ON THE SEQUENCE OF "LUCKY NUMBERS"

A Thesis

Presented to

The Faculty of the Department of Mathematics

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Master of Arts

By

Beverly Roane Campbell

August 1964

APPROVAL SHEET

This thesis is submitted in partial fulfillment of
the requirements for the degree of
Master of Arts

Beverly Roane Campbell
Author

Approved, August 1964:

James E. Shockley
James E. Shockley, Ph.D.

Michael H. Kutner
Michael H. Kutner, M.A.

Hugh B Easler
Hugh B. Easler, M.S.

ACKNOWLEDGMENT

The author wishes to acknowledge aid received from the Computing Center, Johns Hopkins University, Baltimore, Maryland, and the Computational Laboratory, College of William and Mary, Williamsburg, Virginia, as well as to acknowledge the guidance received from Professor James E. Shockley.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The "lucky numbers" and their generating sieve considered in this paper are a variation of those first defined by S. Ulam and further considered by P. Erdös and E. Jabotinsky.

In Chapter I, basic sieving processes--sieve of Eratosthenes, a random sieve and a general sieving process--are discussed as a basis for the introduction of the "lucky number" sieve in the same chapter.

In Chapter II, general formulae are derived for the "lucky number" sieve and certain asymptotic approximations are listed as did Erdös and Jabotinsky.

In Chapter III, the experimental information obtained by the use of the 7094 and 1620 IBM computers is discussed.

The Appendices contain information concerning the computer program for the generation of the "lucky numbers" as well as graphical and tabular representations of the experimental results obtained.

ON THE SEQUENCE OF "LUCKY NUMBERS"

INTRODUCTION

Sieving processes are used considerably in number theory. The prime numbers are obtained by the sieve of Erathosthenes, which was devised more than two thousand years ago and is still essentially unaltered. The study of prime numbers has been an experimental as well as a theoretical investigation. Many of the facts that have been proved began as conjectures, based on inspection of an actual series of primes. Therefore to attempt to clarify properties of the primes that are consequences of the sieving process, David Hawkins [7] considered a random sieve and noted that a theorem analagous to the Prime Number Theorem is a common feature of sequences of numbers generated by sieves of a certain type. It should be noted that the random sieve preserves only the general features of the sieve of Eratosthenes. Thus it has been hypothesized that there must exist other sieves, having the same general characteristics as those of the sieve of Eratosthenes but differing in the details of sieving. This is not to say the yield will be the prime numbers, but the sequence of numbers so obtained would have some distinctive property. It is conjectured that such studies would be indicative of the properties of prime numbers which are a result of the definition of the sieve of Eratosthenes as opposed to those resulting from the definition of a prime number. It is further theorized that other sieves may indicate whether some problems concerning the primes are undecidable. Thus in 1956, Stanislaw

M. Ulam [5] and his associates at the Los Alamos Scientific Laboratory published limited results of yet another sieving process. Ulam termed the results of this sieving process "lucky numbers".

Experimental information was obtained for the "lucky numbers" less than 48,600 and excerpts were published by Ulam and others as well as some conjectures concerning the similarity of the properties of the "lucky numbers" as compared with those of the prime numbers. In the publication by Ulam and others, it was noted that this particular sieving process had been discussed some years ago with Paul Erdös.

In 1957, Paul Erdös and Eri Jabotinsky [4] considered a variation of the "lucky number" sieving process, that is, the same definition for the sieving process but omitting the integer one from the initial sequence and forcing the integer two to be the first "lucky number". This increased the similarity between the "lucky numbers" and the prime numbers and rendered the experimental information obtained by Ulam incorrect insofar as the actual values of the "lucky numbers" and only indicative of the other properties. Erdös and Jabotinsky conjectured and verified several asymptotic approximations for the nth "lucky number"; specifically verifying that as the prime number, p_k, is asymptotic to $k \cdot \log k$ so is the "lucky number", a_k, asymptotic to $k \cdot \log k$. The more precise approximations obtained by Erdös and Jabotinsky showed that for large "k" the "lucky number", a_k, is strictly greater than the corresponding prime number.

W. E. Briggs [1] investigated the variations which can be produced in the asymptotic approximations for the nth term by varying the "lucky number" sieve. Thus this investigation is not a direct continuation of the properties of the "lucky numbers" as reformulated in the publication by Erdös and Jabotinsky.

We will use the variation of the "lucky number"sieve used by Erdös and Jabotinsky. In Chapter II some asymptotic approximations due to Erdös and Jabotinsky will be derived and others will be listed. In Chapter III the significance of the experimental results obtained by high-speed computers will be interpreted. In Appendix A information obtained by use of the 7094 and 1620 IBM computers will be given in tabular and graphical form for the "lucky numbers" less than 131,000 which were computed by the use of the previously mentioned 7094 computer. The information given will include the density of the "lucky numbers" in specific intervals, the frequency of gaps varying in length from two units to ninety-two units inclusive and the error which the asymptotic approximations produce for known "lucky numbers".

In this work it is assumed the reader is familiar with the notational symbols "o", "O", and "$\sim$" as defined in An Introduction to the Theory of Numbers by Hardy and Wright.

CHAPTER I

ON BASIC SIEVING PROCESSES

The sieve of Eratosthenes, which produces the prime numbers, is relatively simple. From its rigidly ordered procedure it would seem possible to find a formula for the exact number of primes in any given interval or a formula expressing a function of "n" that would give a unique prime for every integral value of "n". In actuality mathematicians have not been able to do either.

In the sieve of Eratosthenes we begin with the sequence of natural numbers greater than one, that is:

$$A_1 = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots\}.$$

We let $a_1 = 2$ and form A_2 by removing all multiples of a_1 from A_1, thus obtaining: $A_2 = \{3, 5, 7, 9, 11, \ldots\}$. The number a_1 is called a <u>sieving number</u>. We now let $a_2 = 3$ be the second sieving number and form A_3 by deleting all multiples of a_2 from A_2, obtaining: $A_3 = \{5, 7, 11, \ldots\}$. If we continue in this manner and construct $A_1, A_2, A_3, \ldots, A_{n-1}$, we then let the (n-1)st sieving number be a_{n-1}, the first number in A_{n-1}. We now eliminate all multiples of a_{n-1} from A_{n-1} and obtain A_n. Thus the sequence of sieving numbers:

$$A = \{a_1, a_2, a_3, \ldots, a_{n-1}, \ldots\}$$

is the sequence of prime numbers. When we actually use this process a finite interval is necessarily considered and the procedure continues until the sieving number is so large that no multiples of it exist in the given finite sequence. The finite interval has been extended so that at present the computer division at Los Alamos Scientific Laboratory has a magnetic tape on which ninety million prime numbers are recorded. Using the above procedure and obtaining tables of prime numbers has enabled mathematicians to conjecture correctly the existence of the Prime Number Theorem which states that if "n" is a positive number, then the number of primes before it is asymptotic to n divided by the natural logarithm of n. This was not proved for approximately a century following its conjecture.

The difficulties of the Prime Number Theorem are related to the irregular manner in which the primes are distributed. David Hawkins [7], noting that the Prime Number Theorem does no more than state a statistical average, was led to consider a "statistical" model of the prime number distribution. This model is termed by Hawkins a random sieve.

Random Sieving Process

In a random sieve as in the sieve of Eratosthenes we begin with the sequence of natural numbers greater than one, that is: $B_1 = \{2, 3, 4, 5, 6, 7, 8, \ldots\}$ and we let $b_1 = 2$ be the first sieving number. Then some random way is devised so that each number following two has a probability of one-half of being omitted. Thus one-half of the integers will be eliminated in forming B_2 from B_1 as occurred in the first sieving operation of the sieve of Eratosthenes. However, each time the above operation is performed a different B_2 will be obtained. Now let b_2, the first number in B_2, be the second sieving number. We will again devise some random way such that each number following b_2 in the sequence B_2 will have a probability of $1/b_2$ of being omitted. If we continue in this manner we will construct B_1, B_2, B_3, ..., B_{n-1}. Then we define the $(n-1)$st sieving number as b_{n-1}, the first number in B_{n-1}, and we obtain B_n from B_{n-1} in the above manner with a sieving probability of $1/b_{n-1}$. Then the sequence $B = \left\{b_1, b_2, b_3, \ldots, b_{n-1}, \ldots\right\}$ of sieving numbers is termed the "sequence of random primes" by Hawkins.

In both sieves previously mentioned, the first number, not previously eliminated, in each sequence becomes a sieving number and eliminates a proportion of the remaining numbers equal to its reciprocal. However, the random sieve produces a different set of numbers each time it is used, while the set of prime numbers is invariant. Hawkins noted that the resemblance of the two sieves tends to intensify as they are increased in length, but the random sieve preserves only the general features of the sieve of Eratosthenes. The abnormalities of the latter are averaged out by randomizing them. In the sieve producing the prime numbers every sieving process except the first depends entirely on the sequences produced by the previous sieving processes and at every corresponding point the random sieve makes probability choices partially determined by its earlier statistical behavior. Hawkins also proved fairly readily that the Prime Number Theorem holds for "random primes", thus showing how parallel the two sieves are and impling that the random sieve can be taken as a criterion of normality. Hawkins further noted that the random sieve serves as a "statistical model" for the "lucky number" sieve, which at the time of the publication by Hawkins generated "lucky numbers" according to the notation of Ulam. This does not imply that the Prime Number Theorem is true for the "lucky numbers".

General Sieving Process

As previously stated in the introduction we will use the notation of Erdős and Jabotinsky in our work with the "lucky numbers" and the sieve by which the "luckies" are generated. We will first consider the general sieving process which with certain restrictions becomes the "lucky number" sieve. We begin with a sequence of natural numbers:

$$A^{(1)} = \left\{ a_1^{(1)}, a_2^{(1)}, a_3^{(1)}, \ldots, a_k^{(1)}, \ldots \right\}$$

and an integer b_1. We form

$$A^{(2)} = \left\{ a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, \ldots, a_k^{(2)}, \ldots \right\}$$

by deleting all $a_i^{(1)}$ from $A^{(1)}$ where "i" is of form $(1 + mb_1)$ with $m = 0, 1, 2, \ldots$. For example, if

$$A^{(1)} = \left\{ 2, 5, 3, 4, 7, 1, 15, 25, 30, 31, \ldots \right\}$$

and $b_1 = 4$, then $A^{(2)}$ is formed by deleting the first element of $A^{(1)}$ and every fourth element thereafter. Thus $A^{(2)}$ as obtained from the above example would be:

$$A^{(2)} = \left\{ 5, 3, 4, 1, 15, 25, 31, \ldots \right\} .$$

We now consider the integer b_2, which may be given initially or be calculated from $A^{(1)}$ or $A^{(2)}$ by some method given originally.

Again $A^{(3)}$ is formed from $A^{(2)}$ by deleting all $a_i^{(2)}$ where $i = 1 + mb_2$ $(m = 0, 1, 2, \ldots)$. We continue this process and obtain the resultant sequence

$$A = \left\{ a_1^{(1)}, a_1^{(2)}, a_1^{(3)}, \ldots, a_1^{(k)}, \ldots \right\}$$

generated by the general sieving process. Thus the resultant sequence is composed of the first elements of each subsequent sequence $A^{(i)}$ $(i = 1, 2, 3, \ldots)$. In the previously mentioned example we obtain $A = \{2, 5, \ldots\}$. The distinct difference between this sieve and that of Eratosthenes is that in the construction of $A^{(n)}$ from $A^{(n-1)}$ we remove elements with a specific index rather than removing the multiples of the sieving number.

Lucky Number Sieving Process

We will obtain the "lucky number" sieve from the general sieving process by first restricting the sequence $A^{(1)}$. The elements of $A^{(1)}$ will be defined by $a_k^{(1)} = 1 + k$ $(k = 1, 2, \ldots)$, that is the sequence of natural numbers greater than one. Let the first sieving number, b_1, be $b_1 = a_1 = a_1^{(1)} = 2$. We obtain $A^{(2)}$ from $A^{(1)}$ by deleting all terms from $A^{(1)}$ of form $a_{1+mb_1}^{(1)}$ $(m = 0, 1, 2, \ldots)$. Thus we delete $a_1^{(1)}, a_3^{(1)}, a_5^{(1)}, \ldots$ and then rename the remaining terms to obtain

$$A^{(2)} = \left\{ a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, \ldots, a_k^{(2)}, \ldots \right\},$$

which is specifically $A^{(2)} = \{3, 5, 7, 9, 11, 13, 15, 17, 19, \ldots\}$. Then let $b_2 = a_2 = a_1^{(2)} = 3$ be the second sieving number and obtain $A^{(3)}$ from $A^{(2)}$ by omitting all terms from $A^{(2)}$ of form $a_{1+mb_2}^{(2)}$ $(m = 0, 1, 2, \ldots)$ and renaming the remaining terms to obtain

$$A^{(3)} = \left\{a_1^{(3)}, a_2^{(3)}, a_3^{(3)}, \ldots, a_k^{(3)}, \ldots\right\},$$

which is specifically $A^{(3)} = \left\{5, 7, 11, 13, 17, 19, \ldots\right\}$. We continue this process and construct $A^{(1)}, A^{(2)}, A^{(3)}, \ldots, A^{(n)}$ defining $b_n = a_n = a_1^{(n)}$. Thus to obtain $A^{(n+1)}$ from $A^{(n)}$ we delete all terms from $A^{(n)}$ of form $a_{1+mb_n}^{(n)}$ $(m = 0, 1, 2, \ldots)$ and rename the remaining elements to form $A^{(n+1)}$. Note at each step we have $a_k = b_k$. Thus the resultant sequence of sieving numbers is:

$$A = \left\{a_1, a_2, a_3, \ldots, a_n, \ldots\right\} = \left\{a_1^{(1)}, a_1^{(2)}, a_1^{(3)}, \ldots, a_1^{(n)}, \ldots\right\}.$$

This is called the sequence of lucky numbers. The restrictions on the general sieving process necessary to produce the "lucky number" sieve do not affect the distinct difference between this and the prime number sieve. To generate the "lucky numbers" we remove elements with an index dependent on the sieving number and re-index each sequence rather than removing terms which are themselves multiples of the sieving number. Thus in the "lucky number" sieve the actual value of the term in the sequence is not considered. We consider only the index. However, as in the prime number sieve and in the random number sieve, we remove from each sequence a proportion of the numbers equal to the reciprocal of the sieving number in obtaining the subsequent sequence.

It should be noted that although the first eight "lucky numbers" are also prime numbers the two sieves are not consistly the same. The first exception is $a_9 = 25$ and further

$p_8 = 19$ is not one of the "lucky number" sequence. Thus we have definite similarities in the sieving processes, but each has its own characteristic yield.

CHAPTER II

ON FORMULAE FOR "LUCKY NUMBERS"

Before investigating the asymptotic properties of the "lucky number" sieve, we will develop several less precise estimates, using a restricted general sieve, which formulate a_k in terms of b_i $(i = 1, 2, \ldots, k-1)$. Thus we will not use the precise "lucky number" sieve, that is $a_k = b_k$ and $a_k^{(1)} = 1 + k$ for the initial estimates. We will impose these restrictions: 1) $b_i \geq 2$ and 2) the sequence $\{b_i\}$ is a non-decreasing sequence of integers.

In the following we shall denote by $\lceil x \rceil$ the smallest integer greater than or equal to x. We wish to first consider the generation of the sequence $A^{(i+1)}$ from the sequence $A^{(i)}$. It will be shown that:

$$(1) \qquad a_k^{(i+1)} = a^{(i)}_{\lceil c_i k \rceil} \qquad \text{where } c_i = \frac{b_i}{b_i - 1} .$$

It follows easily that if $2 < k \leq b_i - 1$ then $\lceil c_i k \rceil = k + 1$. This and the use of induction imply $\lceil c_i k \rceil = \lceil c_i (k-1) \rceil + 1$ for $k = 2, 3, \ldots, b_i - 1$. When $k = b_i$ we have $\lceil c_i b_i \rceil = \lceil c_i (b_i - 1) \rceil + 2$ since $\lceil c_i (b_i - 1) \rceil + 2 = b_i + 2$ and $b_i + 2 \geq c_i b_i > b_i + 1$. Thus for k less than b_i we would have deleted only $a_1^{(i)}$ from $A^{(i)}$ and we have $a_1^{(i+1)} = a_2^{(i)} = a^{(i)}_{\lceil c_i \rceil}$ followed by $a_k^{(i+1)} = a^{(i)}_{\lceil c_i k \rceil}$ for $k = 2, 3, \ldots, b_i - 1$. When $k = b_i$ we have an increase of two for $\lceil c_i k \rceil$ over the previous value $\lceil c_i (k-1) \rceil$ since $a_1^{(i)}$ and $a_{1+b_i}^{(i)}$ will have been deleted from $A^{(i)}$. Similarly if $b_i < k < 2b_i - 1$ we have $[c_i k] = [c_i(k-1)] + 1$, while if $k = 2b_i - 1$ then we have $[c_i k] = [c_i(k-1)] + 2$. Continuing in this manner it is easy to see that if n is a positive integer and $nb_i - (n-1) < k < (n+1)b_i - n$ then $[c_i k] = [c_i(k-1)] + 1$ while $k = nb_i - (n-1)$ implies $[c_i k] = [c_i(k-1)] + 2$. Thus in general for any k such that $nb_i - (n-1) < k < (n+1)b_i - n$, we have $a_k^{(i+1)} = a^{(i)}_{[c_i k]} = a^{(i)}_{[c_i(k-1)] + 1}$ and for any $k = nb_i - (n-1)$ we have the following:

$$a_k^{(i+1)} = a^{(i)}_{[c_i k]} = a^{(i)}_{[c_i(k-1)] + 2} .$$

Therefore this process reindexes the remaining elements of $A^{(i)}$ to form $A^{(i+1)}$.

Now we see by repeating (1) i times we obtain:

$$(2) \qquad a_k^{(i+1)} = a_{k'}^{(1)} \quad \text{with } k' = [c_1[c_2[\ldots[c_i k]\ldots]]].$$

The previous results are true provided $\{b_i\}$ is a non-decreasing sequence and $b_i \geq 2$. To obtain our initial or "zero-step" estimate for a_k we now consider a sieve which has defining properties similar to the "lucky number" sieve, namely $a_k^{(1)} = \lambda + k$ and $a_k = a_1^{(k)}$. Using (2) we see that when $i + 1 = k$ we have the following:

$$a_k = a_{i+1} = a_1^{(i+1)} = a_{k'}^{(1)} \quad \text{where } k' = [c_1[c_2[\ldots[c_{k-1}]\ldots]]] .$$

The formula "$a_1 = \lambda + 1$, $a_k = \lambda + k'$" is termed by Erdős and Jabotinsky the "explicit formula for the zero-step". By the notation $[\ ]$ we know $c_j m \leq [c_j m] < 1 + c_j m$ and applying this to the zero-step formula we obtain the following estimate for a_k $(k \geq 2)$:

$$(3) \qquad \lambda + \prod_{i=1}^{k-1} c_i \leq a_k < \lambda + 1 + \sum_{s=1}^{k-1} \Big(\prod_{i=1}^{s} c_i \Big).$$

Since $c_i > 1$ it follows by a simple induction that $\sum_{s=1}^{k-2} \Big(\prod_{i=1}^{s} c_i \Big)$

$\leq (k-2) \prod_{i=1}^{k-1} c_i$ where when $k = 2$ the sum is defined to be zero. Then by (1) we see that $\lambda + \prod_{i=1}^{k-1} c_i \leq a_k < \lambda + 1 + \prod_{i=1}^{k-1} c_i + \sum_{s=1}^{k-2} \left(\prod_{i=1}^{s} c_i \right) \leq \lambda + 1 + \prod_{i=1}^{k-1} c_i(1 + k - 2) = \lambda + 1 + (k-1) \prod_{i=1}^{k-1} c_i$. We will call this the "zero-step estimate for a_k" $(k \geq 2)$.

Consideration of the actual calculation of the value of k' in the zero-step formula shows an increase of one for successive brackets, $[\;]$, and then increases of two and then more. We wish to formulate a more precise estimate for a_k. Thus for fixed k, if a positive integer exists such that $b_{k-t} - 1 < t$ we let Q be the *smallest* such integer t. The first Q-1 brackets give an increase by one, that is $[c_{k-1}] = 2$, $[c_{k-2} \cdot 2] = 3$. Thus if no such t exists, each bracket gives an increase of one. The inductive assumption becomes $[c_{k-t}[c_{k-t+1}[\ldots]]] = t + 1$. Then we see $[c_{k-t-1}[c_{k-t}[\ldots]]] = [c_{k-t-1}(t+1)] = t + 2$ since $b_{k-(t+1)} \geq t + 1$ where $(t + 1)$ is a value less than or equal to Q. For every $q \leq Q$, we have the following "explicit formula for the *one-step*":

$$a_k = \lambda + [c_1[c_2[\ldots[c_{k-q} \cdot q]\ldots]]] .$$

When $q = Q$ "the explicit formula for the one-step" leads to the second estimate for a_k, "the one-step estimate":

(4) $$\lambda + Q \prod_{i=1}^{k-Q} c_i \leq a_k < \lambda + k \prod_{i=1}^{k-Q} c_i .$$

We see the second inequality of (4) follows from the use of $[c_j m] < 1 + c_j m$ which gives $a_k < \lambda + 1 + \sum_{s=1}^{k-Q-1} \left(\prod_{i=1}^{s} c_i \right) + Q \prod_{i=1}^{k-Q} c_i$ and from the fact that $1 + \sum_{s=1}^{k-Q-1} \left(\prod_{i=1}^{s} c_i \right) < (k-Q) \prod_{i=1}^{k-Q} c_i$. We state (4) in the more compact form:

(5) $$a_k = \lambda + (k - \theta(k-Q)) \prod_{i=1}^{k-Q} c_i \quad \text{where } 0 < \theta \leq 1.$$

We have considered the zero-step and the one-step formulae. Now we wish to consider an "m-step" ($m = 1, 2, \ldots$) or "multistep formula" for a_k. We will define q_m as the <u>smallest</u> integer for which $m(b_{k-q_m} - 1) < mq_m - \sum_{i=0}^{m-1} q_i$ where $q_0 = 0$. We will show by mathematical induction that:

$$a_k = \lambda + \lceil c_1 \lceil c_2 \lceil \ldots \lceil c_{k-q_m} \left\{ mq_m - \sum_{i=0}^{m-1} q_i \right\} \rceil \ldots \rceil \rceil \rceil .$$

Note when $m = 1$ we have the explicit formula for the one-step with $q_1 = Q$. The inductive assumption becomes:

$$a_k = \lambda + \lceil c_1 \lceil c_2 \lceil \ldots \lceil c_{k-q_n} \left\{ nq_n - \sum_{i=0}^{n-1} q_i \right\} \rceil \ldots \rceil \rceil \rceil .$$

The inductive assumption and the fact that $(q_{n+1} - q_n)$ is the number of brackets which give increases of $n+1$ imply the proof when $m = n + 1$, since we will have $(q_{n+1} - q_n)(n + 1) + nq_n - \sum_{i=0}^{n-1} q_i$ for the value of $\lceil c_{k-q_{n+1}+1} \lceil \ldots \lceil c_{k-q_n} \left\{ nq_n - \sum_{i=0}^{n-1} q_i \right\} \rceil \ldots \rceil$. Thus $a_k = \lambda + \lceil c_1 \lceil c_2 \lceil \ldots \lceil c_{k-q_{n+1}} \left\{ (n+1)q_{n+1} - \sum_{i=0}^{n} q_i \right\} \rceil \ldots \rceil \rceil \rceil$, which completes the induction.

Analogous to the manner of obtaining the one-step estimate for a_k from the explicit one-step formula, we obtain the "m-step" estimate from the explicit formula for the "m-step", that is:

$$\lambda + \left(\prod_{i=1}^{k-q_m} c_i \right) \left(mq_m - \sum_{i=0}^{m-1} q_i \right) \leq a_k < \lambda + \left(\prod_{i=1}^{k-q_m} c_i \right) \left(mq_m - \sum_{i=0}^{m-1} q_i \right) + R$$

with $R = 1 + \sum_{s=1}^{k-q_m-1} \left(\prod_{i=1}^{s} c_i \right) < (k - q_m) \left(\prod_{i=1}^{k-q_m} c_i \right)$. Thus we have the following restatement of the multi-step estimate for a_k:

$$a_k = \lambda + \left(\prod_{i=1}^{k-q_m} c_i \right) \left(mq_m - \sum_{i=0}^{m-1} q_i + \theta(k - q_m) \right), \quad 0 < \theta \leq 1. \tag{6}$$

Another result which will be needed in later work with the "lucky number" sieve is one which holds under the restrictive assumption that $\lim_{k \to \infty} \frac{b_k}{k} = \infty$. We will show that:

(7) $$\frac{a_k}{k} = (1 + o(1)) \prod_{i=1}^{k} c_i.$$

From (5) we have $a_k = \lambda + (k - \theta(k - Q)) \prod_{i=1}^{k-Q} c_i$. Now $\lim_{k\to\infty} \frac{b_k}{k} = \infty$ implies from $b_{k-Q} < Q + 1$ that $\lim_{k\to\infty} \frac{k - Q}{Q + 1} = 0$. Thus $k \geq Q + 1$ implies $\frac{k - Q}{k} = o(1)$ and we have $\frac{\lambda}{k} + \theta\left(\frac{k - Q}{k}\right) \prod_{i=1}^{k-Q} c_i = o(1) \prod_{i=1}^{k-Q} c_i$. Therefore $\frac{a_k}{k} = (1 + o(1)) \prod_{i=1}^{k-Q} c_i$. Now to prove (7) we need only prove $\prod_{i=k-Q+1}^{k} c_i = 1 + o(1)$. This follows by proving:

(8) $$\sum_{i=k-Q+1}^{k} d_i = o(1) \quad \text{where } d_i = \frac{1}{b_i - 1}.$$

We prove (8) by first choosing ϵ such that $0 < \epsilon < \frac{1}{2}$ and noting $\sum_{i=k-Q+1}^{k} d_i = \sum_{i=k-Q+1}^{[\epsilon k]} d_i + \sum_{i=[\epsilon k]+1}^{k} d_i$. We show that:

$$\sum_{i=k-Q+1}^{[\epsilon k]} d_i \leq \sum_{i=k-Q+1}^{[\epsilon k]} d_{k-Q+1} \leq \frac{[\epsilon k]}{Q - 1} < \frac{\epsilon k}{k - \epsilon k} < 2\epsilon.$$

The first inequality holds since Q is the smallest integer such that $b_{k-Q} - 1 < Q$ implying $b_{k-j} - 1 > j$ for $j < Q$ and $b_{k-(Q-1)} - 1 \geq Q - 1$. From $\frac{k - Q}{k} = o(1)$ we have $\frac{k - Q + 1}{k} = o(1)$ and we know that for sufficiently large k, $k - Q + 1 < \epsilon k$. This implies the second inequality. The third follows from the fact that $0 < \epsilon < \frac{1}{2}$. Now $\lim_{k\to\infty} \frac{b_k}{k} = \infty$ implies $\lim_{k\to\infty} \frac{[\epsilon k]}{b_{[\epsilon k]}} = 0$. Thus $\sum_{i=[\epsilon k]+1}^{k} d_i < \frac{k}{b_{[\epsilon k]}}$ and $\lim_{k\to\infty} \frac{1}{\epsilon} \frac{\epsilon k}{b_{[\epsilon k]}} = \frac{1}{\epsilon} \lim_{k\to\infty} \frac{[\epsilon k]}{b_{[\epsilon k]}} = \frac{1}{\epsilon} \cdot 0$ implies $\lim_{k\to\infty} \frac{k}{b_{[\epsilon k]}} = 0$.

Therefore this proves (8). Now to show $\prod_{i=k-Q+1}^{k} c_i = 1 + o(1)$ we define $n = k - Q + 1$ and consider $\prod_{i=n}^{k} c_i$ as a symmetric function $\prod_{i=n}^{k} c_i = \prod_{i=n}^{k} (1 + d_i)$, that is:

$\sum_{i=1}^{n}(1 + d_i) = 1 + (d_n + d_{n+1} + \dots + d_k) + (d_n d_{n+1} + \dots + d_{k-1} d_k)$ $+ \dots + (d_n d_{n+1} \dots d_k) \leq 1 + (d_n + \dots + d_k) + d_n^2(d_{n+2} + \dots + d_k)$ $+ \dots + (d_n)^{k-n}(1) \leq 1 + (d_n + \dots + d_k)(1 + d_n d_n^2 + \dots + d_n^{k-n}) =$ $1 + (d_n + \dots + d_k)\frac{1}{1 - d_n} = 1 + (\sum_{i=n}^{k} d_i)\frac{d_n}{d_n - 1}$. Therefore by (8) we have $\prod_{i=n}^{k} c_i = 1 + o(1)$.

The previous work gives an indication of the type of proof required to obtain asymptotic estimates for the "lucky numbers". Based on this work, Erdős and Jabotinsky listed several asymptotic estimates, most of them without proof. In addition, no information is given to indicate how "good" the approximations are for specific "lucky numbers" or how rapidly the "o" functions tend toward zero.

The formulae listed by Erdős and Jabotinsky [4] are:

(9) $a_k \sim k \cdot \log k$

(10) $a_k = k\log k + \frac{1}{2}k(\log\log k)^2 + (2 - \gamma)k\log\log k + o(k\log\log k)$.

In the next chapter we describe how these estimates compare with the actual values obtained from the "lucky number" sieve.

CHAPTER III

SIGNIFICANCE OF COMPUTER PROGRAM RESULTS

The "lucky numbers" between 1 and 131,000 generated on the 7094 IBM computer were used in various other computer programs to obtain experimental information concerning the yield of the "lucky number" sieve in this given finite interval.

A comparison between the number of prime numbers between 1 and 131,000 and the number of "lucky numbers" showed the "lucky numbers" to be less dense than the primes. However, they varied in the same manner; that is, both were more dense in the initial intervals considered and then declined in density in a similar manner as shown by Table I, page 24.

A yet unverified conjecture concerning the prime numbers is whether the limit inferior of $(p_{k+1} - p_k)$ is less than infinity. It has been conjectured that a similar statement is true for the "lucky numbers"; that is, $\liminf (a_{k+1} - a_k) < \infty$. Thus we considered the values of $(a_{k+1} - a_k)$. From Table II, page 26, we see for the "lucky numbers" between 1 and 131,000 that $a_{k+1} - a_k = 6$ is the most frequent difference and that $a_{k+1} - a_k = 92$ is the largest difference for the "lucky numbers" between 1 and 131,000. Further the majority of the differences computed were less than 30. Thus it would appear that the conjecture $\liminf (a_{k+1} - a_k) < \infty$ may be justified.

Experimental values were also obtained for $\frac{a_{k+1} - a_k}{\log k}$ since it has been shown for the prime numbers that lim sup $\frac{p_{k+1} - p_k}{\log k} = \infty$ and has been similarly conjectured for the "lucky numbers". However, in the experimental work the largest value attained by $\frac{a_{k+1} - a_k}{\log k}$ is 9.8349560. As is shown in Figure 4, page 37, the results, other than the exceptional one above, obtained for $\frac{a_{k+1} - a_k}{\log k}$ vary from 0.2 to 7.0 and develop no trend. Since the value of $\frac{a_{k+1} - a_k}{\log k}$ for the "lucky numbers" between 1 and 131,000 is less than ten and no apparent trend developed, it would seem to indicate that lim sup $\frac{a_{k+1} - a_k}{\log k} = \infty$ may be unfounded.

The initial asymptotic approximation, $a_k \sim k \cdot \log k$, obtained by Erdős and Jabotinsky is consistently less than the corresponding value of a_k for the "lucky numbers" between 1 and 131,000. However, the percentage of the difference between $k \cdot \log k$ and a_k, though less initially, leveled to approximately twenty-five percent of the value of a_k for the sample considered. Figure 2, page 27, gives a graphical representation of this at indicative indices. However, this information was obtained for each "lucky number". Thus $k \cdot \log k$ does not appear to be a close approximation since it maintains a constant deviation of approximately twenty-five percent from the actual value of a_k.

Other asymptotic approximations were stated which incorporated a "o" term. We considered experimentally the estimates from formulae (7) and (10): $\frac{a_k}{k \prod_{i=1}^{k} c_i} - 1 = o(1)$ and $\frac{a_k}{k \log\log k} - \frac{\log k}{\log\log k}$

$- \frac{1}{2}\log\log k - (2 - \gamma) = o(1)$. The later was the refinement stated by Erdős and Jabotinsky and supposedly the better approximation. However, it was stated with the details of proof suppressed and as can be noted from Figure 3, page 33, it is very irregular initially. Although a downward trend to zero develops, we find in the interval investigated that it remains greater than 0.9. Now $\frac{a_k}{k \prod_{i=1}^{k} c_i} - 1 = o(1)$ is regular initially and begins at approximately 0.6, showing a definite downward trend to zero immediately. However, it levels to a value greater than 0.2 and less than 0.23 with a slight but very slow trend to zero. Thus the least refined of the asymptotic approximations appears to be the better of those investigated experimentally, at least on the interval from 1 to 131,000.

BIBLIOGRAPHY

1. Briggs, W. E. *Prime-like Sequences Generated by a Sieve Process.* Duke Mathematical Journal, Vol. 30, No. 2 (June, 1963), pp. 297-311.

2. Cramér, Harald. *On the Order of Magnitude of the Difference Between Consecutive Prime Numbers.* Acta Arithmetica, Vol. 2 (1937), pp. 23-28.

3. David, Yosef. *On a Sequence Generated by a Sieving Process.* Riveon le Matematika, Vol. 11 (1957), pp. 26-31.

4. Erdös, Paul and Eri Jabotinsky. *On Sequences Generated by a Sieving Process.* Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Vol. 61 (1958), pp. 115-128.

5. Gardiner, Vernon, R. Lazarus, N. Metropolis and S. Ulam. *On Certain Sequences of Integers Defined by Sieves.* Mathematics Magazine, Vol. 29 (January-February, 1956), pp. 117-122.

6. Hardy, G. H. and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford: The Clarendon Press, 1960.

7. Hawkins, David. *The Random Sieve.* Mathematics Magazine, Vol. 31 (September-October, 1957), pp. 1-5.

8. Hawkins, David and W. E. Briggs. *The Lucky Number Theorem.* Mathematics Magazine, Vol. 31 (September-October, 1957), pp. 81-84.

9. Lehmer, D. N. *List of Prime Numbers from 1 to 10,006,721.* Washington: Carnegie Institution, 1914.

APPENDICES

APPENDIX A

PROGRAM FOR THE COMPUTATION OF "LUCKY NUMBERS" THROUGH 131,000 ON THE 7094 IBM COMPUTER

```
      XEQ
      LIST 8
      LABEL
C     LUCKY NUMBERS
      DIMENSION  NN(10000), NF(1000), NC(1000)
      MAX = 10000
      MAXIA = 1000
      MQ = 1
C     GENERATE SECOND SEQUENCE
      DO 10 JA = 1, MAX
      MQ = MQ + 2
10    NN(JA) = MQ
      IA = 1
49    MFAC = NN(1)
37    IF (MFAC - MAX)  26, 36, 36
26    IF (NN(MFAC + 1))  20, 30, 40
40    NF(IA) = MFAC
      IC = IA
      ML = IA
      NN(1) = 0
      MFIA = MFAC + 1
      LRET = 0
```

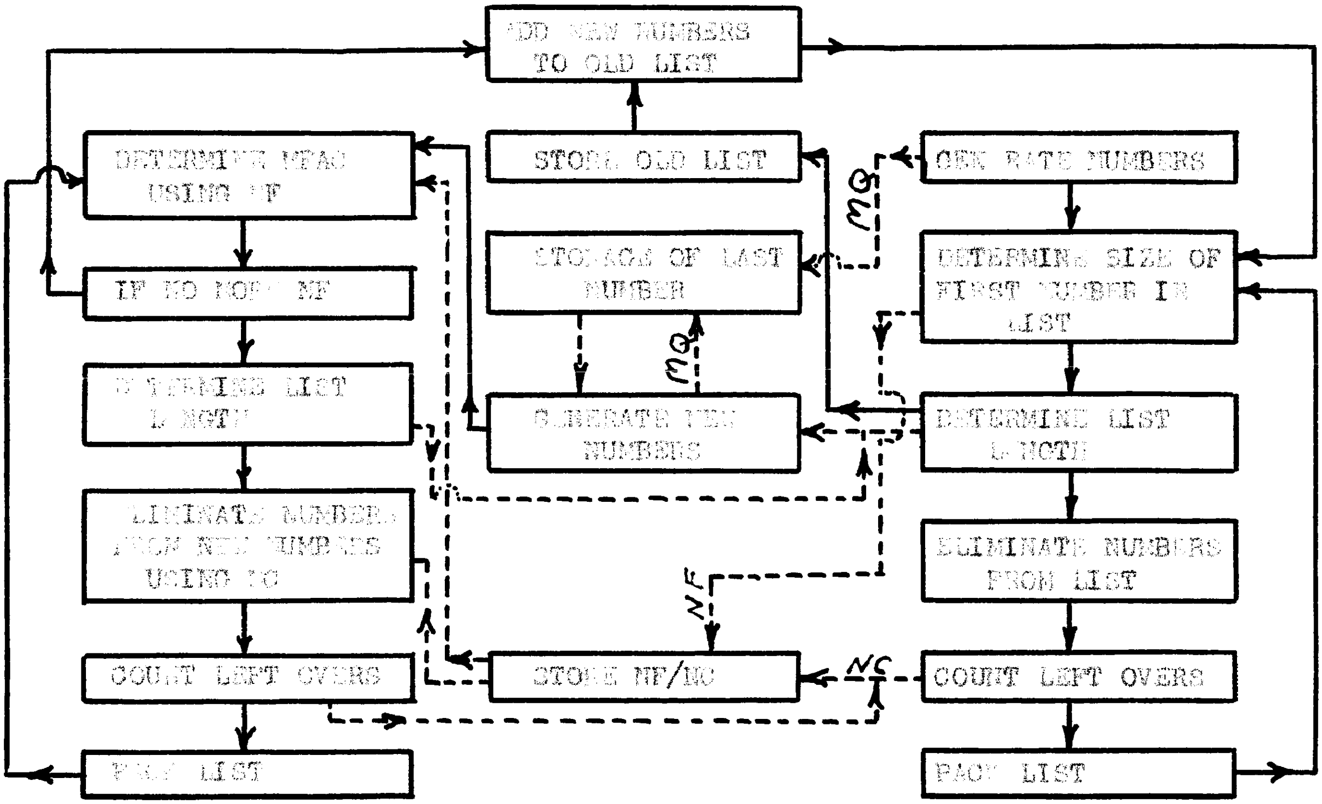
```
C   ELIMINATE NUMBERS
45  DO 60 JB = MFIA, MAX, MFAC
    MR = JB
    IF (NN(JB)) 20, 31, 41
41  NN(MR) = 0
    MRET = 0
60  CONTINUE
    MR = MR + MFAC
C   COUNT LEFT-OVERS
31  DO 70 KA = 1, MAX
    MRES = MR - MFAC + KA
    IF (MRES - MAX) 27, 27, 32
27  IF (NN(MRES))  20, 32, 70
70  CONTINUE
32  NC(IC) = KA - 1
    IF(MRET) 20, 244, 144
C   PACK THE LIST
244 IF(LRET)  20, 44, 344
344 MS = MSLI
    GO TO 345
44  MS = 0
345 MSI = MS + 1
    DO 80 KB = MSI, MAX
    IF (NN(KB)) 20, 80, 43
43  MS = MS + 1
    NN(MS) = NN(KB)
```

```
      IF(KB - MS)  20, 80, 81
81    NN(KB) = 0
80    CONTINUE
      MRET = 1
144   IF(LRET)  20, 50, 155
50    IF(IA - MAXIA)  51, 36, 36
51    IA = IA + 1
      GO TO 49
C     ADD MORE NUMBERS
30    MSJB = MS + 1
      DO 90 JC = MSJB, MAX
      MQ = MQ + 2
      IF (MQ - 131000) 90, 91, 91
90    NN(JC) = MQ
C     ELIMINATE UNLUCKY NUMBERS FROM NEW NUMBERS
91    LRET = 1
      JD = 0
      MSLI = MS
155   JD = 0
      IF (JD - ML)  157, 157, 156
157   MFAC = NF(JD)
      IC = JD
      MR = MSLI - NO(JD) + MFAC
      IF (MR - MAX) 193, 193, 31
193   IF (NN(MR))  20, 31, 93
93    NN(MR) =0
```

```
     MRET = 0
     MR = MR + MFAC
     IF(MR - MAX) 193, 193, 31
156  IF(MQ - 131000) 49, 36, 36
20   WRITE OUTPUT TAPE 6, 98
98   FORMAT (1H1, 18H NEGATIVE VALUE)
36   MLL = ML + 1
     WRITE OUTPUT TAPE 6, 101, IA,IC,I,JB,JC,JD,KA,KB,LRET,MRET
    1MAXIA,MAX,MFAC,MFIA,JLL,JM,MQ,MRES,MR,MSJB,MS,MSLI
101  FORMAT (1H1 / (1H, I11) )
     WRITE OUTPUT TAPE 6, 99, MQ,MS,ML,(I,NN(I),NF(I),NC(I),I=1,ML)
99   FORMAT (1H13X3HMQ=, I8, 3X3HMS=, I8, 3X3HML=, I8//
    18X3H(I)6X5HNN(I)6X15HNF(I)6X5HNC(I)/(1H, 4I11) )
    WRITE OUTPUT TAPE 6, 100, (I, NN(I), I=MLL, MS)
100  FORMAT (1H, 2I11)
97   CALL EXIT
     END
```

FIGURE

FLOW DIAGRAM OF PROGRAM TO GENERATE "LUCKY NUMBERS" ON THE IBM 7090



NC - Number of numbers after the last number to be eliminated

NF - "lucky numbers"

MQ - Size of last number generated

NFAC - The first number in the sequence to which new numbers are being added.

APPENDIX B

TABLE I

DENSITY OF "LUCKY NUMBERS" AND PRIMES IN CORRESPONDING INTERVALS THROUGH 131,000

	LUCKIES IN INTERVAL N		PRIMES IN INTERVAL N[1]	
N	No. in Interval	Total thru N	No. in Interval	Total thru N
1-2,000	252	252	303	303
2,000-4,000	205	457	247	550
4,000-6,000	193	650	233	783
6,000-8,000	183	833	224	1007
8,000-10,000	181	1014	222	1229
10,000-12,000	177	1191	209	1438
12,000-14,000	175	1366	214	1652
14,000-16,000	174	1540	210	1862
16,000-18,000	171	1711	202	2064
18,000-20,000	168	1879	198	2262
20,000-22,000	162	2041	202	2464
22,000-24,000	166	2207	204	2668
24,000-26,000	159	2366	192	2860
26,000-28,000	164	2530	195	3055
28,000-30,000	166	2696	190	3245
30,000-32,000	157	2853	187	3432
32,000-34,000	163	3016	206	3638
34,000-36,000	160	3176	186	3824
36,000-38,000	152	3528	193	4017
38,000-40,000	162	3490	186	4203
40,000-42,000	154	3644	189	4392
42,000-44,000	155	3799	187	4579
44,000-46,000	157	3956	182	4761
46,000-48,000	154	4110	185	4946
48,000-50,000	158	4268	187	5133
50,000-52,000	151	4419	186	5319
52,000-54,000	158	4577	181	5500
54,000-56,000	152	4729	183	5683
56,000-58,000	151	4880	190	5873
58,000-60,000	149	5029	184	6057
60,000-62,000	152	5181	175	6232
62,000-64,000	153	5334	181	6413
64,000-66,000	151	5485	178	6591
66,000-68,000	146	5631	183	6774
68,000-70,000	154	5785	161	6935

---

[1] D. N. Lehmer, List of Prime Numbers from 1 to 10,006,721, Washington, 1914.

TABLE I (continued)

N	No. in Interval	Total thru N	No. in Interval	Total thru N
70,000-72,000	146	5931	193	7128
72,000-74,000	149	6080	173	7301
74,000-76,000	150	6230	183	7484
76,000-78,000	149	6379	178	7662
78,000-80,000	147	6526	175	7837
80,000-82,000	149	6675	180	8117
82,000-84,000	148	6823	173	8190
84,000-86,000	144	6967	172	8362
86,000-88,000	152	7119	181	8543
88,000-90,000	144	7263	170	8713
90,000-92,000	143	7406	175	8888
92,000-94,000	147	7553	183	9071
94,000-96,000	147	7700	182	9253
96,000-98,000	146	7846	166	9419
98,000-100,000	139	7985	173	9592
100,000-102,000	147	8132	174	9766
102,000-104,000	151	8283	167	9933
104,000-106,000	144	8427	173	10106
106,000-108,000	146	8573	168	10274
108,000-110,000	138	8711	179	10453
110,000-112,000	152	8863	167	10620
112,000-114,000	135	8998	169	10789
114,000-116,000	151	9149	175	10964
116,000-118,000	146	9295	171	11135
118,000-120,000	139	9434	166	11301
120,000-122,000	138	9572	174	11475
122,000-124,000	149	9721	176	11651
124,000-126,000	138	9859	167	11818
126,000-128,000	146	10005	161	11979
128,000-130,000	145	10150	172	12151
130,000-131,000	73	10223	85	12236

TABLE II

NUMBER OF GAPS OF LENGTH K BETWEEN SUCCESSIVE "LUCKY NUMBERS"

Length (K) of Gap	No. of Gaps of Length K
1	1
2	1087
4	1082
6	1665
8	749
10	698
12	1178
14	542
16	571
18	629
20	222
22	256
24	418
26	159
28	159
30	245
32	66
34	71
36	97
38	47
40	42
42	60
44	36
46	27
48	30
50	16
52	15
54	9
56	5
58	3
60	10
62	0
64	4
66	7
68	3
70	3
72	2
74	2
76	2
78	2
80	0
82	0
84	0
86	0
88	1
90	0
92	1

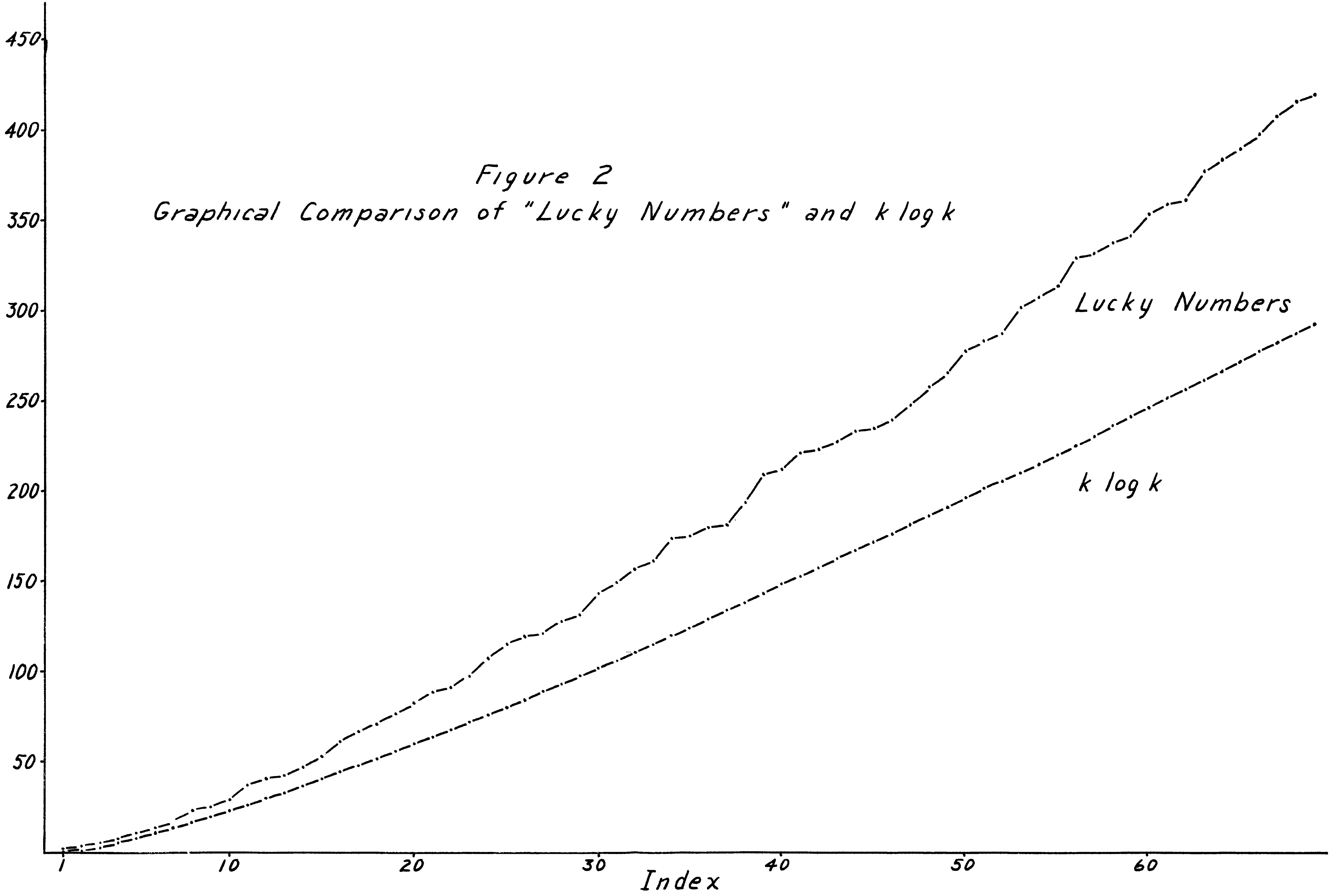


Figure 2
Graphical Comparison of "Lucky Numbers" and k log k

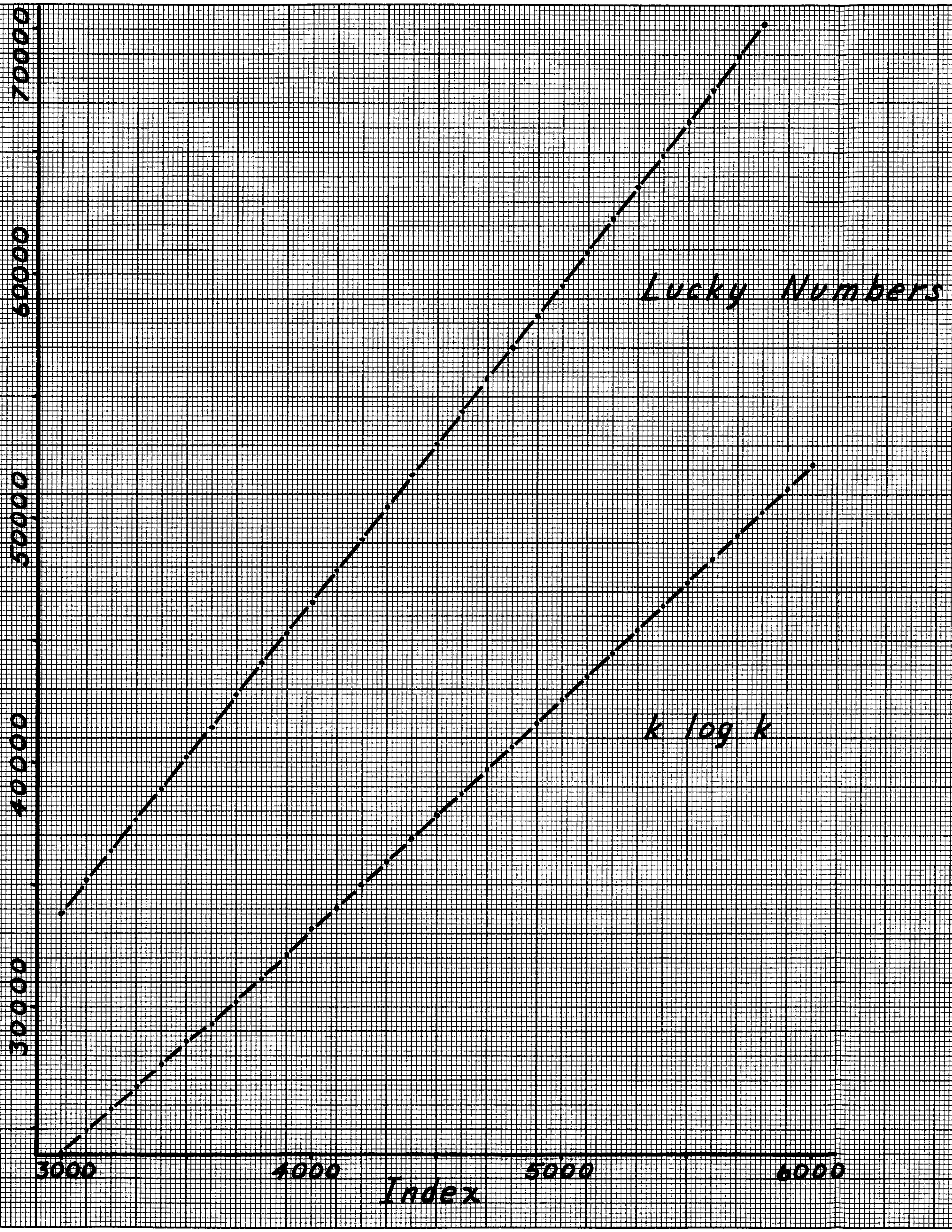


Figure 2 (Continued)

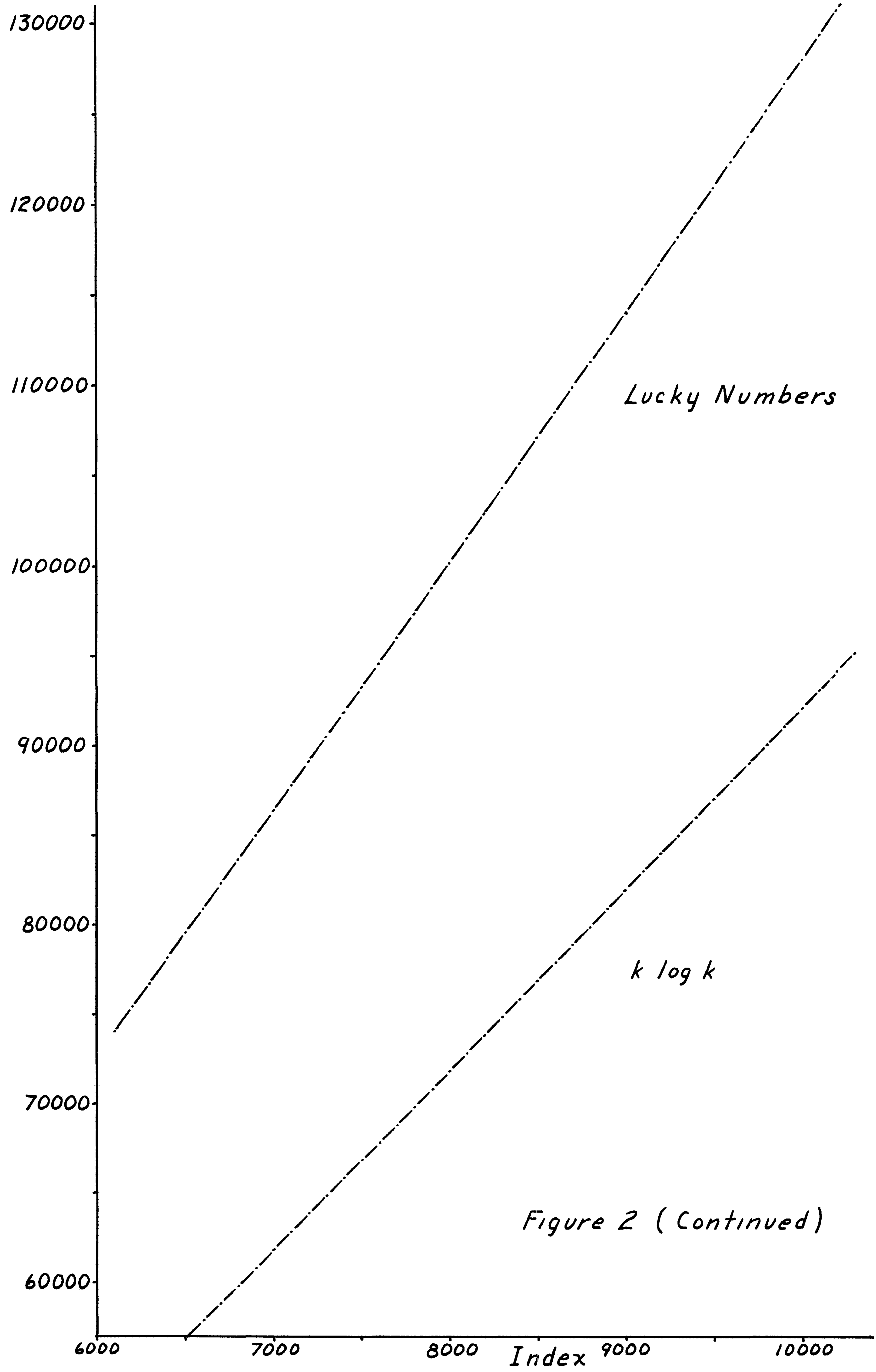


Figure 2 (Continued)

TABLE III

LIST OF "LUCKY NUMBERS" FOR INDICES 1-69

Index	Lucky Number
1	2
2	3
3	5
4	7
5	11
6	13
7	17
8	23
9	25
10	29
11	37
12	41
13	43
14	47
15	53
16	61
17	67
18	71
19	77
20	83
21	89
22	91
23	97
24	107
25	115
26	119
27	121
28	127
29	131
30	143
31	149
32	157
33	161
34	173
35	175
36	179
37	181
38	193
39	209
40	211
41	221
42	223
43	227
44	233
45	235
46	239
47	247
48	257
49	265
50	277
51	283
52	287
53	301
54	307
55	313
56	329
57	331
58	337
59	341
60	353
61	359
62	361
63	377
64	383
65	389
66	397
67	407
68	415
69	419

TABLE IV

LIST OF "LUCKY NUMBERS" FOR INDICES 5000-5070

Index	Lucky Number
5000	59549
5001	59557
5002	59561
5003	59591
5004	59597
5005	59605
5006	59609
5007	59611
5008	59617
5009	59639
5010	59651
5011	59653
5012	59677
5013	59711
5014	59723
5015	59741
5016	59761
5017	59767
5018	59785
5019	59801
5020	59837
5021	59863
5022	59905
5023	59911
5024	59933
5025	59941
5026	59951
5027	59953
5028	59965
5029	59969
5030	60011
5031	60043
5032	60047
5033	60077
5034	60083
5035	60085
5036	60097
5037	60107
5038	60113
5039	60121
5040	60137
5041	60149
5042	60157
5043	60167
5044	60175
5045	60179
5046	60181
5047	60187
5048	60203
5049	60233
5050	60235
5051	60239
5052	60287
5053	60295
5054	60299
5055	60307
5056	60325
5057	60337
5058	60343
5059	60347
5060	60361
5061	60367
5062	60377
5063	60397
5064	60421
5065	60427
5066	60443
5067	60467
5068	60473
5069	60481
5070	60491

TABLE V

LIST OF "LUCKY NUMBERS" FOR INDICES 10150-10223

Index	Lucky Number
10150	129977
10151	130007
10152	130013
10153	130045
10154	130057
10155	130093
10156	130105
10157	130109
10158	130133
10159	130157
10160	130163
10161	130193
10162	130223
10163	130247
10164	130253
10165	130271
10166	130273
10167	130277
10168	130291
10169	130303
10170	130333
10171	130345
10172	130361
10173	130375
10174	130381
10175	130403
10176	130411
10177	130423
10178	130435
10179	130439
10180	130447
10181	130451
10182	130465
10183	130471
10184	130487
10185	130493
10186	130499
10187	130501
10188	130511
10189	130517
10190	130553
10191	130555
10192	130567

Index	Lucky Number
10193	130585
10194	130589
10195	130591
10196	130603
10197	130631
10198	130663
10199	130645
10200	130649
10201	130663
10202	130675
10203	130709
10204	130721
10205	130747
10206	130763
10207	130777
10208	130795
10209	130817
10210	130823
10211	130847
10212	130853
10213	130885
10214	130903
10215	130913
10216	130931
10217	130933
10218	130945
10219	130961
10220	130963
10221	130975
10222	130981
10223	130991

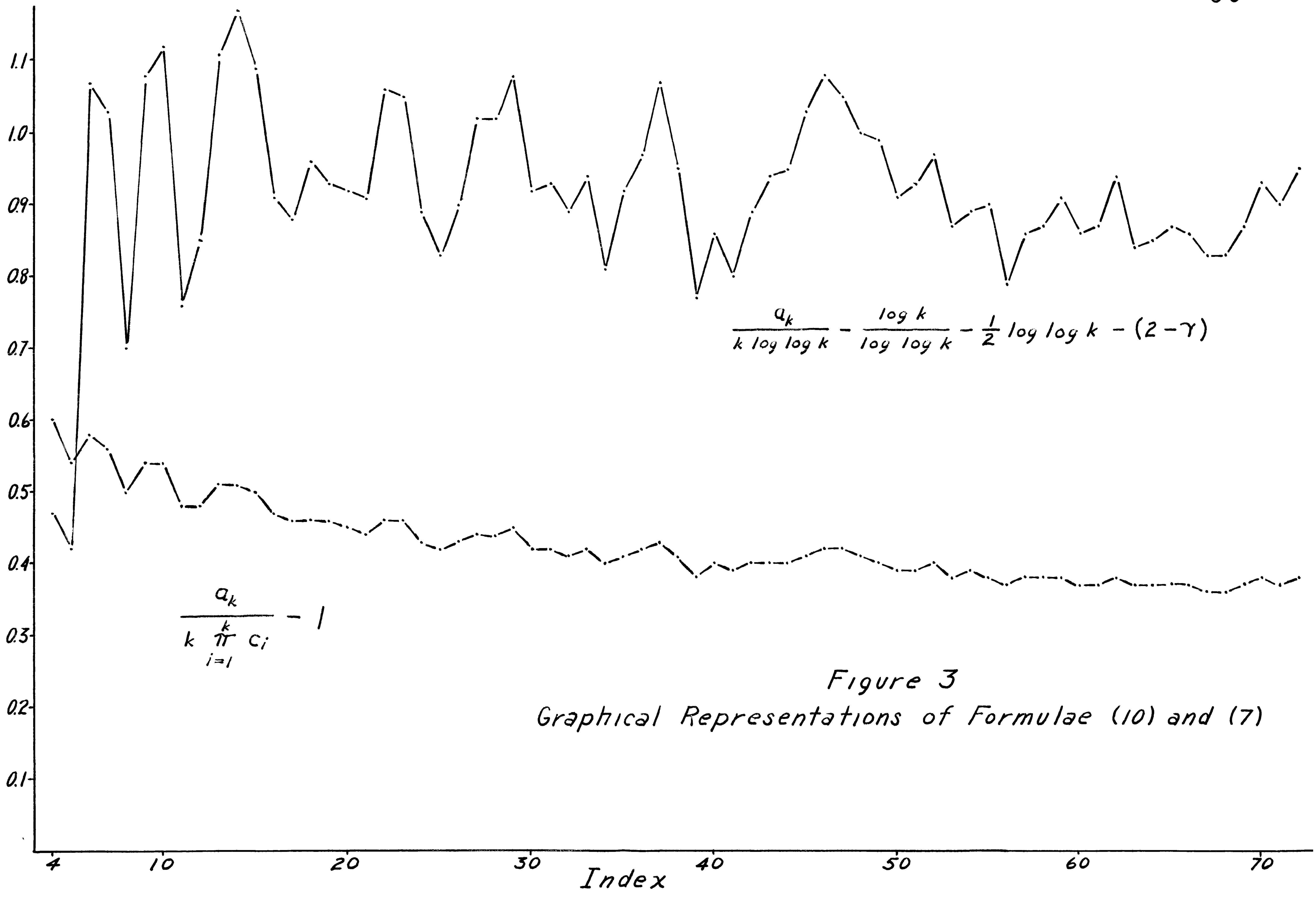


Figure 3
Graphical Representations of Formulae (10) and (7)

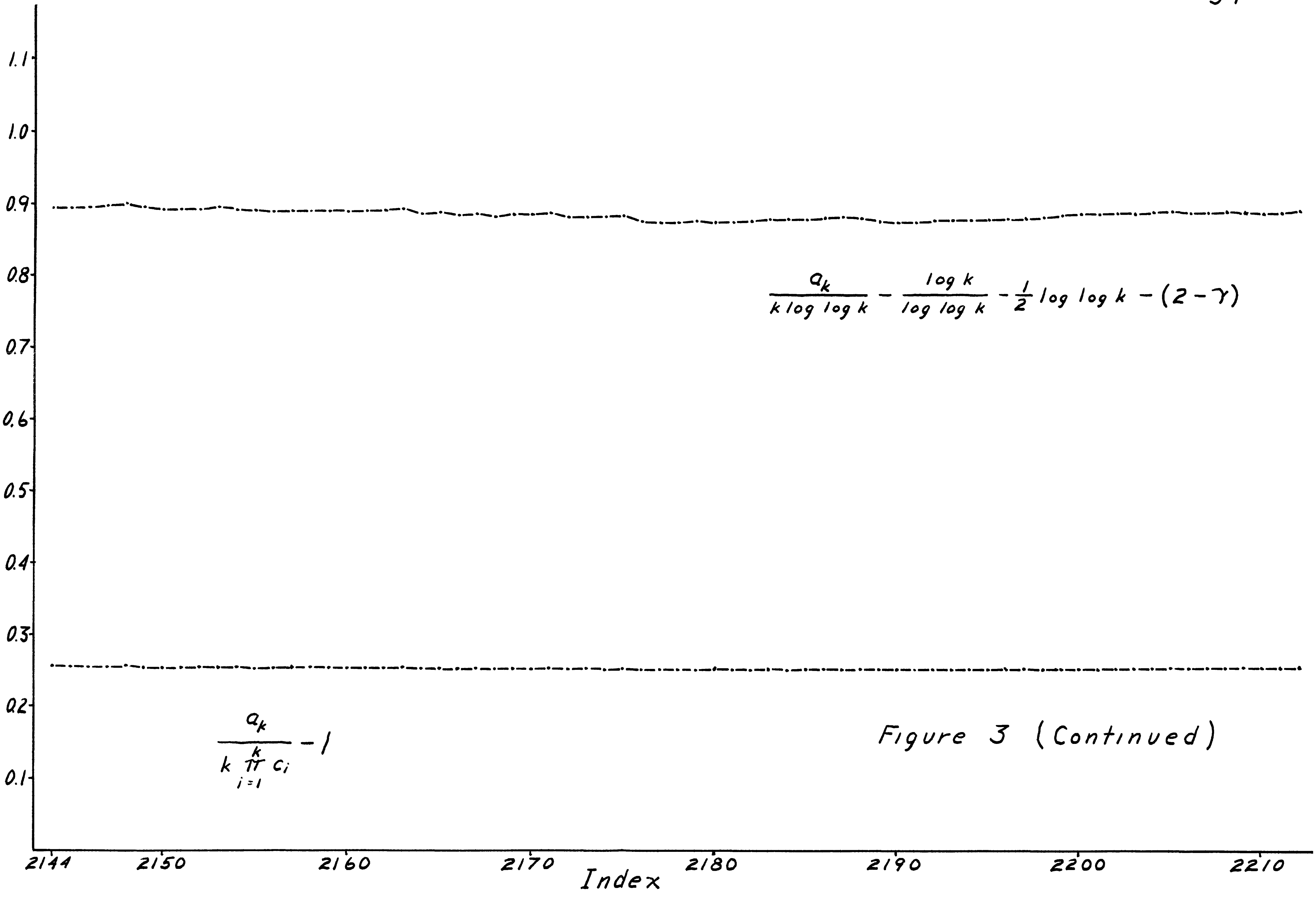


Figure 3 (Continued)

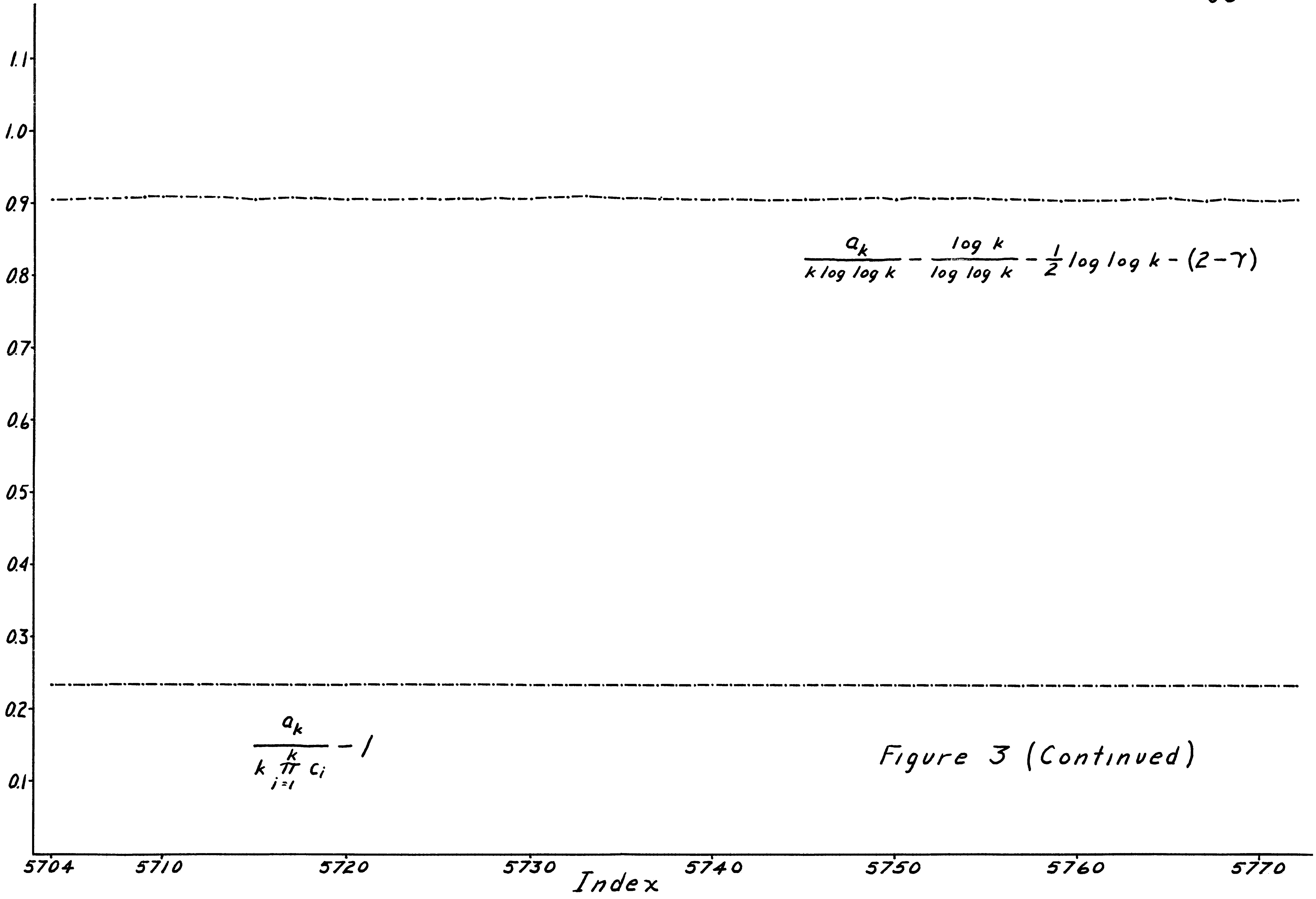


Figure 3 (Continued)

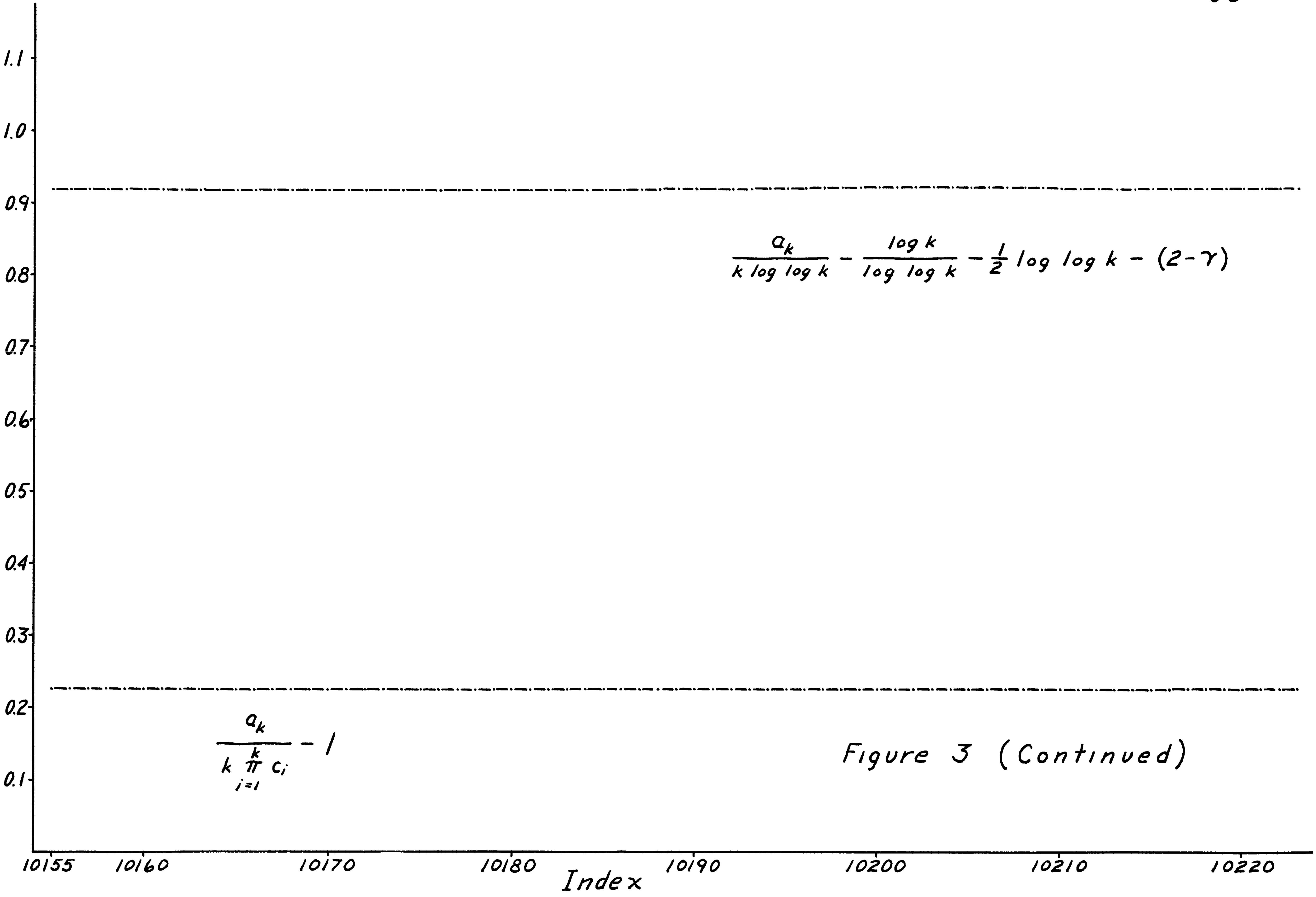


Figure 3 (Continued)

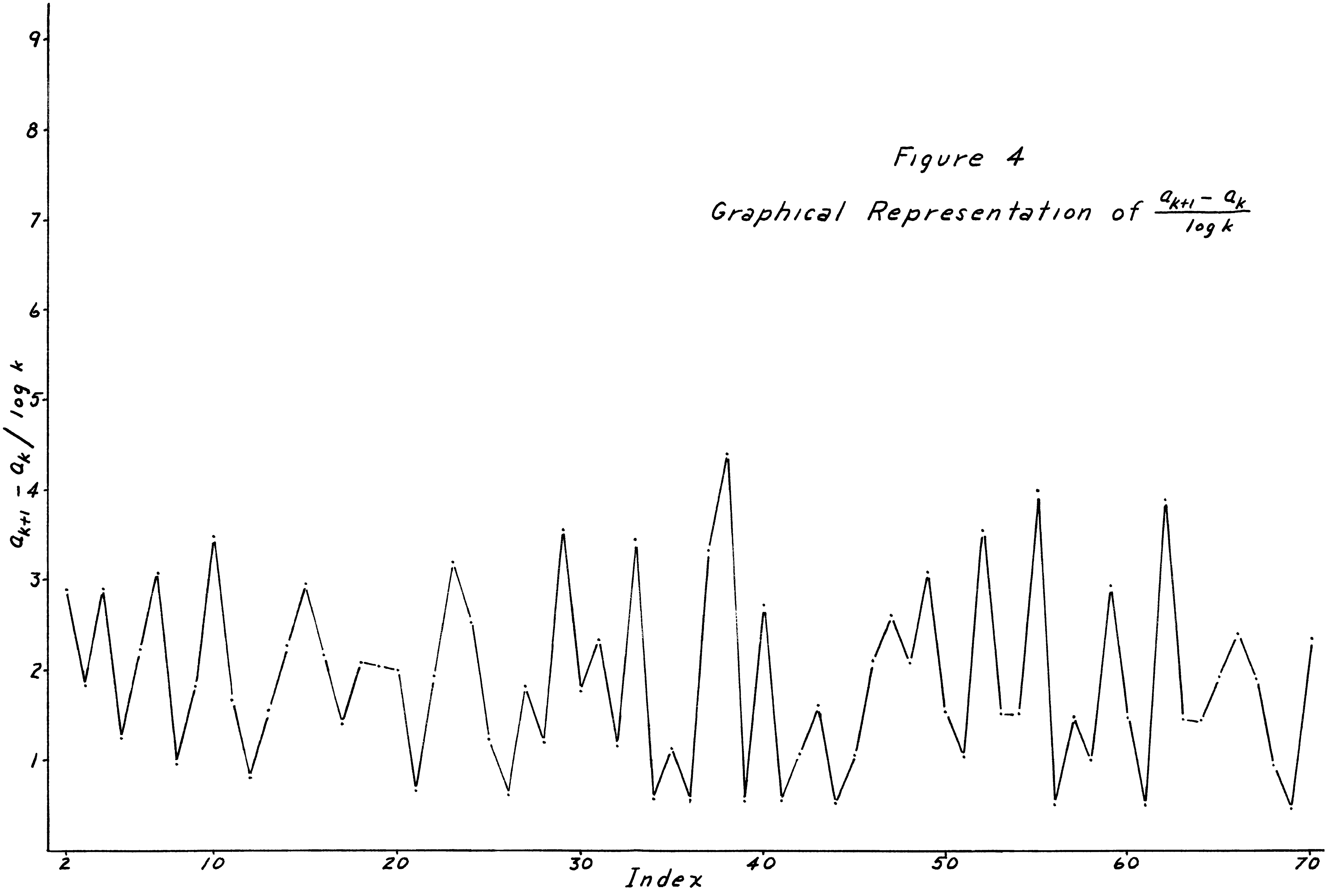


Figure 4

Graphical Representation of $\frac{a_{k+1} - a_k}{\log k}$

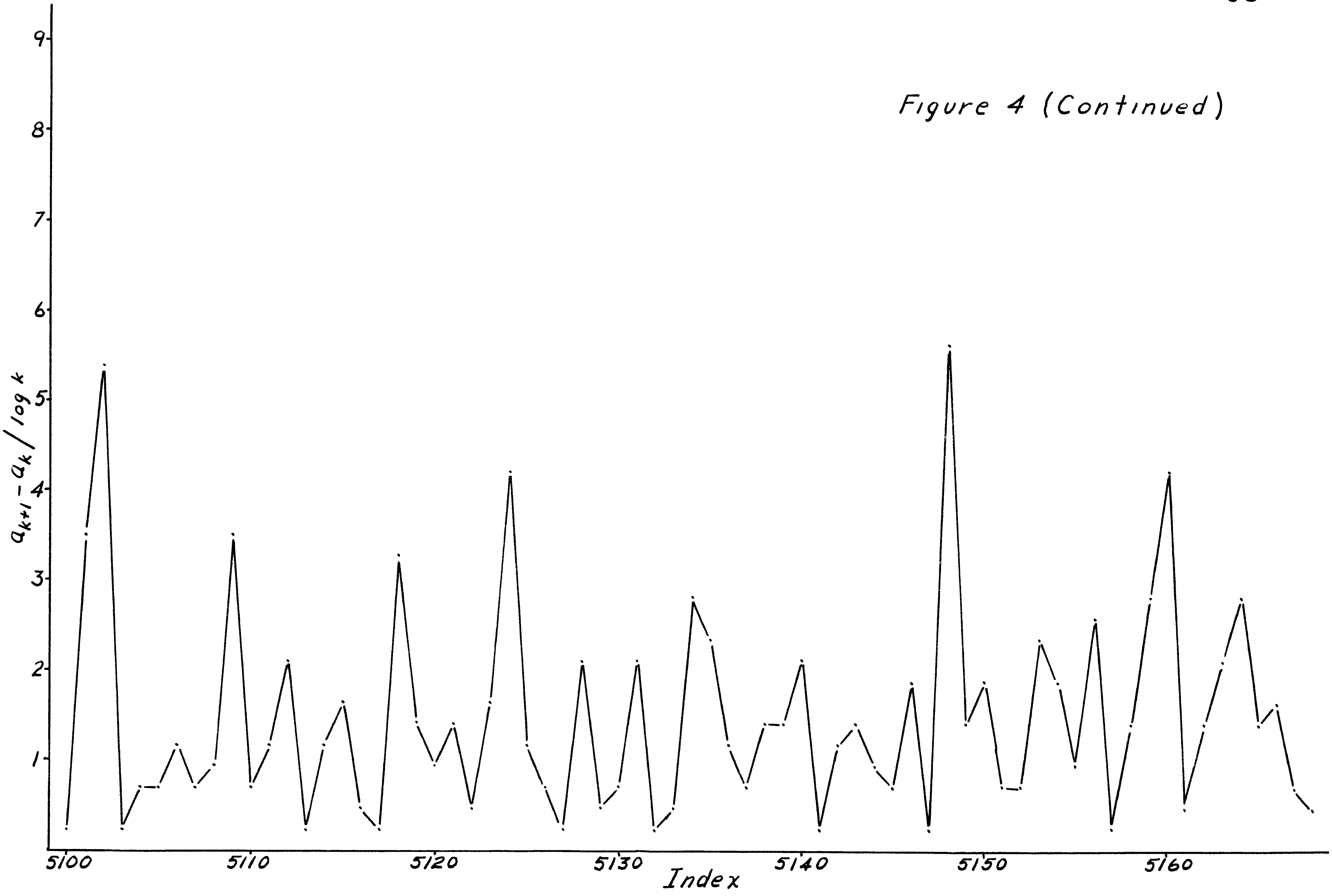


Figure 4 (Continued)

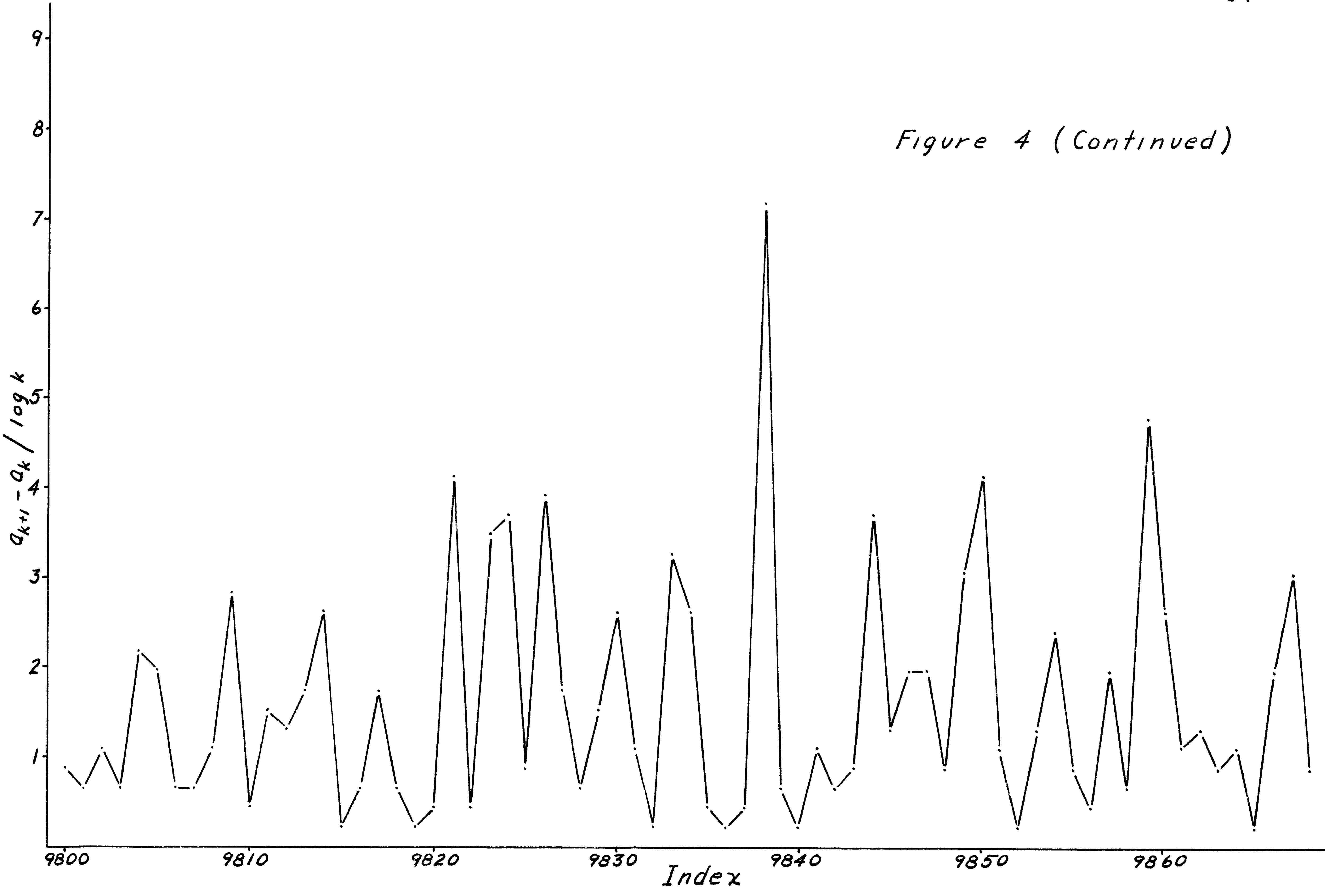


Figure 4 (Continued)