The dimension of affine-invariant fractals

1988 J. Phys. A: Math. Gen. 21 L121

(http://iopscience.iop.org/0305-4470/21/3/002)

J. Phys. A: Math. Gen. **21** (1988) L121-125. Printed in the UK

LETTER TO THE EDITOR

The dimension of affine-invariant fractals

K J Falconer and D T Marsh

School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK

Received 24 September 1987

Abstract. A generic formula for the Hausdorff and box counting dimensions of self-affine fractals is presented. Extensions to the non-linear situation and to repellers in discrete dynamical systems are discussed.

A wide variety of natural, physical and mathematical objects may be conveniently regarded as fractals. Some of these have a self-similar structure, in the sense that locally they appear to be a scaled-down copy of the whole object. These include the von Koch 'snowflake' curve and Julia sets in dynamical systems. Other fractals are self-similar in a statistical sense: the spatial probability distribution of the object locally is the same as that of the whole object, but scaled by an appropriate factor. This class includes Brownian motion, percolation models, the Ising model of magnetism and some models of turbulence (Mandelbrot 1982).

Many fractals, however, are better regarded as self-affine, either strictly or in a statistical sense. In this case locally the fractals are images of the whole object under certain distorting or non-isotropic transformations. Examples include mountain skylines, flames, ferns, fractal viscous fingering and the diffusion-limited aggregation model (see Stanley and Ostrowsky 1986). Strictly self-affine fractals have been used in image reconstruction (Barnsley *et al* 1987) and fractal interpolation (Barnsley 1986).

It is of interest to calculate or estimate the basic parameters of a fractal, such as the Hausdorff dimension. Here we give a mathematical treatment of the strictly self-affine case, i.e. for fractals that are invariant under a collection of contracting affine transformations.

Suppose that the transformations $S_1, \ldots, S_k$ on $\mathbb{R}^n$ are contractions, i.e. there exist constants c_i with $0 < c_i < 1$ such that $|S_i(x) - S_i(y)| \leq c_i|x-y|$ for all $x, y \in \mathbb{R}^n$. By a well known result (see Hutchinson 1981, Barnsley and Denko 1985), there is a unique non-empty compact set F which is *invariant* for the S_i, i.e. such that

$$F = \bigcup_{i=1}^{k} S_i(F). \tag{1}$$

For instance, the familiar 'middle-third' Cantor set is invariant for the contractions $S_1(x) = x/3$ and $S_2(x) = (x+2)/3$.

Let J_r denote the collection of all r-digit sequences formed using the integers $1, \ldots, k$; let $J = \bigcup_{r=1}^{\infty} J_r$ be the set of all such finite sequences and let J_∞ be the corresponding set of infinite sequences. For each $\boldsymbol{i} = (i_1, \ldots, i_r) \in J_r$, we write $S_{\boldsymbol{i}}$ for the composition of the contractions $S_{i_1} \circ S_{i_2} \circ \ldots \circ S_{i_r}$.

If E is any non-empty, bounded set such that $S_i(E)\subset E$ for $1\leq i\leq k$, then the invariant set is given by

$$F=\bigcap_{r=1}^{\infty}\bigcup_{i\in J_r} S_i(E). \tag{2}$$

Thus, taking S_1 and S_2 as before, and E as the unit interval, this intersection gives the 'usual' construction of the middle-third Cantor set. Alternatively, the single points

$$x_i=\bigcap_{r=1}^{\infty}(S_{i_1}\circ\dots\circ S_{i_r})(E) \tag{3}$$

where $\boldsymbol{i}=(i_1, i_2, \dots)$, are defined independently of E, and

$$F=\bigcup_{i\in J_\infty} x_i. \tag{4}$$

Typically, the invariant sets F are fractals and it is of interest to calculate the dimension of F (see Mandelbrot 1982, Falconer 1985). If the S_i are *similarities* (i.e. $|S_i(x)-S_i(y)|=c_i|x-y|$ for $1\leq i\leq k$) then the Hausdorff dimension, $\dim_h$, and box counting dimension, $\dim_b$, are often both equal to the unique value of s satisfying

$$\sum_{i=1}^{k} c_i^s=1. \tag{5}$$

In particular this is true if the components $S_i(F)$ in (1) are 'essentially disjoint' (see Hutchinson 1981, Moran 1946). Thus the Cantor set has dimension $\log 2/\log 3$ ($c_1=c_2=\frac{1}{3}$) and the dimensions of the von Koch curve and Menger sponge may be obtained similarly. Equation (5) also gives the dimension in 'almost all' cases where the $S_i(F)$ overlap (Falconer 1987).

In this letter we discuss the situation where the S_i are affinities. An *affinity*, or *affine transformation*, is a composition of a linear mapping (always contracting in our case) and a translation. Thus $S_i(x)=T_i(x)+a_i$ where $a_i\in\mathbb{R}^n$ and T_i is a non-singular linear mapping on $\mathbb{R}^n$. It is convenient to regard the T_i as fixed and let $F(\boldsymbol{a})$ denote the invariant set for the mappings $S_i=T_i+a_i$, where $\boldsymbol{a}=(a_1,\dots,a_k)$, so that

$$F(\boldsymbol{a})=\bigcup_{i=1}^{k}(T_i(F(\boldsymbol{a}))+a_i). \tag{6}$$

Bedford (1984) and McMullen (1984) have calculated $\dim F(\boldsymbol{a})$ in some special cases with $T_1=\dots=T_k$ and differing a_i. Unfortunately, their results depend on the a_i in a rather unstable way. Some general aspects of self-affine fractals are discussed in Mandelbrot (1985, 1986). Here we present a formula which gives $\dim F(\boldsymbol{a})$ for almost all $\boldsymbol{a}\in\mathbb{R}^{nk}$, in the sense that the exceptional parameters have zero nk-dimensional volume.

The *singular values* $\alpha_1\geq\alpha_2\geq\dots\geq\alpha_n$ of a linear mapping T may be thought of in two ways: they are the lengths of the principal semi-axes of the ellipsoid $T(B)$ where B is the unit ball in $\mathbb{R}^n$, or they are the positive square roots of the eigenvalues of T^*T (with T^* the adjoint of T). For $0\leq s\leq n$ define the *singular value function*

$$\phi^s(T)=\alpha_1\alpha_2\dots\alpha_{m-1}\alpha_m^{s-m+1} \tag{7}$$

where m is the integer such that $m-1<s\leq m$. If T is contracting and non-singular, then $\phi^s(T)$ is continuous and strictly decreasing in s, and for fixed s is easily seen to be submultiplicative, i.e. $\phi^s(TU)\leq\phi^s(T)\phi^s(U)$. With T_i as above, and writing $T_{\boldsymbol{i}}=T_{i_1}\circ\dots\circ T_{i_r}$, it follows that the rth level sums $\Sigma_r^s=\Sigma_{\boldsymbol{i}\in J_r}\phi^s(T_{\boldsymbol{i}})$ are also submultiplicative

($\Sigma_{r+q}^s \leq \Sigma_r^s \Sigma_q^s$), so that $(\Sigma_r^s)^{1/r}$ converges, for each s, to a value that is strictly decreasing in s. Assume that $\lim_{r\to\infty}(\Sigma_r^n)^{1/r} \leq 1$. Then there is a unique positive s, which we denote by $d(T_1, \ldots, T_k)$, such that

$$\lim_{r\to\infty}\left(\sum_{i\in J_r} \phi^s(T_i)\right)^{1/r} = 1. \tag{8}$$

Equivalently, using the asymptotic behaviour as $r\to\infty$ of Σ_r^s,

$$d(T_1, \ldots, T_k) = \inf\left\{s: \sum_{i\in J} \phi^s(T_i) \text{ is convergent}\right\}. \tag{9}$$

Our principal claim is that, for almost all $\boldsymbol{a} \in \mathbb{R}^{nk}$,

$$\dim_h(F(\boldsymbol{a})) = \dim_b(F(\boldsymbol{a})) = d(T_1, \ldots, T_k) \tag{10}$$

where $F(\boldsymbol{a})$ is given by (6).

We outline the derivation of (10); full technical details may be found in Falconer (1988). A straightforward covering argument shows that $\dim_h F(\boldsymbol{a}) \leq d(T_1, \ldots, T_k)$ for *all* $\boldsymbol{a}$: choose a large ball B with $S_i(B) = T_i(B) + a_i \subset B$ for all i. Then for each r, the set $F(\boldsymbol{a})$ is covered by the collection of ellipsoids $\bigcup_{i\in J_r} S_i(B)$. But $S_i(B)$ is contained in a rectangular parallelepiped of sides $2|B|\alpha_1, \ldots, 2|B|\alpha_n$ where $\alpha_1, \ldots, \alpha_n$ are the singular values of T_i. If m is the least integer greater than or equal to s, we may divide such a parallelepiped into at most

$$(4|B|)^{n-m+1}(4|B|\alpha_1/\alpha_m)(4|B|\alpha_2/\alpha_m)\ldots(4|B|\alpha_{m-1}/\alpha_m) \leq (4|B|^n)\alpha_1\ldots\alpha_{m-1}\alpha_m^{1-m} \tag{11}$$

cubes of side α_m, and these coverings give the upper estimates for $\dim_h F(\boldsymbol{a})$. A modification of this argument, selecting cubes of 'roughly the same size', shows that $\dim_b F(\boldsymbol{a}) \leq d(T_1, \ldots, T_k)$.

The opposite inequality is obtained using the potential-theoretic characterisation of Hausdorff dimension (see Falconer 1985): if F supports a mass distribution ν with $0 < \nu(F) < \infty$ such that

$$\int\int |x-y|^{-s}\, \mathrm{d}\nu(x)\, \mathrm{d}\nu(y) < \infty \tag{12}$$

then $\dim F \geq s$. If $s < d(T_1, \ldots, T_k)$, so that $\Sigma_{i\in J_r} \phi^s(T_i) \to \infty$ as $r\to\infty$, it is possible to construct a mass distribution μ on J_∞ such that $\mu(B_i) \leq \phi^s(T_i)$ for $\boldsymbol{i} = (i_1, \ldots, i_r) \in J_r$ where B_i is the 'cylinder' $\{(i_1, \ldots, i_r, j_{r+1}, j_{r+2}, \ldots): 1 \leq j_{r+1}, j_{r+2}, \ldots \leq k\}$ in J_∞. With x_i as in (3), but now dependent on the parameter $\boldsymbol{a}$, the mapping $\boldsymbol{i} \mapsto x_i$ transfers μ to a mass distribution on $F(\boldsymbol{a})$ for each such $\boldsymbol{a}$. A certain amount of calculation, valid at least on the assumption that $\|T_i\| < \frac{1}{3}$ for $1 \leq i \leq k$, gives

$$\left\langle \int_{J_\infty}\int_{J_\infty} |x_i - x_j|^{-s}\, \mathrm{d}\mu(\boldsymbol{i})\, \mathrm{d}\mu(\boldsymbol{j}) \right\rangle < \infty \tag{13}$$

where $\langle\ \rangle$ denotes the average as $\boldsymbol{a}$ ranges over a region in parameter space. The crucial factor here is that, if $\boldsymbol{i} \neq \boldsymbol{j}$, then the vectors $x_i - x_j$ are distributed in a non-singular manner as $\boldsymbol{a}$ varies. Hence for almost all $\boldsymbol{a} \in \mathbb{R}^{nk}$ the mass distribution μ transferred to $F(\boldsymbol{a})$ has finite energy, so that $\dim_h \geq s$.

There are two situations where $\dim_h F(\boldsymbol{a}) < d(T_1, \ldots, T_k)$ (though, by the above, they only occur for exceptional values of $\boldsymbol{a}$). First, if the components $S_i(F)$ making up F in (1) overlap substantially, a reduction in dimension may occur. Second, even if the components $S_i(F)$ with $i \in J_r$ are disjoint for each r, they can align in such a way that unusually efficient coverings are possible in the definitions of dimension. (This second case cannot occur if the S_i are similarities, whence $d(T_1, \ldots, T_k)$ is given by (5).) A consequence of the second situation is that $\dim_H F(\boldsymbol{a})$ need not vary continuously with $\boldsymbol{a}$ even if no overlapping occurs.

In figure 1, we illustrate the invariant set for triples of mappings $S_i = T_i + a_i$ $(i = 1, 2, 3)$ with the same T_i in each case, but with different values of a_i. The value $d(T_1, T_2, T_3)$, which gives the dimension in all three cases (assuming that we have not been extremely unlucky in the choice of the a_i) was computed as being 1.43 ± 0.01 using (8).

It is natural to extend the generic formulae for dimension to the situation where the S_i are non-linear contractions. The special case where the S_i are conformal (angle-preserving), generalising the self-similar rather than the self-affine case, was analysed by Ruelle (1982, 1983).

Assume that the S_i are of differentiability class C^2 on a region $E \subset \mathbb{R}^n$, with $\|DS_i(x)\| \leqslant c_i < 1$ for $x \in E$, where the linear mapping $DS_i(x)$ is the derivative of S_i

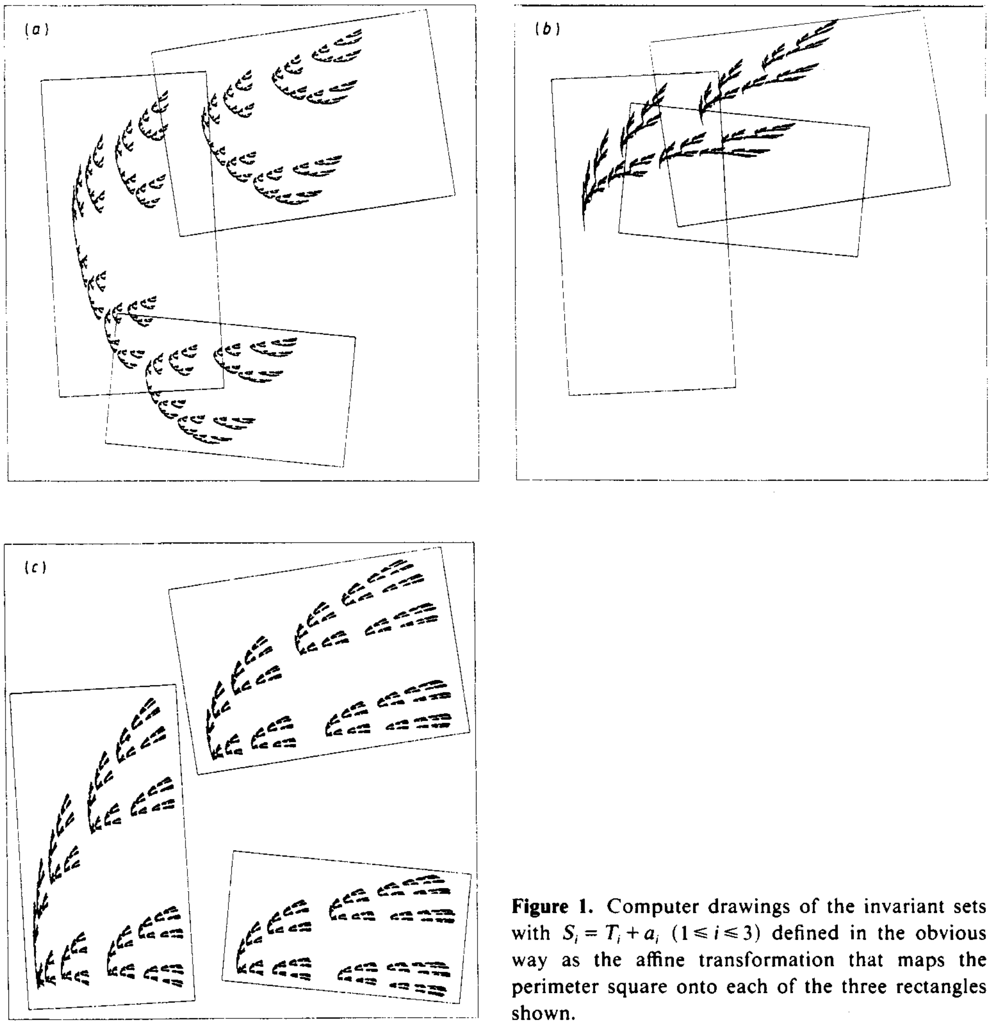

Figure 1. Computer drawings of the invariant sets with $S_i = T_i + a_i$ $(1 \leqslant i \leqslant 3)$ defined in the obvious way as the affine transformation that maps the perimeter square onto each of the three rectangles shown.

evaluated at x. The counterparts of (8) and (9) are

$$d(S_1, \ldots, S_k) = \left\{ s: \lim_{r \to \infty} \left(\sum_{i \in J_r} \phi^s(\mathrm{D}S_i(x)) \right)^{1/r} = 1 \right\} \tag{14}$$

$$= \inf \left\{ s: \sum_{i \in J_r} \phi^s(\mathrm{D}S_i(x)) < \infty \right\} \tag{15}$$

where $S_i = S_{i_1} \circ \ldots \circ S_{i_r}$. Here, x is chosen to be any point of E; because of the contractive properties of the derivative, the particular point chosen does not affect the convergence rates that define $d(S_1, \ldots, S_k)$. As in the linear case, (14) and (15) will not always give the Hausdorff dimension of the invariant set F associated with the S_i, although always providing an upper bound. However, if the S_i are dependent on a sufficiently general set of parameters (so that for each pair $\boldsymbol{i}, \boldsymbol{j} \in J_\infty$ the points $x_i - x_j$ are distributed in a non-singular manner as the parameters are varied) then we would expect that $\dim_h F = \dim_b F = d(S_1, \ldots, S_k)$ at almost all points in the parameter space.

We may extend this further to obtain a generic formula for the dimension of repellers of certain discrete dynamical systems. Suppose that $f: \mathbb{R}^n \to \mathbb{R}^n$ is a C^2 mapping such that there are open regions E and $E_1, \ldots, E_k$ such that the restriction of f to each E_i is a strictly expansive bijection onto E (for example, in figure 1(c), think of a function which maps each rectangle onto the square). This situation arises naturally as a Markov partition for an expansive function (see Ruelle 1983). Taking the restriction of f^{-1} to E_i as the contraction S_i, we see that the invariant set for the S_i is a repeller F for f. Moreover, in terms of (14) and (15), we may replace x by the fixed points of $S_{i_1,\ldots,i_r}$ corresponding in a unique way to a fixed point of the rth iterate $f^{(r)}$; again the critical value of s is unaffected because of the geometric convergence rates. Hence under these circumstances, (15) becomes

$$\inf \left\{ s: \sum_{r=1}^{\infty} \sum_{x \in \mathrm{fix}(f^{(r)})} \phi^s((\mathrm{D}f^{(r)}(x))^{-1}) < \infty \right\}. \tag{16}$$

This number is always at least $\dim_h F$ and may be expected to give the actual value in the general situation.

DTM thanks the SERC for financial support.

References

Barnsley M F 1986 *Constr. Approx.* **2** 303–29

Barnsley M F and Denko S 1985 *Proc. R. Soc.* A **399** 243–75

Barnsley M F, Reuter L H and Sloan A D 1987 to appear

Bedford T 1984 *PhD Thesis* University of Warwick

Falconer K J 1985 *The Geometry of Fractal Sets* (Cambridge: Cambridge University Press)

—— 1987 *J. Stat. Phys.* **47** 123–32

—— 1988 *Math. Proc. Camb. Phil. Soc.* to appear

Hutchinson J E 1981 *Indiana Univ. Math. J.* **30** 713–47

Mandelbrot B B 1982 *The Fractal Geometry of Nature* (San Francisco: Freeman)

—— 1985 *Phys. Scr.* **32** 257–60

—— 1986 *Fractals in Physics* ed L Pietronero and L Tosatti (Amsterdam: Elsevier) pp 3–28

McMullen C 1984 *Nagoya Math. J.* **96** 1–9

Moran P A P 1946 *Proc. Camb. Phil. Soc.* **42** 15–23

Ruelle D 1982 *Ergodic Theor. Dyn. Syst.* **2** 99–107

—— 1983 *Progress in Physics* vol 7 (Basel: Birkhauser) pp 351–77

Stanley H E and Ostrowsky N (ed) 1986 *On Growth and Form* (Dordrecht: Martinus Nijhoff)