PERGAMON

Neural Networks 15 (2002) 57–67

Neural Networks

www.elsevier.com/locate/neunet

Contributed article

Space-filling curves and Kolmogorov superposition-based neural networks

David A. Sprecher[a,*], Sorin Draghici[b]

[a]Department of Mathematics, University of California, Santa Barbara, CA 93106, USA
[b]Department of Computer Science, Wayne State University, Detroit, MI 48202, USA

Received 22 January 2001; revised 20 July 2001; accepted 20 July 2001

Abstract

Kolmogorov superpositions and Hecht-Nielsen's neural network based on them are dimension reducing. This dimension reduction can be understood in terms of space-filling curves that characterize Kolmogorov's functions, and the subject of this paper is the construction of such a curve. We construct a space-filling curve with Lebesgue measure 1 in the unit square, $[0,1]^2$, with approximating curves Λ_k, $k = 1,2,3,\ldots$, each with 10^{2k} rational nodal-points whose order is determined for each k by the linear order of their image-points under a nomographic function $y = \alpha_1 \psi(x_1) + \alpha_2 \psi(x_2)$ that is the basis of a computable version of the Kolmogorov superpositions in two dimensions. The function $\psi:[0,1] \to [0,1]$ is continuous and monotonic increasing, and α_1, α_2 are suitable constants. The curves Λ_k are composed of families of disjoint closed squares of diminishing diameters and connecting joins of diminishing lengths as $k \to \infty$.

Keywords: Space-filling curves; Peano curves; Kolmogorov superpositions; Superpositions; Nomographic functions; Hecht-Nielsen neural network; Neural networks

1. Introduction

Referring to the Nomenclature, consider the two-dimensional version of the Kolmogorov (1957) superpositions:

$$f(x_1, x_2) = \sum_{q=0}^{4} \Phi_q(y_q), \tag{1}$$

in which an arbitrary real-valued continuous function $f:\mathscr{E}^2 \to \mathscr{R}$ is computed with continuous functions $\Phi_q:\mathscr{E} \to \mathscr{R}$. The arguments:

$$y_q = \alpha_1 \psi(x_1 + qa) + \alpha_2 \psi(x_2 + qa) \tag{2}$$

are fixed nomographic functions that are independent of f; The function $\psi:\mathscr{E} \to \mathscr{E}$ is monotonic increasing and continuous, and α_1, α_2 and a are suitable constants (Sprecher, 1996, 1997). In a pioneering paper, Hecht-Nielsen (1987) linked the n-dimensional version of Kolmogorov superpositions to computer architecture by interpreting them as a four-layer architecture of a feedforward neural network, as represented schematically in Fig. 1. This architecture consists of two pairs of *nonlinear* $\to$ *linear* layers, the first pair constituting a dimension-dependent and otherwise fixed hidden layer, and the second pair constituting the output layer in which an arbitrary target function f is implemented. The nonlinear activation functions of the units in the hidden layer may therefore be implemented in hardware and hard-coded into the network.

The functions in Eq. (2) are the most basic nonlinear continuous functions that can be used to replace a pair of variables (x_1, x_2) with new variables y_q, with the result that the computation of the two-variable function $f(x_1, x_2)$ is carried out exclusively in terms of the five one-variable functions $\Phi_q(y_q)$. This dimension-reduction is accomplished with the binary operation ' + ' in Eq. (2), and correspondingly in the hidden layer of the network. Here we examine the underlying topological properties of these functions that enable this dimension reduction, and we begin with the general statement that there exist continuous curves passing through all points of a square or a subset thereof, thereby inducing a linear order on an infinite dense set of points (x_1, x_2). The question of existence of such curves followed Cantor's 1878 work that demonstrated, among other things, that there is a one–one relationship between the points of a square and the points of an interval. The first *space-filling curve*, as these curves are called, was discovered by Peano in 1890, and mathematicians since have followed his

* Corresponding author. Tel.: +1-805-568-0681.
E-mail address: sprecher@math.ucsb.edu (D.A. Sprecher).

0893-6080/02/$ - see front matter
PII: S0893-6080(01)00107-1

Nomenclature

$\mathscr{E}$	Unit interval [0,1]
$\mathscr{E}^2$	Cartesian product $[0,1] \times [0,1]$
ξ	Nomographic mapping $\mathscr{E}^2 \to \mathscr{E}$
λ	Mapping $\mathscr{E} \to \mathscr{E}^2$
Λ	$\{\lambda(z) \in \mathscr{E}^2 : z \in \mathscr{E}\}$
d_k	Rational numbers $\Sigma_{r=1}^{k} i_r 10^{-r}$, $i_r = 0,1,2,\ldots 9$
$\mathbf{d}_k$	Rational point $(d_{k,1}, d_{k,2})$
z_k	Image point $\xi(\mathbf{d}_k)$ of $\mathbf{d}_k$
$\mathrm{dist}(\mathbf{d}_1, \mathbf{d}_2)$	Euclidean distance
$\lvert T_k \rvert$	Length of interval T_k
$\mathbf{x}$	Point (x_1, x_2)
b_k	Series $\Sigma_{r=k+1}^{\infty} 10^{-2^r+1}$

discovery with a detailed study of their properties as well as the construction of a rich variety of space-filling curves (Sagan, 1991). Each of the functions in Eq. (2) determines a space-filling curve with the specific properties that are required to obtain Eq. (1). The subject of this paper is the construction of a space-filling curve characterizing the functions (Eq. 2), and in turn illuminating the algorithm leading to the Kolmogorov superpositions and Hecht-Nielsen's neural network.

Following Kolmogorov's original strategy, Eq. (1) is established through the construction of suitable families of disjoint closed squares of increasing refinement for each value of q, each covering the unit square $\mathscr{E}^2$ except for narrow horizontal and vertical strips. Each function (Eq. 2) is such that it maps its families of squares one–one onto families of disjoint closed intervals.[1] The diameters of the squares and the widths of the gaps separating them tend asymptotically to zero, and so do the lengths of the corresponding intervals and their separating gaps. When each family of squares is appropriately joined pair wise at each stage, they form curves (chains) that converge to a space-filling curve. We are able to refer to such chains as curves because a square can be mapped onto its diagonal. The constructions in this paper are for the case $q = 0$ in Eq. (2).

Section 1 is devoted to a rehearsal of those properties of the function y_0 that are necessary for the analytic construction of a space-filling curve Λ in Section 2. This construction culminates in a lemma and four theorems—all proved in Section 4. Section 3 contains a geometric construction of approximating curves Λ_k. MatLab was used to generate computer graphs of Λ_0 and Λ_1. These graphs clearly reveal the non-linearity of the function y_0. An examination of the Kolmogorov (1957) original functions $y_r = \psi_{1,r}(x_1) + \psi_{2,r}(x_2)$, $r = 0,1,\ldots,4$ suggests that these also generate similar space filling curves, but this needs yet to be verified.

[1] Details of the specific constructions can be found in Sprecher (1996, 1997).

2. The nomographic function $\xi(\mathbf{x})$

Our starting point is the nomographic function:

$$\xi(\mathbf{x}) = \alpha_1 \psi(x_1) + \alpha_2 \psi(x_2) \tag{3}$$

with ψ as defined in the Appendix, having rational values at the points:

$$d_k = \sum_{r=1}^{k} i_r 10^{-r},$$

$i_r = 0,1,2,\ldots 9$ and $k = 1,2,3,\ldots$. The normalized constants α_1, α_2, also defined in the Appendix, are rationally independent, so that if we set $z_k = \xi(\mathbf{d}_k)$ and $z'_k = \xi(\mathbf{d}'_k)$, then for fixed k, $z_k = z'_k$ *if and only if* $\mathbf{d}_k = \mathbf{d}'_k$. Consequently, $\mathbf{d}_k = \xi^{-1}(z_k)$ is well defined for fixed k, and listing the image point $z_k = \xi(\mathbf{d}_k)$ in increasing order: $0 = z_k^1 < z_k^2 < z_k^3 < \ldots < z_k^{10^{2k}}$ results in a corresponding sequence $\mathbf{d}_k^1, \mathbf{d}_k^2, \mathbf{d}_k^3, \ldots, \mathbf{d}_k^{10^{2k}}$ that is indexed through $\xi^{-1}(z_k^\omega) = \mathbf{d}_k^\omega$, so that $\mathbf{d}_k^{\omega+1}$ is the immediate successor of $\mathbf{d}_k^\omega$.[2] Associated with each grid-point is a closed square:

$$S_k(\mathbf{d}_k^\omega) = [d_{k,1}^\omega, e_{k,1}^\omega] \times [d_{k,2}^\omega, e_{k,2}^\omega], \tag{4}$$

where:

$$e_{k,p}^\omega = d_{k,p}^\omega + \tfrac{8}{9} 10^{-k}. \tag{5}$$

For fixed k, these squares are separated by horizontal and vertical strips of widths $\frac{1}{9} 10^{-k}$, and the function (Eq. 3) maps each square $S_k(\mathbf{d}_k^\omega)$ onto a closed interval:

$$T_k(z_k^\omega) = [z_k^\omega,\ z_k^\omega + 8b_k], \tag{6}$$

where:

$$b_k = \sum_{r=k+1}^{\infty} 10^{-2^r+1}. \tag{7}$$

[2] Warning: The fact that $\mathbf{d}_{k,1}^{\omega+1} = (d_{k,1}^{\omega+1}, d_{k,2}^{\omega+1})$ is the successor of $\mathbf{d}_k^\omega = (d_{k,1}^\omega, d_{k,2}^\omega)$ does not imply that $d_{k,p}^{\omega+1}$ is the successor of $d_{k,p}^\omega$. In the case $k = 1$, for example, $\mathbf{d}_1^{11} = (\frac{1}{10}, 0)$ is the successor of $\mathbf{d}_1^{10} = (0, 9/10)$, yet 0 is not the successor of 9/10. Here the superscript ω serves only to indicate the association of an ordered-pair $(d_{k,1}^\omega, d_{k,2}^\omega)$ with $\mathbf{d}_k^\omega$.

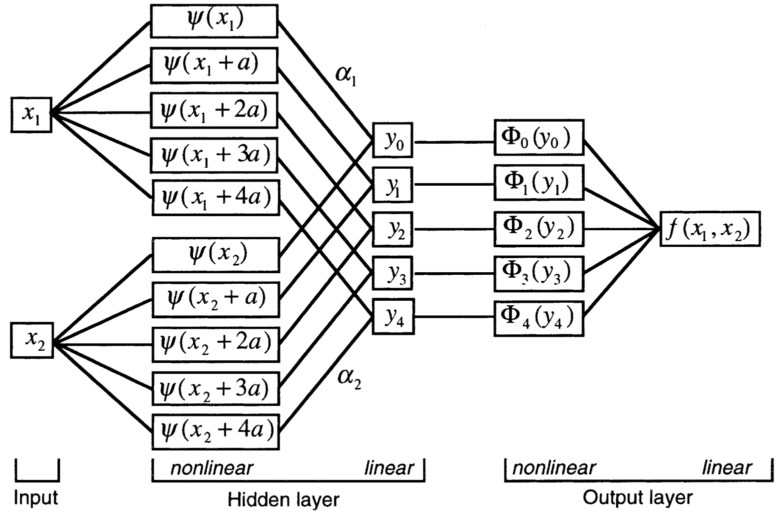

Fig. 1. Hecht-Nielsen's neural network for $n = 2$.

Since:

$$\tfrac{8}{9}10^{-k} = 8\sum_{r=k+1}^{\infty} 10^{-r},$$

an application of Eq. (A1) to Eq. (5) shows that $\psi(e_{k,p}^{\omega}) = \psi(d_{k,p}^{\omega}) + 8b_k$, so that $\xi(e_k^{\omega}) = z_k^{\omega} + 8b_k < z_k^{\omega+1}$. Like each system of squares, the intervals are pair wise disjoint for each value of k.

Now, given arbitrary squares $S_k(\mathbf{d}_k)$ and $S_m(\mathbf{d}_m)$ with $m > k$ and corresponding intervals $T_k(z_k)$ and $T_m(z_m)$, then either $S_m(\mathbf{d}_m) \cap S_k(\mathbf{d}_k) = \varnothing$ or $S_m(\mathbf{d}_m) \subset S_k(\mathbf{d}_k)$, and likewise $T_m(z_m) \cap T_k(z_k) = \varnothing$ or $T_m(z_m) \subset T_k(z_k)$. It follows directly from Eqs. (4)–(7) that the inclusion of squares $S_m(\mathbf{d}_m) \subset S_k(\mathbf{d}_k)$ implies the inclusion of corresponding intervals $T_m(z_m) \subset T_k(z_k)$. *The converse, however, is not true*, as the following example shows.

Example 1. Let $\mathbf{d}_1 = (0,0)$ and $\mathbf{d}_2 = (0, \frac{9}{10^2})$. Then according to Eq. (A1) and Eq. (3), $z_1 = \xi(\mathbf{d}_1) = 0$ and:

$$z_2 = \xi(\mathbf{d}_2) = \alpha_1\psi(0) + \alpha_2\psi(\tfrac{9}{10}) = \frac{\alpha_2}{2\cdot 10},$$

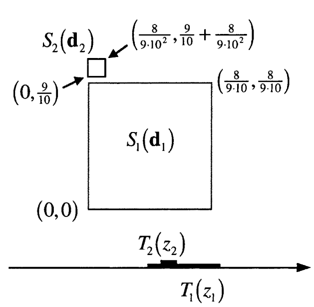

Fig. 2. The case $T_2(z_2) \subset T_1(z_1)$ and $S_2(\mathbf{d}_2) \not\subset S_1(\mathbf{d}_1)$.

and this gives intervals:

$$T_1(z_1) = \left[0,\ 8\sum_{r=2}^{\infty} 10^{-2^r+1}\right]$$

$$T_2(z_2) = \left[\frac{\alpha_2}{2\cdot 10},\ \frac{\alpha_2}{2\cdot 10} + 8\alpha_2\sum_{r=3}^{\infty} 10^{-2^r+1}\right]$$

with corresponding squares:

$$S_1(\mathbf{d}_1) = \left[0,\ \tfrac{8}{9\cdot 10}\right]\times\left[0,\ \tfrac{8}{9\cdot 10}\right]$$

$$S_2(\mathbf{d}_2) = \left[0,\ \tfrac{8}{9\cdot 10^2}\right]\times\left[\tfrac{9}{10^2},\ \tfrac{9}{10^2} + \tfrac{8}{9\cdot 10^2}\right]$$

(Fig. 2). The inclusion $T_2(z_2) \subset T_1(z_1)$ follows from the inequality:

$$\frac{\alpha_2}{2\cdot 10} + 8\sum_{r=3}^{\infty} 10^{-2^r+1} < 8\sum_{r=2}^{\infty} 10^{-2^r+1},$$

yet clearly $S_1(\mathbf{d}_1) \cap S_2(\mathbf{d}_2) = \varnothing$.

To complete the preparatory constructions of this section, let $U_k(z_k^{\omega})$ designate for $\omega = 1,2,\ldots 10^{2k} - 1$ the open interval separating intervals $T_k(z_k^{\omega})$ and $T_k(z_k^{\omega+1})$:

$$U_k(z_k^{\omega}) = (z_k^{\omega} + 8b_k, z_k^{\omega+1}); \tag{8}$$

For $\omega = 10^{2k}$ let:

$$U_k(z_k^{10^{2k}}) = (z_k^{10^{2k}} + 8b_k, 1). \tag{9}$$

Then:

$$\mathscr{E} = \bigcup_{\omega=1}^{10^{2k}} T_k(z_k^{\omega}) \cup U_k(z_k^{\omega}). \tag{10}$$

3. The curve Λ

We begin with the observations that every infinite sequence of nested intervals $T_{k_1}(z_{k_1}) \supset T_{k_2}(z_{k_2}) \supset \dots$ determines a unique point:

$$z = \bigcap_{r=1}^{\infty} T_{k_r}(z_{k_r})$$

as well as a unique sequence of squares $S_{k_1}(\mathbf{d}_{k_1})$, $S_{k_2}(\mathbf{d}_{k_2})$,…. This sequence may contain a nested sequence $S_{k_s}(\mathbf{d}_{k_s}) \supset S_{k_{s+1}}(\mathbf{d}_{k_{s+1}}) \supset \dots$ defining a unique image point of z for some integer $s \geq 1$: We shall show that the sequence $\mathbf{d}_{k_1}$, $\mathbf{d}_{k_2}$, $\mathbf{d}_{k_3}$,… converges also when no such nested sequence exists. When $z \in \mathscr{E}$ is not the infinite intersection of intervals $T_{k_r}(z_{k_r})$ then there is an integer k and a sequence of nested intervals $U_k(z_k) \supset U_{k+1}(z_{k+1}) \supset \dots$ such that:

$$z = \bigcap_{r=0}^{\infty} U_{k+r}(z_{k+r}).$$

The existence of a space-filling curve associated with Eq. (3) is derived from the following property of this function:

Lemma 1. Let an integer k be given. If $|d_{k,2} - d'_{k,2}| \geq 10^{-s}$ for some integer $1 \leq s \leq k-1$, then:

$$|z_k - z'_k| > \frac{6}{10^{2^s+2}}.$$

From this lemma we have:

Theorem 1.

1. $\max_\omega \operatorname{dist}(\mathbf{d}_k^{\omega+1}, \mathbf{d}_k^{\omega}) < C \cdot 10^{-\log_2(k-1)}$, where $C = \sqrt{2} \cdot 10^{\log_2 \log_2 10}$.
2. If $T_{k_{r+1}}(z_{k_{r+1}}) \subset T_{k_r}(z_{k_r})$ then $\operatorname{dist}(\mathbf{d}_{k_{r+1}}, \mathbf{d}_{k_r}) < \sqrt{2} \cdot 10^{-k_r}$.
3. If $U_{k+1}(z_{k+1}) \subset U_k(z_k)$ then $\operatorname{dist}(\mathbf{d}_{k+1}\boldsymbol{d}_k) < C \cdot 10^{-\log_2(k-1)+1}$.

Theorem 2.

1. If:

   $$z = \bigcap_{r=1}^{\infty} T_{k_r}(z_{k_r}),$$

   then the corresponding sequence $\mathbf{d}_{k_1}$, $\mathbf{d}_{k_2}$, $\mathbf{d}_{k_3}$… converges.
2. If:

   $$z = \bigcap_{r=0}^{\infty} U_{k+r}(z_{k+r})$$

   for some integer k then the corresponding sequence $\mathbf{d}_{k+r}$, $\mathbf{d}_{k+r+1}$, $\mathbf{d}_{k+r+2}$,… converges.

We now define the sets:

$$\mathscr{K}_1 = \left\{ z \in \mathscr{E} : \ z = \bigcap_{r=1}^{\infty} T_{k_r}(z_{k_r}) \text{ and } \bigcap_{r=s}^{\infty} S_{k_r}(\mathbf{d}_{k_r}) \neq \emptyset \text{ for some } s \right\}$$

$$\mathscr{K}_2 = \left\{ z \in \mathscr{E} : \ z = \bigcap_{r=1}^{\infty} T_{k_r}(z_{k_r}) \text{ and } \bigcap_{r=s}^{\infty} S_{k_r}(\mathbf{d}_{k_r}) = \emptyset \text{ for every } s \right\}.$$

$$\mathscr{K}_3 = \left\{ z \in \mathscr{E} : \ z = \bigcap_{r=0}^{\infty} U_{k+r}(z_{k+r}) \text{ for some } k \right\}$$

and state:

Theorem 3.

1. If $z, z' \in \mathscr{K}_1$, then $z \neq z'$ if and only if $\lambda(z) \neq \lambda(z')$;
2. If $z, z' \in \mathscr{K}_2$ or $z, z' \in \mathscr{K}_3$ then $z \neq z'$ if $\lambda(z) \neq \lambda(z')$.

With this and the observation that $\mathscr{E} = \mathscr{K}_1 \cup \mathscr{K}_2 \cup \mathscr{K}_3$ we can now define Λ: Consider the mapping $\lambda : \mathscr{E} \to \mathscr{E}^2$, specified as follows:

$$\lambda(z) = \bigcap_{r=1}^{\infty} S_{k_r}(\mathbf{d}_{k_r}) \tag{11}$$

when $z \in \mathscr{K}_1$

$$\lambda(z) = \lim_{r \to \infty} \mathbf{d}_{k_r} \tag{12}$$

when $z \in \mathscr{K}_2$

$$\lambda(z) = \lim_{s \to \infty} \mathbf{d}_{k+s} \tag{13}$$

when $z \in \mathscr{K}_3$.

We now have:

Theorem 4. The direct image $\Lambda = \{\lambda(z) \in \mathscr{E}^2 : z \in \mathscr{E}\}$ of λ is a space-filling curve.

This implies, of course, that the mapping $\lambda : \mathscr{E} \to \mathscr{E}^2$ is continuous (Sagan, 1991).

4. Approximating curves

We begin with:

Definition 1. Let a point $z \in \mathscr{E}$ be a given. The curve $\xi^{-1}(z) = \{\mathbf{x} \in \mathscr{E}^2 : \xi(\mathbf{x}) = z\}$ is called a *level-curve* of the function ξ.

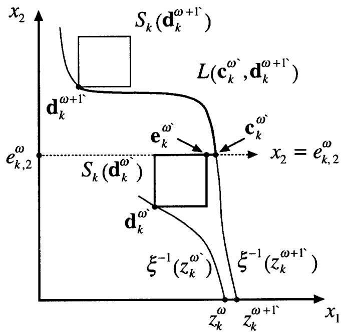

Fig. 3. Anatomy of a join $J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1})$.

Because the function ψ is continuous and monotonic increasing and the constants α_1 and α_2 are positive, $\xi^{-1}(z)$ is a continuous monotonic decreasing curve for fixed z, with initial and end points on the boundary of $\mathscr{E}^2$. For fixed k and $\omega = 1,2,3,\ldots$ we now connect each square $S_k(\mathbf{d}_k^{\omega})$ to its immediate successor $S_k(\mathbf{d}_k^{\omega+1})$ with a *join* to form a continuous chain (*approximating curve*) originating at (0,0) and ending at (1,1), in such a way that no two joins have common contact points: Our application of the term *successor* to squares is to be understood in terms of the order induced on the grid-points $\mathbf{d}_k^{\omega}$ through $\xi^{-1}(z_k^{\omega}) = \mathbf{d}_k^{\omega}$, and no ambiguity is likely to arise out of this informal use. Consulting Fig. 3, let the point of intersection of the horizontal line $x_2 = e_{k,2}^{\omega}$ with the level-curve $\xi^{-1}(z_k^{\omega+1})$ be designated $\mathbf{c}_k^{\omega} = (c_{1,k}^{\omega}, d_{2,k}^{\omega+1})$. Designate by $I(\mathbf{e}_k^{\omega}, \mathbf{c}_k^{\omega}]$ the half-open line segment connecting the points $\mathbf{e}_k^{\omega}$ and $\mathbf{c}_k^{\omega}$, and by $L(\mathbf{c}_k^{\omega}, \mathbf{d}_k^{\omega+1})$ the open level-curve segment connecting the points $\mathbf{c}_k^{\omega}$ and $\mathbf{d}_k^{\omega+1}$.

Definition 2. For fixed k, the *join* $J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1}) \subset \mathscr{E}^2$ is the set:

$$J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1}) = I(\mathbf{e}_k^{\omega}, \mathbf{c}_k^{\omega}] \cup L(\mathbf{c}_k^{\omega}, \mathbf{d}_k^{\omega+1})$$

when $\omega = 0, 1, 2, \ldots, 10^{2k} - 1$ and:

$$J_k(\mathbf{e}_k^{10^{2k}}, \mathbf{d}_k^{10^{2k}+1}) = D(\mathbf{e}_k^{10^{2k}}, 1]$$

when $\omega = 10^{2k}$.

We note that for fixed k, the half-closed interval $D(\mathbf{e}_k^{10^{2k}}, 1]$ connects the points:

$$\mathbf{e}_k^{10^{2k}} = \left(\sum_{r=1}^{k} 9 \cdot 10^{-r} + \tfrac{8}{9} 10^{-k}, \sum_{r=1}^{k} 9 \cdot 10^{-r} + \tfrac{8}{9} 10^{-k}\right)$$
$$= \left(1 - \tfrac{1}{9} 10^{-k}, 1 - \tfrac{1}{9} 10^{-k}\right)$$

and (1,1). Thus, $J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1})$ connects the pair of squares $S_k(\mathbf{d}_k^{\omega})$ and $S_k(\mathbf{d}_k^{\omega+1})$ for $\omega = 0, 1, 2, \ldots, 10^{2k} - 1$, whereas $J_k(\mathbf{e}_k^{10^{2k}}, \mathbf{d}_k^{10^{2k}+1})$ connects the last square $S_k(\mathbf{d}_k^{10^{2k}})$ to the point (1,1).

Definition 3. The kth *approximating curve* $\Lambda_k \subset \mathscr{E}^2$ is the set:

$$\Lambda_k = \bigcup_{\omega=1}^{10^{2k}} S_k(\mathbf{d}_k^{\omega}) \cup J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1}).$$

Clearly, each curve Λ_k originates at (0,0) and terminates at (1,1). The anatomy of approximating curves is as follows: For a given k, consider a curve Λ_k and a pair of squares $S_k(\mathbf{d}_k^{\omega})$ and $S_k(\mathbf{d}_k^{\omega+1})$. This curve has 10^{2k} nodal points $\mathbf{d}_k^{\omega}$ that are also nodal points of the approximating curve Λ_{k+1} that has, in addition, $10^{2k}(10^2 - 1)$ new nodal points not lying on Λ_k. Some of the squares $S_{k+1}(\mathbf{d}_{k+1}^{s})$ associated with the new nodal points and their joins $J_{k+1}(\mathbf{e}_{k+1}^{s}, \mathbf{d}_{k+1}^{s+1})$ form a new curve segment connecting the given squares, thereby replacing the existing join $J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1})$. Now, a square $S_k(\mathbf{d}_k^{\omega})$ contains the 9^2 squares $S_{k+1}(\mathbf{d}_{k+1}^{s_1})$, $S_{k+1}(\mathbf{d}_{k+1}^{s_2}), \ldots$ (see the note in the Appendix). Let the squares $S_{k+1}(\mathbf{d}_{k+1}^{s})$ occupying the lower left and upper right corners of $S_k(\mathbf{d}_k^{\omega})$ be $S_{k+1}(\mathbf{d}_{k+1}^{u})$ and $S_{k+1}(\mathbf{d}_{k+1}^{v})$, respectively. Some other squares $S_{k+1}(\mathbf{d}_{k+1}^{s})$ belonging to new nodal points are part of the curve-segment connecting these squares, and according to Theorem 1(b), these nodal points are such that $\mathrm{dist}(\mathbf{d}_k^{\omega}, \mathbf{d}_{k+1}^{s}) < \sqrt{2} \cdot 10^{-k}$.

The path along which the points $\mathbf{d}_1^{\omega}$ are connected is presented in Fig. 4; Fig. 5 presents the path obtained when joining the points $\mathbf{d}_2^{\omega}$, and Fig. 6 presents the details of this path in the interval $[0,0.5] \times [0,0.5]$. These graphs have been obtained for the value $\alpha_1 = 1$ and the truncated value $\alpha_2 = \frac{1}{10} + \frac{1}{10^3} + \frac{1}{10^7}$. The graphing of the actual curves with their joins and squares is beyond the scope of our program. We note in passing that other interesting space-filling curves can be obtained by replacing α with constants:

$$\alpha_p = \frac{p}{10} + \alpha$$

for values $p = 0,1,\ldots,8$.

When writing a code to implement such mappings, great care must be taken to avoid representation problems. The main cause of such problems is the fact that certain rational numbers cannot be represented exactly in base 2 using a finite number of digits. As a consequence, standard implementations of functions such as floor and ceiling may provide wrong results. One way to avoid this type of problem is to modify the algorithm in such a way that all computations are done with integers. However, such modifications are outside the scope of the present paper.

Note: An alternative definition of Λ_k is obtained by mapping in one of several ways each square $S_k(\mathbf{d}_k^{\omega})$ onto its diagonal $D_k(\mathbf{d}_k^{\omega})$ and defining the kth approximating curve to be:

$$\Lambda_k = \bigcup_{\omega=1}^{10^{2k}} D_k(\mathbf{d}_k^{\omega}) \cup J_k(\mathbf{e}_k^{\omega}, \mathbf{d}_k^{\omega+1}).$$

For a similar construction and use of terminology see Sagan (1991).

We now define mappings $\lambda_k : \mathscr{E} \to \mathscr{E}^2$ as follows:

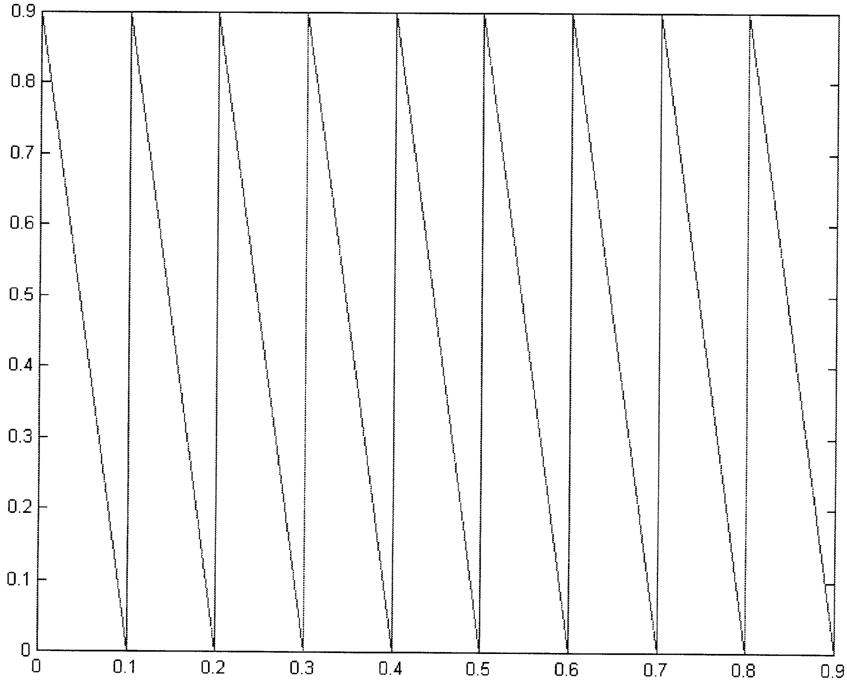

Fig. 4. The path of joining the points $\mathbf{d}_1^\omega$.

Definition 4. For $k = 1,2,3,\ldots,$

$$\lambda_k(z_k^\omega) = \mathbf{d}_k^\omega$$

$$\lambda_k(z_k^\omega + 8b_k) = \mathbf{e}_k^\omega$$

$$\lambda_k : T_k(z_k^\omega) \to S_k(\mathbf{d}_k^\omega)$$

$$\lambda_k : U_k(z_k^\omega) \overset{1-1}{\leftrightarrow} J_k(\mathbf{e}_k^\omega, \mathbf{d}_k^{\omega+1})$$

According to this definition, each mapping λ_k is such that:

$$\lambda_k : [z_k^\omega, z_k^{\omega+1}) \to S_k(\mathbf{d}_k^\omega) \cup J_k(\mathbf{e}_k^\omega, \mathbf{d}_k^{\omega+1})$$

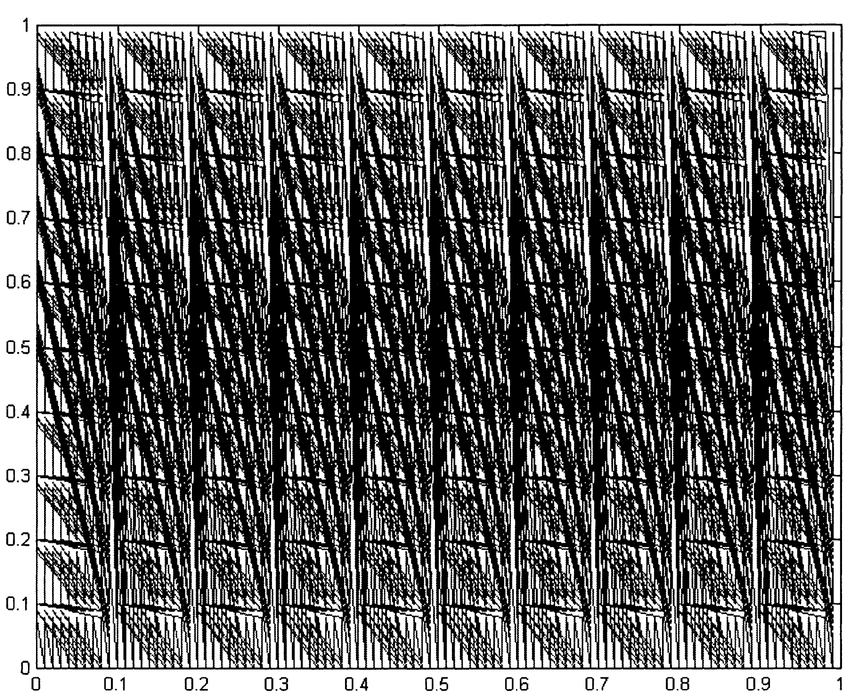

Fig. 5. The path of joining the points $\mathbf{d}_2^\omega$.

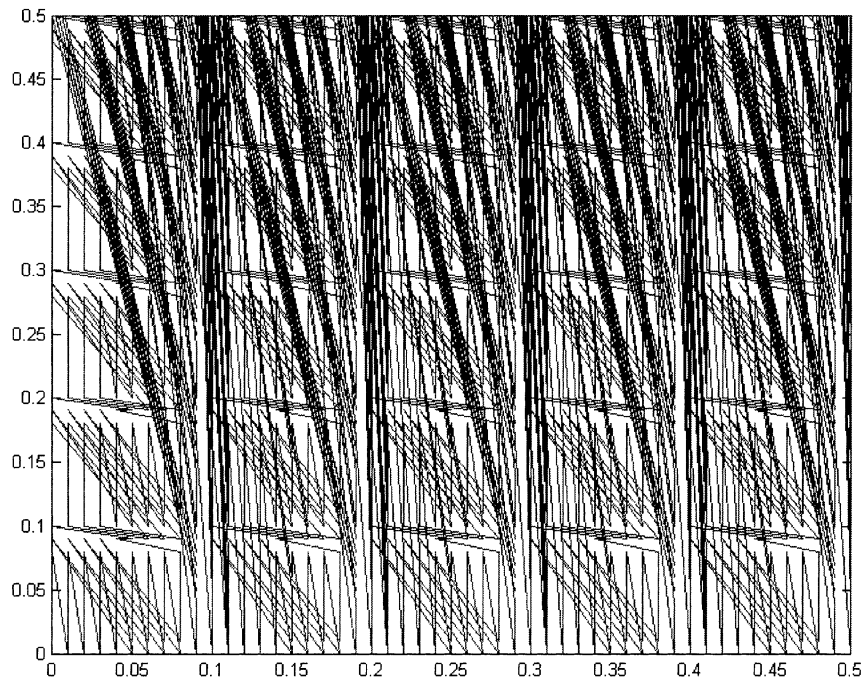

Fig. 6. Detail of the path connecting the points $\mathbf{d}_2^\omega$.

for $\omega = 1,2,3,\ldots$, and this results in mappings $\lambda_k : \mathscr{E} \to \Lambda_k$. We observe that the only square intersecting a given level curve $\xi^{-1}(z_k^\omega)$ is $S_k(\mathbf{d}_k^\omega)$, and that there are squares $S_m(\mathbf{d}_m^t)$ with $m > k$ intersecting $I(\mathbf{e}_k^\omega, \mathbf{c}_k^\omega]$ or $D(\mathbf{e}_k^{10^{2k}}, 1]$. That each join has finite length is already guaranteed by the monotonicity of ψ. Fig. 7 depicts the four possible relative positions of an arbitrary pair of squares $S_k(\mathbf{d}_k^\omega)$ and $S_k(\mathbf{d}_k^{\omega+1})$, and the details of configurations A and D are given in Fig. 8(a) and (b), respectively. Each join is composed of a horizontal line segment and a monotonic decreasing curve, and an examination shows that:

$$|J_k(e_k^\omega, d_k^{\omega+1})| \le \begin{cases} |d_{k,1}^{\omega+1} - d_{k,1}^\omega| + |d_{k,2}^{\omega+1} - d_{k,2}^\omega| + \frac{1}{9}10^{-k} \text{ (position } A) \\ |d_{k,1}^{\omega+1} - d_{k,1}^\omega| + |d_{k,2}^{\omega+1} - d_{k,2}^\omega| + \frac{8}{9}10^{-k} \text{ (position } D) \end{cases}$$

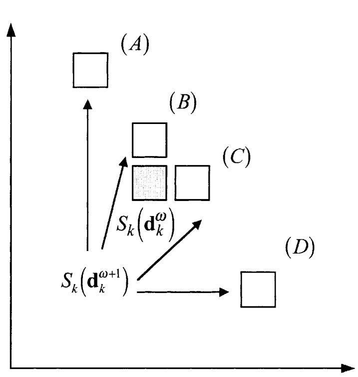

Fig. 7. $S_k(\mathbf{d}_k^\omega)$ and the relative points of $S_k(\mathbf{d}_k^{\omega+1})$.

and:

$$|J(e_k^{10^{2^k}}, d_k^{10^{2^k}+1})| < \tfrac{2}{9}10^{-k}.$$

Clearly, positions B and C in Fig. 7 need no elaboration, and with Theorem 1 we therefore have:

Corollary 1. For fixed k, $\max_r |J_k(\mathbf{e}_k^\omega, \mathbf{d}_k^{\omega+1})| \le C|10^{-\log_2(k-1)} + \frac{8}{9}10^{-k}$.

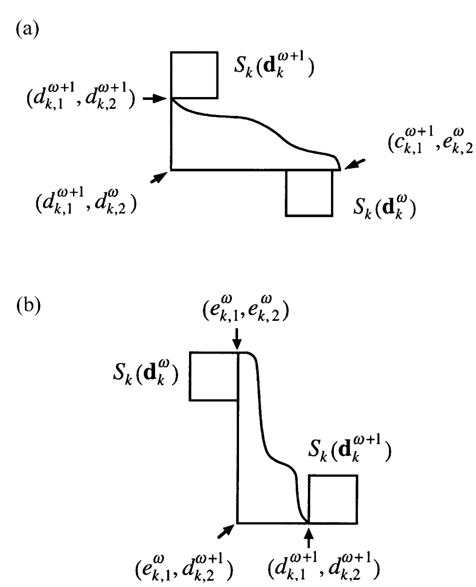

Fig. 8. (a) Detail of configuration A in Fig. 7. (b) Detail of configuration B in Fig. 7.

Theorem 5. Λ is the limit of the sequence of approximating curves Λ_k as $k \to \infty$.

5. Proofs

Proof of Lemma. 1: Let $k \geq 2$ be fixed and let an integer $1 \leq s \leq k-1$ be given. Then:

$$|\psi(d_{2,k}) - (d'_{2,k})| \geq \frac{1}{10^{2^s-1}}$$

according to the Appendix, and we consider the inequality:

$$|z_{s,k} - z'_{s,k}| \geq \frac{1}{\alpha+1}\left|[\psi(d_{1,k}) - \psi(d'_{1,k})] + \frac{\alpha}{10^{2^s-1}}\right|$$
$$= \frac{1}{\alpha+1}\left|A_{1,k} + \frac{\alpha}{10^{2^s-1}}\right|. \tag{14}$$

We seek a lower bound on $|z_{s,k} - z'_{s,k}|$ as a function of s, and toward this end we consider the formula:

$$A_{1,k} = \sum_{r=1}^{s} \frac{(\tilde{\imath}_{1,r} - \tilde{\imath}'_{1,r})}{2^{m_r}10^{2^{r-m_r}-1}} + \sum_{r=s+1}^{k} \frac{(\tilde{\imath}_{1,r} - \tilde{\imath}'_{1,r})}{2^{m_r}10^{2^{r-m_r}-1}}$$
$$= B_{1,s} + B_{s+1,k}$$

obtained from Eq. (A1). Using the lower positive bound:

$$|B_{1,s}| \geq \frac{1}{10^{2^s-1}},$$

the inequality:

$$\left|B_{1,s} + \frac{\alpha}{10^{2^s-1}}\right| \geq \left|\frac{1}{10^{2^s-1}} - \frac{\alpha}{10^{2^s-1}}\right| > \frac{1-\alpha}{10^{2^s-1}}$$

tells us that the right side of Eq. (14) cannot attain a minimum value unless that $B_{1,s} \equiv 0$, implying, in particular, that $|i_{1,r} - i'_{1,r}| = 0$ for all values $1 \leq r \leq s$. Hence, we consider the inequality:

$$|z_{s,k} - z'_{s,k}| \geq \frac{1}{\alpha+1} \min_{B_{s+1,k}} \left|B_{s+1,k} + \frac{\alpha}{10^{2^s-1}}\right|. \tag{15}$$

instead of Eq. (14). In view of the inequalities:

$$\min_{B_{s,2}} |z_{s,2} - z'_{s,2}| \geq \min_{B_{s,3}} |z_{s,3} - z'_{s,3}| \geq \ldots \geq \min_{B_{s,k}} |z_{s,k} - z'_{s,k}|$$

we assume without loss of generality that $k \geq 7$. We now note that the terms of $B_{s+1,k}$ closest to:

$$\frac{\alpha}{10^{2^s-1}}$$

are of the form:

$$\frac{1}{10^{2^s-1}}\left(\frac{1}{2^u} + \ldots + \frac{1}{2^v}\right),$$

and it is a simple exercise to verify that:

$$\left|\alpha - \frac{1}{2^u} - \ldots - \frac{1}{2^v}\right| > \left|\frac{1}{10} + \frac{1}{10^3} - \frac{1}{2^4} - \frac{1}{2^5}\right| = \frac{7}{10^3} \tag{16}$$

and consequently, when $s \leq k-4$ we obtain from an examination of $B_{s+1,k}$ that:

$$|z_{s,k} - z'_{s,k}| > \frac{1}{\alpha+1}\left[\frac{7}{10^{2^s-1}\cdot 10^3} - \frac{1}{10^{2^{s+6}-1}} \times \left(8 + \frac{1}{2} + \ldots + \frac{1}{2^{k-s-6}}\right)\right] > \frac{6}{10^{2^s+2}}. \tag{17}$$

When $k-4 < s \leq k-1$ it is clear that the lower bound in Eq. (16) cannot be attained, but inequality Eq. (17) remains true.

Proof of Theorem 1.

1. By Definition 1, the successor $\mathbf{d}_k^{\omega+1}$ of a given grid point $\mathbf{d}_k^{\omega}$ is determined by that number $z_{k+1}^{\omega} = \xi(\mathbf{d}_k^{\omega+1})$ for which:

   $$|z_k^{\omega+1} - z_k^{\omega}| = \min_r |z_k^r - z_k^{\omega}|, \tag{18}$$

   and therefore dist$(\mathbf{d}_k^{\omega+1}, \mathbf{d}_k^{\omega})$ is determined by the function ξ. Assuming that $\mathbf{d}_k^{\omega}$ is given, we seek a lower bound on s, $1 \leq s \leq k-1$, such that the minimum in Eq. (18) can be attained under the conditions of Lemma 1. The inequality:

   $$|z_k^{\omega+1} - z_k^{\omega}| \leq \frac{1}{2^{k-1}\cdot 10}, \tag{19}$$

   that follows from Lemma A1 leads to the inequality:

   $$\frac{6}{10^{2^s+2}} \leq \frac{1}{2^{k-1}\cdot 10}. \tag{20}$$

   which is trivially true for all values $1 \leq k \leq 9$; for values $k \geq 10$ a direct calculation that it is satisfied when:

   $$s \geqq \log_2(k-1) - \log_2\log{}_2 10. \tag{21}$$

   We now observe from Lemma 1 that $|d_{k,1} - d'_{k,1}| \leq 10^{-s}$ and we conclude that $|d_{k,p} - d'_{k,p}| < 10^{\log_2(k-1) - \log_2\log_2 10}$ for $p = 1,2$ and (a) follows.
2. The inclusion $T_{k_{r+1}}(z_{k_{r+1}}) \subset T_{k_r}(z_{k_r})$ gives the inequalities:

   $$z_{k_r} \leq z_{k_{r+1}} < z_{k_r} + 8b_{k_r} < z_{k_r} + 8\cdot 10^{-2^{k_r}+1}$$

   and according to construction, there exists a number $z'_{k_{r+1}} = z_{k_r}$, and hence:

   $$|z_{k_{r+1}} - z_{k_r}| = |z_{k_{r+1}} - z'_{k_{r+1}}| < 8\cdot 10^{s_{r+1}}. \tag{22}$$

   Furthermore, there are rational numbers $d_{k_r,1} = d'_{k_{r+1},1}$

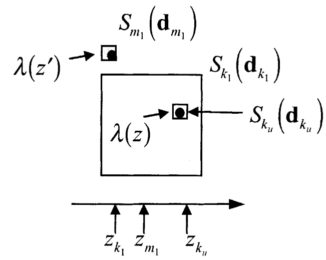

Fig. 9. Showing that if $\lambda(z) \neq \lambda(z')$ then $z \neq z'$.

and $d_{k_r,2} = d'_{k_{r+1},2}$, and with the supposition that:

$$|d_{k_{r+1},2} - d_{k_r,2}| \geq \frac{1}{10^s},$$

this enables us to emulate the argument employed to establish Lemma 1. This leads us to the inequality:

$$|z_{k_{r+1}} - z_{k_r}| \geq \frac{6}{10^{2^s+2}}$$

which, in turn, gives in place of Inequality (20) the inequality:

$$\frac{6}{10^{2^2+2}} < \frac{8}{10^{2^{k_r}-1}}.$$

An inspection shows that this inequality is satisfied when $s \geq k_r$, and (b) follows.

3. The inclusion $U_{k+1}(z_{k+1}) \subset U_k(z_k)$, gives the inequalities $z_k^\omega + 8b_k < z_{k+1} < z_{k+1} + 8b_{k+1} < z_k^{\omega+1}$, and we have the inequality:

$$|z_k^{\omega+1} - z_k^\omega| \leq \frac{1}{2^{k-1} \cdot 10}.$$

Also from the construction we know that there is a point $z'_{k+1} = z_k^{\omega+1}$, so that:

$$|z'_{k+1} - z_{k+1}| < |z_k^{\omega+1} - z_k^\omega| \leq \frac{1}{2^{k-1} \cdot 10}.$$

We apply Lemma 1 to the difference $|z'_{k+1} - z_{k+1}|$ and have:

$$|z'_{k+1} - z_{k+1}| \geq \frac{6}{10^{2^{s+1}+2}}$$

and instead of inequality (Inequality 20) we arrive at:

$$\frac{6}{10^{2^{s+1}+2}} \leq \frac{1}{2^k \cdot 10}.$$

Solving this for s verifies the assertion made.

Proof of Theorem 2.

1. Since the condition of the theorem entails a nested sequence of intervals, Theorem 1(b) applies. Hence, given any number $\epsilon > 0$, there is an integer k_ϵ such that:

$$\mathrm{dist}(\mathbf{d}_{k_u}, \mathbf{d}_{k_v}) < \sqrt{2}\sum_{r=u}^{v} 10^{-k_r} \leq \sqrt{2}\sum_{r=u}^{v} 10^{-r}$$

$$\leq \sqrt{2}\sum_{r=u}^{\infty} 10^{-r} = \frac{\sqrt{2}}{9} 10^{-u+1} < \epsilon,$$

for all $k_v > k_u > k_\epsilon$, thereby showing that the sequence $\mathbf{d}_{k_1}, \mathbf{d}_{k_2}, \mathbf{d}_{k_3}\ldots$ is a Cauchy sequence.

2. Theorem 1(c) applies here, so that:

$$\mathrm{dist}(\mathbf{d}_{k+u}, \mathbf{d}_{k+v}) < C\sum_{r=u}^{v} 10^{-\log_2(k+r-1)+1}$$

$$\leq C\sum_{r=u}^{v} (k+r-1)^{-3}.$$

for a sufficiently large value of u, and hence the sequence $\mathbf{d}_{k_1}, \mathbf{d}_{k_2}, \mathbf{d}_{k_3}\ldots$ is also in this case a Cauchy sequence.

Proof of Theorem 3.

1. Let:

$$z = \bigcap_{r=1}^{\infty} T_{k_r}(z_{k_r})$$

and

$$z' = \bigcap_{s=1}^{\infty} T_{m_s}(z_{m_s})$$

have corresponding image points:

$$\lambda(z) = \bigcap_{r=1}^{\infty} S_{k_r}(\mathbf{d}_{k_r})$$

and:

$$\lambda(z') = \bigcap_{s=1}^{\infty} S_{m_s}(\mathbf{d}_{m_s}),$$

and suppose that $\lambda(z) \neq \lambda(z')$. Then we can select integers k_1 and m_1 so that $S_{k_1}(\mathbf{d}_{k_1}) \cap S_{m_1}(\mathbf{d}_{m_1}) = \emptyset$, and we suppose that $z_{k_1} < z_{m_1}$. If $k_1 < m_1$, then we can select an integer k_u such that $k_u \geq m_1$ and:

$$z = \bigcap_{r=u}^{\infty} T_{k_r}(z_{k_r})$$

(Fig. 9). There is a point z'_{k_u} among the points of the set $\{z_{k_l}\}$ such that $z'_{k_u} = z_{m_1}$, and from Sprecher (1996), Lemma 2, we have the inequalities:

$$z_{m_1} - z_{k_u} \geq z'_{k_u} - z_{k_u} \geq \frac{1}{10^{2(2^{k_u}-1)}}.$$

Since $z \in T_{k_u}(z_{k_u})$ we see from Eq. (6) that $z_{k_u} \leq z \leq z_{k_u} + 8b_{k_u}$, and a simple calculation shows that:

$$8b_{k_u} < \frac{1}{10^{2(2^{k_u}-1)}}.$$

Since $z_{m_1} \le z'$ we have the inequalities:

$$z \le z_{k_u} + 8b_{k_u} < z_{k_u} + 8b_{k_u} < \frac{1}{10^{2(2^{k_u}-1)}} \le z_{m_1} \le z',$$

as was to be shown.

Conversely, suppose that $z \ne z'$. Then there is an integer $u > 0$ such that $T_{k_u}(z_{k_u}) \cap T_{m_u}(z_{m_u}) = \emptyset$, implying that $S_{k_u}(\mathbf{d}_{k_u}) \cap S_{m_u}(\mathbf{d}_{m_u}) = \emptyset$ and so:

$$\lambda(z) = \bigcap_{r=u}^{\infty} S_{k_r}(\mathbf{d}_{k_r}) \ne \bigcap_{s=u}^{\infty} S_{m_s}(\mathbf{d}_{m_s}) = \lambda(z').$$

2. Let $z, z' \in \mathscr{K}_2$ and suppose $\lambda(z) = \lim_{r\to\infty} \mathbf{d}_{k_r}$, $\lambda(z') = \lim_{s\to\infty} \mathbf{d}_{h_s}$. Suppose further that $\lambda(z) \ne \lambda(z')$, and let $\mathrm{dist}(\lambda(z), \lambda(z')) = \eta$. We deduce from Corollary 1(b) that for any integer u,

$$\mathrm{dist}(\lambda(z), \mathbf{d}_{k_u}) + \mathrm{dist}(\lambda(z'), \mathbf{d}_{h_u}) < \sqrt{2}\sum_{r=u}^{\infty}(10^{-k_r} + 10^{-h_r}),$$

and we select u sufficiently large, so that:

$$\begin{aligned}&\mathrm{dist}(\lambda(z), \mathbf{d}_{k_u}) + \mathrm{dist}(\lambda(z'), \mathbf{d}_{h_u})\\ &\quad + \mathrm{diameter}(S_{k_u}(\mathbf{d}_{k_u})) + \mathrm{diameter}(S_{h_u}(\mathbf{d}_{h_u}))\\ &< \sqrt{2}\sum_{r=u}^{\infty}(10^{-k_r} + 10^{-h_r}) + \sqrt{2}\frac{8}{9}(10^{-k_u} + 10^{-h_u}) < \eta.\end{aligned}$$

Finally, when $z, z' \in \mathscr{K}_3$ we simply replace respectively k_u and h_u with $k+u$ and $h+u$ in these inequalities and use Corollary 1(c). In either case we can guarantee that $S_{k_u}(\mathbf{d}_{k_u}) \cap S_{m_u}(\mathbf{d}_{m_u}) = \emptyset$, and this enables us to repeat the first argument in (1) to derive that $z \ne z'$.

Proof of Theorem 4. For a given point $a \in \mathscr{E}$ we define the δ-neighborhood $N_\delta(a) \cap \mathscr{E} = \{z \in \mathscr{E} : |z - \mathbf{a}| < \delta\}$ and the corresponding ϵ-neighborhood $N_\epsilon(\lambda)a)) \cap \mathscr{E}^2 = \{\mathbf{x} \in \mathscr{E}^2 : \mathrm{dist}(\mathbf{x}, \lambda(a)) < \epsilon\}$. We establish the pointwise continuity of the mapping $\lambda : \mathscr{E} \to \mathscr{E}^2$ by showing that for any point $a \in \mathscr{E}$ and number $\epsilon > 0$ there exists a number $\delta = \delta(\epsilon, a)$ such that if $z \in N_\delta(a) \cap \mathscr{E}$ then $\lambda(z) \in N_\epsilon(\lambda(a)) \cap \mathscr{E}^2$.

Consider an arbitrary point $z \in \mathscr{E}$. Then for any integer k, $\lambda(a)$ can be obtained as the limit of a (convergent) infinite series whose largest term does not exceed 10^{-c_k}, where c_k has one of the values a_k, b_k, or k. Given points $a, z \in T_k(z_k) \cup U_k(z_k)$, then $\mathrm{dist}(\lambda(a), \lambda(z))$ can be estimated with the triangle inequality $\mathrm{dist}(\lambda(a), \lambda(z)) \le \mathrm{dist}(\lambda(a), \mathbf{d}_{k_s}) + \mathrm{dist}(\mathbf{d}_{k_s}, \lambda(z))$ and such infinite series derived from Lemma 2, and Theorems 2 and 4. Therefore, given a number $\epsilon > 0$, we can find an integer k so large that if $a, z \in N_\delta(a) \subset [T_k(z_k) \cup U_k(z_k)]$ then $\lambda(z) \in N_\epsilon(\lambda(a))$. The proof is conveniently divided into the following cases:

1. $a \in \mathscr{K}_1 \cup \mathscr{K}_2$ and $a = \bigcap_{r=1}^{\infty} \mathring{T}_{k_r}(z_{k_r})$, where $\mathring{T}_{k_r}$ designates the interior of T_{k_r};
2. $a \in \mathscr{K}_3$.
3. α is an endpoint, $a = z_k$ or $a = z_k + 8b_k$, of some interval T_k.

We omit the laborious and routine computations as these are similar to the arguments employed in preceding proofs.

To complete the proof, it remains only to show that Λ has positive Lebesgue measure and, in fact, that its Lebesgue measure is 1. To this end, consider the set:

$$\Gamma = \{\mathbf{x} \in \mathscr{E}^2 : \mathbf{x} = \bigcap_{r=1}^{\infty} S_{k_r}(\mathbf{d}_{k_r})\}.$$

That this set has Lebesgue measure 1, follows from the fact that $\Lambda \supset \Gamma$. A proof of this can be found in Sprecher (1966). It follows directly from the fact that ψ is monotonic increasing and singular (i.e., ψ has a vanishing derivative almost everywhere).

Proof of Theorem 5. Let a number $\epsilon > 0$ be given, then for each point $a \in \mathscr{E}$ there is an integer k_ϵ and a neighborhood $N_\epsilon(\lambda(a))$ such that $\Lambda_k \subset N_\epsilon(\lambda(a)) \cap \mathscr{E}$ for $k > k_\epsilon$. The system of neighborhoods $N_\epsilon(\lambda(a))$ forms an infinite open covering of $\mathscr{E}$, and since $\mathscr{E}$ is compact, there is a finite sub-covering of neighborhoods $N_\epsilon(\lambda(a_1))$, $N_\epsilon(\lambda(a_2)), \ldots, N_\epsilon(\lambda(a_s))$ corresponding to integers $k_1, k_2, \ldots, k_s$ such that $\Lambda_k \subset N_\epsilon(\lambda(a_r))$ for $k \ge k_r$, $r = 1,2,\ldots,s$. Hence, $\Lambda_k \subset \bigcup_{r=1}^{s} N_\epsilon(\lambda(a_r))$ for $k \ge k_s$, and we conclude that the approximating curves Λ_k lie within an ϵ-neighborhood of Λ.

Appendix A

The following is an abbreviated decimal version of the construction of the function ψ in Sprecher (1996) for the case $n = 2$.

Definition A1. Let $\langle i_1 \rangle = [i_1] = 0$ and for $r > 1$ let:

$$\langle i_r \rangle = \begin{cases} 0 & \text{when} \quad i_r = 0, 1, \ldots, 8 \\ 1 & \text{when} \quad i_r = 9 \end{cases}$$

$$[i_r] = \begin{cases} 0 & \text{when} \quad i_r = 0, 1, \ldots, 7 \\ 1 & \text{when} \quad i_r = 8, 9 \end{cases}$$

$$\tilde{i}_r = i_r - 8\langle i_r \rangle$$

$m_1 = 0$

$$m_r = \langle i_r \rangle \left(+ \sum_{s=1}^{r-1} [i_s] \times [i_{s+1}] \times \ldots \times [i_{r-1}] \right)$$

for $r = 2,\ldots,k$.

The function ψ is defined for $k = 1,2,3,\ldots$ at the points:

$$d_k = \sum_{r=1}^{k} \frac{i_r}{10^r},$$

$i_r = 0,1,\ldots,9$ through the formula:

$$\psi(d_k) = \sum_{r=1}^{k} \frac{\tilde{i}_r}{2^{m_r} 10^{2^{r-m_r}-1}}, \tag{A1}$$

$\psi(1) = 1.$

So defined at the dense set of rational points of $\mathscr{E}$, ψ is monotonic increasing with a continuous extension $\psi : \mathscr{E} \xrightarrow{1-1} \mathscr{E}$, also designated ψ. Immediate consequences of Eq. (A1) are the following two lemmas:

Lemma A1. If:

$$|d_k - d'_k| \le \frac{1}{10^k}$$

then:

$$|\psi(d_k) - \psi(d'_k)| \le \frac{1}{2^{k-1}10}.$$

Lemma A2. If:

$$|d_k - d'_k| \ge \frac{1}{10^s}$$

for some integer $s \le k$ then:

$$|\psi(d_k) - \psi(d'_k)| \ge \frac{1}{10^{s^s-1}}.$$

The last lemma follows from the observation that 10^{-2^s+1} is the smallest possible of the increments $2^{-\nu}10^{-2^{s-\nu}+1}$ of ψ at step s, where $0 \le \nu \le s-1$.

The constants used in Eqs. (2) and (3) are defined as follows:

$$\alpha = \sum_{r=1}^{\infty} 10^{-2^r+1}$$

$$\alpha_1 = \frac{1}{\alpha+1},$$

$$\alpha_2 = \frac{\alpha}{\alpha+1},$$

Note: For readers who wish to work with Eq. (A1) we mention that for fixed k, the values $\psi(d_k)$ increase in uniform increments 10^{-2^r+1} as the index i_r increases through the range 0,1,…,8. Otherwise the increments of $\psi(d_k)$ are of the form $2^{-\nu}10^{-2^{s-\nu}+1}$, with the values of s and ν being determined by strings of digits 8 and 9 immediately proceeding i_r in the decimal representation $d_k = 0.i_1i_2\ldots i_{r-1}i_r\ldots i_k$. It is the purpose of the expression m_r to count the lengths of such strings. Referring to Eq. (4) in Section 1, an examination of Eq. (A1) shows that $S_m(\mathbf{d}_m) \subset S_k(\mathbf{d}_k)$ if and only if the following three conditions hold for $p = 1,2$:

$$d_{m,p} = d_{k,p} + \sum_{r=k+1}^{m} i_r 10^{-r}, \tag{i}$$

$$i_{k+1,p} \ne 9, \tag{ii}$$

$$i_{k+1,p}, i_{k+2,p}, \ldots \ne 8,\ldots,8,9,\ldots, \tag{iii}$$

meaning that the string of digits starting with $i_{k+1,p}$ cannot consist of digits 8 followed by digits 9 only.

Further elaboration can be found in Sprecher (1996, 1997).

References

Hecht-Nielsen, R. (1987). Kolmogorov's mapping neural network existence theorem. *Proceedings of the IEEE International Conference on Neural Networks* (pp. 11–13). New York: IEEE Press.

Kolmogorov, A.N. (1957). On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Dokl. Akad. Nauk SSSR*, *114*, 953–956; *Trans. Am. Math. Soc*. *2*(28), 55–59.

Sagan, H. (1991). *Space-filling curves*. New York: Universitext, Springer-Verlag.

Sprecher, D. A. (1966). On the structure of representations of continuous functions of several variables as finite sums of continuous functions of one variable. *Proc. Am. Math. Soc.*, *17* (1), 98–105.

Sprecher, D. A. (1996). A numerical implementation of Kolmogorov's superpositions. *Neural Networks*, *9* (5), 765–772.

Sprecher, D. A. (1997). A numerical implementation of Kolmogorov's superpositions II. *Neural Networks*, *10* (3), 447–457.