with $\epsilon>0$ a fixed constant, after some value of $m=m^*$, will have an x-y characteristic which is asympotically less than -1, as shown in Figure 3, and will satisfy Equation (11). The condition that an error source model have an α-function which is asymptotically greater than 1 is sufficient to guarantee convergence of $E(G_n)$. A necessary condition for convergence is obtained in Section VII.

Since $P(0^m|1)$ is always a monotonically decreasing function of m, the slope of $P(0^m|1)$ will always be negative, implying $\alpha(x)$ will be a non-negative function of x. The α-function for the harmonic series equals unity for all values of x, and so if $P(0^m|1)<1/m^{(1+\epsilon)}$ for all $m>m^*$, the α-function will fall in the region $\alpha(x)>1$, for $x>x^*= \log_e m^*$.

III. The α-Function of Some Special Models

The preceding considerations suggest that real channels might be expected to yield α-functions which asymptotically take on values greater than 1. Examination of the α-function of some models proposed to represent real channels will bring to light some interesting behavior.

Binary Symmetric Channel

The simplest model and the one most commonly used to represent error sequences is the Binary Symmetric Channel (BSC). If $p = 1-q$ is the probability of an error of either type ($1\to0$ or $0\to1$), $P(0^m|1) = q^m$ and $y = \log_e P(0^m|1)=m \log_e q$. Since $x = \log_e m$, $m = e^x$ and $y = e^x \log_e q$, we have

$$\alpha(x) = -\frac{dy}{dx} = (\log_e \frac{1}{q}) e^x \tag{15}$$

The factor $\log_e 1/q$ is positive and constant, and so the BSC has an exponential α-function.

Pareto Model

Berger and Mandelbrot[3] proposed a model with an EGD given by a Pareto distribution, i.e., they let

$$P(0^m|1) = \frac{1}{m^\theta} \tag{16}$$

where θ is a positive constant. Computation of $\alpha(x)$ for this case yields the value θ.

The authors of [3] claim that for small m, the Pareto distribution is a good approximation to actual measured distributions. From experimental measurements they found values of $\theta = 0.5$ as typical. Sussman [11] also found values of $\theta = 0.11$ and 0.3 to represent some channel measurements.

The previous discussion showed channels with $\lim_{x\to\infty} \alpha(x) < 1$ do not have finite recurrence times, since the series $\sum_m P(0^m|1)$ diverges. Berger and Mandelbrot resolved this problem by letting θ take on a new constant value greater than unity at some point $m = m^*$; the point m^* becomes a parameter of the model. In our terminology,

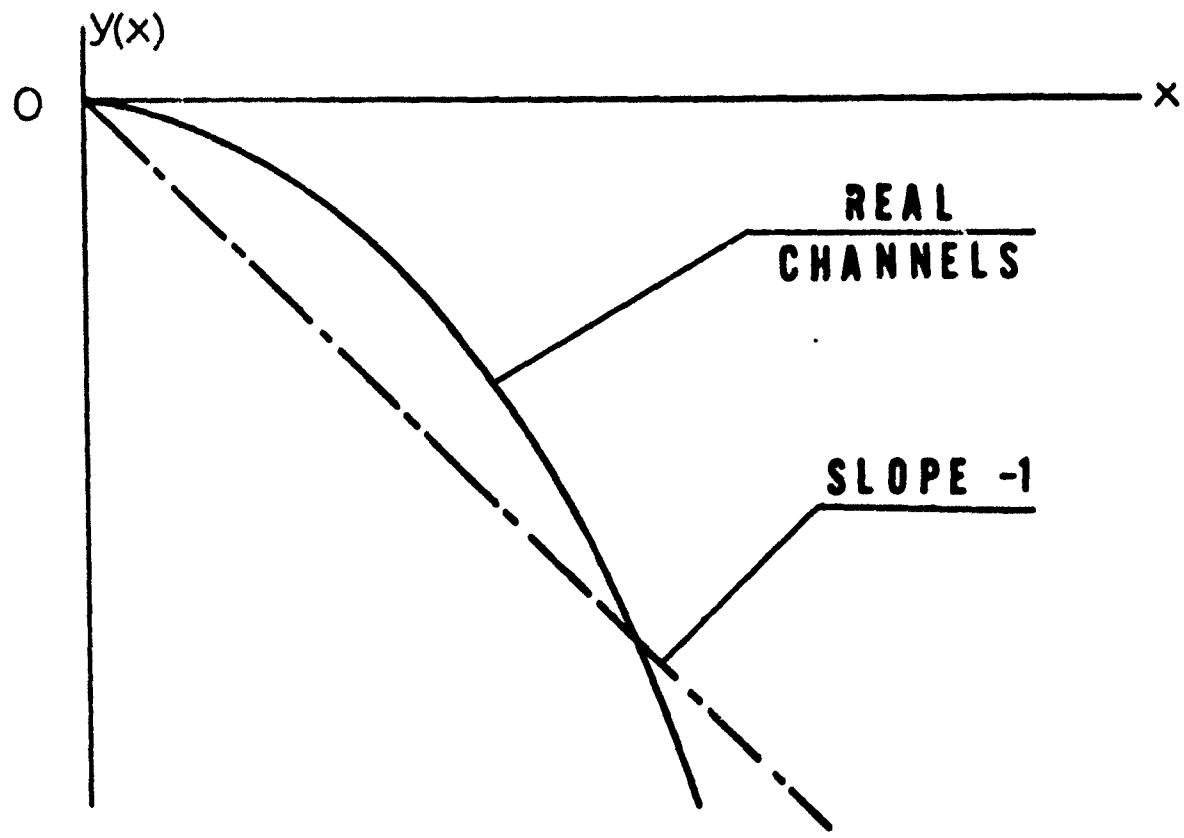


FIG. 3: REAL CHANNEL GAP DISTRIBUTION

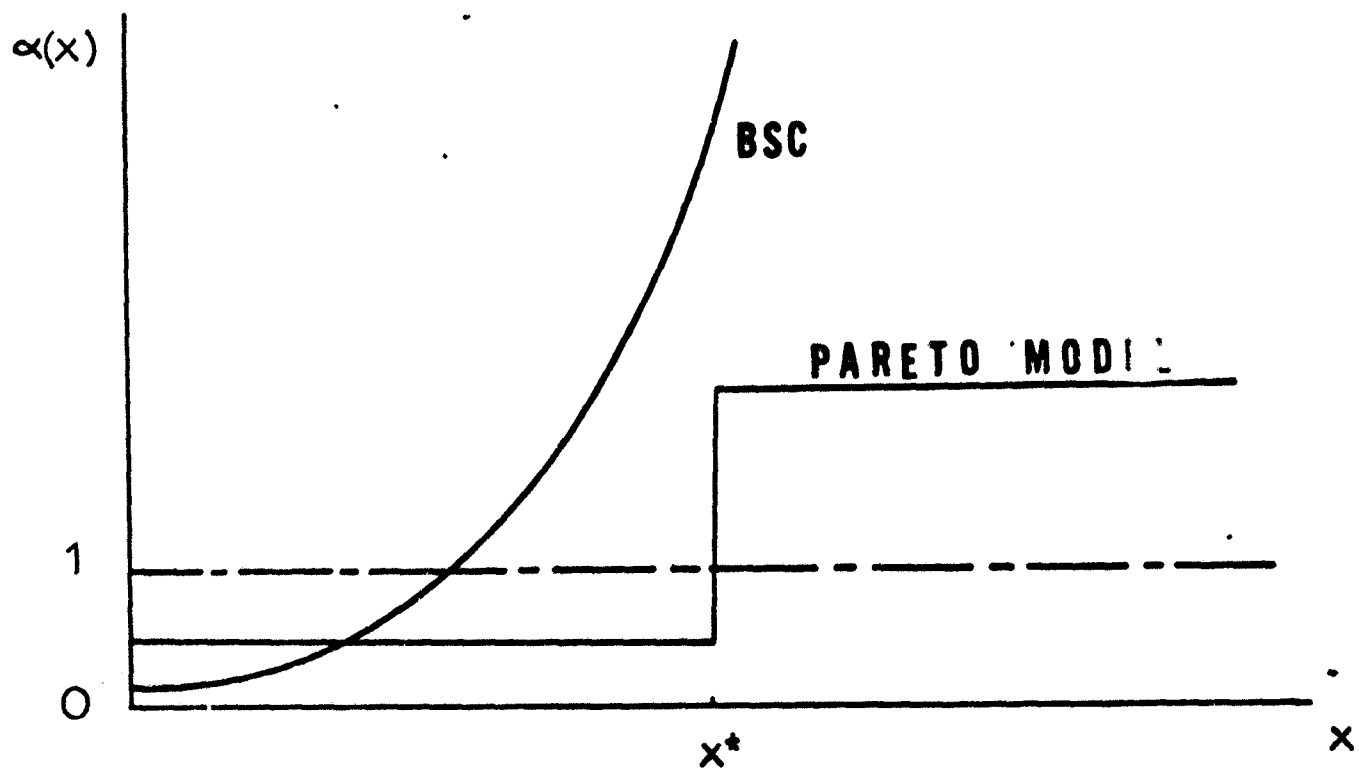


FIG. 4: TRUNCATED PARETO DISTRIBUTION

they assume that for values of x greater than $x^* = \log_e m^*$, α takes on a constant value greater than 1. This is illustrated in Figure 4 along with the α-function of the harmonic series and the BSC.

Finite-State Markov Models

A state model for the binary, discrete-time error process is a set of states together with a mapping φ, from these states onto the set $\{0,1\}$. The function φ^{-1} brings about a partitioning of the state space: those states which are mapped into 0 and those which are mapped into 1.

Gilbert [1] initiated the application of finite-state Markov Models to the representation of error sequences by proposing a model composed of a good state "G", which is error-free, and a bad state "B", in which the channel has only probability h to be in error. If the state sequence is represented by the discrete-time process $\{Z_i\}$, then the transition probabilities between states are defined by

$$\begin{aligned} \Pr\{Z_n \in G | Z_{n-1} \in B\} &= p \; ; \\ \Pr\{Z_n \in B | Z_{n-1} \in G\} &= P. \end{aligned} \tag{17}$$

The model is shown in Figure 5A.

Since even in the bad state there is a probability of having no error, the mapping φ cannot be directly applied. However, the Gilbert model can be transformed into a three-

state Markov chain as shown in Figure 5B; the new states are called G, B_0 and B_1. The function φ mapping the states onto the error sequence, can be defined as:

$$\varphi(G) = 0, \ \varphi(B_0) = 0 \text{ and } \varphi(B_1) = 1 .$$

Thus only B_1 is an error state.

Fritchman [8] extended Gilbert's results by studying the general case of finite-state models (Figure 6) with k error-free states and N-k error states. For such models he showed that the EGD can be written as

$$P(O^m|1) = \sum_{i=1}^{k} f(i)\lambda_i^{\,m}, \tag{18}$$

where λ_i are the eigenvalues of the matrix of transition probabilities among the k error-free states and f(i) is a function of the transition probabilities among all states. Ordering the set λ_i by decreasing magnitude, i.e., $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \ . \ . \ . \geq |\lambda_k|$, for large m and aperiodic chains we get asymptotically

$$P(O^m|1) \sim f(1)\lambda_1^{\,m}. \tag{19}$$

Consequently

$$y = \log_e P(O^m|1) \sim \log_e f(1) + m \log_e \lambda_1 \tag{20}$$

since $m = e^x$,

$$y = \sim e^x \log_e \lambda_1 + \log_e f(1) \tag{21}$$

and

$$\alpha = -\frac{dy}{dx} \sim \left(\log_e \frac{1}{\lambda_1}\right) e^x \tag{22}$$

This demonstrates that, regardless of the number of states, an aperiodic finite-state model always yields an α-function which is asymptotically exponential as shown in Figure 7.

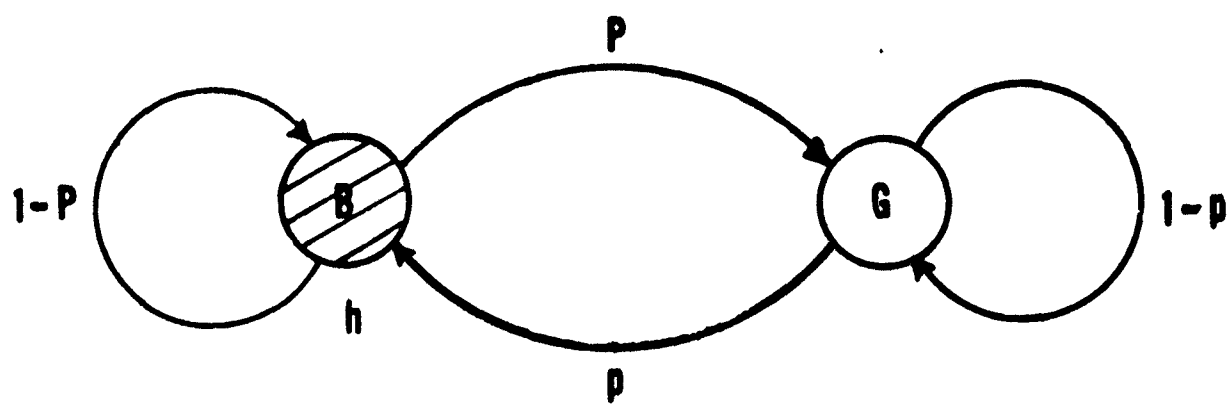
Renewal Process

A Renewal Process can be viewed as a series of trials in which the probability of success, i.e., no-error, at a certain trial is solely a function of the number of successes since the last failure, i.e., error. When an error finally occurs, the process "starts" all over again, thus giving rise to the name renewal. A renewal process is uniquely defined by the EGD.

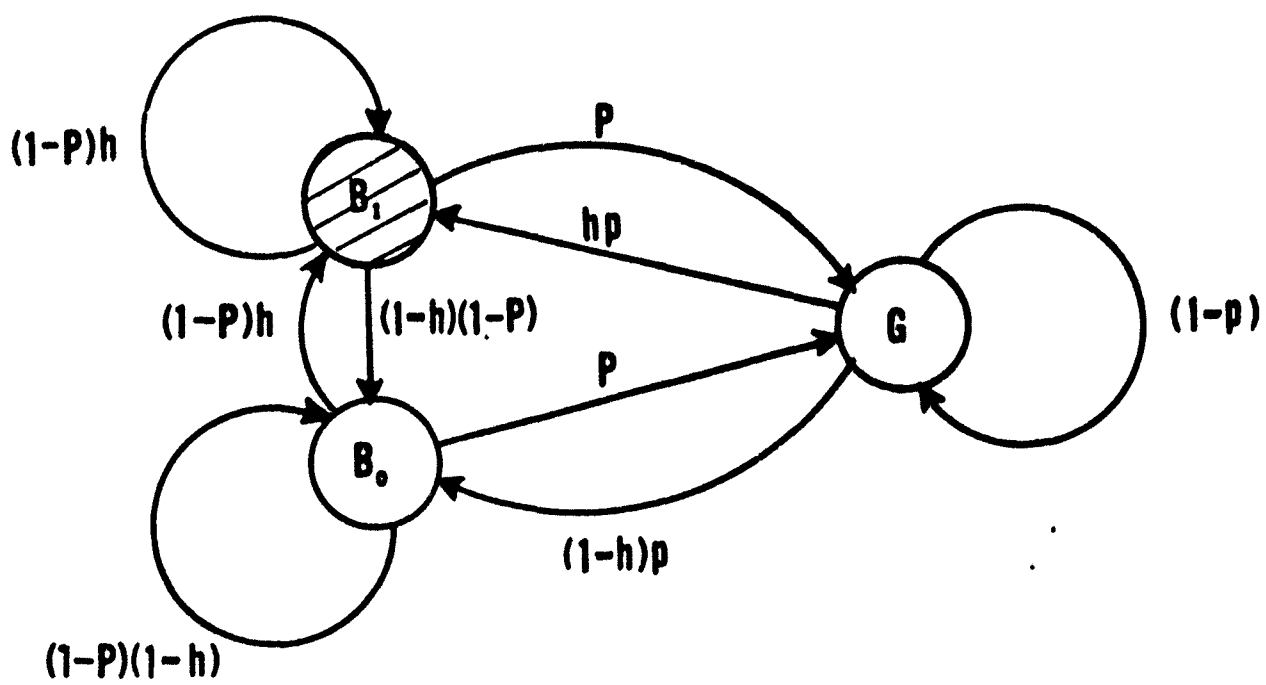
As noted in the discussion of the Pareto model, Berger and Mandelbrot proposed that the EGD be described by a Pareto distribution. In this model they as well as Sussman further assumed statistical independence between successive error gaps, thereby defining a renewal process. It is obvious that this renewal process cannot be modeled by a finite-state Markov chain for, as already demonstrated, that would lead to an α-function which is asymptotically exponential. We next show that a denumerably infinite-state Markov model, termed a slowly spreading chain, does allow modeling of this renewal process as well as more general ones.

V. Slowly Spreading Markov Chains

Renewal processes have received considerable attention in the context of point processes, see e.g., Smith [12]. The development of discrete-time renewal processes is however meager, except in relation to a class of denumerably infinite-state Markov chains which Kemeny [13] calls slowly spreading chains of the first kind. Representation



A



B

FIG. 5: GILBERT MODEL

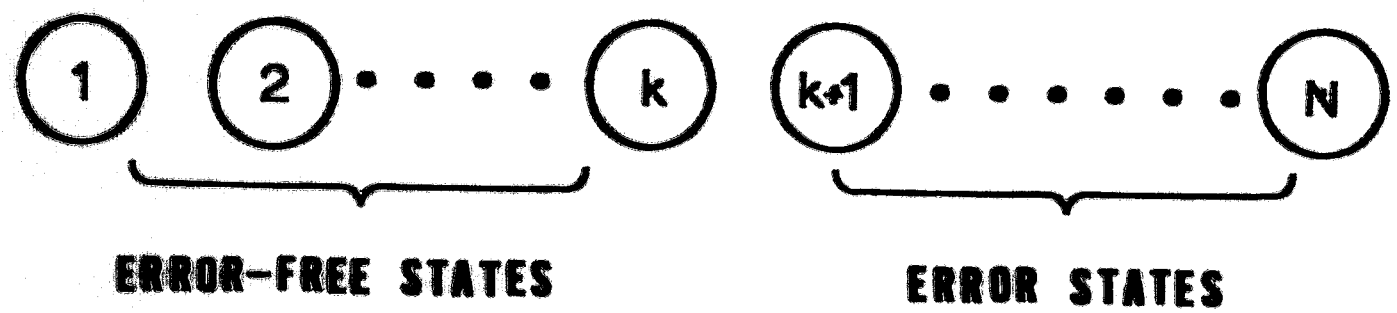


FIG. 6: FRITCHMAN'S ONE-ERROR STATE MODEL

of the general, discrete renewal process via slowly spreading Markov chains is now considered.

The denumerable state space is labeled by a non-negative integer and the following mapping relates the chain states to the states {0,1} of the renewal process:

$$\varphi(i) = \begin{cases} 1 & \text{if } i = 0 \\ 0 & \text{otherwise} \end{cases} . \tag{23}$$

If the discrete process $\{Z_n\}$ represents the state sequence, the transition probabilities can be defined as

$$\begin{aligned} \Pr\{Z_n = i \mid Z_{n-1} = i-1\} &= \hat{p}_i, \\ \Pr\{Z_n = 0 \mid Z_{n-1} = i\} &= \hat{q}_i. \end{aligned} \tag{24}$$

All other transitions have zero probability of occurrence. This slowly spreading chain is shown in Figure 8.

Letting

$$\beta_i \triangleq \prod_{k=1}^{i} \hat{p}_k , \tag{25}$$

it is assumed $\beta_i \neq 0$, i.e., all states are communicating. If Q_{on} is defined as the probability of returning to state 0 before reaching state n, then

$$\begin{aligned} Q_{on} &= \hat{q}_1 + \hat{p}_1\hat{q}_2 + \hat{p}_1\hat{p}_2\hat{q}_3 + - - - - - + \hat{p}_1\hat{p}_2 \bullet \bullet \hat{p}_{n-1}\hat{q}_n \\ &= \hat{q}_1 + \beta_1\hat{q}_2 + \beta_2\hat{q}_3 + - - - - + \beta_{n-1}\hat{q}_n \\ &= \sum_{i=o}^{n-1} \beta_i \hat{q}_{i+1} \quad , \text{ where } \beta_o \triangleq 1 \end{aligned} \tag{26}$$

The individual terms of the series can be interpreted

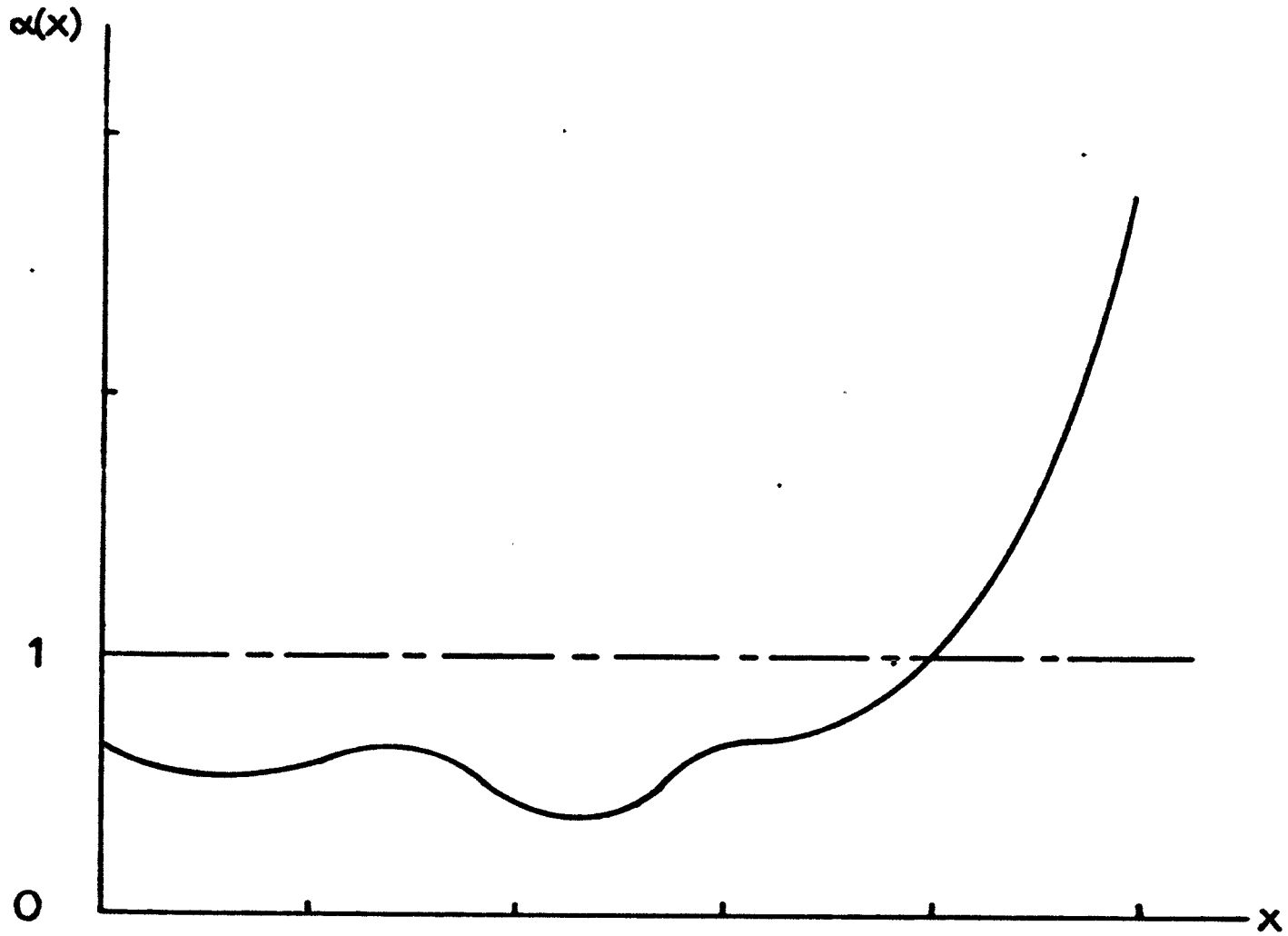


FIG. 7: ASYMPTOTICALLY EXPONENTIAL α FOR FINITE-STATE MODELS

as follows:

$$\beta_i \hat{q}_{i+1} = \left(\begin{array}{c}\text{Probability of reaching}\\ \text{state i from state 0}\end{array}\right) x \left(\begin{array}{c}\text{Probability of returning}\\ \text{from state i to state 0}\end{array}\right). \tag{27}$$

Thus β_i can be interpreted as the probability of reaching state i when the process is making a round trip to state 0. Also, since $\hat{q}_k = 1 - \hat{p}_k$ and $\beta_k = \beta_{k-1} \hat{p}_k$, substitution into the preceding expression yields $Q_{on} = 1 - \beta_n$. Similarly the probability of returning to any state i before reaching a state $n > i$, Q_{in}, can be expressed as

$$Q_{in} = 1 - \frac{\beta_n}{\beta_i} \quad ; \quad \text{for } i = 0, 1, 2, \ldots. \tag{28}$$

Here and throughout the paper $\beta_o \triangleq 1$. If h_{ii} is defined as the probability of eventually returning to state i, then for a slowly spreading chain

$$h_{ii} = \lim_{n\to\infty} Q_{in} = 1 - \frac{1}{\beta_i} \lim_{n\to\infty} \beta_n \; ; \text{ for } i = 0, 1, 2, \ldots \tag{29}$$

A denumerable chain is said to be recurrent if the probability of eventually returning, h_{ii}, is unity. For the slowly spreading chain this is equivalent to requiring $\lim \beta_n = 0$ as $n\to\infty$.

Although certain, the return to state i can take a very large number of steps. To insure that on the average this number of steps is finite it is necessary for the chain to meet the more restrictive ergodic condition which follows. Let $\bar{M}_{oo}$ be the expected time for the first return to state 0. Then from Figure 8 it is seen that

$$\overline{M}_{oo} = \hat{q}_1 + 2\hat{p}_1\hat{q}_2 + 3\hat{p}_1\hat{p}_2\hat{q}_3 + \text{------}$$

$$= \hat{q}_1 + 2\beta_1\hat{q}_2 + 3\beta_2\hat{q}_3 + \text{------}$$

$$= \sum_{k=o}^{\infty} (1 + k)(\beta_k - \beta_{k+1}) \tag{30}$$

$$= \sum_{k=o}^{\infty} \beta_k .$$

State 0 and hence all states[1] will be ergodic if and only if

$$\sum_{k=o}^{\infty} \beta_k < \infty . \tag{31}$$

Referring to the transformation of chain states into the states 0 and 1, given by equation (23), it is clear that

$$\beta_i = P(0^i|1), \tag{32}$$

and

$$P(0^i 1|1) = \beta_i - \beta_{i+1} . \tag{33}$$

---

1 Defining the rth moment of the mean recurrence time by $b_k^{(r)} = E(T_k^r)$ where T_k is the random duration needed to go from state k back to state k. Kemeny, Snell and Knapp [14] have shown that if $b_o^{(r)} < \infty$ for some state S_o, then $b_j^{(r)} < \infty$ for all states j. Consequently the conditions which guarantee that the rth moment of state 0 is finite are the same for all states. We make use of this reasoning throughout the remainder of the paper.

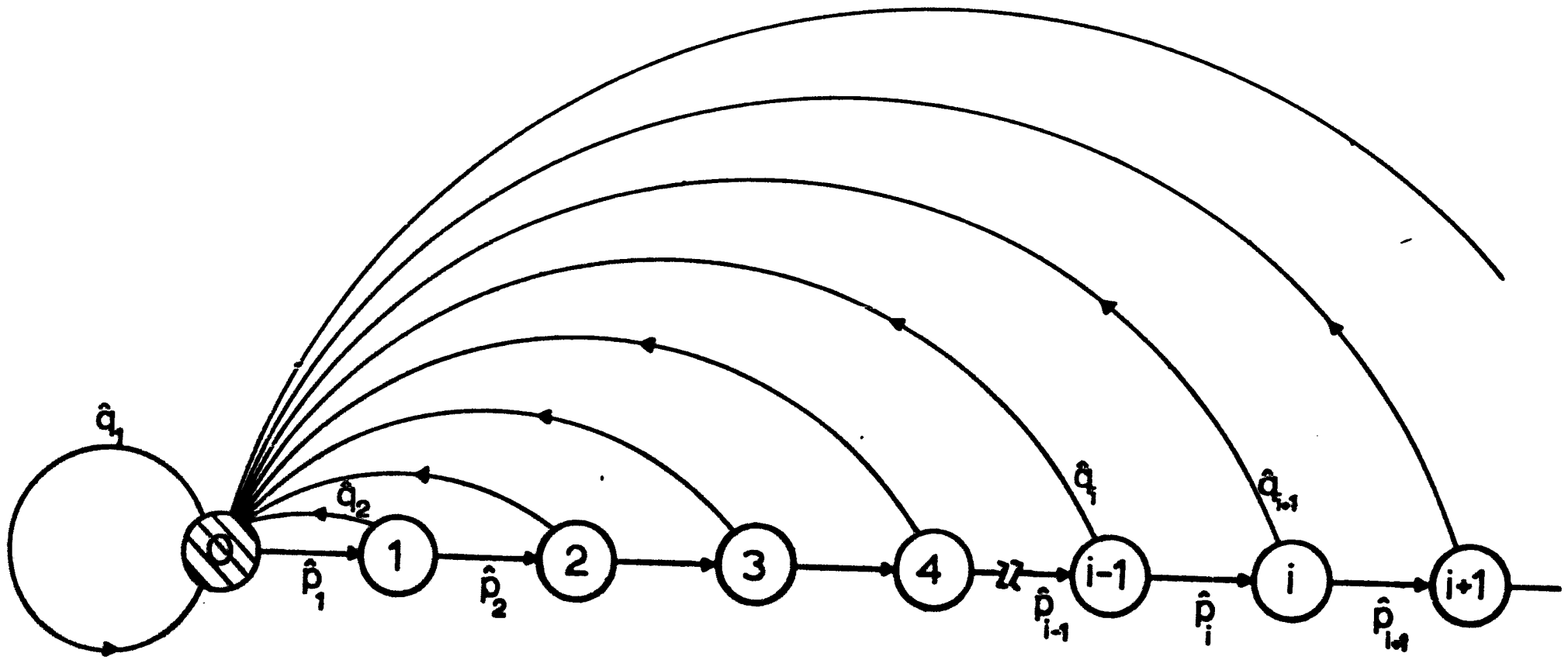


FIG. 8: SLOWLY SPREADING CHAIN

In other words β_i is the EGD of the chain and this model allows the specification of a renewal process having any desired EGD or equivalently any desired α-function. The only constraint on the specification of β_i is that $\sum_{i=o}^{\infty} \beta_i < \infty$.

The slowly spreading model includes as special cases the renewal models previously discussed. For example if we let $\beta_i = (1-p)^1$ the resulting model is just the BSC having error rate p. Alternatively if we set $\beta_i = 1/i^\theta$ the result is the Pareto model. Between these two extremes a renewal process with almost any desired α-function can be specified.

VI Conditions on α for Ergodicity of Degree r.

The preceding discussion has centered around chains with finite mean recurrence times or equivalently chains with ergodic state probabilities. This is actually the weakest ergodic condition which can be imposed With T_o denoting the 0-state recurrence time. Kemeny, Snell and Knapp [14] define a chain to be ergodic of degree r if $b_o^{(r)} = E(T_o^r) < \infty$ but $b_o^{(r+1)} = \infty$. They prove the notable fact that a finite-state Markov chain is ergodic of degree infinity. Therefore more restrictive conditions can be imposed on a chain by requiring the first r moments of the 0-state recurrence time, T_o, to be finite, and $E(T_o^{r+1}) = \infty$.

For a slowly spreading Markov chain to be ergodic of at least degree 1, it was shown in the preceding section that $\sum_{i=0}^{\infty}\beta_i = \sum_{i=0}^{\infty} P(0^i|1) < \infty$. The conditions imposed on $\alpha(x)$ by ergodicity of degree r are now considered. The rth moment of the 0-state recurrence time $b_0^{(r)}$ is given by

$$
\begin{aligned}
b_0^{(r)} &= E[T_0^{\,r}] = E[(m+1)^r] \\
&= \sum_{m=0}^{\infty} (m+1)^r P(0^m 1|1) \\
&= \sum_{m=0}^{\infty} (m+1)^r [P(0^m|1) - P(0^{m+1}|1] \qquad (34) \\
&= \sum_{m=1}^{\infty} [(m+1)^r - m^r] P(0^m|1) \\
&= \sum_{m=1}^{\infty} \{[m^r + \frac{rm^{r-1}}{1!} + \frac{r(r-1)m^{r-2}}{2!} + ---] - m^r\} P(0^m|1),
\end{aligned}
$$

so that for large m

$$E(T_0^{\,r}) \sim \sum_{m=1}^{\infty} m^{r-1} P(0^m|1). \qquad (35)$$

If the rth moment is required to be finite, then clearly

$$\sum_{m=1}^{\infty} m^{r-1} P(0^m|1) < \infty. \qquad (36)$$

There are thus two conditions for ergodicity of degree r

namely

$$\sum_{m=1}^{\infty} m^{r-1} P(0^m|1) < \infty , \tag{37}$$

$$\sum_{m=1}^{\infty} m^{r} P(0^m|1) = \infty . \tag{38}$$

These conditions can be directly related to the asymptotic behavior of the α-function. First observe that if $m^{r-1}P(0^m|1) < 1/m$ for all values of $m > m^*$, where m^* is any positive integer then condition (a) is met, i.e., the rth moment is finite if

$$P(0^m|1) < \frac{1}{m^r} . \tag{39}$$

Taking the natural logarithm of both sides of this expression and using the logarithmic variables x and y, differentiating and changing sign, gives the result that if $\alpha(x)$ is greater than r for all values of $x > x^* = \log_e m^*$, then rth moment exists. Alternately if

$$P(0^m|1) > \frac{1}{m^{r+1}} , \tag{40}$$

then the (r+1)th moment does not exist. Sufficient conditions to guarantee ergodicity of degree r are therefore

$$r \le \alpha(x) < r+1. \tag{41}$$

On the $\alpha(x)$ diagram the asymptotic behavior of $\alpha(x)$ can be categorized as shown in Figure 9. An α-function, which for x greater than some x^*, always remains between lines r and r+1, corresponds to a process having ergodicity of degree r. Processes whose α-functions do not asymptote to finite values will have infinite ergodic degree. For example processes having α-functions which are asymptotically

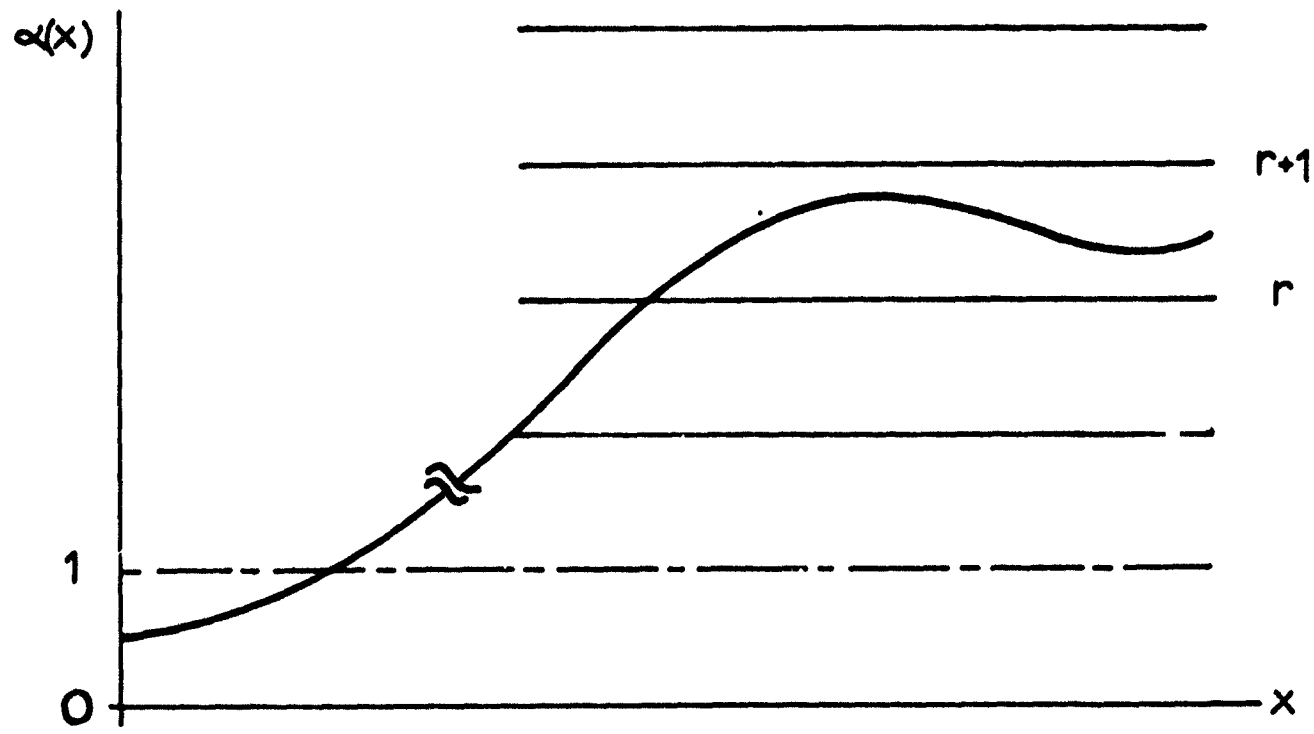


FIG. 9: α-FUNCTION AND ERGODICITY DEGREE

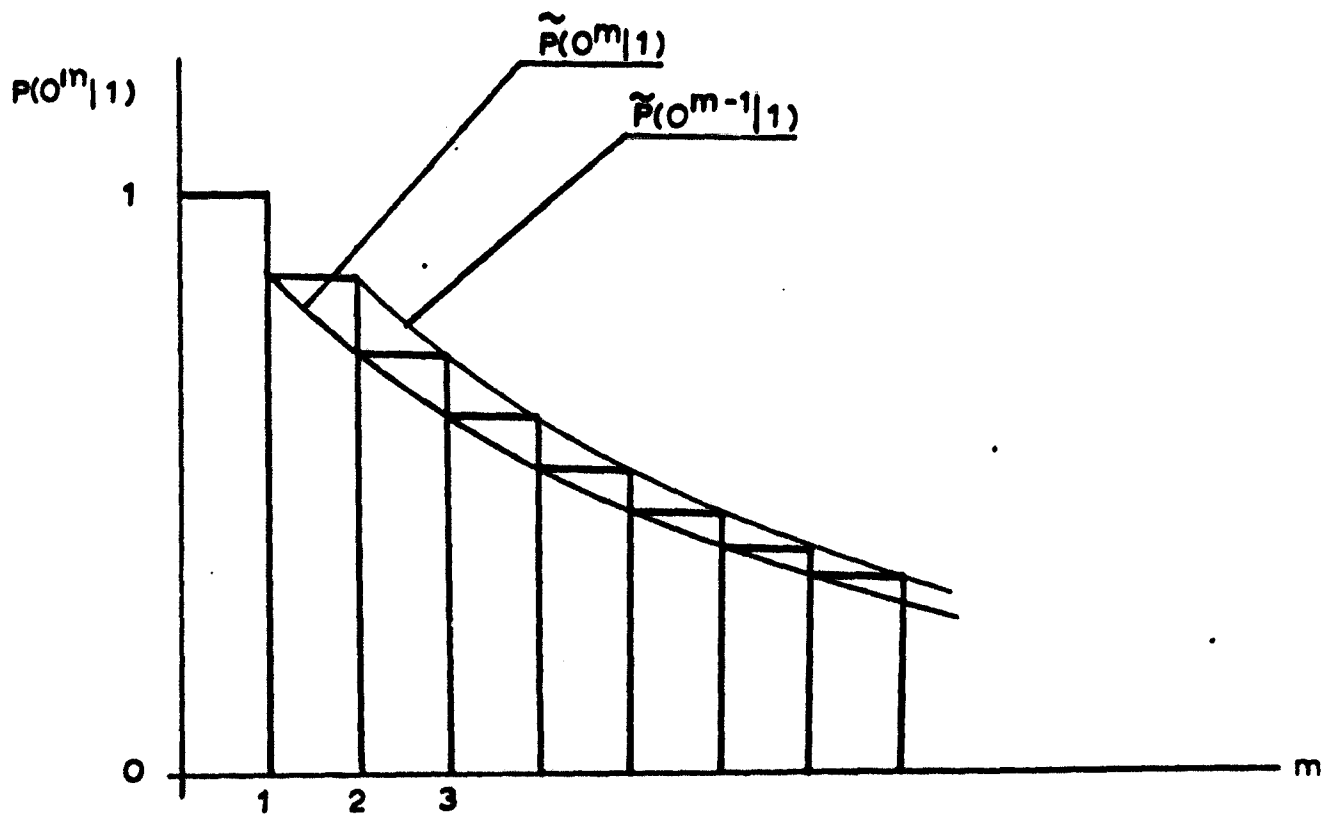


FIG. 10: BOUNDS FOR THE MEAN RECURRENCE TIME

logarithmic or exponential have infinite ergodic degree. Processes having α-functions which asymptotically oscillate cannot be classified by this procedure. The classification of such processes will be considered in Section VIII.

VII Relation Between $\alpha(x)$ and the Recurrence Time

In Section I it was shown that

$$E(G_n) = \sum_{m=o}^{\infty} P(0^m|1) = R_t \tag{42}$$

From the plot of $P(0^m|1)$ vs. m shown in Figure 10 it is observed that R_t is the area contained in the rectangles of height $P(0^m|1)$ and of unit width. $\tilde{P}(0^m|1)$ is defined to be a continuous function of m which has the same values as $P(0^m|1)$ when m is a positive integer. Referring to Figure 10 the area in the first rectangle is always unity. Therefore the area in the first rectangle plus the area under $\tilde{P}(0^m|1)$ for values of $m > 1$, will be a lower bound on the total area in all the rectangles, i.e.,

$$1+ \int_1^{\infty} \tilde{P}(0^m|1)\,dm < R_t . \tag{43}$$

As indicated in the figure $\tilde{P}(0^{m-1}|1)$ is an upper bound on the rectangles. Therefore the area in the first rectangle plus the area in the second rectangle plus the area under $\tilde{P}(0^{m-1}|1)$ for $m > 2$ will give an upper bound on the area in all the rectangles, i.e.,

$$R_t \leq 1 + P(0|1) + \int_2^{\infty} \tilde{P}(0^{m-1}|1)\,dm . \tag{44}$$

But

$$\int_2^{\infty} \tilde{P}(0^{m-1}|1)\,dm = \int_1^{\infty} \tilde{P}(0^m|1)\,dm$$

and so

$$1 + \int_1^{\infty} \tilde{P}(0^m|1)\,dm < R_t < 1 + P(0|1) + \int_1^{\infty} \tilde{P}(0^m|1)\,dm. \tag{45}$$

The upper and lower bounds differ only by $P(0|1)$ a quantity less than 1. Since it can be expected that $R_t > 10$ for real channels and more usually $R_t > 10^3$, R_t is closely approximated by

$$R_t \cong 1 + \frac{P(0|1)}{2} + \int_1^{\infty} \tilde{P}(0^m|1)\,dm. \tag{46}$$

The integral on the right-hand side can now be expressed in terms of the logarithmic variables x and y. In the x-y plane $\tilde{P}(0^m|1) = e^y$ and $dm = e^x dx$, so that

$$\int_1^{\infty} \tilde{P}(0^m|1)\,dm = \int_0^{\infty} e^{(y+x)}\,dx. \tag{47}$$

To help in understanding the meaning of the right-hand integral a new function A(x) is defined as

$$A(x) \triangleq \int_0^x [\alpha(\mu) - 1]\,d\mu \tag{48}$$

This is just the area between $\alpha(x)$ and 1 for values between 0 and x. Recall that $\alpha(x) = -dy/dx$ which upon substitution into the integral yields

$$A(x) = y(0) - y - x$$

or equivalently

$$y + x = y(0) - A(x). \tag{49}$$

Therefore

$$\int_0^\infty e^{(y+x)}dx = e^{y(0)} \int_0^\infty e^{-A(x)}dx \cdot \tag{50}$$

Now $x = 0$ when $m = 1$ and so $y(0) = \log_e P(0|1)$ and $P(0|1) = e^{y(0)}$. The mean recurrence time is therefore bounded as

$$1 + P(0|1) \int_0^\infty e^{-A(x)}dx < R_t < 1 + P(0|1) + P(0|1)\int_0^\infty e^{-A(x)}dx, \tag{51}$$

and is closely approximated by

$$R_t \cong 1 + \frac{P(0|1)}{2} + P(0|1) \int_0^\infty e^{-A(x)}dx. \tag{52}$$

Since both the upper and lower bounds are infinite when the integral is infinite and finite when it is finite, a necessary condition for the convergence of R_t is that

$$\int_0^\infty e^{-A(x)}dx < \infty \,. \tag{53}$$

This is equivalent to requiring that the area $A(x)$, between $\alpha(x)$ and 1, be greater than $\epsilon > 0$ for all values of x greater than some value x^*, i.e.,

$$A(x) > \epsilon > 0 \text{ for } x > x^*. \tag{54}$$

Moreover observing that this condition is satisfied if the previous sufficient condition is satisfied, i.e.,

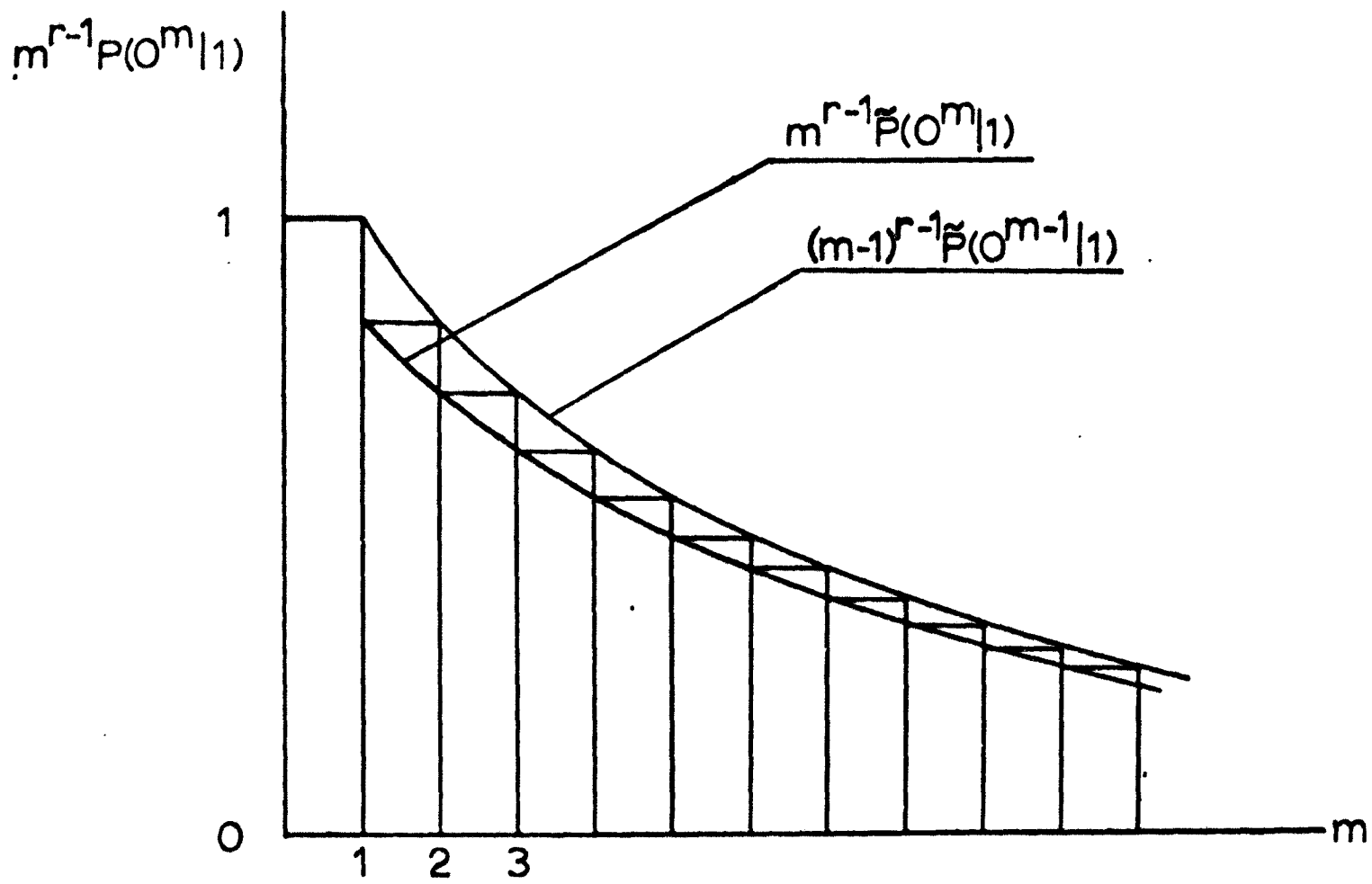


FIG. 11: BOUNDS ON $m^{r-1}P(0^m|1)$

that $\alpha(x) > 1$ for $x > x^*$, the above condition on $A(x)$ is both necessary and sufficient.

VIII A Necessary Condition on α For Ergodicity of Degree r.

Recall in the discussion of Section VI that ergodicity of degree r required that the conditions given by equation (40) be satisfied. This lead to a sufficient condition on α which guaranteed ergodicity of degree r, namely, that $r < \alpha(x) < r+1$. Following a procedure similar to the one used in the preceding section, which lead to a necessary and sufficient condition on α for ergodicity of degree 1, a necessary and sufficient condition on α for ergodicity of degree r will now be obtained.

Referring to Figure 11 the area in the rectangles is equal to

$$\sum_{m=1}^{\infty} m^{r-1} P(0^m|1) . \tag{55}$$

Once again the upper and lower bounds differ only by $P(0|1)$, a quantity having value less than unity. If

$$\int_1^{\infty} m^{r-1} \tilde{P}(0^m|1)\, dm < \infty , \tag{56}$$

both upper and lower bounds are finite, but if this integral is infinite, then so are both bounds. The convergence of this integral is a necessary condition for ergodicity of at least degree r.

Defining the auxiliary function $A_r(x)$ such that

$$A_r(x) \triangleq \int_0^x [\alpha(\mu) - r]\, d\mu \quad , \tag{57}$$

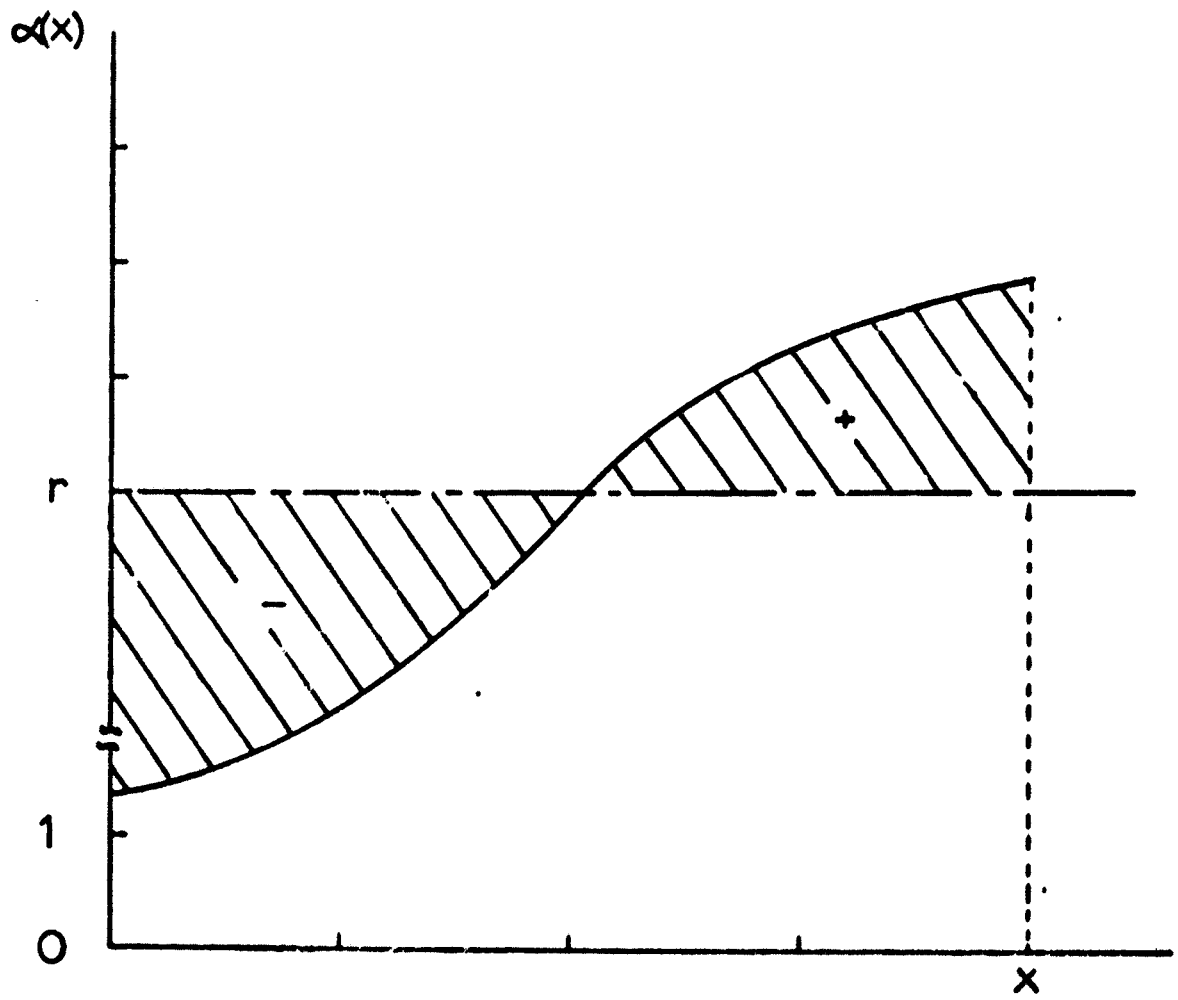


FIG. 12: REPRESENTATION OF $A_r(x)$

it is observed that $A_r(x)$ is the area between the function $\alpha(x)$ and the line r along the x-axis between 0 and x, as shown in Figure 12. The necessary condition for the existence of the rth moment of the recurrence time and hence ergodicity of at least degree r, is that

$$\int_0^\infty e^{-A_r(x)} dx < \infty \quad . \tag{58}$$

Consequently necessary and sufficient conditions for ergodicity of degree r are

$$A_r(x) > 0 \quad \text{for } x > x_1^* ,$$

$$A_{r+1}(x) < 0 \text{ for } x > x_2^* \tag{59}$$

where x_1^* and x_2^* are any positive real numbers.

Summarizing, the $\alpha(x)$ diagram can be used to determine the degree of ergodicity of a process as follows. If the area between $\alpha(x)$ and each of the lines 1, 2, ..., r is positive, but the area between $\alpha(x)$ and the line r+1 is negative, then the process is ergodic of degree r. In order to guarantee a finite mean recurrence time R_t, the process must be ergodic of at least degree 1. For processes that satisfy this condition, R_t can be closely approximated in terms of the area A(x) between $\alpha(x)$ and 1.

IX Data Analysis

To gain a better understanding of the behavior of the α-function of real channels three troposcatter channels are examined.[2] We were particularly interested in the behavior of $\alpha(x)$ for large values of x, i.e. for large m, necessitating the examination of very long error sequences. The largest sequence examined was approximately 7×10^9 bits which in comparison to previous analyses is indeed long. The three channels are labeled C1, C2, and C3. Some general features of these channels follow:

	C1	C2	C3
Bits per second	307,000	614,000	1,000,000
Error rate	2.355×10^{-2}	1.347×10^{-5}	2.949×10^{-5}
Duration of run (min)	0.41	102	119.76
Total number of bits	7,635,632	378,266,112	7,125,322,352
Total number of errors	179,828	50,936	210,153
Mean recurrence time	42.46	74263	33905
Longest gap length	10088	11,306,822	3,585,916

The α-function for these three channels is shown in Figures 13-15. Channels C1 and C2 appear to exhibit exponential or nearly exponential asymptotic behavior which suggests they might be represented by finite-state Markov models. Channel C3, however, exhibits a most unusual behavior. Its α-function continues to vary from very small values to large values and back again. It appears to be barely ergodic, though the area A(x) between $\alpha(x)$ and the line x = 1 is still positive for $m > 10^6$. Such a channel clearly cannot be represented

[2] The data was provided by the United States Army Electronic Command, Fort Monmouth, New Jersey, in March 1970.

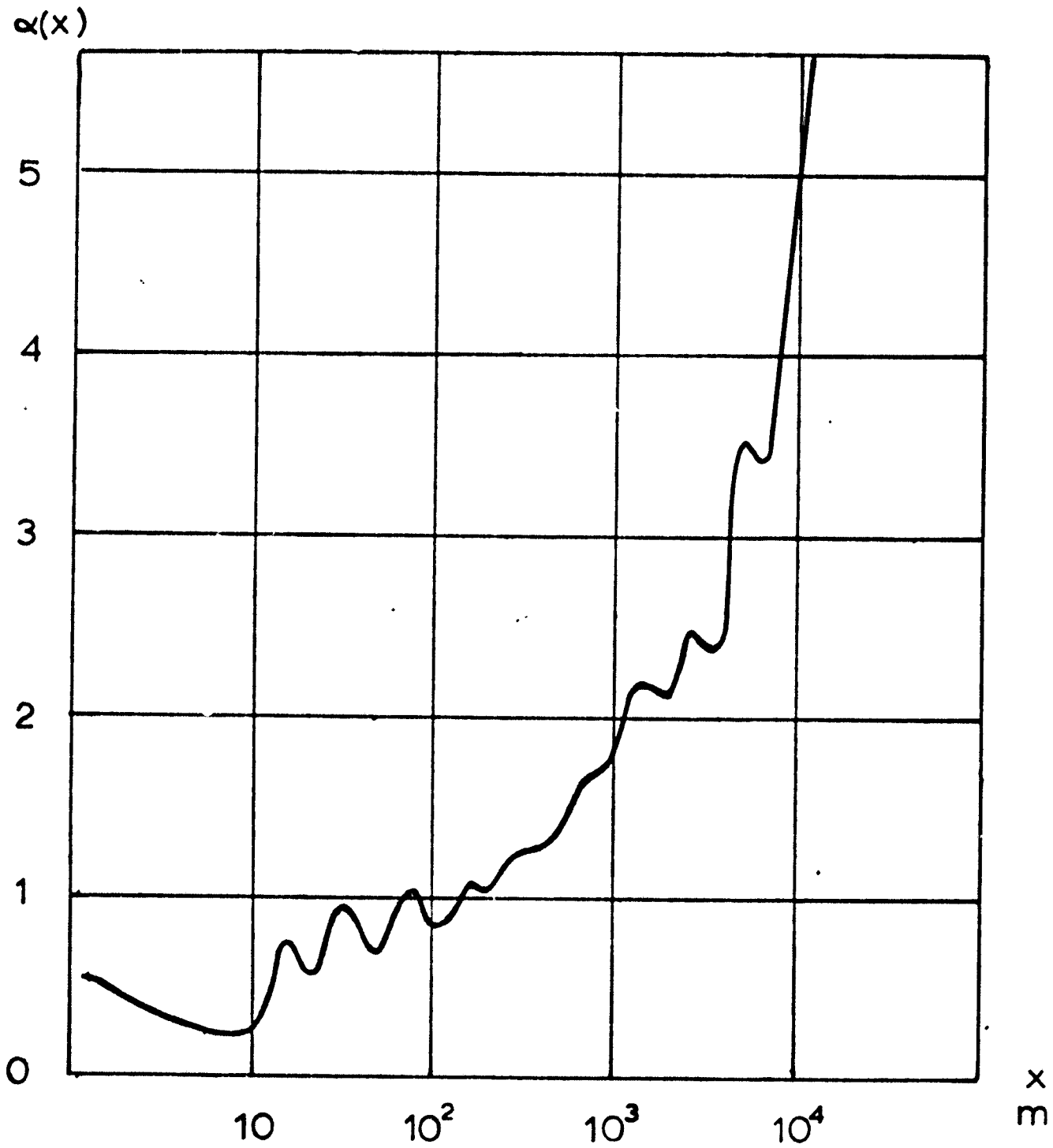


FIG. 13: α-FUNCTION - C_1

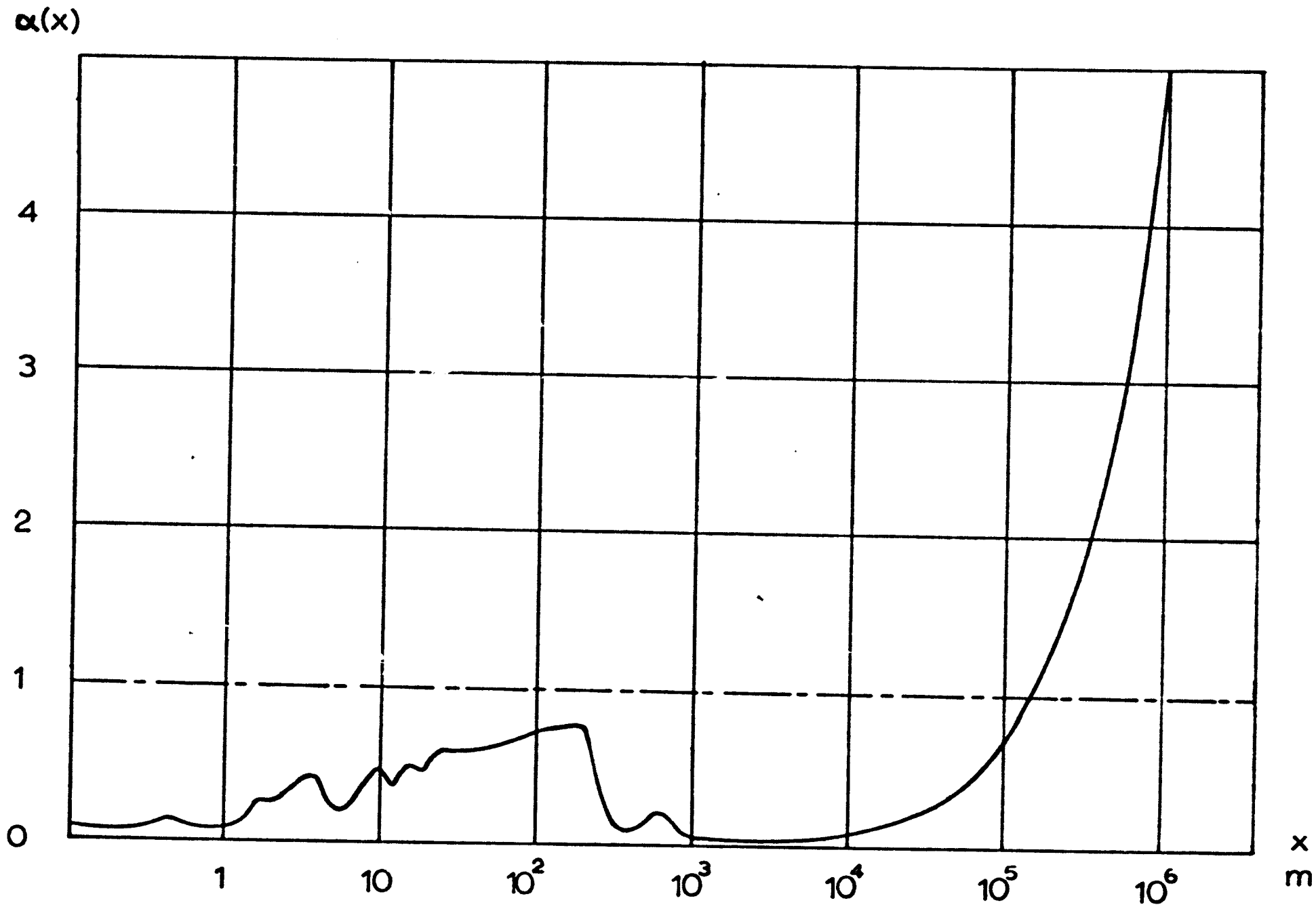


FIG. 14: α-FUNCTION - C_2

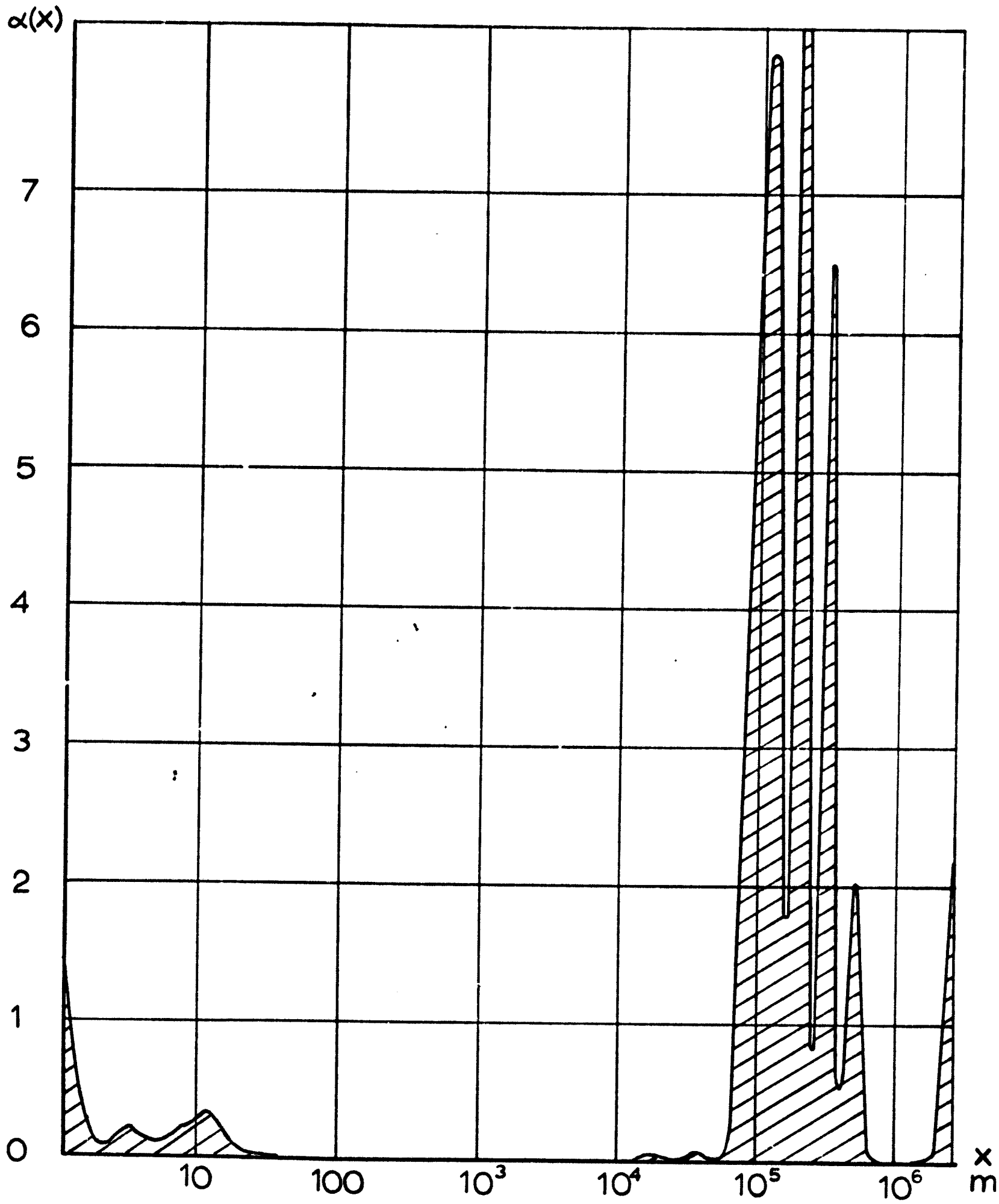


FIG. 15: α-FUNCTION - C_3

by finite-state models or the Pareto model, but could be represented by the slowly spreading chain model discussed in Section V provided of course that error gaps were not correlated.

X. Summary and Conclusions

The α-function is capable of revealing a number of interesting properties of real communication channels and of uncovering some of the underlying behavior of channel models. From the α-diagram of a real channel it is possible to obtain an indication of its degree of ergodicity. Moreover since it was shown that finite-state models lead to α-functions having an exponential asymptotic behavior, the α-function can be used to determine when a finite-state model can or cannot be expected to characterize a channel.

As a consequence of our investigation of the α-function of models previously developed to characterize real channel behavior, it was discovered that certain α-function behavior could not be described, even though real channels can exhibit such behavior as evidenced by channel C3 investigated in Section IX. This lead to the consideration of a slowly spreading Markov model capable of exhibiting almost any α-function behavior desired. A model consisting of two coupled, infinite-state, slowly spreading Markov chains, which is capable of accounting for a wide variety of properties of real channels, has been developed by the authors. This will be a subject of a future paper.

Finally it should be pointed out that though the present analysis has been directed toward the representation of error sequences, the results are applicable to any binary, discrete-time random processes; for example those generated by binary data sources.

1. E. N. Gilbert, "Capacity of a Burst-Noise Channel," BSTJ, Vol. 39, pp. 1253-1266; September, 1960.

2. P. Mertz, "Model of Error-Burst Structure in Data Transmission," Proc. National Electronics Conference, Vol. 16, pp. 232-240; October 1960.

3. J. M. Berger and B. Mandelbrot, "A New Model for Error Clustering in Telephone Circuits, " IBM J. Research and Development, Vol. 7, pp. 224-236; July 1963.

4. E. O. Elliott, "A Model of the Switched Telephone Network for Data Communications," BSTJ, Vol. 44; January 1965.

5. J. J. Metzner, "An Interesting Property of Some Infinite-State Channels," IEEE Transaction on Information Theory, pp. 370-372; April 1965.

6. S. Berkovits, E. L. Cohen and N. Zierler, "A Model for Digital Error Distributions," First IEEE Annual Communication Convention, pp. 103-111; June 1965.

7. P. A. W. Lewis and D. R. Cox, "A Statistical Analysis of Telephone Circuit Error Data," IEEE Transaction on Communication Technology, COM-14, pp. 382-389; August 1966.

8. B. D. Fritchman, "A Binary Channel Characterization Using Partitioned Markov Chains," IEEE Transaction on Information Theory, IT-13, pp. 221-236; April 1967.

9. N. Muntner and J. K. Wolf, "Predicted Performances of Error-Control Techniques Over Real Channels," IEEE Transactions on Information Theory, IT-14, pp. 640-650; September 1968.

10. R. H. McCullough, "The Binary Regenerative Channel," BSTJ Vol. 47, pp. 1713-1735; October 1968.

12. S. Sussman, "Communication Channel Attributes as Related to Error Control, " First IEEE Annual Communication Convention Proceedings, pp. 5-13 June 1965.

13. J.G. Kemeny, "Slowly Spreading Chains of the First Kind," Journal of Mathematical Analysis and Applications, Vol. 15, pp. 295-310; August 1966.

14. J. G. Kemeny, J. L. Snell and A. W. Knapp Denumerable Markov Chains, Princeton, Van Nostrand, pp. 269-273, 1966.

15. W. L. Smith, "Renewal Theory and Its Ramifications," J.R. Statist. Soc. B, Vol. 20, No. 2, pp. 243-302; 1958

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Computer Science Center
University of Maryland, College Park, Md. 20742

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
A CRITICAL STATISTIC FOR CHANNELS WITH MEMORY

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*
J-P. A. Adoul, B.D. Fritchman and L.N. Kanal

6. REPORT DATE
June 3, 1971

7a. TOTAL NO. OF PAGES
39

7b. NO. OF REFS
15

8a. CONTRACT OR GRANT NO.
AFOSR 71-1982

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
Technical Report TR - 161

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT
Distribution of this document is unlimited

11. SUPPLEMENTARY NOTES
Accepted for publication in IEEE Transactions on Information Theory

12. SPONSORING MILITARY ACTIVITY
Information Sciences, AFOSR, Air Force Systems Command, 1400 Wilson Blvd., Arlington, Virginia 22209

13. ABSTRACT

We present a new descriptive statistic for channels with memory and show its utility (a) in evaluating and comparing existing models for such channels, and (b) as a theoretical tool in defining the error-gap distribution characteristics of real channels. We demonstrate that certain kinds of real channel behavior cannot be adequately described by previously proposed models and offer an example of a better model which includes many of the earlier models as special cases.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Communication Channels with Memory						
Channel Modeling						
Critical Statistic						
Slowly Spreading infinite-state Markov Chains						
Finite State Models						
Pareto Models						

UNCLASSIFIED

Security Classification