JOURNAL OF COMPUTER AND SYSTEM SCIENCES **49**, 605–619 (1994)

On Hausdorff and Topological Dimensions of the Kolmogorov Complexity of the Real Line*

JIN-YI CAI[†]

Department of Computer Science, State University of New York at Buffalo, Buffalo, New York 14260

AND

JURIS HARTMANIS[‡]

Department of Computer Science, Cornell University, Ithaca, New York 14853

Received July 21, 1991; revised July 21, 1992

We investigate the Kolmogorov complexity of real numbers. Let K be the Kolmogorov complexity function; we determine the Hausdorff dimension and the topological dimension of the graph of K. Since these dimensions are different, the graph of the Kolmogorov complexity function of the real line forms a fractal in the sense of Mandelbrot. We also solve an open problem of Razborov using our exact bound on the topological dimension.

1. INTRODUCTION

We investigate the Kolmogorov complexity of real numbers. We show that, from a computational point of view, the real line is not a set of points without individual distinguishing characteristics, but rather, the real numbers and their complexity form a very complex object.

We consider a well-defined function that assigns to each real number its *Kolmogorov complexity* and determine the Hausdorff dimension and the topological dimension of the graph of the Kolmogorov complexity of the real line. In particular, the Hausdorff dimension is strictly greater than the topological dimension. Thus the graph of the Kolmogorov complexity of the real line forms a *fractal* in the sense of Mandelbrot [M].

* A preliminary version of the paper titled "The Real Line is a Fractal" has appeared in "The Structure in Complexity Theory Conference 1989."

[†] Research supported by NSF Grant CGR-8709818, while the author was at Yale Unversity.

[‡] Research supported by NSF Grants CCR-8520597 and CCR-8823053.

0022-0000/94 $6.00

In Section 2, we give some motivations for our results and state all the necessary definitions. In Section 3, we deal with the Hausdorff dimension of the Kolmogorov complexity function of the real line. We prove matching upper and lower bounds for the "fibre set" of points with Kolmogorov complexity equal to a. It has Hausdorff dimension exactly a. We then determine the Hausdorff dimension of the graph of K to be exactly 2.

In Section 4, we deal with the topological dimension of the same graph. We show that the topological dimension is exactly 1. Our determination of the topological dimension of the graph of the Kolmogorov complexity of the reals solves an interesting open problem by Razborov [Ra] who asked what relationship can there be between $K(x)$ and $K(K(x))$? He asked, for instance, is it true that there exist easily computable functions f from [0, 1] to [0, 1], such that, $K(K(x)) \leqslant f(K(x))$, for all $0 \leqslant x \leqslant 1$? In particular, is it true that $K(K(x)) \leqslant K(x)$ for all $0 \leqslant x \leqslant 1$?

We show that the answer to Razborov's question is negative, and it follows easily from our exact bound on the topological dimension. In fact we give a stronger statement in Theorem 4.4.

Our work grew out of an attempt to formulate a theory of computational complexity over the reals. In our approach, for a wide class of functions such as any non-constant polynomial with rational coefficients, or, analytic functions with uniformly computable coefficients, the complexity of a real number is invariant under the transformation. Thus, the nature of the complexity graph is preserved by finite iterations of such computable functions.

Previous work on recursive real analysis has adopted a different approach, where the objects one deals with are necessarily countable. This has the advantage of being more true to recursion theory, but much of "continuous mathematics" is rendered inapplicable. Recently Blum, Shub, and Smale considered a complexity theory of real numbers [BSS], where the emphasis is on the algebraic operational costs. They make the assumption that every real number has unit complexity. On the other hand, their approach brings more classical mathematics into the picture.

We believe an approach similar to the finitary Kolmogorov complexity should be pursued. Such an approach has both the advantage of being more realistic in differentiating the complexity between easily computable numbers, on the one hand, and intractable ones, on the other, and much of classical mathematics, especially non-trivial analysis and topology, are inherently applicable.

The search for appropriate complexity theories for real number computations is a challenging task. We do not claim to have found the ultimate approach; it is doubtful whether such a single approach exists that is "right" for all. However, we hope that this paper may help to stimulate a systematic investigation to the nature of the underlying computational domains in real computations.

Some of our results in Section 3 were independently obtained by Staiger and others. As one referee pointed out, although motivated by different issues, Besicovitch had pursued a highly related notion that can in fact be used to give separate proofs of some of our results in Section 3. We will discuss this in more detail at the end of that section. For related work please see [L2, Ry, St1, St2, Be].

2. Definitions and Preliminaries

What is randomness? And what is a random object? Surely a large object with any easily distinguishable patterns, or one which can be generated by any well specified short procedure, should not be considered random. The *Kolmogorov complexity* $K(x)$ of a binary string x is defined to be the information content of x, i.e., the size in bits of the smallest input string—program—which will cause a fixed universal Turing machine to produce x. (The choice of the fixed universal Turing machine introduces at most an additive constant in the value of $K(x)$, which asymptotically can be ignored. We shall fix one universal machine once and for all.) The notion of Kolmogorov complexity was due to Solomonoff, Kolmogorov, and Chaitin [S, K, C]. There have been quite a few variations of the original notion of Kolmogorov complexity, most notably by Chaitin and Levin on self-delimiting Kolmogorov complexity [C, L], and the resource-bounded versions, such as polynomial time/space bounded Kolmogorov complexity [Ba, *H*]. However, the result of this paper is *robust*, in the sense that any and all such definitions lead to the same conclusion—the computational line is a fractal. For definiteness, we adopt the classical definition through out this paper.

To define the complexity of a real number x, we consider any reasonable representation of x, such as its binary expansion. We take the n-bit prefix x_n of x and consider its normalized Kolmogorov complexity $K(x_n)/n$. It should be clear that the choice of which particular enumeration scheme to represent x (for instance, ternary, decimal, or continued fraction) is of no significance, asymptotically speaking, as long as the conversion between them is computable. (If we are using polynomial time bounded Kolmogorov complexity, then we should require polynomial time conversion algorithms, which certainly exist for those we mentioned.) Now we define the complexity of x as

$$K(x) = \lim_{n \to \infty} K(x_n)/n.$$

We denote the graph of the function K by Γ_K.

A technical note. When the limit does not exist, we may take any reasonable value, such as the arithmetic mean of upper and lower limit[1]:

$$K(x) = \tfrac{1}{2}(\liminf_{n \to \infty} K(x_n)/n + \limsup_{n \to \infty} K(x_n)/n).$$

It follows from the definition that Γ_K has perfect scaling properties:

$$K(rx+s) = K(x), \qquad \forall x \in \mathbf{R},\ r, s \in \mathbf{Q},\ r \neq 0.$$

In fact this scaling property can be significantly strengthened to arbitrary polynomials (or even analytic functions) with (uniformly) computable coefficients. To see

[1] Such arbitrariness is perhaps disquieting; however, the reassuring fact is that it leads to the same theory no matter how one extends the definition, as we will see.

this, we first note that the zero set of any such function f (and therefore that of its derivative f') is discrete in the domain of its definition and consists of computable numbers (in the sense of Turing). Thus, modulo a discrete set of points, where $K(x) = 0 = K(f(x))$, the function f is locally monotonic with a non-zero derivative. This enables us to prove $K(x) = K(f(x))$, for all x. As a consequence of this scaling property, we will only consider the function K as defined on the unit interval $I = [0, 1]$.

We will investigate the Hausdorff and topological dimensions of the graph of K. A general reference on dimension theory can be found in [HW].

DEFINITION 2.1. Given a set S in a metric space X, and any real number $p \geqslant 0$, let $\varepsilon > 0$ and

$$m_p^\varepsilon(S) = \inf \sum_{i \geqslant 1} \delta(S_i)^p,$$

where $S = \bigcup_{i=1}^{\infty} S_i$ is any decomposition of S in a countable number of subsets of diameter $\delta(S_i)$ less than ε and the superscript p denotes exponentiation. Let

$$m_p(S) = \lim_{\varepsilon \to 0} m_p^\varepsilon(S);$$

$m_p(S)$ is called the p-dimensional (Hausdorff) measure of S.

We observe that the limit in the definition exists (including infinity ∞), since $m_p^\varepsilon(S)$ is monotonic non-decreasing as $\varepsilon \to 0$. We also note that $p < q$ and $m_p(S) < \infty$ imply that $m_q(S) = 0$.

DEFINITION 2.2. Given a set S in a metric space X, the Hausdorff dimension of S, $\dim_H(S)$, is the supremum of all real numbers p such that $m_p(S) > 0$.

Clearly the above definition of the Hausdorff dimension of S can be equivalently stated in terms of the limit

$$\liminf_{\varepsilon \to 0} \sum_{i \geqslant 1} \delta(D_i)^p,$$

where the infimum takes over all countable coverings of S by open (or closed) discs $\mathcal{O} = \{D_i \mid i \geqslant 1\}$, with $\sup_{i \geqslant 1} \delta(D_i) \leqslant \varepsilon$. In what follows, we will use the notion of a covering to compute the Hausdorff dimension.

As an example, it is well known that the (classical) Cantor set $\mathscr{C}$ has Hausdorff dimension log 2/log 3. This can be seen intuitively by the following family of *finite* coverings for $\mathscr{C}$ inductively defined. $\mathcal{O}_1$ consists of a single interval $[0, 1]$; $\mathcal{O}_k$ consists of all the intervals that are the first or the last third of any interval in $\mathcal{O}_{k-1}$. (Although a rigorous proof of equality can be given along this line, the existence of such a cover only shows that $\dim_H(\mathscr{C}) \leqslant \log 2/\log 3$. Note also that in general a *countable* cover is used instead of a finite one.)

We now define the notion of the topological dimension of a space X. It turns out that there are three commonly used concepts of dimension in the literature. Although for more general spaces they do not necessarily agree, they do agree on all separable metric spaces (spaces with a countable dense subset.) Since this is the case for our investigation (subspaces of Euclidean space) we will give just one definition of the topological dimension, also known as the Urysohn–Menger (small inductive) dimension.

DEFINITION 2.3. Given a metric space X,

1. $\dim_T(X) = -1$, if $X = \varnothing$;

2. $\dim_T(X) \leqslant n$, if for every $p \in X$ and open set U containing p there is an open set V satisfying

$$p \in V \subset U \qquad \text{and} \qquad \dim_T(\partial V) \leqslant n-1.$$

3. $\dim_T(X) = n$, if $\dim_T(X) \leqslant n$ and $\dim_T(X) \nleqslant n-1$.

4. $\dim_T(X) = \infty$, if $\dim_T(X) \nleqslant n$ for all n.

We note that when X is a subspace, say of an Euclidean space, the topology on X is the induced topology. If X is everywhere dense, then the boundary of an open set in X, $\partial_X(O \cap X)$, equals $\partial O \cap X$.

As an example, any non-empty finite or countable space is zero-dimensional. Any subset of the real line that does not contain any interval also has topological dimension zero. And as a consequence of the Brouwer fix-point theorem, the Euclidean n-space has topological dimension n [Br]. (The non-trivial part is to show that $\dim_T(\mathbf{R}^n) \nleqslant n-1$).

Note that the topological dimension of a space is always an integer (if it is finite). It is known that the Hausdorff dimension is always greater than or equal to the topological dimension. Mandelbrot defined a space to be a fractal if they do not agree; i.e., X is called a fractal if $\dim_H(X) > \dim_T(X)$. For our set Γ_K, the graph of K, we will establish just that: $\dim_H(\Gamma_K) = 2$ and $\dim_T(\Gamma_K) = 1$.

3. THE HAUSDORFF DIMENSION OF Γ_K

The main theorem in this section is the following.

THEOREM 3.1. *For any numbers $0 \leqslant a < b \leqslant 1$, the Hausdorff dimension of the set*

$$\Gamma_K \cap ([0, 1] \times [a, b])$$

is $1 + b$.

An immediate corollary is

COROLLARY 3.2. *The Hausdorff dimension of the graph* $\dim_H(\Gamma_K) = 2$.

We first investigate the "fibre sets" $F_a = \{x \in [0, 1] \mid K(x) = a\}$, for $0 \leqslant a \leqslant 1$. We will show that $\dim_H(F_a) = a$, from which the main theorem will follow.

LEMMA 3.3. *Almost all points in* $[0, 1]$ *have complexity* 1; *i.e.,* F_1 *has full Lebesgue measure.*

The proof is a simple counting argument, which we shall omit here. The next lemma is in fact implied by the more general Lemma 3.6 (where the proof is independent of this). But we include a separate proof sketch here since it introduces in the simple setting a method which will be generalized later in the proof of Theorem 4.1. It also has a corollary that will be used in Section 4.

LEMMA 3.4. *For all a,* $0 \leqslant a \leqslant 1$, F_a *is an uncountable infinite set.*

Proof. For notational simplicity we assume that $0 < a < 1$. (The case $a = 1$ is implied by Lemma 3.3.) To exhibit a real number x with complexity a, we first take a random string as the initial segment of x, so long that the normalized complexity is "pushed" above a. Then we append any "simple" string such as all zeros, so long that the normalized complexity is "pushed" below a. Now we repeat the process, with ever smaller oscillation. The number x defined by this infinite sequence of bits clearly has complexity a. Moreover, if we used simple strings such as all ones, in addition to all zeros, it is clear there are uncountably many points in F_a. ∎

A consequence of this lemm and the scaling property noted in Section 2 is Corollary 3.5.

COROLLARY 3.5. *The graph* Γ_K *is everywhere dense in the unit square.*

Consider the Cantor set $\mathscr{C}$ again. We claim that the fibre set F_c, where $c = \log 2/\log 3$ contains "almost all" points of $\mathscr{C}$. It follows that $\dim_H(F_c) \geqslant c$, for $c = \log 2/\log 3$. First we have to clarify the meaning of "almost all" here, as the Cantor set itself has Lebesgue measure zero. Intuitively the notion of a "random" Cantor set point should be clear, as points in $\mathscr{C}$ are represented by ternary numbers with 0 or 2 as its bits. This can be formalized as follows: Define a map e from the Cantor set $\mathscr{C}$ onto the unit interval $[0, 1]$ that is one-to-one, except on a countable subset of $\mathscr{C}$. Furthermore, modulo a countable subset the map e is an isomorphism between the measure space $\mathscr{C}$ endowed with the c-dimensional Hausdorff measure and the unit interval with the Lebesgue measure. The map can be defined by a sequence of "expansion" as follows: first map the points $\frac{1}{3}$ and $\frac{2}{3}$ to $\frac{1}{2}$ and expand the two intervals $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$ linearly onto $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, respectively. Then recursively expand the remaining two intervals exactly the same way, ad infinitum. It can be shown rigorously that all claims of the map e are satisfied. Now every $x \in \mathscr{C}$ certainly has complexity no more than $c = \log 2/\log 3$; in order to obtain $\lfloor \log_2 3 \cdot n \rfloor$ bits in a binary expansion we need no more than n bits asymptotically. On the other hand, just as in Lemma 3.3, a "random" point of the Cantor set (i.e., "almost all" under the c-dimensional Hausdorff measure) has complexity exactly c.

The above discussion is capable of generalization to an arbitrary a.

LEMMA 3.6. *For any a, $0 \leqslant a \leqslant 1$, the fibre set F_a has dimension at least a.*

We observe that there is nothing special about $\frac{1}{3}$ and $\frac{2}{3}$ in the Cantor set construction. One can easily construct generalized Cantor sets. Let $\{p_n/q_n\}$ be a recursive sequence of rational numbers so that $0 < p_n < q_n$ and $\log p_n/\log q_n \to a$, for the given real number a. Such a sequence certainly exists. One constructs a generalized Cantor set where in the nth step, we delete the middle $q_n - p_n$ subintervals each of length $1/q_n$ of the length of intervals obtained in the $(n-1)$th step. It can be shown that almost all points (under the a-dimensional Hausdorff measure) of the generalized Cantor set are contained in F_a, and thus the latter has dimension at least a. The lemma follows.

On the other hand, we claim that for $0 \leqslant a \leqslant 1$, and any $\varepsilon > 0$, $\dim_H(F_a) \leqslant a + \varepsilon$. And, hence, taking the limit, we have

LEMMA 3.7. *For any a, $0 \leqslant a \leqslant 1$, the fibre set F_a has dimension at most a.*

Proof. Let $x \in F_a$ and $1/k < \varepsilon$. Consider the family of closed intervals,

$$\{[m/2^n, (m+1)/2^n] \mid K(m) \leqslant (a+1/k)n\}, \qquad n = 1, 2, ...,$$

where $K(m)$ is the Kolmogorov complexity of the binary number m. Observe that for x with $\liminf_{n \to \infty} K(x(n))/n < a + 1/k$, where $x(n)$ is the n-place binary expansion of x, x is covered by infinitely many intervals in the above family. However, the number of intervals of length $1/2^n$ in the above family is bounded by $2^{(a+1/k)n+1}$, and thus the series

$$\sum_{n=1}^{\infty} 2 \cdot 2^{(a+1/k)n} \left(\frac{1}{2^n}\right)^{a+\varepsilon}$$

converges. Therefore its tail can be made arbitrarily small, and the tail corresponds to a countable covering of the set F_a with arbitrarily small diameter. ∎

We note that the preceding proof actually proved more, namely that $\dim_H(\bigcup_{0 \leqslant y \leqslant a} F_y) \leqslant a$, for all a. Combining the above two lemmas, we have Theorem 3.8.

THEOREM 3.8. *For any a, $0 \leqslant a \leqslant 1$, the fibre set F_a has dimension exactly a.*

We now turn to Theorem 3.1. We need the following technical result about Hausdorff dimension. For completeness we will include a proof here. See also Corollary 7.12 in [F].

THEOREM 3.9. *If for any y, $0 \leqslant y \leqslant 1$, a "fibre set" $F_y \subseteq I$ is defined and has Hausdorff dimension at least h, then* $\dim_H(\bigcup_{0 \leqslant y \leqslant 1} (F_y \times \{y\})) \geqslant 1 + h$.

Proof. Without loss of generality, we consider any countable covering of the set $\bigcup_{0 \leqslant y \leqslant 1} (F_y \times \{y\})$ by squares, $\mathscr{S} = \{[a_i, a_i + \delta_i] \times [b_i, b_i + \delta_i] \mid i \geqslant 1\}$. We note that the covering $\mathscr{S}$ naturally induces a covering for each fibre set F_y, $\{[a_i, a_i + \delta_i] \mid i \geqslant 1 \text{ and } b_i \leqslant y \leqslant b_i + \delta_i\}$. Fix any $\varepsilon > 0$, define a modified "characteristic" function for each square,

$$\chi_i(y) = \begin{cases} \delta_i^{h-\varepsilon} & \text{if } \quad b_i \leqslant y \leqslant b_i + \delta_i \\ 0 & \text{otherwise.} \end{cases}$$

Since each χ_i is non-negative, it follows from the monotone convergence theorem that

$$\int_0^1 \sum_{i=1}^{\infty} \chi_i(y)\, dy = \sum_{i=1}^{\infty} \int_0^1 \chi_i(y)\, dy = \sum_{i=1}^{\infty} \delta_i^{1+h-\varepsilon}.$$

We only need to show that the integral on the left approaches infinity (uniform over all coverings) as $\delta = \sup_{i \geqslant 1} \delta_i \to 0$. This follows from Ergorov's theorem. We can show directly as follows: For any M large and integer n, define $S_n = \{y \mid \inf \sum_i \delta_i^{h-\varepsilon} \geqslant 2M$, where the infimum takes over all countable coverings of F_y by intervals of lengths δ_i, and $\sup \delta_i \leqslant 1/n\}$. Since each F_y has Hausdorff dimension at least h, S_n forms a monotone non-decreasing sequence of sets with limit $\bigcup_n S_n = [0, 1]$. Then by the continuity of Lebesgue measure, $\lim_{n \to \infty} \mu(S_n) = 1$. We choose n sufficiently large such that $\mu(S_n) > \frac{1}{2}$, and

$$\int_0^1 \sum_{i=1}^{\infty} \chi_i(y)\, dy \geqslant \int_{S_n} \sum_{i=1}^{\infty} \chi_i(y)\, dy \geqslant M. \quad \blacksquare$$

We note that in Theorem 3.9, one can replace the interval $0 \leqslant y \leqslant 1$ by any other non-trivial interval. It follows that

$$\dim_H(\Gamma_K \cap ([0, 1] \times [a, b])) \geqslant \dim_H(\Gamma_K \cap ([0, 1] \times [b - \varepsilon, b])) \geqslant 1 + b - \varepsilon,$$

for all $\varepsilon > 0$. On the other hand, it follows from the remark after Lemma 3.7,

$$\dim_H(\Gamma_K \cap ([0, 1]) \times [a, b])) \leqslant \dim_H\left(\left(\bigcup_{0 \leqslant y \leqslant b} F_y\right) \times [a, b]\right) \leqslant 1 + b.$$

Theorem 3.1 follows.

As one referee pointed out, although motivated by different issues, Besicovitch [Be] had some remarkable resuls that are highly related to what has been presented in this section, in particular, Lemma 3.6.

Let us define the set $E_p = \{x : \limsup_{n \to \infty} [\#$ of ones in n-bits expansion of $x]/n = p\}$. Then Besicovitch proved that E_p has Hausdorff dimension $H(p) = -p \log_2 p - (1-p) \log_2(1-p)$. Although the Besicovitch set E_p is not a subset of some fibre set F_a, "almost all" points of E_p indeed are, for $a = H(p)$. This follows exactly the same reasoning as in Lemma 3.6, where the Cantor set was used.

Instead of using the Cantor set, it is now possible to use Besicovitch's result. Consider the subset of the Besicovitch set E_p consisting of only those points with Kolmogorov complexity that are equal to $H(p)$. "Almost all" points of E_p belong to this subset; thus it has Hausdorff dimension $H(p)$. As the function $H(p) = -p \log_2 p - (1-p) \log_2(1-p)$ is a one to one mapping from $[0, \frac{1}{2}]$ onto $[0, 1]$, it follows that our *fibre set* F_a has Hausdorff dimension at least a, for all a, namely the lower bound in Lemma 3.6.

Addendum

The referee in his detailed report suggested that we look for "a simultaneous generalization of the Besicovitch and Cai–Hartmanis theorems, which preserves the spirit and the technical content of both." He gave the following definition A, and then he stated the following theorem B and conjecture C.[2]

DEFINITION A. A real function f on A is r-expansive at x if there is a function g from strings to strings: (1) if $\sigma \subset \tau$, where σ and τ are initial segments of the binary expansion of some $y \in A$, then $g(\sigma) \subset g(\tau)$; (2) for all $y \in A$, $\lim_{\sigma \to y} g(\sigma) = f(y)$; and (3) $\limsup_{\sigma \to x}(|\sigma|/|g(\sigma)|) = r$. f is r-expansive on A if f is r-expansive at x for all $x \in A$.

THEOREM B. *Let A be a set of positive Lebesgue measure, and let f be r-expansive on A. Then the Hausdorff–Besicovitch dimension of $f(A)$ is at most r; moreover, this dimension is attained if f and g are one–one functions.*

Conjecture C. Let A be a set of Hausdorff–Besicovitch dimension s, and let f be r-expansive on A. Then the Hausdorff–Besicovitch dimension of $f(A)$ is at most rs; moreover, this dimension is attained if f and g are one–one functions.

He outlined how one can derive the results in Theorem 3.8 using his notion of r-expansiveness. Essentially the derivation goes as follows. For $y \in A$, we will consider the initial segments of the binary expansion of y as "programs" used in the Kolmogorov complexity. We will consider only programs that are "extensional" in the sense that the requirement (1) in Definition A is satisfied. Note that for any x with $K(x) = r$, there exist encodings (e.g., using relative Kolmogorov complexity of successive segments) that satisfy this extensionality. Let $A_{f,r}$ be the set of those y where the recursive function f is r-expansive. Then $F_r \subseteq \bigcup_{f \text{rec}} \{f(A_{f,r})\}$. It follows that $\dim_H(F_r) \leqslant r$, since in this countable union each set $A_{f,r}$ has dimension at most one and, thus, each set $f(A_{f,r})$ has dimension at most r. Moreover, the equality $\dim_H(F_r) = r$ holds. This follows from a limit argument similar to that of Lemma 3.6 and the second part of Theorem B and Conjecture C. Thus one can construct an appropriate A and one–one functions f and g such that f is r-expansive on A, $\dim_H(A) = 1$, and $f(A) \subseteq F_r$. This completes the derivation of the estimate on the Hausdorff dimension of the fibre sets as in Theorem 3.8: $\dim_H(F_r) = r$.

[2] These statements are taken from the referee's report with minor modifications.

Regarding his conjecture C the referee thinks "it would be nice to have such a theorem." We prove the conjecture in the remaining part of this addendum.

For a finite string $\sigma \in \{0, 1\}^*$, let $|\sigma|$ denote its length, let $\boldsymbol{\sigma}$ denote the real number between 0 and 1 corresponding to σ, i.e., $\boldsymbol{\sigma} = \sigma/2^{|\sigma|}$, let $I(\sigma)$ denote the interval $[\boldsymbol{\sigma}, \boldsymbol{\sigma} + 1/2^{|\sigma|})$, and let $|I|$ denote the length of the interval I, so $|I(\sigma)| = 1/2^{|\sigma|}$. Note that $\forall x \in I(\sigma)$, the $|\sigma|$th place binary expansion of x, denoted by $\sigma_{|\sigma|}(x)$, is just σ.

We fix any $r' > r$, $s' > s$, and any $\varepsilon > 0$. Our goal is to construct a covering $\mathscr{D}$ of $f(A)$ with arbitrary small diameter $\delta(\mathscr{D})$, such that

$$\sum_{I \in \mathscr{D}} |I|^{r's'} < \varepsilon.$$

Since for $x \in A$, $\limsup(|\sigma|/|g(\sigma)|) = r$, where $\sigma = \sigma_n(x)$ and the limit is with $n \to \infty$, we have $n(x)$, such that for all $n \geqslant n(x)$, $|g(\sigma_n(x))| > |\sigma_n(x)|/r' = n/r'$.

Define an infinite sequence $n_1 < n_2 < \cdots < n_k < \cdots$, such that $\forall k$, a cover $\mathscr{C}_k$ of A exists, $\mathscr{C}_k = \{I(a_{kj}) : j \geqslant 1\}$, satisfying $|a_{kj}| \geqslant n_k$ and

$$\sum_{j \geqslant 1} |I(a_{kj})|^{s'} < \frac{\varepsilon}{2^k}.$$

We claim that such a covering exists because $\dim_H(A) < s'$, and we can assume that the cover is of this form. In fact, since $\dim_H(A) < s'$, for the given ε, k, and n_k, there exists a covering $\mathscr{C}'_k = \{[\alpha_{kj}, \beta_{kj}) : j \geqslant 1\}$, such that, $\beta_{kj} - \alpha_{kj} < 1/2^{n_k}$ and,

$$\sum_{j \geqslant 1} (\beta_{kj} - \alpha_{kj})^{s'} < \frac{\varepsilon}{2^{k+s'+1}}.$$

To obtain our cover $\mathscr{C}_k$ from $\mathscr{C}'_k$, we will replace each $[\alpha, \beta) = [\alpha_{kj}, \beta_{kj})$ by at most two intervals which cover it and which have a combined length at most twice that of $[\alpha, \beta)$.

Let N be the least integer such that there exists an integer u, $\alpha \leqslant u/2^N < \beta$. For the least N such a u is unique; thus $(u-1)/2^N < \alpha$ and $(u+1)/2^N \geqslant \beta$. Let N_1 (and N_2, respectively) be the largest integer such that $u/2^N - \alpha < 1/2^{N_1}$ (and $\beta - u/2^N \leqslant 1/2^{N_2}$, respectively.) Then clearly $N_1, N_2 \geqslant N$, and $1/2^{N_1+1} \leqslant u/2^N - \alpha$ and $1/2^{N_2+1} < \beta - u/2^N$, by the maximality of N_1 and N_2.

Thus, the length of the interval $I = [(2^{N_1-N} \cdot u - 1)/2^{N_1}, 2^{N_1-N} \cdot u/2^{N_1})$ is $1/2^{N_1}$, which is at most $2(u/2^N - \alpha)$; i.e., it is at most twice the length of $[\alpha, u/2^N)$. Call this length A. Similarly the length of $J = [2^{N_2-N} \cdot u/2^{N_2}, (2^{N_2-N} \cdot u + 1)/2^{N_2})$ is at most twice the length of $[u/2^N, \beta)$. Call this length B. Note that $A + B = \beta - \alpha$. Now,

$$\begin{aligned} |I|^{s'} + |J|^{s'} &\leqslant 2^{s'}(A^{s'} + B^{s'}) \\ &\leqslant 2^{s'+1} \max(A^{s'}, B^{s'}) \\ &\leqslant 2^{s'+1} \cdot (A+B)^{s'}. \end{aligned}$$

Let $\mathscr{C}_k$ consist of I_j and J_j for each $[\alpha_{kj}, \beta_{kj})$ in $\mathscr{C}'_k$, we obtain

$$\sum_{j\geqslant 1} (|I_j|^{s'} + |J_j|^{s'}) < \frac{\varepsilon}{2^k}.$$

The claim is proved.

Now we define a covering $\mathscr{D}$ of $f(A)$:

$$\mathscr{D} = \{I(g(a)) \mid \exists k,\ I(a)\in\mathscr{C}_k, \text{ and } \exists x\in A,\ x\in I(a), \text{ and } n_k \geqslant n(x)\}.$$

We claim that $\mathscr{D}$ is a covering of $f(A)$.

Let $x\in A$, then $\exists k$, $n_k\geqslant n(x)$. Consider $\mathscr{C}_k$ which covers A. So $\exists I(a)\in\mathscr{C}_k$ and $x\in I(a)$. Now $I(g(a))\in\mathscr{D}$ and $f(x)\in I(g(a))$. The first part is clear, as witnessed by k and x. Since $x\in I(a)$, $a=\sigma_{|a|}(x)$ and $f(x)\in I(g(\sigma_n(x)))$ for all n by the extensionality property of g. By setting $n=|a|$, $f(x)\in I(g(\sigma_{|a|}(x)))=I(g(a))$.

We now estimate $|I(g(a))|$ for any interval in $\mathscr{D}$. For $I(g(a))\in\mathscr{D}$, $\exists k$ and $x\in A$, $n_k\geqslant n(x)$, $x\in I(a)$, and $I(a)\in\mathscr{C}_k$. By the definition of $\mathscr{C}_k$, $|a|\geqslant n_k\geqslant n(x)$. As $a=\sigma_{|a|}(x)$, since $x\in A$ and by the definition of $n(x)$,

$$|g(a)| = |g(\sigma_{|a|}(x))| > \frac{|\sigma_{|a|}(x)|}{r'} = \frac{|a|}{r'}.$$

Hence,

$$|I(g(a))| = \frac{1}{2^{|g(a)|}} < \frac{1}{2^{|a|/r'}} = |I(a)|^{1/r'}.$$

Therefore,

$$\sum_{\mathscr{D}} |I(g(a))|^{r's'} < \sum_{k=1}^{\infty} \sum_{j\geqslant 1} |I(a_{kj})|^{s'} < \varepsilon.$$

As we can make n_1 arbitrarily large, so that the diameter $\delta(\mathscr{D})$ is arbitrarily small, this proves that $\dim_H(f(A))\leqslant r's'$. But as $r'>r$, $s'>s$ are arbitrary, we have shown that $\dim_H(f(A))\leqslant rs$.

In the case of one–one functions, we just apply the above theorem to f^{-1}, and equality follows. This completes the proof of Conjecture C and concludes this addendum.

4. The Topological Dimension of Γ_K

In this section we prove the following theorem.

Theorem 4.1. *The topological dimension of the set Γ_K is* 1.

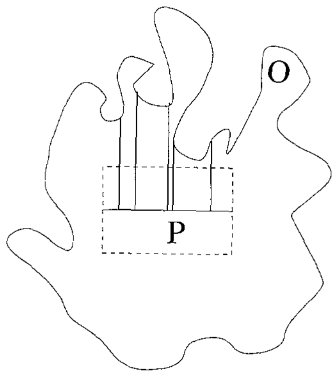

FIG. 1. The function l.

COROLLARY 4.2. *The graph Γ_K of the Kolmogorov complexity function K is a fractal in the sense of Mandelbrot.*

The proof of Theorem 4.1 has two parts; we show that $\dim_T(\Gamma_K) \leqslant 1$ and $\dim_T(\Gamma_K) \neq 0$.

It is easy to show that $\dim_T(\Gamma_K) \leqslant 1$. Given any point $p \in \Gamma_K$, we need to find an arbitrarily small neighborhood of p such that its boundary has topological dimension zero. This can be accomplished by a square $(a, a') \times (b, b') \ni p$, where $K(a), K(a') \notin [b, b']$. Thus the boundary of the square in the subspace Γ_K is a part of the fibre sets F_b and $F_{b'}$, which certainly has dimension zero, for it does not contain any interval.

We show next that the topological dimension of Γ_K is not zero. In fact, we show, for all $p \in \Gamma_K$ and any sufficiently small open neighborhood O of p, that the boundary $\partial_{\Gamma_K} O \neq \varnothing$. Recall that, by Lemma 3.5, $\partial_{\Gamma_K} O = \partial O \cap \Gamma_K$ as Γ_K is everywhere dense in $[0, 1]^2$.

Suppose that $p = (p_x, p_y) \in O$. Either $p_y < 1$ or $p_y > 0$. Without loss of generality we assume that $p_y < 1$, and $O \subseteq [0, 1] \times [0, 1)$. Take a small square $[a, a'] \times [b, b'] \subset O$ centered at p. We define a function l:

$$l(x) = \inf\{y \mid y > p_y \text{ and } (x, y) \in \partial O\} \qquad \text{for} \quad a \leqslant x \leqslant a'.$$

Surely, $p_y < b' < l(x) \leqslant 1$ (see Fig. 1).

LEMMA 4.3. *Except on a countable subset of $[a, a']$, the function l satisfies*

$$\liminf_{z \to x} l(z) = l(z).$$

Proof. We first observe that for all $x \in (a, a']$, since ∂O is closed, $\liminf_{z \to x^-} l(z) \geqslant l(x)$. Similarly, for all $x \in [a, a')$, $\liminf_{z \to x^+} l(z) \geqslant l(x)$. Let

$$J_n^- = \{a < x \leqslant a' \mid \liminf_{z \to x^-} l(z) > l(x) + 1/n\},$$

and $J^- = \bigcup_{n>1} J_n^-$. Similarly,

$$J_n^+ = \{a \leqslant x < a' \mid \liminf_{z \to x^+} l(z) > l(x) + 1/n\},$$

and $J^+ = \bigcup_{n>1} J_n^+$. Finally,

$$J = J^- \cup J^+ \supseteq \{a \leqslant x \leqslant a' \mid \liminf_{z \to x} l(z) > l(x)\}.$$

We claim that J is a countable set. Clearly, since a dual argument applies, it suffices to show that for each n and $1 \leqslant m \leqslant n$, the set $J_{n,m}^- = J_n^- \cap l^{-1}(((m-1)/n, m/n])$ is countable.

For all $x \in J_{n,m}^-$, since $(m-1)/n < l(x)$ and $\liminf_{z \to x^-} l(z) > l(x) + 1/n$, $\exists \varepsilon_x > 0$, such that

$$\inf\{l(z) \mid x - \varepsilon_x < z < x\} \geqslant l(x) + 1/n > m/n.$$

Thus $(x - \varepsilon_x, x) \cap J_{n,m}^- = \varnothing$. It follows that $\{(x - \varepsilon_x, x) \mid x \in J_{n,m}^-\}$ is a pairwise disjoint class of open intervals. Since

$$\sum_{x \in J_{n,m}^-} \varepsilon_x \leqslant a' - a \leqslant 1,$$

$J_{n,m}^-$ must be countable, and hence so is the set J. ∎

We write $J = \{a_1, a_2, ...\}$. Now, to complete the proof of Theorem 4.1, we can exhibit a point on the intersection of Γ_K and ∂O. The idea is to construct binary sequence in stages as in Lemma 3.4, approximating a "moving target" value which converges. Specifically, at stage i, we take the value $\inf l(z)$, where the infimum takes over the small interval $[m/2^n, (m+1)/2^n]$ defined by the binary number m which, as a binary string, was constructed up to the previous stage $i-1$. Then we "push" the normalized Kolmogorov complexity closer (up or down) to this infimum, by appending hard or easy strings. Meanwhile, we avoid one more exceptional point a_i from J by a positive distance (starting with 00 or 11). As the nested intervals shrink, it defines a unique number $x \notin J$. Therefore, $\liminf_{z \to x} l(z) = l(x)$. On the other hand, the "moving target" clearly converges to $\liminf_{z \to x} l(z)$. Thus the construction yields $K(x) = \liminf_{z \to x} l(z) = l(x)$. The proof of Theorem 4.1 is completed.

We remark that Theorem 4.1, as well as Theorem 3.1, are valid no matter how one extends the definition of $K(x)$ for x, where $\lim_{n \to \infty} K(x_n)/n$ does not exist. For instance, for the point x we exhibited in the proof above, the limit $\lim_{n \to \infty} K(x_n)/n$ in fact exists.

Theorem 4.1 can be quite a powerful tool in the study of Kolmogorov complexity of the reals. We indicate this by a simple solution to a problem of Razborov: What relationship can there be between $K(x)$ and $K(K(x))$? He asks in particular, for instance, is it true that $K(K(x)) \leqslant K(x)$ for all $0 \leqslant x \leqslant 1$?

The answer is negative. We simply consider the line $Y = X + \varepsilon$ within the unit square $[0, 1] \times [0, 1]$. As Γ_K intersects both triangular regions formed by the line $Y = X + \varepsilon$ (Γ_K is everywhere dense in the unit square), it is impossible to have the boundary of these triangular open sets, namely the line, disjoint from Γ_K, by Theorem 4.1. Hence, there exists t, such that $K(t) = t + \varepsilon > t$. Then, since the fiber set F_t is nonempty, there exists x, such that $K(x) = t$ and, thus, $K(K(x)) > K(x)$. Of course this argument can be generalized.

THEOREM 4.4. *For any differentiable function f from $[0, 1]$ to $(0, 1)$ neither $K(K(x)) \leqslant f(K(x))$ nor $K(K(x)) \geqslant f(K(x))$ is true for all x.*

ACKNOWLEDGMENTS

The authors thank Professor Wolfgang Fuchs for pointing out a simplification of a proof in an earlier draft. We thank Professors Allen Back, David Henderson, and Tom Rishel for interesting discussions on dimension theory and analysis. The authors also thank Sasha Razborov for discussions on his questions regarding the relationship of $K(x)$ and $K(K(x))$. Last but not least, we wish to thank the anonymous referee for providing a detailed report, for telling us his notion of r-expansiveness, his theorem, and his conjecture. We also wish to thank him for pointing out the related work of Besicovitch [Be] and the reference in Falconer's book [F].

REFERENCES

[Ba] Y. M. BARZDIN, Complexity of programs to determine whether natural numbers not greater than n belong to a recursively enumerable set, *Soviet Math. Dokl.* **9** (1968), 1251–1254.

[Be] A. S. BESICOVITCH, On the sum of digits of real numbers represented in the dyadic systems, *Math. Ann.* **34** (1934), 321–330.

[Bl] P. BLANCHARD, Complex analytic dynamic systems on the Riemann sphere, *Bull. Amer. Math. Soc.* **11** (1984), 85–141.

[Br] J. BROUWER, Über den natürlichen Dimensionsbegriff, *J. Math.* **142** (1913), 146–152.

[C1] G. J. CHAITIN, On the length of programs for computing finite binary sequences, *J. Assoc. Comput. Math.* **13** (1966), 547–569.

[C2] G. J. CHAITIN, A theory of program size formally identical to information theory, *J. Assoc. Comput. Math.* **22** (1975), 329–340.

[F] K. FALCONER, "Fractal Geometry—Mathematical Foundations and Applications," Wiley, New York, 1990.

[H] J. HARTMANIS, Generalized Kolmogorov complexity and the structure of feasible computation, *in* "Proceedings, 24th Annual Symposium on Foundations of Computer Science, 1983," pp. 439–445.

[HW] W. HUREWICZ AND H. WALLMAN, "Dimension Theory," Princeton Univ. Press, Princeton, NJ, 1974.

[K] A. N. KOLMOGOROV, Three approaches to the quantitative definition of information, *Problems Inform. Transmission* **1**, No. 1 (1965), 1–7.

[L1] L. A. LEVIN, Laws of information conservation (non-growth) and aspects of the foundation of probability theory, *Problems Inform. Transmission* **10** (1974), 206–210.

[L2] L. A. LEVIN, Homogeneous measures and polynomial time invariants, *in* "Proceedings, 29th Annual Symposium on Foundations of Computer Science, 1988," pp. 36–41.

[LV] M. LI AND P. M. B. VITANYI, Two decades of applied Kolmogorov complexity: in memoriam Andrei Nikolaevich Kolmogorov 1903–1987, *in* "Proceedings, Structure in Complexity Theory, 3rd Annual Conference, 1988," pp. 80–101.

[M] B. MANDELBROT, "The Fractal Nature of Geometry," Freeman, San Francisco, 1983.

[Ra] A. A. RAZBOROV, personal communication.

[Ry] B. YA. RYABKO, Noiseless encoding of combinatorial sources, Hausdorff dimension and Kolmogorov complexity, *Problems Inform. Transmission* **22**, No. 3 (1986), 16–26. [Russian]

[Sm] S. SMALE, Differentiable dynamic systems, *Bull. Amer. Math. Soc.* **73** (1987), 747–817.

[So] R. J. SOLOMONOFF, A formal theory of inductive inference, Parts 1, 2, *Inform. and Control* **7** (1964), 1–22, 224–254.

[St1] L. STAIGER, "Complexity and Entropy, MFCS81," Lecture Notes in Computer Science, Vol. 118, pp. 508–514, Springer-Verlag, New York/Berlin, 1981.

[St2] L. STAIGER, Kolmogorov complexity and Hausdorff dimension, manuscript.