SSC-N-310,12/86, 2/87

MAXIMUM INFORMATION ENTROPY-BAYESIAN ANALYSIS
FOR AVAILABILITY STUDIES OF
ACCELERATED LIFE TESTS OF SSC MAGNETS

ELY SHRAUNER

SSC/CDG and WASHINGTON UNIVERSITY

Introduction and Objective

We previously outlined an analysis to quantify the probabilistic confidence for specified availabilities obtainable from accelerated life tests on SSC magnets.[1,2] That analysis was based on a form of Bayes' theorem; thereby involving certain subjectivities and ambiguities.[3] The objective of the present analysis is to outline and illustrate a method of refinement by which those subjectivities and ambiguities may be reduced systematically to an absolute mathematical minimum.

Bayesian Prior Probabilities

Bayesian statistical analysis is peculiar in that it requires prior probability measures on the independent variable. In our case prior probabilities of the possible trial failure rates for the individual SSC magnets are introduced to compute the probabilistic confidence that the actual effective failure rate corresponding to a specified system availability is supported by the evidence gained from accelerated life tests. The Bayesian prior probability distribution quantifies prior information known about the system to be analysed. Its specification is (notoriously) susceptible of introducing subjective bias beyond those conditions that are explicitly required by the objective data (of prior information). It is necessary that the specification of these prior probabilities correctly comprehend all the objectively known and relevant prior data, and include no effects beyond that supported by the data.

In our previous analysis this subjectivity was manifested in our choosing for illustrative purposes a microcanonical distribution of prior probabilities for trial values of the effective failure rate λ. The microcanonical distribution represents the prior information, or ignorance, about the trial λ's by restricting them to a cutoff region $\lambda_c \leqslant \lambda \leqslant \lambda_0$, and no other condition; i.e., probabilities are uniformly distributed within this cutoff region.

Maximum Information Entropy Method

In the present analysis this subjectivity is eliminated in the sense that the unique prior probability distribution obtained contains the effects of all the information from the test data, and maximizes the ignorance, or absent information not specifically supported by the data. Our method is to derive the unique prior distribution which variationally maximizes its information-theoretical entropy subject to the constraints that account for all the objective data resulting from accelerated life tests. The information-theoretical entropy represents the information that is absent from our prior probability distribution; or it represents the uncertainity included.
This information entropy would be zero for a distribution that exactly specifies complete knowledge of the state of the system; such as a distribution that specifies with unit probability one value of λ and zero probability all other values.

The distribution which maximizes the information entropy subject to the data is of the canonical form; *as opposed to the microcanonical form chosen in our previous analysis.* The empirical forms of both distributions are strikingly similar, and their effects under the integrals in our analysis are very similar; so that our previous choice of the simple microcanonical distribution for heuristic purposes was gratifyingly astute.

Analysis

As was the case for our previous analysis, the data are of the form of probabilistic evidence that a sample of N magnets survive y equivalent operational years of accelerated life testing with no more than n failures. We then compute the probability, or confidence, that these test results would imply that the actual effective single-magnet failure rate λ is less than a value λ' that is specified by the required stationary availability of the whole SSC system.

The probability that a sample of N magnets will survive y equivalent years of operational cycling (about y weeks of testing) with no more than n failures is

$$(1) \qquad P(n,N,y|\lambda) = \sum_{\ell=0}^{n} \binom{N}{\ell} [p(\lambda,y)]^{N-\ell} [1-p(\lambda,y)]^{\ell}$$

where $p(\lambda,y)$ is the probability that an individual magnet will survive y equivalent operational years of cycling with a (trial) failure-rate function labelled by λ. Under our assumed exponential failure hypothesis $p(\lambda,y) \cong \exp(-\lambda y)$. A more realistic model of the failure function could be inserted into Eq.(1). The important point is that Eq.(1) gives the probability of a sample of N (identical) magnets surviving y equivalent operational years with $\leqslant$ n failures under the assumption of a definite given failure law with a definite given effective failure rate λ. But this effective failure rate is what we want to infer from the test results, namely, that a sample of N magnets survives y equivalent operational years of testing with $\leqslant$ n failures.

We quantify our confidence that the actual effective constant failure rate λ is less than λ' may be inferred on evidence of an N-magnet sample surviving y equivalent operational years of (test) cycles with no more than n failures as

$$(2) \qquad R(\lambda \leqslant \lambda' | n,N,y,c) = \frac{\int_0^{\lambda'} P(n,N,y|\lambda)\, P_c(\lambda)\, d\lambda}{\int_0^{\infty} P(n,N,y|\lambda)\, P_c(\lambda)\, d\lambda}$$

In this expression we have assumed a continuous distribution of conceivable values in the domain of the integration variable λ. The function $P_c(\lambda)$ weights the probability distribution of the variable on this continuum on the evidence of all prior information, and the label c in the arguements of R indicates its functional dependence on this prior probability distribution. The numerator integrates over all possible values of λ up to the hypothetical test value λ' the probability of the outcome of our accelerated life test on condition of the prior distribution. The denominator integrates the same function over all possible values of λ. The ratio R is the probability that the value to be inferred from our test results lies in the interval $0 \leqslant \lambda \leqslant \lambda'$, subject to the prior distribution $P_c(\lambda)$.

The expectation of a given set of test data on the assumption of the variational trial distribution of prior probabilities $P_c(\lambda)$ is

$$(3) \qquad D(n,N,y) = \int P(n,N,y|\lambda) d\lambda P_c(\lambda)$$

The information entropy of this trial distribution is

$$(4) \qquad I[P_C(\lambda)] = -\int P_C(\lambda)d\lambda \ln P_C(\lambda)$$

The variational maximum of I subject to the constraints of the data D and the normalization condition is obtained as the zero of the functional derivative with respect to $P_C(\lambda)$ of the quantity

$$(5) \qquad \Lambda[P_C(\lambda),\alpha_0,\alpha_1] = I[P_C(\lambda)] - \alpha_0 - \alpha_1 D[P_C(\lambda)]$$

$$= \int P_C(\lambda)d\lambda[-\ln P_C(\lambda) - \alpha_0 - \alpha_1 P(n,N,y|\lambda)]$$

where α_0 and α_1 are Lagrange undetermined multipliers.

Setting to zero the functional derivative

$$(6) \qquad 0 = \frac{\delta\Lambda}{\delta P_C(\lambda)} = -\ln P_C(\lambda) - 1 - \alpha_0 - \alpha_1 P(n,N,y|\lambda)$$

we can solve for

$$(7) \qquad P_C(\lambda) = \exp[-(1+\alpha_0) - \alpha_1 P(n,N,y|\lambda)]$$

We see that $P_C(\lambda) = P(\alpha_1,n,N,y|\lambda)$ depends on n,N,y as well as λ. This is in addition to the assumption of the exponential failure function $p(\lambda,y)$.

The Lagrange multipliers are determined as solutions of the constraint equantions

$$(8) \qquad 1 = -\frac{\partial\Lambda}{\partial\alpha_0} = \int P_C(\lambda)d\lambda = e^{-(1+\alpha_0)}\int e^{-\alpha_1 P(n,N,y|\lambda)}d\lambda$$

giving

$$(9) \qquad e^{1+\alpha_0} = \int e^{-\alpha_1 P(n,N,y|\lambda)}d\lambda = Z[P_C(\lambda),\alpha_1]$$

and

$$(10) \qquad D = -\frac{\partial\Lambda}{\partial\alpha_1} = \int P(n,N,y|\lambda)e^{-\alpha_1 P(n,N,y|\lambda)}Z^{-1}d\lambda$$

$$= -\frac{\partial}{\partial\alpha_1}\ln Z[P_C(\lambda),\alpha_1]$$

We will have no need to determine α_0 or α_1 explicitly.[4]

The resulting prior probability distribution which maximizes the information entropy subject to data given in the form of expectations over the distribution is the canonical distribution

(11) $$P_c(\lambda) = \frac{e^{-\alpha_1 P(n,N,y|\lambda)}}{\int e^{-\alpha_1 P(n,N,y|\lambda)} d\lambda}$$

It is readily seen that in our canonical distribution α_1 plays the role of the usual Lagrange multiplier that in statistical mechanics is called $\beta = 1/k_B T$ where T is the temperature and k_B is Boltzmann's constant; $P(n,N,y|\lambda)$ plays the role of the Hamiltonian, a dynamical operator defined on the domain of variables λ and parameterized by n,N,y. The data D(n,N,y) plays the role of the average energy , the expectation of the Hamiltonian-like $P(n,N,y|\lambda)$; which is equivalent to specifying the temperature-like α_1.

The Canonical Distribution __________

We want to see what the canonical distribution looks like on the λ domain. The binomial distribution $P(n,N,y|\lambda)$ can be put in the form

(12) $$P(n,N,y|\lambda) = e^{-(N-n)\lambda y}\binom{N}{n}[1 - A(N,n)e^{-\lambda y} + O(e^{-2\lambda y})]$$

where A(N,n) is readily computed and is of the order of n. So $P(n,N,y|\lambda)$ is smaller than .05 =exp(-3) when

(13) $$\lambda y > [3 + \ln \binom{N}{n}]/(N\text{-}n) \equiv \lambda_c y$$

$P(n,N,y|\lambda)$ decreases exponentially for larger values of λy . For small λy we have

(14) $$P(n,N,y|\lambda) \xrightarrow[\lambda y \to 0]{} 1$$

The canonical distribution $P_c(\lambda) = Z^{-1}\exp[-\alpha_1 P(n,N,y|\lambda)]$ starts from $P_c(\lambda) = Z^{-1}\exp(-\alpha_1)$ at $\lambda y = 0$ and $ZP_c(\lambda)$ grows until it is effectively unity for $\lambda y > 2\lambda_c(N,n)y$. For given N the effective cutoff $\lambda_c(N,n)$ grows with increasing n. Thus, the canonical distribution behaves as the curves shown in Fig. 1. It is clear from the curves that the canonical distribution is quite similar to the microcanonical distribution chosen for our previous considerations, especially when α_1 is large. Large α_1 corresponds to low temperature in our statistical mechanics analogy. The principal distinctions are the soft,and n-dependent cutoff in the canonical distribution versus the hard, and n-independent cutoff in the microcanonical case.

The canonical and microcanonical distributions of prior probabilities perform very similarly under the integrals of our analysis. The microcanonical distribution

(15) $$P_c(\lambda) = \text{constant} > 0\ , \quad \lambda \geqslant \lambda_c\ ,$$
$$= 0\ , \quad \lambda < \lambda_c\ ,$$

with λ_c determined from Eq.(13) is a good working approximation for the canonical distributioin of Eq.(11). This microcanonical distribution was shown in Ref. 3 to afford a readily useable basis for estimating quantitatively the probabilistic confidence in specified availability levels on evidence of results of the accelerated magnet life tests as proposed in Refs. 2 . The detailed analysis is described and the results are summarized in Ref. 3.

An observation that will be of use for future refinements of our maximum information entropy-Bayesian analysis is to be made of the fact that with the canonical distribution $P_c(\lambda)$, as derived from our functional maximumization of the information entropy and shown in Eqs.(7) and (11), the expression for the confidence $R(\lambda \leqslant \lambda'|n,N,y,c)$ in Eq.(2) becomes

$$(16) \qquad R(\lambda \leqslant \lambda'|n,N,y,c) = \frac{-\int_0^{\lambda'} P_c(\lambda) \ln P_c(\lambda)\, d\lambda}{-\int_0^{\infty} P_c(\lambda) \ln P_c(\lambda)\, d\lambda} = I[P_c(\lambda),\lambda']/I[P_c(\lambda)] \;;$$

which is the fraction of the missing information in the distribution $P_c(\lambda)$ in the subspace $\lambda < \lambda'$ relative to that in the whole space.

Results of Present Analysis

The results of the present analysis for $n = 0$ can be calculated analytically, and are shown in Fig. 2. for the cases of $N = 100$ sample magnets andseveral values of the parameter $\alpha_1 = 3$ to 20. The results of the present analysis for $n = 1$ and 2, calculated numerically for $N = 100$ and several values of α_1 are shown in Figs. 3. and 4. The similarity of the results of the present analysis with the corresponding results of the previous analysis are evident. Interpolations and extrapolations in the canonical distribution cases follow the same rules derived previously for the microcanonical distribution results to good approximation.

References and Footnotes ___________

1. Accelerated life tests are necessary to assure that the SSC magnets can be expected with specified confidence levels to perform at their specified reliability, availability and lifetime. See: SSC Conceptual Design Report, SSC-SR-2020, Sec.5.2.1, p.267; and Report of Task Force on SSC Magnet System Test Site, SSC-SR-1001, Sec. III, A., p.9.

2. An accelerated life test protocol for the SSC magnets has been established in: V. Karpenko, correspondence with D. Brown, 7/1/86, and D. Brown, correspondence with V. Karpenko. The proposed test protocol is summarized in E. Shrauner, SSC-N-215, 8/86.

The proposed test protocol allows for sampling in about one week of testing the equivalent of about one year of ordinary operational cycling anticipated for the SSC. Availability requirements on the SSC magnet system are severe, because it is comprised of about 9500 individual magnets in a serial-fault configuration. The individual component magnets must therefore be extremely reliable. A general relation among the stationary availbility A_∞ , the mean time to repair MTTR, and the mean time between failures MTBF is

$$A_\infty = 1/[\,1 + \mathrm{MTTR}/\mathrm{MTBF}\,] \quad .$$

For our purposes MTTR is specified as in Ref. 1. to be about 1 week, or about 1/50 year. This determines MTBF through the above relation for A_∞ as specified in Ref. 1. Because of the serial fault configuration of the component magnets in the system, a mean effective failure rate λ for the individual component magnets (considered to be all the same) can be assigned by:

$$\lambda = 1/(9500\,\mathrm{MTBF}) = [(1/A_\infty)\,\text{-}1]/[9500\,\mathrm{MTTR}] \quad .$$

The assignment of a constant effective failure rate implies that the individual magnets fail for a complex variety of causes. This so-called exponential-failure-function hypothesis is chosen mostly for the simplicity of computation that it allows for obtaining rough estimates. If another failure-function is required for reality it may be incorporated into the present analysis without alteration of the structure laid out here.

The SSC Conceptual Design Report, Ref. 1., specifies A_∞ for the magnet system as required to be 96 percent. On the other hand, it also specifies the availability for the whole SSC as required to be 80 percent; and it might be thought that in the sense in which we are concerned the magnet system is in some way equivalent to the whole SSC. For this reason both these values are considered in Ref.3. as perhaps two extremes of a range of values for A_∞ . The value $A_\infty = .96$ specifies $\lambda = 1/(4560\text{ year})$, and the value $A_\infty = .80$ specifes $\lambda = 1/(760\text{ year})$, as determined through the above relation.

3. Analysis of the probabilistic confidence for specified availability levels obtainable from accelerated life tests on the SSC magnets has been outlined before in: E. Shrauner, SSC-N-215, 8/86; and E. Shrauner, SSC-N-254, 10/86.

4. Neither the microcanonical distribution as used in the earlier analysis nor the canonical distribution used in the present is normalizable on the -domain extended to the open interval $0 \leqslant \lambda$

$\leqslant \infty$. This does not cause trouble because Eq.(2) itself is of the form of a normalization condition. If this were to show itself to be problematic it could be regularized through the introduction of an upper cutoff on the domain of support of the distribution.

Figure Captions

Figure 1a. The Canonical Distribution versus $N\lambda y$ for $N = 100$, $\alpha_l = 5$ and $n = 0, 1, 2, \ldots$, as obtained from Eqs. (7) and (11).

Figure 1b. The Canonical Distribution versus $N\lambda y$ for $N = 100$, $n = 0$, and $\alpha_l = 3, 5, 10, 15, 20$.

Figure 2. The Probabilistic Confidence that the actual effective failure rate λ does not exceed λ' on evidence that a sample of $N = 100$ magnets survives y equivalent operational cycles of accelerated life testing with $n = 0$ failures for several different (large) values of the parameter $\alpha_l = 3$ to 20, plotted versus $N\lambda' y$.

Figure 3. The Probabilistic Confidence that the actual effective failure rate λ does not exceed λ' on evidence that a sample of $N = 100$ magnets survives y equivalent operational cycles of accelerated life testing with not more than $n = 1$ failure for several different (large) values of the parameter $\alpha_l = 3$ to 20, plotted versus $N\lambda' y$.

Figure 4. The Probabilistic Confidence that the actual effective failure rate λ does not exceed λ' on evidence that a sample of $N = 100$ magnets survives y equivalent operational cycles of accelerated life testing with not more than $n = 2$ failures for several different (large) values of the parameter $\alpha_l = 3$ to 20, plotted versus $N\lambda' y$.

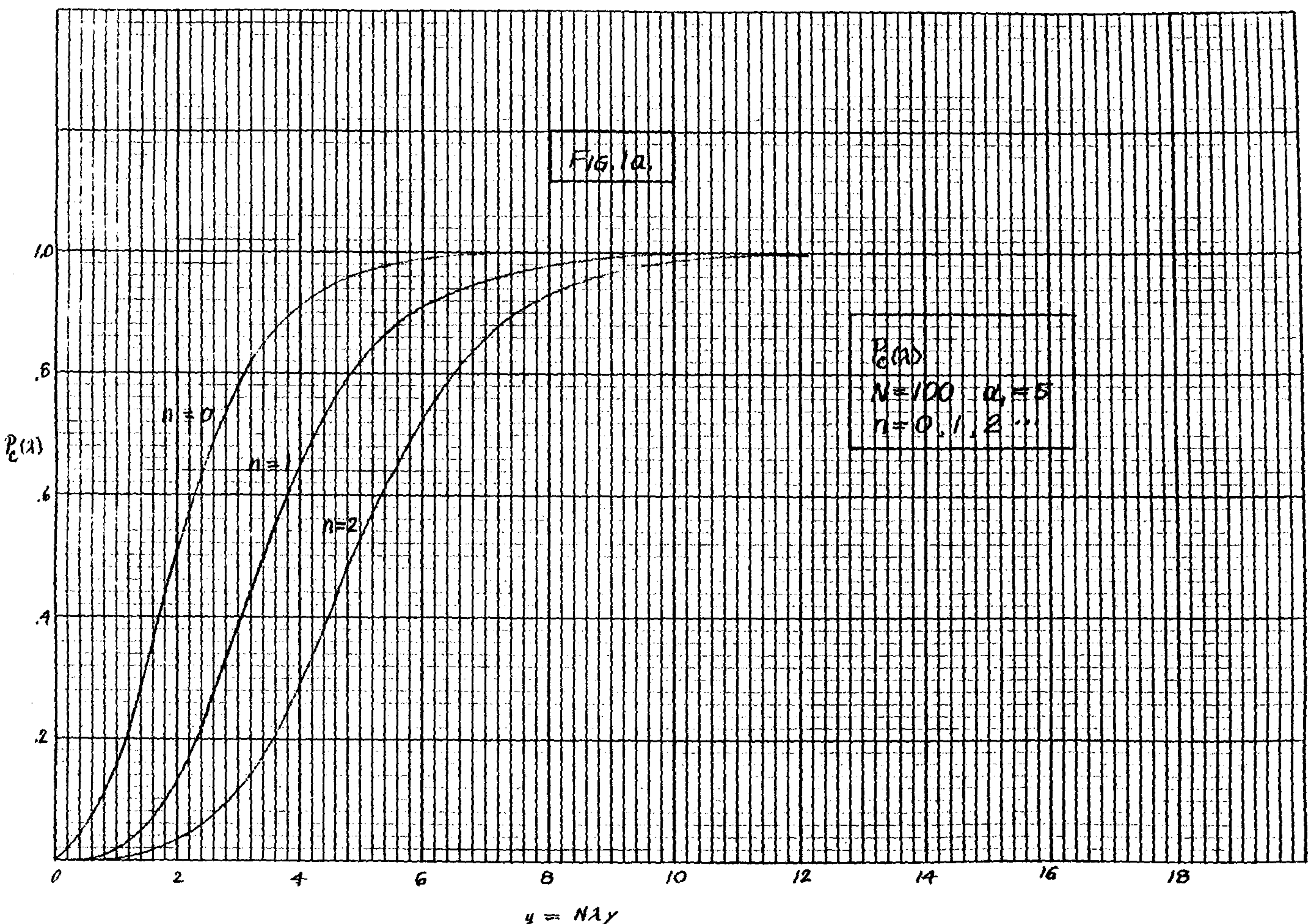

$P_c(\lambda)$
$N = 100$ $a = 5$
$n = 0, 1, 2 \cdots$
$n = 0$
$n = 1$
$n = 2$
1.0
.8
.6
.4
.2
0
2
4
6
8
10
12
14
16
18
$y = N\lambda y$

FIG. 1a.

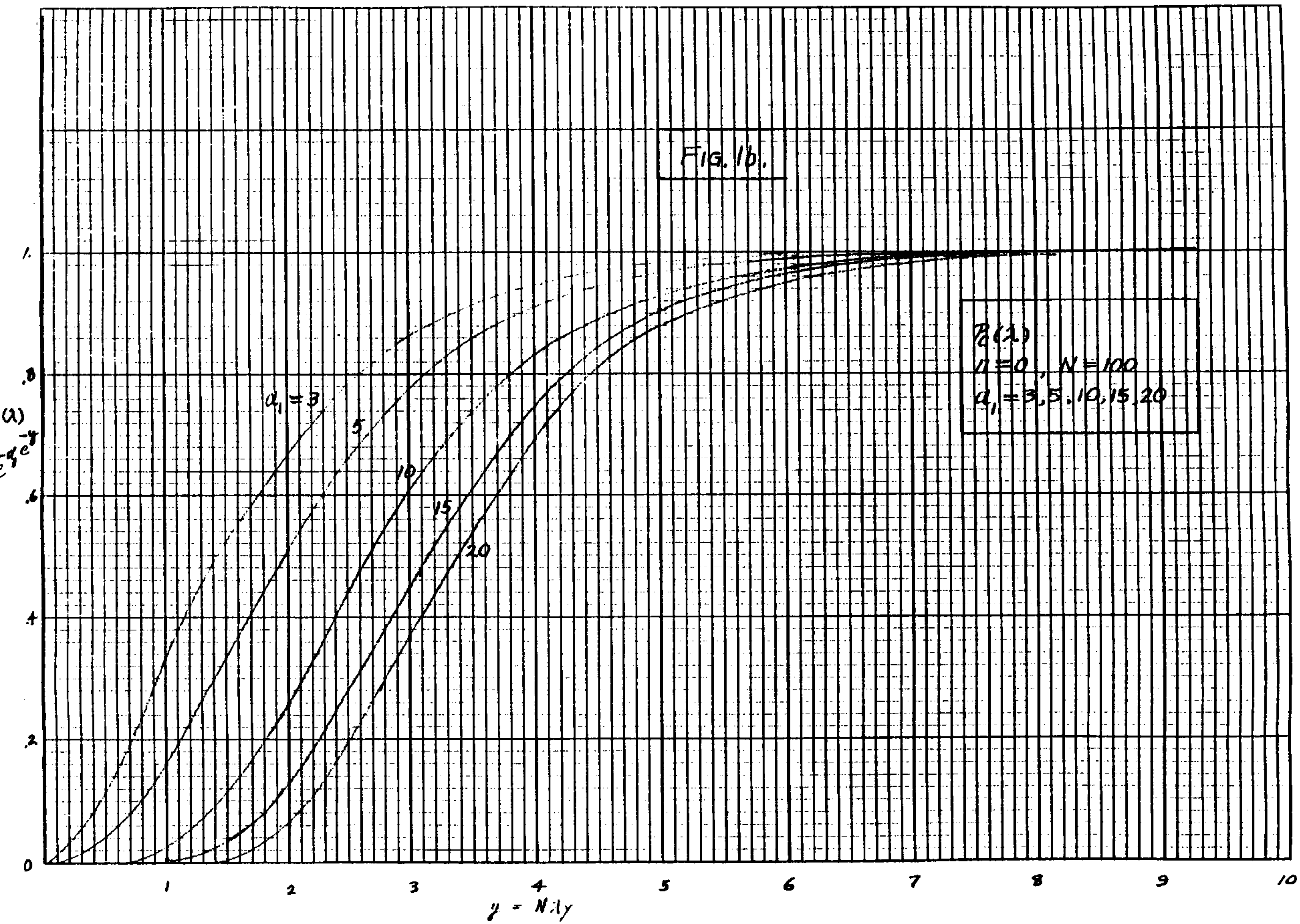
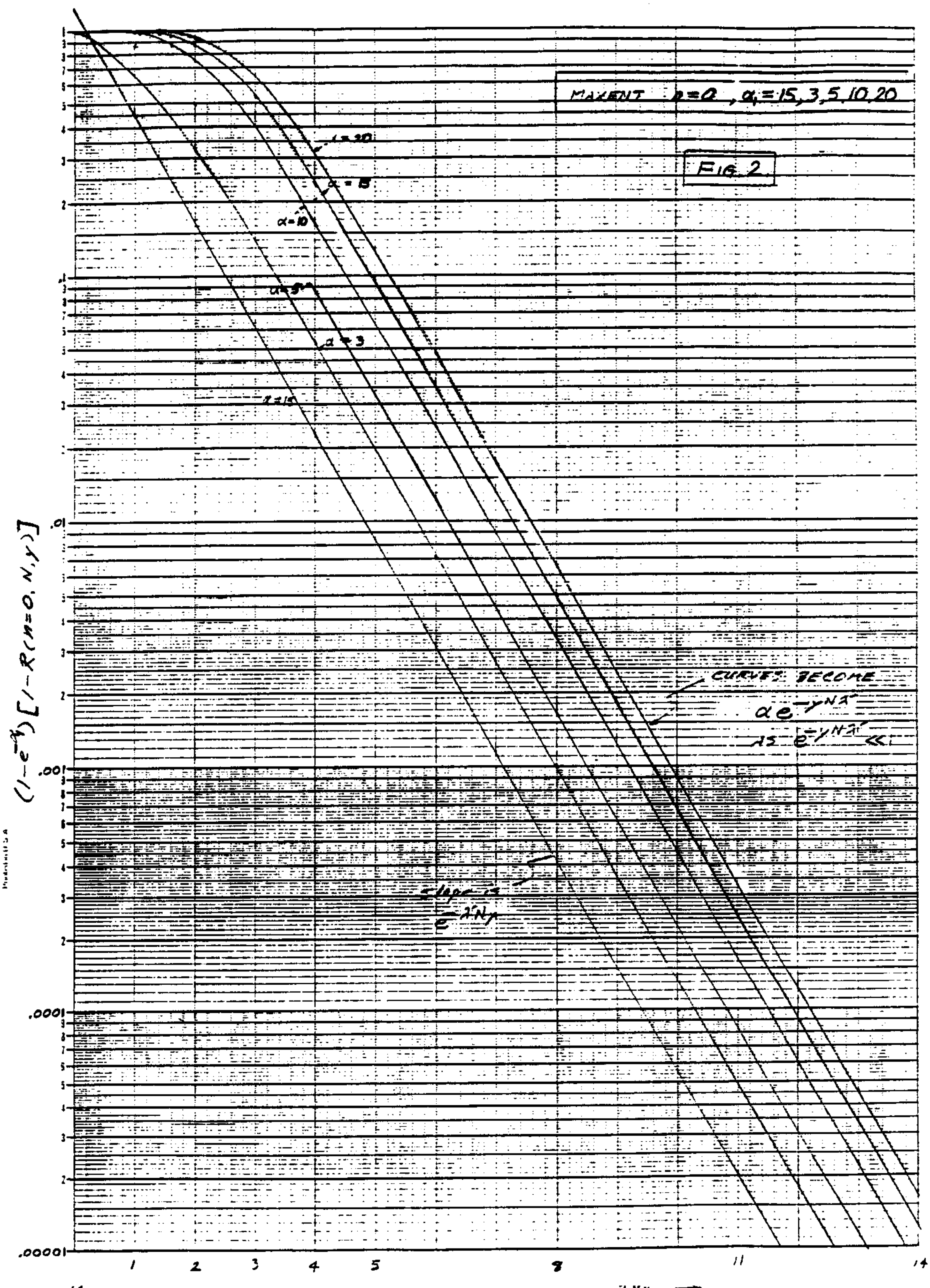
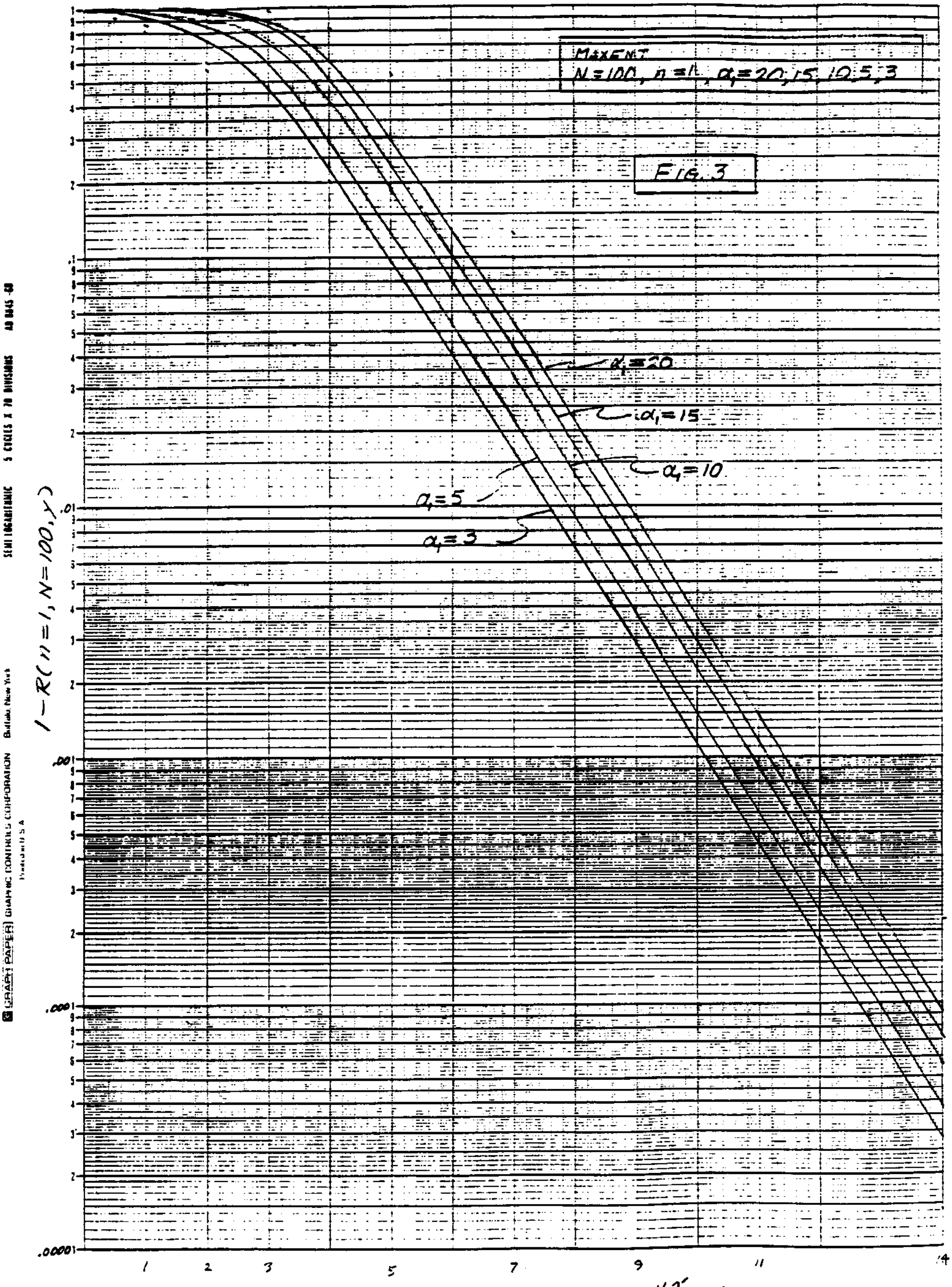
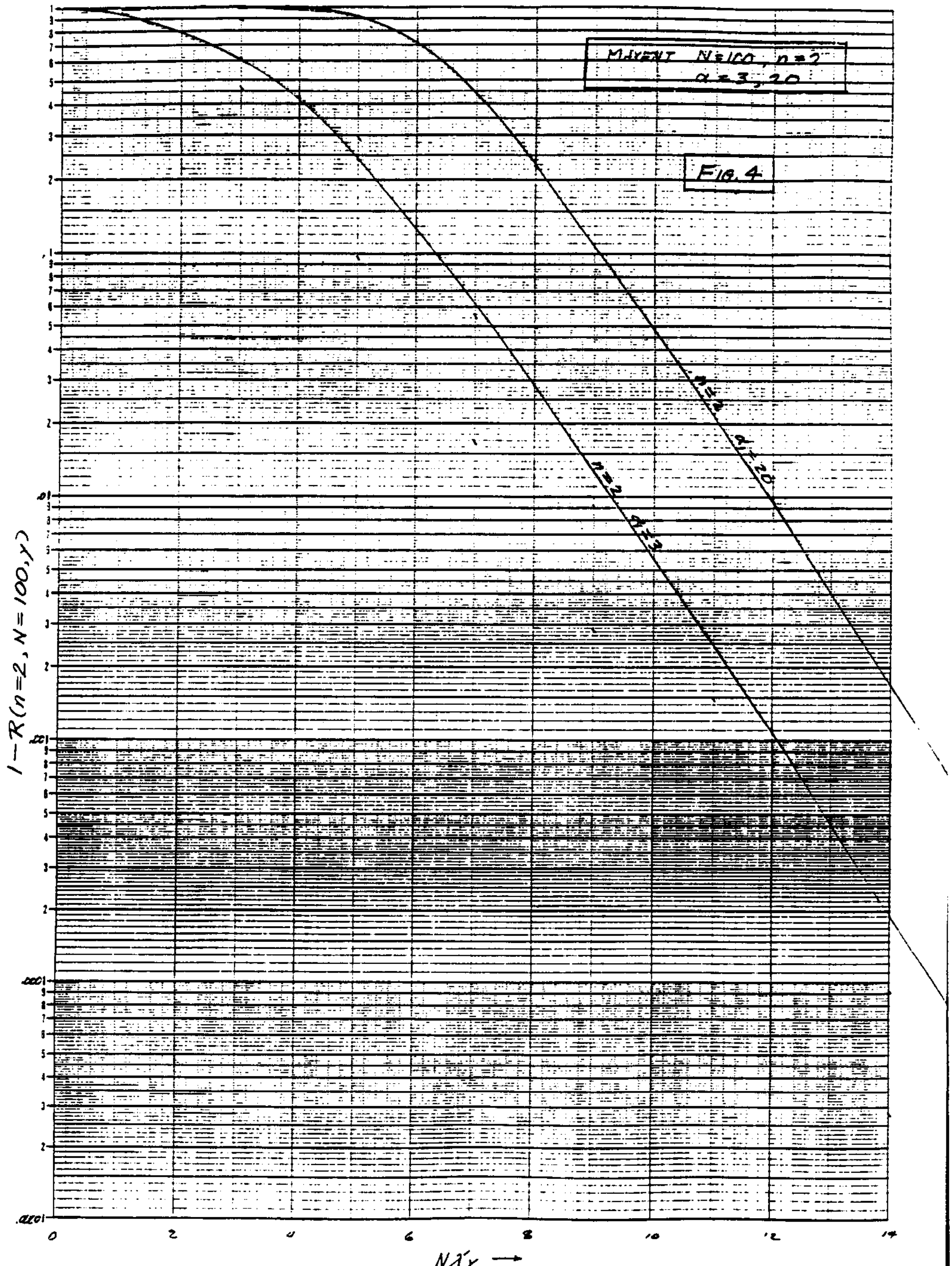


Fig. 1b.
$P_c(\lambda)$
$\eta = 0$, $N = 100$
$\alpha_1 = 3, 5, 10, 15, 20$
$\alpha_1 = 3$
5
10
15
20
0
.2
.4
.6
.8
1.
1
2
3
4
5
6
7
8
9
10
$y = N\lambda y$


MAXENT n=0 , α=15, 3, 5, 10, 20
FIG. 2
α=20
α=15
α=10
α=5
α=3
α=1.5
CURVES BECOME
$\alpha e^{-\gamma N \lambda^2}$
AS $e^{-\gamma N \lambda^2} \ll 1$
slope is $e^{-\lambda^2 N \gamma}$
$(1-e^{-\alpha})[1-R(n=0,N,\gamma)]$
.01
.001
.0001
.00001
1
2
3
4
5
8
11
14
FIG. 2


MAXENT
N = 100, n = 1, α1 = 20, 15, 10, 5, 3
FIG. 3
α1 = 20
α1 = 15
α1 = 10
α1 = 5
α1 = 3
1 − R(n = 1, N = 100, y)
N λ'y →
.01
.001
.0001
.00001
1
2
3
5
7
9
11
14
GRAPHIC CONTROLS CORPORATION Buffalo, New York
Printed in U.S.A.
SEMI LOGARITHMIC 5 CYCLES X 70 DIVISIONS


MAYENT N=100, n=2
α=3, 20
n=2 α=20
n=2 α=3
1−R(n=2, N=100, γ)
Nλx →
0
2
4
6
8
10
12
14
1
.1
.01
.001
.0001
.00001
FIG. 4