Mandelbrot sets for pairs of affine transformations in the plane

1986 J. Phys. A: Math. Gen. 19 1985

(http://iopscience.iop.org/0305-4470/19/11/009)

J. Phys. A: Math. Gen. **19** (1986) 1985-2001. Printed in Great Britain

Mandelbrot sets for pairs of affine transformations in the plane

Edward R Vrscay

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received 14 October 1985

Abstract. For a contractive matrix transformation $\mathbf{A}: \boldsymbol{R}^2 \to \boldsymbol{R}^2$, $\mathbf{A} = a_{ij}$, $a_{ij} \in \boldsymbol{R}$, the set of points $A = \{(\pm\mathbf{I} \pm \mathbf{A} \pm \mathbf{A}^2 \pm \mathbf{A}^3 \pm \ldots)\boldsymbol{e}$ for all sequences of + and $-\}$, where $\boldsymbol{e} = (1, 0)^{\mathrm{T}}$, is a unique attractor of generally fractional Hausdorff dimension. The Mandelbrot set associated with this system is defined as $D = \{a_{ij} \in \boldsymbol{R}^4 : \mathbf{A}$ is contractive, A is disconnected$\}$. The structure of D and its boundary is investigated. Computer approximations of various sections of D are presented with a discussion of the algorithm and principles involved.

1. Introduction

One of the most fascinating structures to emerge from the study of iteration of complex mappings has been the Mandelbrot or M set (Mandelbrot 1980, 1982, 1983). For the quadratic map $R(z) = z^2 - s$, the M set represents the set of parameter values $s \in \boldsymbol{C}$ (the complex field) for which the Julia set $J(R)$ of $R(z)$ is connected (a good review of Julia sets along with numerous references is given by Blanchard (1984); see also Brolin (1965); important concepts are presented in appendix 1). A most significant result of Douady and Hubbard (1982) is that the boundary of the M set of $R(z)$, as infinitely complicated as it appears under repeated magnification, is connected. The morphology of the M set is connected with the well known cascade of period-doubling bifurcations associated with $R(z)$ as well as with universal phenomena (Feigenbaum 1978). Mandelbrot-like sets also appear in parameter spaces associated with Newton's method (Curry *et al* 1983) and its generalisations (Vrscay 1986). We finally mention an application relevant to physics: given an LRC electrical network with impedance function $Z(\omega)$, where ω represents the AC driving frequency, one may construct a set of iterated networks N_k with impedances $Z^k(\omega)$, $k = 1, 2, 3, \ldots$, (Barnsley *et al* 1985). Mandelbrot sets exist in regions of the parameter space $L, R, C \in \boldsymbol{C}$. For example, in the subspace $R = C = 1$, when L is allowed to vary from 1 to $3 + 2^{3/2}$, the impedance $Z^{\infty}(\omega)$ of the infinitely iterated network N_{∞} demonstrates a cascade of bifurcations which eventually evolves into chaotic behaviour.

We emphasise that the M set is a parameter space map of discrete dynamical systems as defined by the iteration $z_{n+1} = R(z_n) = R^n(z_0)$. Each point in this space represents a particular system with characteristic Julia set and attractor of typically fractional Hausdorff dimension, i.e. a fractal (Mandelbrot 1982). Fractal lattices have been objects of vigorous research in the context of critical phenomena and percolation clusters. The determination of dimensionality and spectral density has been the subject of many papers published in this journal.

A Mandelbrot set for pairs of complex linear maps has been investigated by Barnsley and Harrington (1985) (hereafter referred to as BH). They studied the attractors

0305-4470/86/111985+17$02.50

associated with the one-(complex)-parameter *iterated function system* (IFS) (Barnsley and Demko 1985, see also appendix 1) $\{\boldsymbol{C}, T_+, T_-\}$, where

$$T_+(z) = sz + 1 \qquad T_-(z) = sz - 1. \tag{1.1}$$

The attractor for this IFS is given by

$$A(s) = \pm 1 \pm s \pm s^2 \pm s^3 \pm \ldots \qquad \text{for all sequences of } +, -. \tag{1.2}$$

The Mandelbrot set for this IFS is defined as the region in complex parameter space $D = (s \in \boldsymbol{C}: |s| < 1, A(s)$ is disconnected$\}$. The boundary of the set appears to be fractal like; whether or not it is connected is still an open question, however.

In this paper we examine attractors and Mandelbrot sets associated with a generalisation of the IFS in (1.1): a four-real-parameter set of affine transformations $\{\boldsymbol{R}^2, S_+, S_-\}$,

$$S_\pm : \begin{bmatrix} x \\ y \end{bmatrix} \to \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \pm \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad a_{ij} \in R. \tag{1.3}$$

The matrix transformation **A** in (1.3) produces a rotation-shear scaling, where circles are mapped into ellipses, as opposed to the rotation-similitude transformation associated with complex multiplication in (1.1). It will be convenient to parametrise the a_{ij} in terms of scaling factors and shear and rotation angles. In § 2, we examine the attractors associated with this IFS and also develop formulae for the dilatation factors in terms of the above parameters. In § 3, bounds for the regions which contain the boundary ∂D of the Mandelbrot set are obtained. Some computer approximations of Mandelbrot sets are presented. Appendix 1 presents some fundamental concepts of IFS theory as well as its importance and applications in the study of fractal systems. The subject of Julia sets as attractors of IFS is also discussed. The maximum and minimum dilatation factors associated with the affine map **A** in (1.3) are derived in appendix 2 as well as the general evolution of ellipses under this map. Appendix 3 outlines the algorithm employed to determine whether an attractor for this IFS is connected or not.

2. The shear IFS and its attractors

The action of the transformation **A** in (1.3) on a unit square in $\boldsymbol{R}^2$ is illustrated in figures 1(*a*)-(*c*). It will be convenient to adopt other parametrisations of **A** as shown

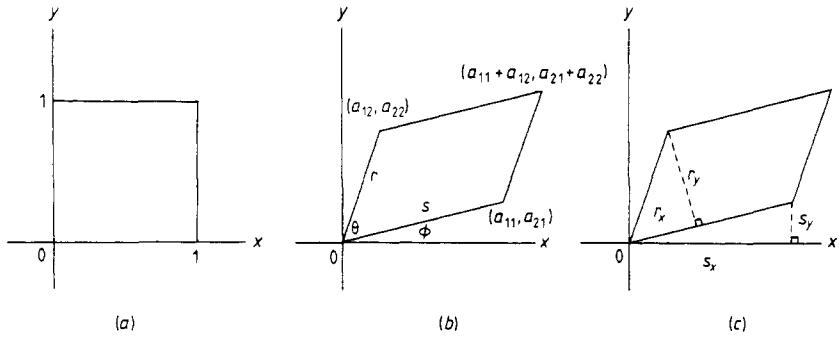

Figure 1. Action of the matrix transformation **A** in (1.3) and (2.1) on the unit square in (*a*): (*b*) polar parametrisation, (*c*) cartesian parametrisation.

in figure 1(*b*)-(*c*): either in the cartesian form $(r_x, r_y, s_x, s_y) \in \boldsymbol{R}^4$, or in polar form $(r, \theta, s, \phi) \in (\boldsymbol{R}^+, \Phi)^2$, where $\Phi = [0, 2\pi]$. Both forms will be used interchangeably and will be represented by the single notation $(\boldsymbol{r}, \boldsymbol{s})$. In these parametrisations, $\boldsymbol{A}$ assumes the forms

$$\mathbf{A} = \begin{bmatrix} s\cos\phi & r\cos(\theta+\phi) \\ s\sin\phi & r\sin(\theta+\phi) \end{bmatrix} \tag{2.1a}$$

$$= \begin{bmatrix} s_x & s^{-1}(s_x r_x - s_y r_y) \\ s_y & s^{-1}(s_x r_y - s_y r_x) \end{bmatrix}. \tag{2.1b}$$

In the polar form, ϕ represents a rotation angle and θ a shear angle. The IFS of Barnsley and Harrington (1985) (henceforth BH IFS) corresponds to the special case $\theta = \pi/2$, $r = s$. The formulae developed below reduce to those of BH for these special parameter values. Finally, let us also mention that the case $a_{12} = 0$ in (1.3) is of special importance in the fractal interpolation scheme of Barnsley (1986).

If we start with any point in the xy plane and calculate the limits of all possible sequences of the two maps in (2.1) then, subject to restrictions on $(\boldsymbol{r}, \boldsymbol{s})$ discussed below, we obtain an attractor $A(\boldsymbol{r}, \boldsymbol{s})$ of the IFS $\{\boldsymbol{R}^2, S_+, S_-\}$. It may thus be written as

$$A(\boldsymbol{r}, \boldsymbol{s}) = (\pm\mathbf{I} \pm \mathbf{A} \pm \mathbf{A}^2 \pm \mathbf{A}^3 \pm \ldots)\boldsymbol{e} \tag{2.2}$$

for all possible sequences of + and −, where $\boldsymbol{e} = (1, 0)^{\mathrm{T}}$ and $\mathbf{I}$ denotes the identity matrix. Attractors corresponding to several values of the parameters $(\boldsymbol{r}, \boldsymbol{s})$ are presented in figures 2–7. The attractor in figure 2 corresponds to $s = (1+\mathrm{i})/2$ in (1.2) whose boundary is the dragon curve of Davis and Knuth (1970).

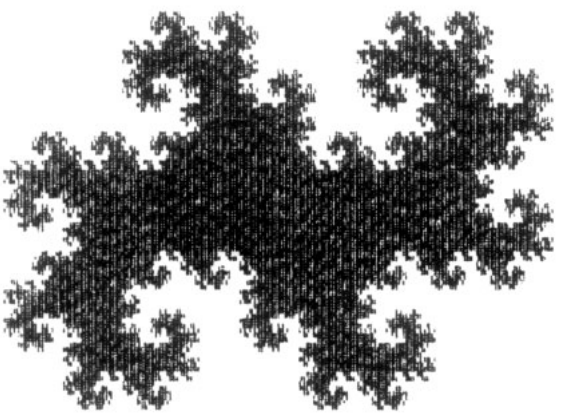

Figure 2. Attractor of IFS in (2.2) for $\theta = \pi/2$, $\phi = \pi/4$, $r = s = 1/\sqrt{2}$. The attractor boundary is the dragon curve of Davis and Knuth (1970). The attractor tiles the xy plane.

The IFS $(\boldsymbol{R}^2, S_+, S_-)$ is *hyperbolic* (or *contractive*, cf appendix 1) if and only if there is a real constant $0 \leqslant s \leqslant 1$ such that

$$|S_\sigma(\boldsymbol{f}) - S_\sigma(\boldsymbol{g})| < s|\boldsymbol{f} - \boldsymbol{g}| \qquad \sigma \in \{+, -\} \quad \text{for all } \boldsymbol{f}, \boldsymbol{g} \in \boldsymbol{R}^2 \tag{2.3}$$

where $||$ denotes the Euclidean metric in $\boldsymbol{R}^2$. Some important properties concerning the attractor follow from this condition:

(a) $A(\boldsymbol{r}, \boldsymbol{s})$ is independent of the initial point (x_0, y_0) used in (1.3),

(b) $A(\boldsymbol{r}, \boldsymbol{s})$ is unique,

(c) $S_+(A) \cup S_-(A) = A$.

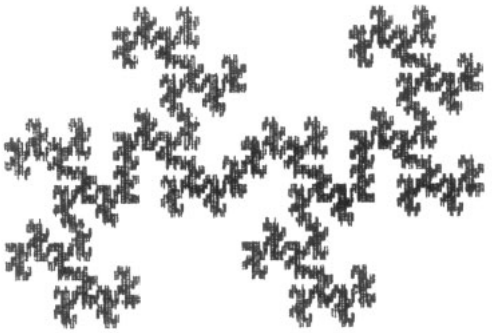

Figure 3. Attractor for $\theta = \pi/2$, $\phi = \pi/4$, $r = s = 0.65$.

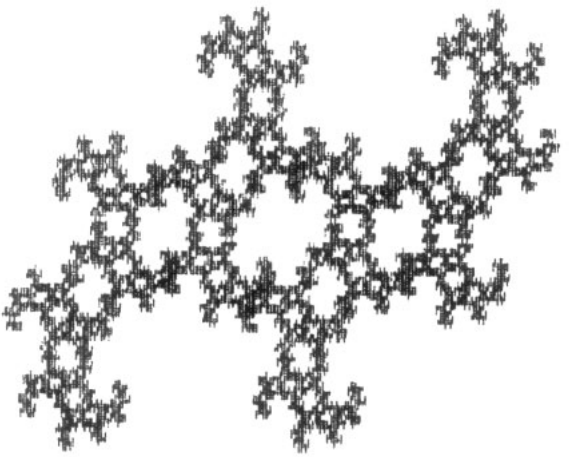

Figure 4. Attractor for $\theta \simeq 1.3090$ (75°), $\phi \simeq 0.6981$ (40°), $r = s = 0.8$.

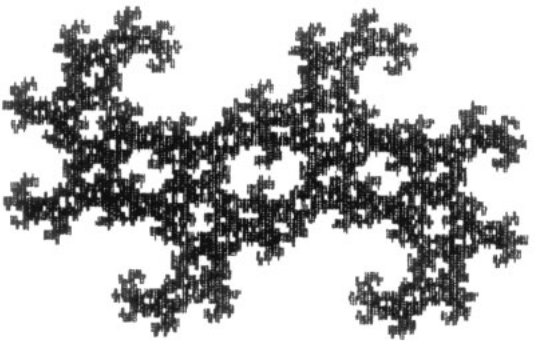

Figure 5. Attractor for $\theta \simeq 1.3090$ (75°), $\phi \simeq -0.6981$ (−40°), $r = s = 0.6$.

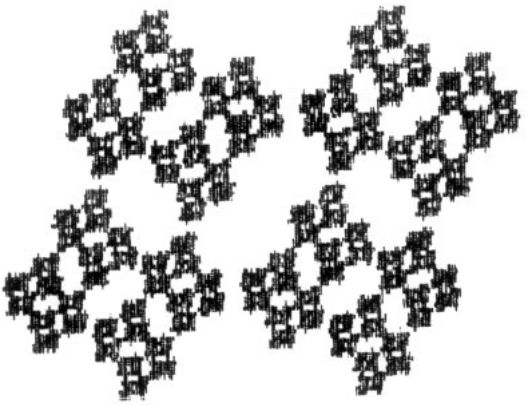

Figure 6. Attractor for $\theta \simeq 1.2217$ (70°), $\phi \simeq 1.2217$ (70°), $r = s = 0.78$.

Figure 7. Attractor for $\theta \simeq 1.3963$ (80°), $\phi \simeq 0.1745$ (10°), $r = 0.6$, $s = 0.7$.

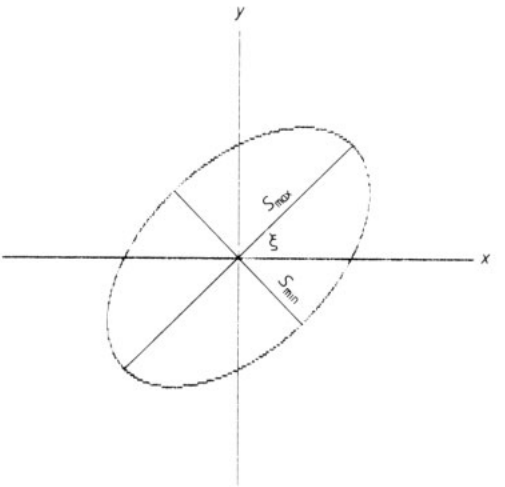

Figure 8. Ellipse produced by action of matrix transformation **A** in (1.3) on the unit circle $x^2 + y^2 = 1$.

Other technical aspects regarding the balanced measure which is supported on $A(r, s)$ will not be discussed here.

In order to determine the conditions for hyperbolicity of this IFS, we consider the non-isotropic dilatation associated with the transformation **A** in (1.3). The action of **A** on the unit circle C produces an ellipse, as shown in figure 8. The major and minor axes of the ellipse correspond to the maximum and minimum dilatation factors, which

we denote as $s_{\max}$ and $s_{\min}$, respectively. These parameters will satisfy the following relation:

$$s_{\min}|f-g| \leq |T_\sigma(f) - T_\sigma(g)| \leq s_{\max}|f-g| \qquad \sigma \in \{+,-\}. \tag{2.4}$$

In terms of the parameters of (2.1a), we have

$$s_{\max} = 2^{-1/2}[r^2+s^2 \pm (r^4+s^4+2r^2s^2\cos 2\theta)^{1/2}]^{1/2}. \tag{2.5}$$

These formulae are derived in appendix 2 and are also given in terms of the original matrix elements a_{ij}. They will be important in the estimation of the location of the boundary of the D set in § 3.

The hyperbolicity condition thus dictates that $s_{\max} < 1$. (For the BH IFS, where $r = s$ and $\theta = \pi/2$, $s_{\max} = s_{\min} = s$. Contractivity requires that $s < 1$.)

The fixed points of the transformations in (1.3), given by $S_+(f_+) = f_+$ and $S_-(f_-) = f_-$ are found to be

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}_\pm = \pm \frac{1}{\det \mathbf{A} - \mathrm{Tr}\,\mathbf{A} + 1} \begin{bmatrix} 1-a_{22} \\ a_{21} \end{bmatrix}. \tag{2.6}$$

The construction of the attractor by our IFS in (1.3) is seen to be a succession of rotation-shear scalings of the plane about these two fixed points.

There also exists a set of two-cycles, t_+ and t_- which satisfies the relations $S_+(t_+) = t_-$, $S_-(t_-) = t_+$. These points are given by

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}_\pm = \pm \frac{1}{\det \mathbf{A} + \mathrm{Tr}\,\mathbf{A} + 1} \begin{bmatrix} 1+a_{22} \\ -a_{21} \end{bmatrix}. \tag{2.7}$$

A comparison with (2.6) reveals that the two-cycles associated with $\mathbf{A}$ are the fixed points for $-\mathbf{A}$.

In the coordinate space $\boldsymbol{R}^2$, the attractors $A(r, s)$ are symmetric with respect to the point (0, 0). In parameter space, $A(r, s) = A(-r, -s)$. Some special cases of attractors are considered below.

(a) If $s_y = r_y = 0$ so that $s = |s_x|$, $r = |r_x|$, then the contribution to (2.2) becomes

$$A(r, s) = \begin{pmatrix} \pm 1 \pm s \pm s^2 \pm \ldots \\ 0 \end{pmatrix} = \begin{pmatrix} A(s) \\ 0 \end{pmatrix} \tag{2.8}$$

where $A(s)$ denotes the BH attractor of (1.2) for $s \in \boldsymbol{R}^+$. Thus $A(r, s)$ lies on the x axis. From hyperbolicity, $s < 1$, and $A(s) \in [-1/(1-s),\ 1/(1-s)]$. If $0 < s < \frac{1}{2}$, then $A(s)$ is a dissection Cantor set obtained from this interval by deleting middle $(1-2s)$th parts of any interval at each stage of construction. For $s = \frac{1}{3}$, the result is the classical ternary Cantor set. (The reader may also envision the Cantor sets produced when S_+ and S_- have different dilatation factors, say $s_+ \neq s_-$.)

(b) If $r_x = s_x = 0$, etc, then the contributions to (2.2) factorise,

$$A(r, s) = \begin{bmatrix} \pm 1 \pm rs \pm r^2s^2 \pm \cdots \\ s(\pm 1 \pm rs \pm r^2s^2 \pm \cdots) \end{bmatrix} = \begin{bmatrix} A(rs) \\ sA(rs) \end{bmatrix}. \tag{2.9}$$

The attractor is a cartesian product of BH attractors which lie on the x and y axes.

(c) If $s_x = r_y = 0$, then $A(r, s) = (\pm 1, sA(r))^{\mathrm{T}}$.

(d) If $s_y = r_x = 0$, then $A(r, s) = (A(s), 0)^{\mathrm{T}}$, as in (a).

3. The Mandelbrot set D and its boundary

The attractor $A(\boldsymbol{r}, \boldsymbol{s})$ is either disconnected or connected and generally has a fractional Hausdorff dimension, i.e. it is a fractal, as defined by Mandelbrot (1982). The attractors of figures 2-5 are connected and those of figures 6 and 7 are disconnected. The attractors in figures 3 and 4 appear to lie near the borderline of connectedness.

We define the Mandelbrot set D of the IFS $\{\boldsymbol{R}^2, S_+, S_-\}$ as

$$D = \{(\boldsymbol{r}, \boldsymbol{s}) \in \boldsymbol{R}^2 \times \boldsymbol{R}^2 : s_{\max} < 1, A(\boldsymbol{r}, \boldsymbol{s}) \text{ is disconnected}\}. \tag{3.1}$$

(Again the cartesian parametrisation will be generally understood, although we feel free to switch to polar form.) The boundary of D, ∂D, separates the regions of four-dimensional parameter space $(\boldsymbol{r}, \boldsymbol{s})$ in which the attractors $A(\boldsymbol{r}, \boldsymbol{s})$ are connected and disconnected. The regions near ∂D are interesting from both the aspect of the complicated nature of ∂D itself as well as the nature of the attractors corresponding to these parameter values. In this region, $A(\boldsymbol{r}, \boldsymbol{s})$ becomes more and more tree-like, i.e. 'barely connected'. Figures 3, 4 and 6 present attractors which correspond to parameter values $(\boldsymbol{r}, \boldsymbol{s})$ near ∂D.

Some bounds on the location of ∂D may be found. A proof of the following procedure is given in theorem 8 of Barnsley and Demko (1985). Suppose that $A(\boldsymbol{r}, \boldsymbol{s})$ is disconnected, i.e. its Hausdorff-Besicovitch (HB) dimension, $\hat{p}(A)$, is less than 1. In fact, $\hat{p}$ is bounded by

$$\min\{2, l\} \leqslant \hat{p}(A) \leqslant u \tag{3.2}$$

where $l = \log \frac{1}{2} / \log(s_{\min})$ and $u = \log \frac{1}{2} / \log(s_{\max})$. Thus $\hat{p}(A) \leqslant 1$ is ensured if $u < 1$, i.e. $A(\boldsymbol{r}, \boldsymbol{s})$ is totally disconnected if $s_{\max} < \frac{1}{2}$. If $l > 2$, then our original supposition that $A(\boldsymbol{r}, \boldsymbol{s})$ is disconnected is false. Hence $A(\boldsymbol{r}, \boldsymbol{s})$ is connected for $s_{\min} < 2^{-1/2}$. The boundary ∂D is thus contained in the four-dimensional region given by

$$B = \{(\boldsymbol{r}, \boldsymbol{s}) : \tfrac{1}{2} < s_{\max} < 1, s_{\min} < 2^{-1/2}\}.$$

We now wish to generate and examine some pictorial representations of the set D, restricting ourselves to two-dimensional slices of the parameter space. The following property provides a practical method of determining whether a given attractor $A(\boldsymbol{r}, \boldsymbol{s})$ is connected or not:

if $T_+(A) \cap T_-(A) = \varnothing$, then A is totally disconnected;
if $T_+(A) \cap T_-(A) \neq \varnothing$, then A is connected.

The proof of these statements is given in BH. A method of constructing computer approximations to D based on the above is outlined in appendix 3. We mention here that the method is a generalisation of a technique developed by Hardin (1985) to study the Mandelbrot set associated with the IFS $\{\boldsymbol{C}, sz+1, s^*z-1\}$, where $s \in \boldsymbol{C}$. It differs from the algorithm employed by Barnsley and Harrington (1985) although both methods are based on the intersection properties given above.

Case 1. $r = s$ with shear angle $0 \leqslant \theta \leqslant \pi$. The dilatation factors are given by $s_{\max} = s(1+|\cos\theta|)^{1/2}$, $s_{\min} = s(1-|\cos\theta|)^{1/2}$, where $s = (s_x^2 + s_y^2)^{1/2}$. The boundary ∂D lies within the annulus $\frac{1}{2}(1+|\cos\theta|)^{-1/2} < s < \min\{[2(1-|\cos\theta|)]^{-1/2}, [1+|\cos\theta|]^{-1/2}\}$. Computer approximations to the D set in the (s_x, s_y) subspace for $\theta = \pi/2$ (the BH IFS), $5\pi/12$, $\pi/3$ and $\theta = \pi/4$ are presented in figures 9-12, respectively. The squares enclose the plane region $(s_x, s_y) \in [-1, 1] \times [-1, 1]$. As θ decreases, more and more of

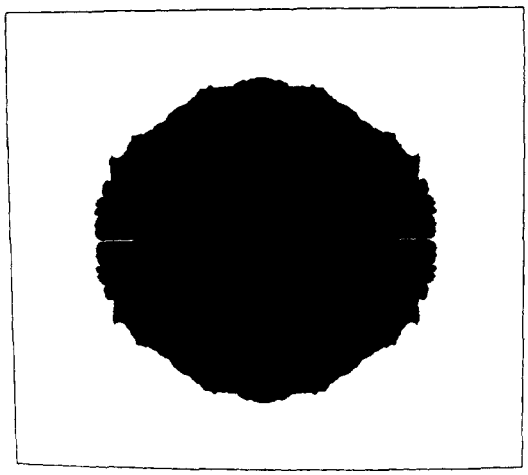

Figure 9. A computer approximation to the Mandelbrot D set for the set of affine maps in (1.3) and (2.1): the region of parameter space $(s_x, s_y) \in [-1, 1] \times [-1, 1]$ for which the attractor is disconnected. Here, $r = s$ with shear angle $\theta = \pi/2$.

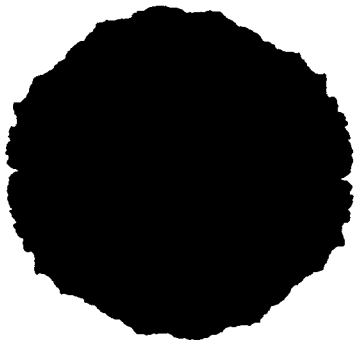

Figure 10. The D set in (s_x, s_y) space for $\theta = 5\pi/12$.

the D set is bounded by the circle $s_{\max} < 1$, the boundary of hyperbolicity. This implies that more attractors of the hyperbolic IFS become disconnected as θ decreases.

The two cases $\theta = 0$ and $\theta = \pi$ (not pictured here) are degenerate—the attractor A consists of two parallel lines passing through the fixed points $(x, y) = \pm(\alpha(1 - s_y), \alpha s_y)$, where $\alpha = (1 - s_x - s_y)^{-1}$, with slopes s_y/s_x. In both cases, the Mandelbrot set consists of the cut disc,

$$D = \{(s_x, s_y): s_x^2 + s_y^2 < \tfrac{1}{2}\} \backslash \{s_x: s_x \in ([-1/\sqrt{2}, -\tfrac{1}{2}] \cup [\tfrac{1}{2}, 1/\sqrt{2}])\}.$$

When $s_y = 0$ and $\frac{1}{2} \leqslant s = |s_x| < 1/\sqrt{2}$, the attractor A is located in the interval $[-1/(1-s), 1/(1-s)]$ on the x axis.

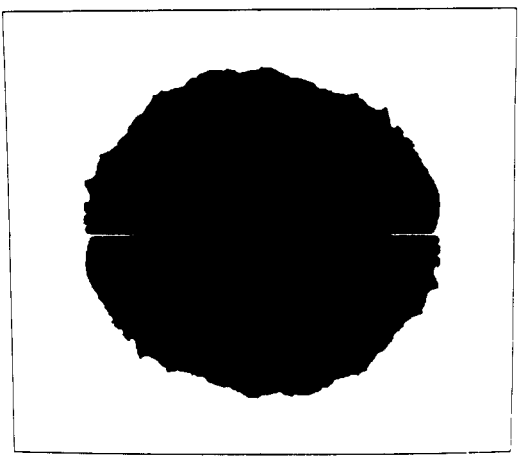

Figure 11. The D set in (s_x, s_y) space for $\theta = \pi/3$.

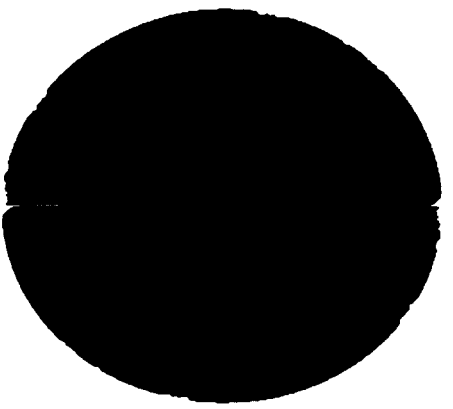

Figure 12. The D set in (s_x, s_y) space for $\theta = \pi/4$.

Case 2. $r = \alpha s$, $0 < \alpha < 1$, with shear angle θ. From appendix 2, the major and minor dilatation factors are given by

$$s_{\substack{\max \\ \min}} = \frac{s}{\sqrt{2}}[(1+\alpha^2) \pm (1+\alpha^4+2\alpha^2 \cos 2\theta)^{1/2}]^{1/2}.$$

The inner and outer radii of the annulus containing ∂D may be computed from these factors. Computer approximations to D in the parameter subspace $(s_x, s_y) \in [-1, 1] \times [-1, 1]$ for $\alpha = 0.9$, 0.85 and 0.8 are presented in figures 13-15, respectively.

4. Concluding remarks

Affine maps of the form (1.3) are rotation-shear scaling (plus translation) transformations whose maximum and minimum dilatation factors may be explicitly computed.

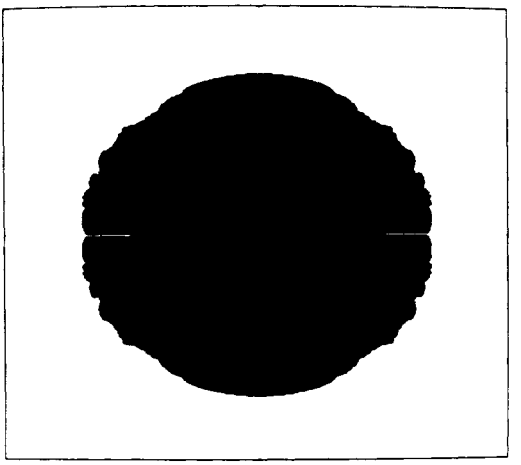

Figure 13. The D set in (s_x, s_y) space: $r = 0.90s$, $\theta = \pi/2$.

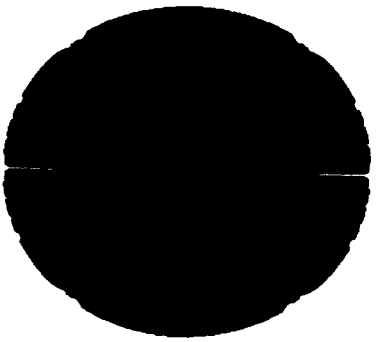

Figure 14. The D set in (s_x, s_y) space: $r = 0.85s$, $\theta = \pi/2$.

In general, the attractors of such iterated function systems are fractals and are either connected or disconnected. The Mandelbrot set for such an IFS may be defined as the region in parameter space for which the IFS is hyperbolic and the attractor is disconnected. The boundary of this set reveals classic fractal-like patterns.

Acknowledgments

The author expresses his sincerest thanks to Professor M F Barnsley for suggesting the problem and to Dr D Hardin for many valuable discussions during the course of this work. The support of a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship for 1984–85 is hereby gratefully acknowledged. The Mandelbrot sets were computed and displayed on Burroughs B22 minicomputers.

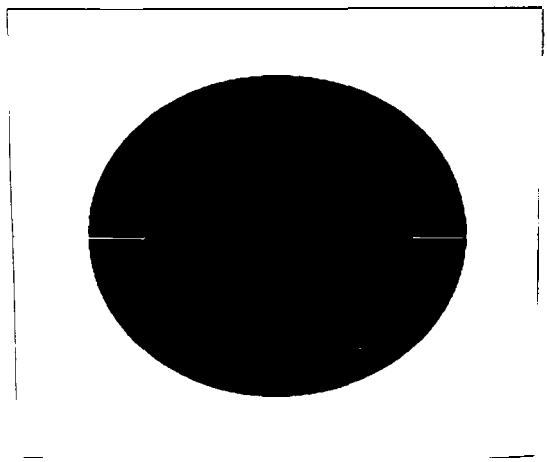

Figure 15. The D set in (s_x, s_y) space: $r = 0.80s$, $\theta = \pi/2$.

Appendix 1. Some remarks on iterated function systems and Julia sets

In this section we briefly outline the important concepts of an iterated function system (IFS) as developed by Barnsley and Demko (1984). Some connections between IFS and Julia sets are made at the end.

Let $\boldsymbol{K}$ be a compact metric space with distance function $d(x, y)$ for $x, y \in \boldsymbol{K}$, and $w = (w_1, w_2, \ldots w_d\}$ be a finite collection of functions $w_i : \boldsymbol{K} \to \boldsymbol{K}$. Here we shall limit our discussion to w_i continuous in $\boldsymbol{K}$. Now treat w as a set-valued function $w(x) : \boldsymbol{K} \to \boldsymbol{K}$, and consider the action of its iterates $\{w^n(x)\}_0^\infty$, where

$$w^0(x) = x \qquad w^n(x) = w(w^{n-1}(x)) \qquad n = 1, 2, 3, \ldots .$$

Thus, the set defined by $w^n(x)$ consists of d^n values, given by the compositions $w_{\sigma_1} w_{\sigma_2} \ldots w_{\sigma_n}(x)$, for all possible sequences $\sigma_i \in \{1, 2, \ldots d\}$, $i = 1, 2, \ldots n$. The *attractor* $A(x)$ corresponding to the above sequence for a given $x \in \boldsymbol{K}$ is defined as

$$A(x) = \lim_{n \to \infty} w^n(x).$$

If $A(x)$ is independent of x, then it is simply referred to as A, the attractor of the IFS $(\boldsymbol{K}, w\}$. It is compact and invariant under w, i.e. $w(A) = A$.

The IFS $\{\boldsymbol{K}, w\}$ is defined as *hyperbolic* if there exists a constant $0 \leq s < 1$ such that

$$d(w_i(x), w_i(y)) \leq s\, d(x, y) \; \forall \, x, y \in \boldsymbol{K}, \; i = 1, 2, \ldots d.$$

In this case, the following important properties exist:

(a) $A(x)$ is independent of x, i.e. $A(x) = A$,

(b) A is unique,

(c) $A = \bigcup_{i=1}^{d} w_i(A)$.

On a computer, the unique attractor for a hyperbolic IFS may be generated for display in two ways.

(i) For a finite number of generations, say N, pick a starting point $x_0 \in \boldsymbol{K}$ ($\boldsymbol{K}$ will usually be $\boldsymbol{R}^2$ or $\boldsymbol{C}$) and calculate all possible d^N points in the set $w^N(x_0)$. This

amounts to calculating all compositions in (A1.1). The recursive procedure of the PASCAL computer language can be exploited here to obtain a reasonably symmetric picture of the attractor. The attractors of figures 2–7 were constructed in this way, with $N \sim 14$.

(ii) Let $\boldsymbol{p} = \{p_1, p_2, \ldots p_d\}$ represent a set of probabilities with each $p_i > 0$ and $\Sigma p_i = 1$. Starting at $x_0 \in K$, calculate the sequence $x_{n+1} = w_{\sigma_n}(x_n)$, $\sigma_n \in \{1, 2, \ldots d\}$, where, at each step, a probability of p_i is associated with the selection of σ_i. Almost always, the random walk sequence (x_n) 'settles' onto the attractor A and wanders over it. Thus, plotting the sequence $\{x_m : m = M, M+1, \ldots N\}$ where, for example, $M = 50$ or 100 and $N = 5000$ or 10 000, yields representations of A quite rapidly. The method is also easily programmable.

The IFS is an effective and informative, albeit mathematically useful, method of constructing fractals and other self-similar systems (Barnsley and Demko 1985, Barnsley *et al* 1985). It replaces their algorithmic construction (with 'generators', see Mandelbrot (1982), for example) with a global procedure involving transformations in a metric space. Moreover, a moment theory of IFS also exists to provide a systematic basis for the inverse problem, i.e. given a system, determine the IFS whose attractor best approximates the system.

As an example, consider the IFS $\{\boldsymbol{C}, w_1, w_2\}$, where $w_1(z) = sz$, $w_2(z) = s(z-1)+1$, $s < 1$ are contractions about the points $z = 0$ and $z = 1$, respectively. The reader may verify that the attractor A must lie in the interval $[0, 1]$. By examining the action of the w_i on this interval for $s = \frac{1}{3}$, one may conclude that A is the classical ternary Cantor set. A similar analysis shows that the IFS $\{\boldsymbol{C}, w_1, w_2, w_3\}$, where

$$w_i = s_i(z - a_i) + a_i \qquad s_i = \tfrac{1}{2},\ a_i \in \boldsymbol{C} \qquad i = 1, 2, 3, \quad a_i \text{ non-collinear} \tag{A1.1}$$

admits the famous Sierpinski gasket (Rammal 1983, Domany *et al* 1983) as its attractor. A computer generated plot of A for a_i constituting the vertices of an equilateral triangle is shown in figure 16. When $s_1 = \frac{3}{5}$ and $s_2 = s_3 = \frac{2}{5}$ in equation (A1.1), the 'Cantor tree' of figure 17 results. The structure of this attractor is similar to the time-evolution patterns encountered in cellular automata (Wolfram 1984). The modified Sierpinski

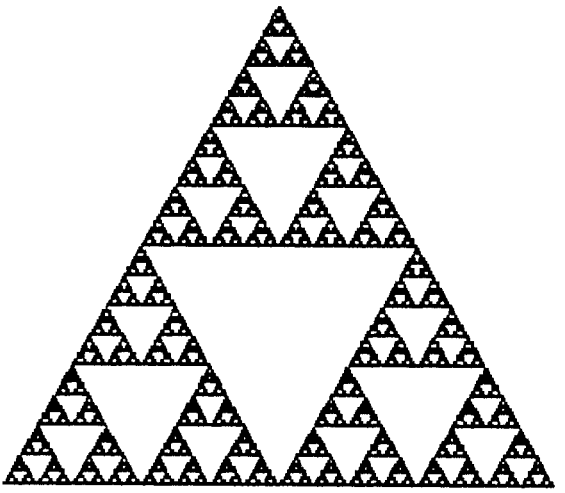

Figure 16. The Sierpinski gasket, an attractor for an iterated system of three contractive maps.

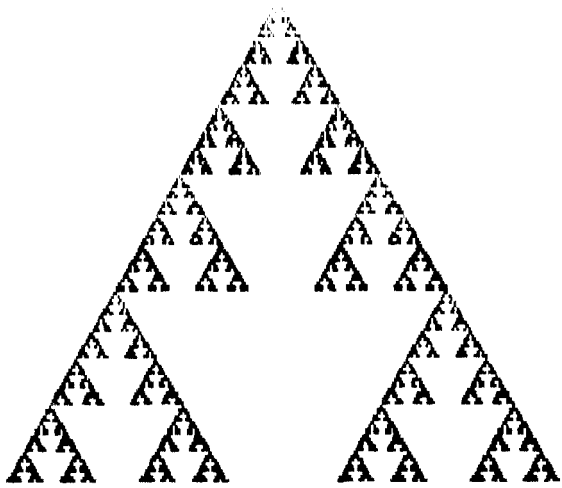

Figure 17. A Cantor tree.

gasket of figure 18 (Hilfer and Blumen 1984) is the attractor of the IFS $\{\boldsymbol{C}, w_i, i = 1, 2, \ldots 6\}$, where the w_i are given in (A1.1), with $s_i = \frac{1}{3}$ and the a_i lie on the vertices and midpoints of the sides of an equilateral triangle.

We conclude this section by outlining the notion of Julia sets as attractors of iterated function systems. Let $R(z)$ be a rational function $R(z) = P(z)/Q(z)$ where $P(z)$ and $Q(z)$ are polynomials with complex coefficients and no common factors and $d \equiv \deg(R) \equiv \max(\deg(P), \deg(Q)) \geqslant 2$. The sequence of iterates $\{R^n\}$ of $R(z)$ is defined by $R^0(z) = z$, $R^{n+1}(z) = R(R^n(z))$, $n = 0, 1, 2 \ldots$. The inverses of $R(z)$ shall be denoted by $R_i^{-1}(z)$, where $i = 1, 2, \ldots d$ enumerates all branches of the inverse. We now consider $R: \bar{\boldsymbol{C}} \to \bar{\boldsymbol{C}}$ where $\bar{\boldsymbol{C}} = \boldsymbol{C} \cup \{\infty\}$ denotes the Riemann sphere with suitably defined

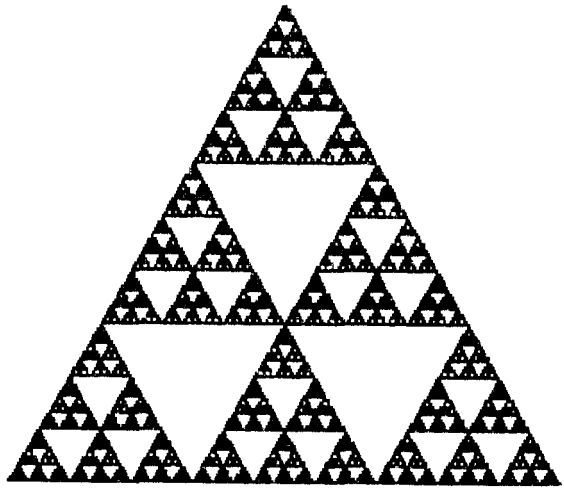

Figure 18. A generalised Sierpinski gasket.

spherical metric. Given a point $z_0 \in \bar{\mathbf{C}}$, the iteration sequence $\{z_n\}_0^\infty$, where $z_{n+1} = R(z_n) = R^{n+1}(z_0)$, defines the *forward orbit* of z_0.

If $R^k(p) = p$ and $R^m(p) \neq p$ for $m < k$, then p is a *fixed point of order k*. The set of distinct points $\{p_i, i = 1, 2, 3, \ldots, k\}$, where

$$p_1 = R(p), p_2 = R(p_2), \ldots, p_k = R(p_{k-1})$$

is termed a k-cycle. If $k = 1$, p is simply called a *fixed point* of $R(z)$. The k-cycle is *attractive, indifferent or repulsive*, depending whether the multiplier $|[R^k(p_i)]'|$ is less than, equal to or greater than one, respectively.

The Julia set $J(R)$ of the rational map $R: \bar{\mathbf{C}} \to \bar{\mathbf{C}}$ is formally defined as the set of $z \in \bar{\mathbf{C}}$ for which the family of maps $R^n(z)$ is not normal, in the sense of Montel (Ahlfors 1979). A more working description is that $J(R)$ is the closure of all repulsive k-cycles of $R(z)$, $k = 1, 2, 3, \ldots$. Its complement, $F = \bar{\mathbf{C}} \backslash J(R)$, the Fatou set, is the set of all $z \in \bar{\mathbf{C}}$ for which the family $R^n(z)$ is equicontinuous in the spherical metric on some neighbourhood of each point of F.

Some important properties of $J(R)$ are listed below:

(a) $J \neq \varnothing$ and J is closed;

(b) J is invariant with respect to R, i.e. $R(J) = J = R^{-1}(J)$;

(c) $J(R) = J(R^m)$, $m = 2, 3, 4, \ldots$;

(d) if J has interior points then $J = \bar{\mathbf{C}}$;

(e) $J(R)$ is compact and non-denumerable. In general, its Hausdorff-Besicovitch dimension is non-integral, whereupon $J(R)$ is a *fractal*, as defined by Mandelbrot (1982).

A simple and illustrative example is afforded by the map $R(z) = z^2$. The unit circle $C = \{z : |z| = 1\}$ is invariant with respect to $R(z)$ and its iterates. All fixed points and cycles of the $R^n(z)$, except $z = 0$ and $z = \infty$, lie on C and are repulsive. C is the Julia set of $R(z)$. The forward orbit of any point in the region $|z| < 1$ is the fixed point $z = 0$. The forward orbit of any point in the region $|z| > 1$ is the point $z = \infty$. The Julia set C may be regarded as a *repeller set* under the action of the forward map $R(z)$. Equivalently, C is the *attractor* for the inverses $R_1^{-1}(z) = +\sqrt{z}$, $R_2^{-1}(z) = -\sqrt{z}$.

For the general quadratic maps $R(z) = z^2 - s$, $s \in \mathbf{C}$, let $w_1(z) = +(z+s)^{1/2}$, $w_2(z) = -(z+s)^{1/2}$. The Julia set $J(R)$ is the attractor for the IFS $\{\mathbf{C}, w_1, w_2\}$ and may be given by

$$J(s) = \sqrt{(s \pm \sqrt{(s \pm \ldots)} \ldots)} \text{ for all sequences of } +, -.$$

Special cases include: (i) $J(0) = C$ (shown above), (ii) $J(2) = [-2, 2]$, (iii) for $s > 2$, $J(s)$ is a Cantor-like set in the interval $[\frac{1}{2} + (s + \frac{1}{4})^{1/2}, \frac{1}{2} - (s + \frac{1}{4})^{1/2}]$.

Appendix 2. Dilatation factors associated with the transformation A

We consider the action of the matrix transformation **A** of (1.3) on a unit circle C which produces an ellipse, as illustrated in figure 8. The major and minor axes represent maximum and minimum dilatation factors, $s_{\max}$ and $s_{\min}$, respectively, which enter into (2.5). These factors are determined below.

Starting with the unit circle centred at the origin, repeated application of the transformation **A** produces a family of ellipses $E^{(i)}(x, y)$ having the general formula

$$E^{(i)}(x, y) = a_i x^2 + b_i x^2 + c_i xy - 1 = 0 \tag{A2.1}$$

with initial conditions $a_0 = b_0 = 1$, $c_0 = 0$. Some straightforward algebra yields the recursion relations,

$$\begin{aligned} a_{i+1} &= a_i u_{11}^2 + b_i u_{21}^2 + c_i u_{11} u_{21} \\ b_{i+1} &= a_i u_{12}^2 + b_i u_{22}^2 + c_i u_{12} u_{22} \\ c_{i+1} &= 2(a_i u_{11} u_{12} + b_i u_{21} u_{22}) + c_i (u_{11} u_{12} + u_{12} u_{21}) \end{aligned} \tag{A2.2}$$

where the u_{ij} are the elements of $\mathbf{A}^{-1}$. These recursion relations will be important for the discussion in appendix 3.

The unit circle transforms to the ellipse $E^{(1)}(x, y)$ whose coefficients are

$$\begin{aligned} a_i &= (s_x^2 r_y^2 + 2 s_x s_y r_x r_y + s_y^2 r_x^2 + s^2 s_y^2) s^{-4} r_y^{-2} \\ b_i &= (s_y^2 r_y^2 - 2 s_x s_y r_x r_y + s_x^2 r_x^2 + s^2 s_x^2) s^{-4} r_y^{-2} \\ c_i &= 2[s_x s_y (r_y^2 - r_x^2) + r_x r_y (s_y^2 - s_x^2) - s^2 s_x s_y] s^{-4} r_y^{-2}. \end{aligned} \tag{A2.3}$$

The major and minor axes of this ellipse are determined by rotating it by an angle of $-\xi$ (cf figure 8). Some analytic geometry reveals that this angle is given by $\tan 2\xi = c_1/(b_1 - a_1)$. The resulting ellipse is $A_1 x^2 + B_1 y^2 = 1$, where

$$A_1, B_1 = \tfrac{1}{2}\{a_1 + b_1 \pm [(a_1 - b_1)^2 + c_1^2]^{1/2}\}. \tag{A2.4}$$

From (A2.3), the major and minor axes are given by

$$s_{\substack{\max \\ \min}} = 2^{-1/2}[(r^2 + s^2) \pm (r^2 + s^4 + 2 r^2 s^2 \cos 2\theta)^{1/2}]^{1/2}. \tag{A2.5}$$

In terms of the original elements a_{ij} of $\mathbf{A}$ in (1.3), these factors are given by

$$s_{\substack{\max \\ \min}} = \sqrt{2}(\det \mathbf{A}) \left(\sum_{i,j=1}^{2} a_{ij}^2 \mp P \right)^{-1/2} \tag{A2.6}$$

where

$$P = (a_{12}^2 + a_{22}^2)^2 + (a_{11}^2 + a_{21}^2)^2 + 2(a_{21}^2 - a_{11}^2)(a_{22}^2 - a_{12}^2) + 8 \prod_{i,j=1}^{2} a_{ij}.$$

Appendix 3. Algorithm for calculating the Mandelbrot set

The method described below exploits the three properties given in appendix 1 concerning the attractor of a hyperbolic IFS. A description will be made with reference to the IFS of (1.3). The algorithm is not the most efficient for computer implementation as there are no shortcuts which exploit any geometric properties of the attractors. An outline of this method is useful from a pedagogical viewpoint.

Rather than observe the action of $\boldsymbol{S} = \{S_+, S_-\}$ on A, we observe their action on a region which contains A. The first step is to consider a circle C_R of radius R, centred at the origin O (since A is symmetric with respect to O) which encloses A. The practical estimation of an R value will be discussed at the end of this appendix. Now consider the evolution of this circle under the repeated action of $\boldsymbol{S}$. After n generations, C_R has evolved into 2^n congruent ellipses, all oriented at an angle ξ_n with respect to the x axis. Their centres correspond to the 2^n descendants $(x_i^{(n)}, y_i^{(n)})$, $i = 1, 2, \ldots 2^n$, of the initial point $(x_1^{(0)}, y_1^{(0)}) = (0, 0)$. The equations of these ellipses are $E_i^{(n)}(x - x_i^{(n)}, y - y_i^{(n)}) = 0$, where the $E^{(n)}(x, y)$ are defined in appendix 2. The major and minor axes of these ellipses are $(s_{\max})^n$ and $(s_{\min})^n$, respectively. The first three generations are shown schematically in figures 19(a) and (b).

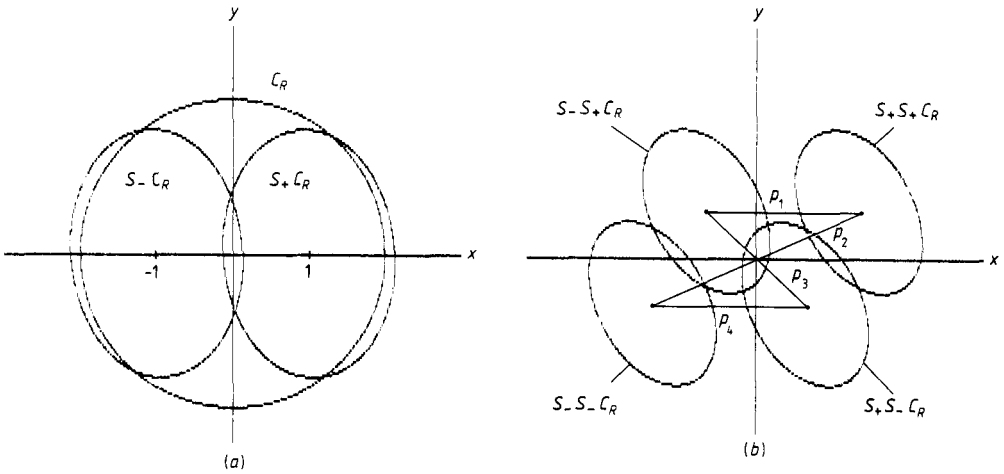

Figure 19. Action of the maps S_+ and S_- of (1.3) on a circle C_R which houses the entire attractor A: (a) the first generation of two ellipses, (b) the second generation of four ellipses.

For each generation n, a part of the attractor A must be located inside each of the 2^n ellipses $E_i^{(n)}$, i.e.

$$A \in \bigcup_{k=1}^{2^n} E_k^{(n)}(x - x_i^{(n)}, y - y_i^{(n)}).$$

We now refer to figures 19(a) and (b) to outline the algorithm. Clearly, if $E_1^{(1)} = S_-(C_R)$ does not overlap $E_2^{(1)} = S_+(C_R)$ then A is disconnected and our search may stop. The converse is not necessarily true, however. If the ellipses overlap, we consider the next generation of four ellipses which evolve from this pair. A is disconnected if

$$S_+(E_i^{(1)}) \cap S_-(E_j^{(1)}) = \varnothing \qquad \forall i, j \in \{1, 2\}.$$

We must check the four pairs of ellipses whose centres are connected by paths p_1 to p_4 in figure 16(b) overlap. The condition for overlap is that the midpoint of p_i lies within either of the ellipses connected. If no pairs overlap, then we conclude that A is disconnected. All pairs which overlap, however, must be further examined in the same way as above. The procedure amounts to exhaustively searching a tree with four branches at each node for overlaps to a prescribed number of generations, N, halting at a node whenever overlapping stops. If all paths become disconnected by the Nth generation, then A is assumed to be disconnected—if any overlaps remain, then A is assumed connected. For the D sets presented in figures 9-15, N was typically equal to 13.

To estimate the radius R of the circle C_R which houses the entire attractor, we compute 2^M descendants of $(x^{(0)}, y^{(0)}) = (0, 0)$, where M is typically 5 or 6, and define

$$r = \max_{1 \leq i \leq M} [(x_i^{(M)})^2 + (y_i^{(M)})^2]^{1/2}.$$

The value r is not expected to be a reasonable estimate of R since it is possible that an image of C_R, namely the ellipse $E^{(M)}$, be stored around each point $(x_i^{(M)}, y_i^{(M)})$, cf

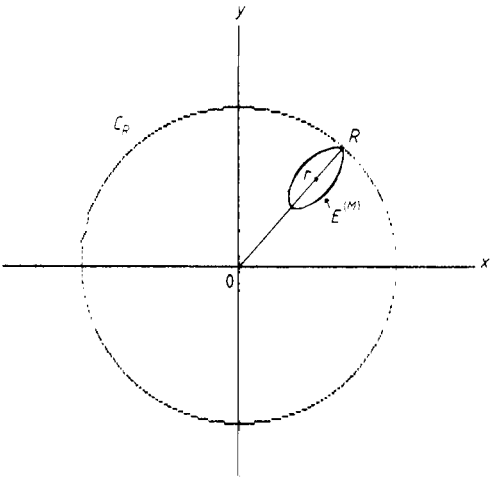

Figure 20. Estimation of the radius of the circle C_R which houses the attractor of the affine IFS.

figure 20. The assumption that its major axis is aligned with the radius vector $\boldsymbol{r}$ yields an upper limit to R. Thus, $r + R(s_{\max})^M = R$ or

$$R = \frac{r}{1-(s_{\max})^M}.$$

Of course, as $M \to \infty$, i.e. as more and more points are calculated, $(s_{\max})^M \to 0$, implying that the distance r becomes a better estimate of R.

References

Ahlfors L 1979 *Complex Analysis* (New York: McGraw-Hill)
Barnsley M F 1986 *J. Construct. Approx.* to appear
Barnsley M F and Demko S G 1985 *Proc. R. Soc.* A **40** 39
Barnsley M F and Harrington A 1985 *Physica* **15D** 421
Barnsley M F, Morley T and Vrscay E R 1985 *J. Stat. Phys.* **40** 39
Blanchard P 1984 *Bull. Am. Math Soc.* **11** 85
Brolin H 1985 *Ark. Mat.* **6** 103
Curry J, Garnett L and Sullivan D 1983 *Commun. Math. Phys.* **91** 267
Davis C and Knuth D E 1970 *J. Rec. Math.* **3** 66-71, 133-49
Domany E, Alexander S, Bensimon D and Kadanoff L P 1983 *Phys. Rev.* B **28** 3110
Douady F and Hubbard P 1982 *C.R. Acad. Sci., Paris* **294** 123-6
Feigenbaum M 1978 *J. Stat. Phys.* **19** 25
Hardin D 1985 *Preprint* Georgia Institute of Technology
Hilfer R and Blumen R 1984 *J. Phys. A: Math. Gen.* **17** L537
Mandelbrot B 1980 *Ann. NY Acad. Sci.* **357** 249
—— 1982 *The Fractal Geometry of Nature* (San Fransisco: Freeman)
—— 1983 *Physica* **7D** 224
Rammal R 1983 *J. Physique* **45** 191
Vrscay E R 1986 *Math. Comput.* **46** 151-79
Wolfram S 1984 *Los Alamos Sci.* Fall 2