[9] — H. C. Griffith, and E. E. Posey, *A characterization of tame curves in 3-space*, Trans. Amer. Math. Soc. 79 (1955), pp. 12-34.

[10] J. Hempel, *A surface in S^3 is tame if it can be deformed into each complementary domain*, Trans. Amer. Math. Soc. 111 (1964), pp. 273-287.

[11] G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc. Colloq. Publ., Vol. XXVIII, Amer. Math. Soc., New York, 1942.

[12] R. L. Wilder, *Topology of Manifolds*, Amer. Math. Soc. Colloq. Publ., Vol. XXXII, Amer. Math. Soc., New York, 1949.

Reçu par la Rédaction 17. 8. 1967

Finitely additive measures and the first digit problem*

by

Richard Bumby and **Erik Ellentuck** (New Brunswick, N. J.)

1. Introduction. In his paper [4], R. S. Pinkham attempts to give a theoretical justification of the remarkable empirical conjecture that

(1) *the proportion of physical constants whose first significant digit lies between* 1 *and* n, *where* $1 \leqslant n \leqslant 9$, *is* $\log_{10}(n+1)$.

His approach consists of imposing 'reasonable' conditions on the distribution $F(x)$ of physical constants so as to yield (1) as a result. Two separate such considerations are given. In the first he argues that if every physical constant were multiplied by some real number $c > 0$, the resulting distribution $F(x/c)$ should agree with $F(x)$ regarding all data concerning first significant digits. This property of $F(x)$ is called scale invariance. Then in ([4], th. 1) it is shown that if the distribution $F(x)$ of physical constants is scale invariant and continuous, then (1) holds. His second argument consists of showing that (1) approximately holds independent of the specific nature of $F(x)$ and depending only on well-known statistical parameters associated with $F(x)$. In ([4], th. 2) bounds on this approximation are estimated in terms of the variation of the density function $f(x)$ associated with $F(x)$.

Our interest in this problem stems from the fact that several investigators have raised such questions as 'what is the probability that a natural number has property φ?' In this context we ask what is the probability that a natural number has a first significant digit which lies between 1 and n, where $1 \leqslant n \leqslant 9$. Our first task consists of giving a 'reasonable' definition of probability for natural numbers. This definition will then be tested against various sets of numbers to see whether it gives results which are in accord with our intuition. Finally, using a modified notion of scale invariance, we compute the probabilities of various sets associated with (1).

We use the following notation. Let $N = \{1, 2, 3, \ldots\}$ be the set of natural numbers, R the real numbers, and R^+ the non-negative real

* This paper is a report on some moonlighting performed by the authors while receiving support from the National Science Foundation via contracts GP 4368 and GP 5786, respectively.

numbers. Natural numbers will be denoted by lower case Latin letters, and real numbers by lower case Greek letters from the beginning of the alphabet. Use $[\alpha]$ for the greatest integer $\leqslant \alpha$. Subsets of N or R will be denoted by upper case Latin letters and sets of such subsets by upper case German letters. 'μ' and 'ν' will be reserved for measures, and sets of measures will be denoted by upper case script letters. $f\colon A \to B$ will have its usual meaning and $f\colon A \sim B$ will mean that f is a one-one function mapping A onto B. We use $|A|$ for the cardinal number of the set A. In our applications it will always be finite. We use $\cap$ for intersection, $\cup$ for union, $-$ for difference, $\emptyset$ for the empty set, and $\mathfrak{P}(A)$ for the set of all subsets (power set) of A. Also let $A+\beta = \{\alpha+\beta\colon \alpha \in A\}$ and $A\beta = \{\alpha\beta\colon \alpha \in A\}$. Intervals in R will be denoted in the usual way, for example $[0,1]$ is the unit closed interval. Intervals in N will receive 'N' as a subscript, thus $[1,2]_N = \{1,2\}$ and $[1,2)_N = \{1\}$. The letter I will be exclusively used to denote intervals. Let λ be the ordinary Lebesgue measure in R.

2. Invariant measures. Let $\mathfrak{A} \subseteq \mathfrak{P}(N)$ be a σ-algebra and $\mu\colon A \to R^+$ a countably additive measure function. If $\mu(N) = 1$, then μ is called a *probability measure*. Now a reasonable condition for the notion 'pick a number at random' is that we are equally likely to pick one number as another, i.e., $\{m\} \in \mathfrak{A}$ and $\mu(\{m\}) = \mu(\{n\})$ for $m, n \in N$. Clearly this condition is incompatible with $\mu(N) = 1$ and countable additivity. Since countable additivity seems least necessary among the preceding conditions, we replace it by finite additivity. But then we may take $\mathfrak{A}$ to be simply an algebra which contains each $\{n\}$ for $n \in N$. However, by a result of Tarski (cf. [5]), every finitely additive measure defined on $\mathfrak{A}$ can be extended to one defined on $\mathfrak{P}(N)$. Thus there is no loss of generality in assuming $\mathfrak{A} = \mathfrak{P}(N)$. For the purpose of this section we define a *measure* to be a function $\mu\colon \mathfrak{P}(N) \to R^+$ which satisfies (i) $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A, B \subseteq N$, $A \cap B = \emptyset$, and (ii) $\mu(N) = 1$. Let $\mathcal{M}$ be the class of all measures. Now finite additivity, $\mu(N) = 1$, and $\mu(\{m\}) = \mu(\{n\})$ for $m, n \in N$ all together imply that $\mu(\{n\}) = 0$ for $n \in N$. Thus we say a measure is *non-atomic* if it satisfies (iii) $\mu(\{n\}) = 0$ for every $n \in N$. Let $\mathcal{N}$ be the class of all non-atomic measures.

By sacrificing countable additivity we have been able to gain something—our measures are defined for all sets. This means that these measures will determine 'integrals' for which all bounded functions will be integrable. Each integral is then a continuous linear functional on the normed space $L_\infty(N)$ of bounded real valued functions on N. Since we wish to emphasize the measures rather than their associated linear functionals, the latter will be denoted $\int \ldots d\mu$. The continuity of these linear functionals follows from

$$\text{(2)} \qquad \operatorname{glb}_n f(n) \leqslant \int f\,d\mu \leqslant \operatorname{lub}_n f(n)\,.$$

From (2) it also follows that $\int 1\,d\mu = 1$. A linear functional L with this property gives a measure μ defined by $\mu(A) = L(\chi_A)$ where $\chi_A(n) = 1$ if $n \in A$, $\chi_A(n) = 0$ if $n \notin A$. One now has an explicit one-one correspondence between measures and certain linear functionals. The set of all measures can be topologized by relativizing the weak* topology on the dual of $L_\infty(N)$. This makes the set of all measures into a compact space. A class of measures will be said to be *closed* if it is closed in this topology.

Let $\mathcal{C} \subseteq \mathcal{M}$ be a class of measures. For any set $A \subseteq N$ define $\mathcal{C}(A) = \{\mu(A)\colon \mu \in \mathcal{C}\}$. We say that A is $\mathcal{C}$-*measurable* if $\mathcal{C}(A)$ contains exactly one element, i.e., if all measures of $\mathcal{C}$ assign the same value to A. It is also useful to define $\mathcal{C}(f) = \{\int f\,d\mu\colon \mu \in \mathcal{C}\}$ for any function f. (Note: $\mathcal{C}(A) = \mathcal{C}(\chi_A)$.) We shall be interested in determining these sets for various classes of measures. The classes that interest us are all *convex*, i.e., if $\mu_1, \mu_2 \in \mathcal{C}$ and $0 \leqslant \alpha \leqslant 1$, then the measure ν defined by $\nu(A) = \alpha\mu_1(A) + (1-\alpha)\mu_2(A)$ is also in $\mathcal{C}$. In this case $\mathcal{C}(f)$ is always an interval. In particular, if we can find $\mu_1, \mu_2 \in \mathcal{C}$ with $\mu_1(A) = 0$, $\mu_2(A) = 1$, then $\mathcal{C}(A) = [0,1]$. Measures of this type are most easily obtained using ultrafilters. Thus in [5] Tarski finds a measure $\mu \in \mathcal{N}$ by taking a non-principal ultrafilter $\mathfrak{U} \subseteq \mathfrak{P}(N)$ and defining

$$\text{(3)} \qquad \mu(A) = 1 \text{ if } A \in \mathfrak{U}\,, \qquad \mu(A) = 0 \text{ if } A \notin \mathfrak{U}\,.$$

To illustrate this let us compute $\mathcal{M}(A)$ and $\mathcal{N}(A)$ for all sets A. Clearly, $\mathcal{M}$ and $\mathcal{N}$ are convex classes of measures. $\mathcal{M}(\emptyset) = 0$, $\mathcal{M}(N) = 1$, and for every other set we can find ultrafilters $\mathfrak{U}_1$ containing A and $\mathfrak{U}_2$ containing $N-A$. Defining μ_1 and μ_2 by (3) gives $\mu_1(A) = 1$, $\mu_2(A) = 0$. Hence $\mathcal{M}(A) = [0,1]$ unless $A = \emptyset$ or $A = N$. If A is finite, $\mathcal{N}(A) = 0$, and if $N-A$ is finite, $\mathcal{N}(A) = 1$. For other sets A, we can construct non-principal ultrafilters to prove $\mathcal{N}(A) = [0,1]$.

Let $\mathfrak{U}$ be an ultrafilter and let μ be the measure defined by (3). If f is any bounded function, we can form $\int f\,d\mu$. This value is easily seen to be characterized by

$$\text{for all } \varepsilon > 0\colon \Big\{x\colon \Big|f(x) - \int f\,d\mu\Big| < \varepsilon\Big\} \in \mathfrak{U}$$

and so qualifies to be called $\mathfrak{U}-\lim f$.

We shall now apply the notion of $\mathfrak{U}$-limit to the problem of constructing measures with given properties. For these constructions we shall assume that we have a fixed non-principal ultrafilter $\mathfrak{U}$.

Let us look once more at the notion 'pick a number at random'. Notice that for $\mu \in \mathcal{N}$, $\mu(N+1) = \mu(N)$. Another reasonable condition for this notion is that the preceding property hold good for every $A \subseteq N$. Thus we say that a measure μ is *translation invariant* if $\mu(A+1) = \mu(A)$ for every $A \subseteq N$. Let $\mathcal{T}$ be the class of all translation invariant measures. Clearly, $\mathcal{T} \subseteq \mathcal{N}$.

It is sometimes useful to use the natural numbers to 'code' any countable set. The given countable set may admit certain mappings which we would expect to be measure preserving (e.g., translations of the integer lattice in some Euclidean space). These mappings are encoded into mappings defined from N to N.

To accommodate this generalization, let $g\colon N\to N$ be any function and let $g^{-1}(A)$ be the complete inverse image of A. A measure μ is called *g-invariant* if $\mu\big(g^{-1}(A)\big)=\mu(A)$ for every $A\subseteq N$. Let $\mathfrak{I}_g$ be the class of all g-invariant measures. If g has no finite orbits, then $\mathfrak{I}_g\subseteq\mathcal{N}$. Subsumed under this definition are $\mathcal{M}$ (take $g(n)=n$) and $\mathcal{T}$ (take $g(n)=n+1$). $\mathfrak{I}_g$ is readily seen to be a closed convex (but possibly empty) class of measures. Let us take a closer look.

Let $g\colon N\to N$ be any function fixed throughout the following discussion and let $f\colon N\to R$ be a bounded function. In order to compute bounds for $\mathfrak{I}_g(f)$, let us define $f^{(m)}$ by

$$f^{(m)}(n)=(1/m)\sum_{k=0}^{m-1}f\big(g^k(n)\big) \tag{4}$$

where g^k is the k-fold iterate of g. Then

THEOREM 1. *$\mathfrak{I}_g$ is non-empty and*

$$\mathfrak{I}_g(f)=[\liminf_m \mathrm{glb}_n f^{(m)}(n),\ \limsup_m \mathrm{lub}_n f^{(m)}(n)]\,.$$

Also

$$\limsup_m \mathrm{lub}_n f^{(m)}(n)=\mathrm{glb}_m \mathrm{lub}_n f^{(m)}(n)\,,$$

and if $\mathfrak{I}_g\subseteq\mathcal{N}$, this is equal to $\mathrm{glb}_m\limsup_n f^{(m)}(n)$. (*And similarly for the lower bound.*)

Proof. We first demonstrate the existence of a measure μ for which

$$\int f\,d\mu=\limsup_m \mathrm{lub}_n f^{(m)}(n)\,.$$

Let $\mu_n^{(m)}$ be the measure for which $\mu_n^{(m)}(A)=k/m$ where k is the number of terms of the sequence $S_n^{(m)}=\langle n, g(n), \ldots, g^{m-1}(n)\rangle$ which are elements of A. Clearly $\mu_n^{(m)}\in\mathcal{M}$ and $f^{(m)}(n)=\int f\,d\mu_n^{(m)}$. We can now choose sequences $m, n\colon N\to N$ such that

$$\lim_s f^{(m(s))}\big(n(s)\big)=\limsup_m \mathrm{lub}_n f^{(m)}(n)$$

while $\lim_s m(s)=\infty$. Using our fixed non-principal ultrafilter $\mathfrak{U}$ we define a measure μ by

$$\mu(A)=\mathfrak{U}\text{-}\lim_s \mu_{n(s)}^{(m(s))}(A)\quad\text{for all } A\subseteq N\,.$$

Since

$$\big|\mu_n^{(m)}\big(g^{-1}(A)\big)-\mu_n^{(m)}(A)\big|\leqslant 1/m\quad\text{and}\quad \lim_s 1/m(s)=0,$$

this gives $\mu\in\mathfrak{I}_g$. The same construction also gives the lower bound so we have shown that $\mathfrak{I}_g(f)$ is at least as big as we claim. To complete the proof it is only necessary to show that

$$\int f\,d\mu\leqslant \mathrm{glb}_m \mathrm{lub}_n f^{(m)}(n)\quad\text{for}\quad \mu\in\mathfrak{I}_g$$

and that lub_n can be replaced by $\limsup_n$ if $\mu\in\mathfrak{I}_g\cap\mathcal{N}$. These quantities are clearly no larger than the attainable upper bound so all conclusions of the theorem follow from these facts. If $\mu\in\mathfrak{I}_g$,

$$\int f\,d\mu=\int f(g)\,d\mu\,.$$

Consequently

$$\int f\,d\mu=\int f^{(m)}d\mu\leqslant \mathrm{lub}_n f^{(m)}(n)\,.$$

The latter may be replaced by $\limsup_n f^{(m)}(n)$ if $\mu\in\mathcal{N}$. Since this holds for all m, the result follows (cf. [3], § 5). q.e.d.

We are particularly interested in the case $g(n)=n+1$. The class $\mathfrak{I}_g$ now becomes the class $\mathcal{T}$ of translation invariant measures. We shall now apply Theorem 1 to various examples to see how well $\mathcal{T}$-measures are in accord with our intuition of randomness.

EXAMPLES. Let $A\subseteq N$ be a set such that for each $n\in N$, A has exactly one element in common with $\{2n, 2n+1\}$. Then $\mu(A)=1/2$ for all $\mu\in\mathcal{T}$. Consequently there are sets, which many logicians would say are non-constructible, that are $\mathcal{T}$-measurable. Likewise, Theorem 1 applies directly to the generalized arithmetic progression $B=\{[\alpha n+\beta]\colon n\in N\}$ for real $\alpha\geqslant 1$, β giving $\mu(B)=1/\alpha$ for every $\mu\in\mathcal{T}$. The proviso $\alpha\geqslant 1$ is necessary to insure that we do not get the same $[\alpha n+\beta]$ from two different n's. In general translation invariance is much more restrictive than simply giving these intuitively correct values on the arithmetic progressions. We can see this by using a modification of the usual Lebesgue technique. Start out with the Boolean algebra of all finite unions of arithmetic progressions endowed with the $1/\alpha$ measure and define inner-outer measures and measurable in the usual way. Let μ be the restriction of outer measure to the measurable sets. Let the set H consist of those even numbers m for which there is an integer n satisfying $4n^2\leqslant m<4n^2+4n+4$ and those odd numbers m for which there is an integer n satisfying $4n^2+4n+4<m<4(n+1)^2$. Since it has the form of our first example, $\mathcal{T}(H)=\{1/2\}$. On the other hand, every generalized arithmetic progression is easily seen to meet both H and its complement. Thus it is not Lebesgue measurable for it has outer measure 1 and inner measure 0.

Let $A\subseteq N$ be a set such that there are intervals $[n, n+m)_N\subseteq A$ for arbitrarily large m. Then by Theorem 1, $\mathcal{T}(A)=[\alpha, 1]$. If $N-A$ has this property, then $\mathcal{T}(A)=[0,\beta]$, and if both do, then $\mathcal{T}(A)=[0,1]$. We can apply this result to the set P which consists of all natural numbers whose decimal expansion begins with a 1, i.e., $P = \bigcup_{k=0}^{\infty}[10^k, 2\cdot 10^k)_N$, to obtain $\mathcal{C}(P) = [0, 1]$. Thus we see that translation invariant measures are inadequate to settle the problem of the distribution of first significant digits. That will be the central problem of the next section.

An easy application of Theorem 1 shows that both the primes and the squares are $\mathcal{C}$-measurable with measure 0. A more interesting case is that of the square free integers S for which $\mathcal{C}(S) = [0, 6/\pi^2]$. We prove this result by a slight extension of the usual discussion of S (cf. [2], ch. 18). The sets $Q_p = \{n\colon n \not\equiv 0 \pmod{p^2}\}$ for various primes p are unions of arithmetic progressions. Thus $\mu(Q_p) = 1-(1/p^2)$ for $\mu \in \mathcal{C}$. Furthermore, the Chinese Remainder theorem tells us that $\mu(Q_{p_1} \cap \ldots \cap Q_{p_k}) = (1-(1/p_1^2))\cdot\ldots\cdot(1-(1/p_k^2))$. Hence $\mu(S) \leqslant \prod_p(1-(1/p^2)) = 6/\pi^2$. It is easy to find intervals of any length in $N-S$, so $\mu(S) = 0$ is possible. In the standard discussion it is shown that $6/\pi^2$ is the asymptotic density of S, i.e., $= \lim_m \mu_1^{(m)}(S)$. Thus $\mathcal{C}(S)$ must be $[0, 6/\pi^2]$.

3. Scale invariance. The main purpose of this section is a clarification of the first digit problem. Although all of the necessary apparatus for such a discussion has been made available in the preceding section, it turns out that a less restrictive notion of measure enhances our presentation. A set $A \subseteq R^+$ is called *discrete* if for each $n \in N$ the set $A \cap [n, n+1)$ contains only finitely many elements. In this case we can form the function $\Psi_A(n) = |A \cap [n, n+1)|$. We say that a discrete set A is *sparse* if Ψ_A is a bounded function. Let $\mathfrak{S}$ be the class of all sparse sets. Every measure $\mu \in \mathcal{M}$ induces a set function $\varphi\colon \mathfrak{S} \to R^+$ which is given by the integral

$$\varphi(A) = \int \Psi_A \, d\mu. \tag{5}$$

It is clear from the properties of the integral that φ is a finitely additive set function and that $\varphi(A) = \mu(A)$ for $A \subseteq N$. Consequently we can think of φ as an extension of μ to a more comprehensive class of sets. We are particularly interested in this extension when $\mu \in \mathcal{C}$. In this case not only is φ translation invariant, but it has a stronger property (cf. lemma 1) which implies translation invariance. For $A, B \subseteq R^+$ we say that B is a *distortion* of A if there exists a function $f\colon A \sim B$ such that $|f(x)-x|$ is a bounded function. It is clear that this notion is an equivalence relation and that if $A \in \mathfrak{S}$, then so does B. We say that $\varphi\colon \mathfrak{S} \to R^+$ is *distortion invariant* if $\varphi(A) = \varphi(B)$ for every distortion B of A. Then we have

LEMMA 1. *Every φ given by* (5) *for $\mu \in \mathcal{C}$ is distortion invariant.*

Proof. Let $A, B \in \mathfrak{S}$ and $f\colon A \sim B$ with $|f(x)-x|$ a bounded function. Since $g(x) = [f(x)]-[x]$ is also bounded, it assumes only finitely many values, say, $g_1, \ldots, g_k$. Partition A into sets $A_i = \{x \in A\colon g(x) = g_i\}$ for $1 \leqslant i \leqslant k$. If $B_i = f(A_i)$, then the B_i's form a partition of B such that $\Psi_{B_i}(n) = \Psi_{A_i}(n-g_i)$ for $1 \leqslant i \leqslant k$. Since $\mu \in \mathcal{C}$, it follows from (5) that $\varphi(A_i) = \varphi(B_i)$ and then by finite additivity that $\varphi(A) = \varphi(B)$. q.e.d.

Conversely we have

LEMMA 2. *If $\varphi\colon \mathfrak{S} \to R^+$ is a finitely additive distortion invariant set function normalized by $\varphi(N) = 1$, then for some $\mu \in \mathcal{C}$ we have $\varphi(A) = \int \Psi_A d\mu$ for every $A \in \mathfrak{S}$.*

Proof. If μ is the restriction of φ to $\mathfrak{P}(N)$, then clearly $\mu \in \mathcal{C}$. For any set $A \in \mathfrak{S}$, Ψ_A is bounded and thus can assume only finitely many values, say $a_1, \ldots, a_k$. Partition N into sets $S_i = \{n \in N\colon \Psi_A(n) = a_i\}$ and A into sets $A_i = \{x \in A\colon x \in [n, n+1)$ and $n \in S_i\}$ for $1 \leqslant i \leqslant k$. Then by the distortion invariance of φ, $\varphi(A_i) = a_i\varphi(S_i) = a_i\mu(S_i)$ and consequently

$$\varphi(A) = \sum_{i=1}^{k} a_i\mu(S_i) = \int \Psi_A d\mu. \qquad \text{q.e.d.}$$

The combined content of lemmas 1 and 2 gives us the natural extension of measures $\mu \in \mathcal{C}$ to all of $\mathfrak{S}$ and justifies the following definition. A *distortion invariant measure* is a finitely additive, distortion invariant set function $\mu\colon \mathfrak{S} \to R^+$, normalized by $\mu(N) = 1$. We identify these measures with the measures from which they arise by (5) and let $\mathcal{C}$ be the class of all distortion invariant measures.

A subset $S \subseteq R^+$ is called *measure inducing* if $S-[0, 1)$ can be expressed as a disjoint union of intervals $S_1, S_2, \ldots$ such that for every $A \subseteq S$ consisting of exactly one point from each S_i, $A \in \mathfrak{S}$, and for every measure $\mu \in \mathcal{C}$, $\mu(A) = 0$. A necessary and sufficient condition that S be measure inducing (cf. th. 1) is that for each interval $I \subseteq R^+$ of length t, I has nonempty intersection with at most $f(t)$ of the S_i's where f is some function such that $\lim_t f(t)/t = 0$. It is sometimes convenient to describe subsets of N as $S \cap N$ where S is a subset of R^+. When S is measure inducing, this description has the following property.

THEOREM 2. *If S is measure inducing and $\mu \in \mathcal{C}$, then for every $a > 0$ we have $\mu(S \cap N) = a\mu(S \cap aN)$.*

Proof. We give separate proofs for (6) and (7) below from which our theorem follows. For S, μ, and a as in our hypothesis

$$\mu(S \cap aN) = n\mu(S \cap naN) \quad \text{for} \quad n \in N. \tag{6}$$

Consider any one of the intervals S_i of which S is composed. The number of elements in $S_i \cap a(nN)$ and $S_i \cap a(nN+k)$, where $0 \leqslant k < n$, differ by at most one and consequently for $f(x) = x+k$, $f(x) \in S_i \cap a(nN+k)$ except for at most one element of $S_i \cap a(nN)$ and $f^{-1}(x) \in S_i \cap a(nN)$ except for at most one element of $S_i \cap a(nN+k)$. Combining these observations for all the S_i and using the measure inducing property of S, we can express $S \cap a(nN) = T_1 \cup T_2$ and $S \cap a(nN+k) = T_3 \cup T_4$ as disjoint unions where $f\colon T_1 \sim T_3$ is a distortion and consequently $\mu(T_1) = \mu(T_3)$, where $\mu(T_2) = \mu(T_4) = 0$. Thus $S \cap a(nN)$ and $S \cap a(nN+k)$ have the same measure. Since $S \cap aN$ is a disjoint union of the $S \cap a(nN+k)$'s for $0 \leqslant k < n$, (6) follows immediately.

Now suppose that $a = p/q$ is a rational number. Then by two applications of (6),

$$a\mu(S \cap aN) = (p/q)\mu\big(S \cap (p/q)N\big) =$$
$$= (1/q)\mu\big(S \cap (1/q)N\big) = (q/q)\mu\big(S \cap (q/q)N\big) = \mu(S \cap N)\,.$$

Thus the statement of our theorem holds for rational a. Finally for $a, \beta > 0$ we show

$$\text{(7)}\qquad a < \beta \quad \textit{implies} \quad \mu(S \cap \beta N) < \mu(S \cap aN)\,.$$

Consider any one of the intervals S_i of which S is composed. Between any consecutive members of βN we can find at least one member of aN. Consequently the function $f(x) = a[x/a]$ assumes values in aN, is one-one when restricted to $S \cap \beta N$, and is a distortion since $-a < f(x) - x \leqslant 0$. Finally we note that $f(x) \in S_i \cap aN$ except for at most one element of $S_i \cap \beta N$. Combining these observations for all S_i and using the measure inducing property of S, we see that we can express $S \cap \beta N = T_1 \cup T_2$ as a disjoint union such that T_1 is a distortion of a subset of $S \cap aN$ and $\mu(T_2) = 0$. This proves (7). Now the functions $(1/a)\mu(S \cap N)$ and $\mu(S \cap aN)$ are both monotone decreasing functions which assume the same value for rational a. Since the former is continuous, they must have the same value for every a, i.e., $\mu(S \cap N) = a\mu(S \cap aN)$. q.e.d.

Taking $S = R^+$ (which is clearly measure inducing), we see that $a\mu(aN) = \mu(N) = 1$, a result which generalizes the example $\{[an]\colon n \in N\}$, $a \geqslant 1$ of the last section. Notice that if $\mu \in \mathcal{C}$, then for every $a > 0$ the set function μ' defined by $\mu'(A) = a\mu(aA)$ is also a member of $\mathcal{C}$. $\mu'(N) = \mu(N) = 1$ which means that if we take a uniformly distributed set such as N, and thin it out by multiplying by $a > 0$, we get $\mu(aN) = (1/a)\mu(N)$. Now we believe that another reasonable condition for the notion 'pick a number at random' is that the preceding thinning property holds for every set $A \in \mathfrak{S}$. This prompts us to define a new class of measures.

A measure $\mu \in \mathcal{C}$ is called *scale invariant with respect to* $a > 0$ if $\mu(A) = a\mu(aA)$ for every $A \in \mathfrak{S}$ (i.e., if μ treats all sets as it must treat N). Let $\mathfrak{S}_a$ be the class of all measures, scale invariant with respect to a. We say that μ is *scale invariant* if it is scale invariant with respect to a for every $a > 0$ and let $\mathcal{S}$ be the class of all scale invariant measures. Existence is guaranteed by the Markov-Kakutani fixed point theorem (cf. [1], p. 456). It is also useful to have an explicit construction for such measures.

If $\mu \in \mathcal{C}$, let $\mu^{(n)}$ be defined by

$$\mu^{(n)}(A) = (1/n)\sum_{k=0}^{n-1} a^k\mu(a^kA)\,,$$

and

$$\nu(A) = \mathfrak{U}-\lim_n \mu^{(n)}(A)$$

for our fixed non-principal ultrafilter $\mathfrak{U}$. ν is then seen to belong to $\mathcal{S}_a$. If $\mu \in \mathcal{S}_\beta$, then $\nu \in \mathcal{S}_\beta \cap \mathcal{S}_a$. This causes ν to be scale invariant with respect to all $a^m\beta^n$ for integers m, n, so a suitable choice of a, β gives measures which are scale invariant for a dense set of positive real numbers. If μ satisfies

$$\text{(8)}\qquad \mu(\delta A) \leqslant \mu(A) \quad \text{for} \quad \delta \geqslant 1\,,$$

then so will the measures constructed from it in the manner just described. A measure which satisfies (8) and is scale invariant for a dense set of positive real numbers is scale invariant for all real numbers.

In number theory, subsets of N are often measured by asymptotic density. If $\mathcal{A}$ denotes the class of measures μ for which

$$\int f\,d\mu \leqslant \mathrm{lub}_n (1/n)\sum_{k=1}^{n} f(k)$$

for all bounded functions f, then these studies can be considered as being an investigation of the classes $\mathcal{A}$ and $\mathcal{A} \cap \mathcal{N} = \mathcal{A} \cap \mathcal{C}$. All measures in $\mathcal{A}$ satisfy (8), and we would like to conjecture the converse. At any rate, the above shows that $\mathcal{A} \cap \mathcal{S} \neq \emptyset$.

We conclude our paper with a discussion of the first digit problem (cf. (1)) for which the following notation will be helpful. Let C be the unit circle in the complex plane and λ the Lebesgue measure on C normalized so that $\lambda(C) = 1$. Use arcs to measure angles centered at the origin so that the complete angle about a point is 1. Measure angles in the counterclockwise sense starting from the x axis. The mapping $e\colon R \to C$ defined by $e(\theta) = \exp(2\pi i\theta)$ maps every interval of length 1 in R onto C in a measure preserving fashion. Now we have the following

THEOREM 3. *Let U be an arc in C, $\beta > 0$ an arbitrary real number and $S = \{x \in R^+\colon e(\log_\beta x) \in U\}$. If $\theta \in R^+$ where $\log_\beta(\theta)$ is irrational and $\mu \in \mathcal{S}_\theta$, then $\mu(S \cap N) = \lambda(U)$.*

Proof. First note that θS is a set which can be expressed in the same way that S is except that its corresponding arc U' is obtained by rotating U through an angle of $2\pi\log_\beta(\theta)$. Let $U^{(k)}$ be obtained from U by rotation through an angle $2k\pi\log_\beta(\theta)$ and observe that this converts S into θ^kS. Since $\log_\beta(\theta)$ is irrational, the multiples $2k\pi\log_\beta(\theta)$ are dense on C (cf. [2], ch. 23) and consequently we can find arcs $U^{(k)}$ as close as we please to any desired position. S, as well as its multiples $\theta^k S$, is measure inducing since $\lim_t(\log_\beta(t)/t) = 0$. Hence $\mu(\theta^k S \cap N) = \theta^k\mu(\theta^k S \cap \theta^k N) = \theta^k\mu(\theta^k(S \cap N)) = \mu(S \cap N)$, the final equality following because $\mu \in S_0$. Let $P^{(k)} = \theta^k S \cap N$, a set of integers which corresponds to the arc $U^{(k)}$, all $P^{(k)}$ having the same measure with respect to μ. Now consider $p, q \in N$ such that $(p/q) < \lambda(U)$. By a simple geometric argument this implies that we can find q arcs $U^{(k)}$, $k \in A$, such that every point of C belongs to at least p of them. Correspondingly every $m \in N$ belongs to at least p of the q sets $P^{(k)}$, $k \in A$. For $n \in N$ let B_n be the set of $m \in N$ which belong to exactly n of the $P^{(k)}$, $k \in A$. Then we have

$$q\mu(P^{(0)}) = \sum_{k \in A} \mu(P^{(k)}) = \sum_{n \geqslant p} n\mu(B_n) \geqslant p .$$

Thus $(p/q) \leqslant \mu(P^{(0)})$. Proceeding in exactly the same way we can show that if $\lambda(U) < (p/q)$, then $\mu(P^{(0)}) \leqslant (p/q)$. If we combine these results, we see that $\mu(S \cap N) = \mu(P^{(0)}) = \lambda(U)$. q.e.d.

Corollary. *If P_n is the set of natural numbers whose first significant digit lies between* 1 *and* n, $1 \leqslant n \leqslant 9$, *and* $\mu \in S$ (*in fact to any* S_0 *where* $\log_{10}\theta$ *is irrational*), *then* $\mu(P_n) = \log_{10}(n+1)$.

Proof. For we can describe $P_n = S_n \cap N$ where S_n is the set of all $x \in R^+$ such that $0 \leqslant e(\log_{10} x) \leqslant \log_{10}(n+1)$. q.e.d.

References

[1] N. Dunford and J. Schwartz, *Linear Operators*. I, Interscience, New York, 1958.

[2] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, 1938.

[3] M. Jerison, *The set of all generalized limits of bounded sequences*, Canad. Journ. Math. 9 (1957), pp. 79-89.

[4] R. S. Pinkham, *On the distribution of first significant digits*, Ann. of Math. Stat. 32 (1961), pp. 1223-1230.

[5] A. Tarski, *Une contribution à la théorie de la mesure*, Fund. Math. 15 (1930), pp. 42-50.

RUTGERS, THE STATE UNIVERSITY
New Brunswick, New Jersey

Reçu par la Rédaction le 28. 8. 1967

Metrizability of trees *

by

Carl Eberhart (Lexington, Ky.)

Introduction. It is a well-known result that dendrites (acyclic Peano continua) can be alternatively defined as metrizable continua in which each pair of points can be separated by a third point. L. E. Ward, in [6], generalized the notion of dendrite by removing the metrizability condition in the second definition, and called such objects trees. He then showed that many properties of dendrites carry over to trees. In this paper we shall be concerned with establishing properties of trees which yield metrizability theorems. The principal results in this connection are I.6, III.1, III.2, and III.5.

I. Separable trees are metrizable. By a *continuum* we mean a compact connected Hausdorff space. A continuum is *hereditarily unicoherent* provided the intersection of any two of its subcontinua is connected. A *tree* is a locally connected hereditarily unicoherent continuum. An *arc* is a continuum with precisely two non-cutpoints.

In Whyburn [7], pp. 88-89, several properties of metric trees (= dendrites) are established. L. E. Ward showed in [6] that a number of these properties carry over to the nonmetric case.

For the rest of this section X will denote a tree. Proposition I.1 is due to Ward.

I.1. Proposition. *For each x and y in X, $[x, y] = \bigcap \{C \mid x, y \in C$ and C is a subcontinuum of $X\}$ is an arc with endpoints x and y.*

Proof. It follows from the hereditary unicoherence of X that $[x, y]$ is the only subcontinuum of X irreducible between x and y. Suppose $z \in (x, y) = [x, y] \setminus \{x, y\}$. If $[x, y] \setminus z$ were connected, then x and y would lie in the same component of $X \setminus z$. But this is impossible since the components of open sets in locally connected continua are continuum-wise connected ([1], p. 110). Hence $[x, y]$ is an arc.

I.2. Proposition. *If C is a component of $X \setminus p$, then $[x, p) = [x, p] \setminus p \subset C$ for each x in C.*

* This research was supported by a Faculty Fellowship from the University of Kentucky.