COMPUTER RENDERING OF FRACTAL CURVES AND SURFACES

LOREN C. CARPENTER

BOEING COMPUTER SERVICES
SEATTLE, WASHINGTON

ABSTRACT

Fractals are a class of highly irregular shapes that have myriad counterparts in the real world, such as islands, river networks, turbulence, and snowflakes. Classic fractals include Brownian paths, Cantor sets, and plane-filling curves. Nearly all fractal sets are of fractional dimension and all are nowhere differentiable.

Previously published procedures for calculating fractal curves employ shear displacement processes, modified Markov processes, and inverse Fourier transforms. They are either very expensive or very complex and do not easily generalize to surfaces. This paper presents a family of simple methods for generating and displaying a wide class of fractal curves and surfaces. In so doing, it introduces the concept of statistical subdivision in which a geometric entity is split into smaller entities while preserving certain statistical properties.

KEY WORDS AND PHRASES: computer graphics, fractals, random curves, terrain models, natural forms.

CR categories: 3.41, 5.12, 5.13, 8.1, 8.2.

INTRODUCTION

Fractal sets have been recently formalized by B. Mandelbrot [7]. He shows that many isolated "monster" sets constructed by mathematicians, mostly as counterexamples, are special cases of a class of objects he calls fractals. Mathematically speaking, a fractal set is a set whose Hausdorff-Besicovitch dimension (fractal dimension) is greater than its topological dimension. The precise definition of Hausdorff-Besicovitch dimension can be found in Ill. As an example, a Brownian path in the plane, which fills the plane, has fractal dimension two (the plane) but topological dimension one (it is still a curve). Sets whose fractal dimension is non-integral are more interesting, since most fractal sets in nature have non-integral dimension.

Before proceeding into a discussion of the fractal dimension of natural forms, we note that all objects in nature have many different topological dimensions depending on the aspects and scale under discussion. For example, a ball of string a mile away is a point, up close it is a three dimensional sphere, closer still it is a one dimensional string, and so on. For any given form there is an outer scale and an inner scale between which one may talk about it having a constant topological (or fractal) dimension. Figures in this paper have an outer scale of the size of the figure and an inner scale determined by a resolution factor dependent on the computation method illustrated.

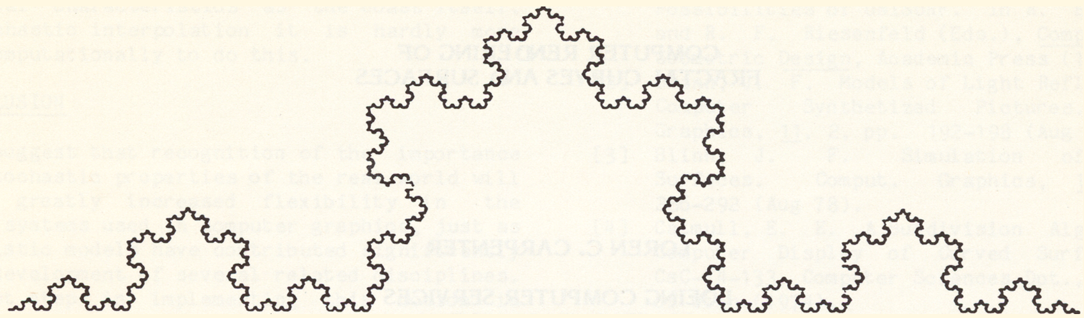
Nature is a rich source of fractals. A coastline is an excellent example. If one measures a coastline using a 100-mile measuring stick, then a lo-mile measuring stick, and then a l-mile measuring stick, the measured length grows apparently without bound [7, 81. The lack of convergence indicates the coastline has a fractal dimension greater than one. In fact, most are around 1.3. This phenomenon of unbounded increase has been discovered in river outlines, mountain skylines, tree branches, blood vessels, the entire Earth's surface, and many other forms. Another property of fractal sets is that of self similarity. Self similarity refers to the property of a form that remains unaltered through changes in scale. Self affine similarity additionally allows translation. The term self similar is often used to describe both cases. Some fractals are exactly self similar like a Koch curve (Figure 1). Others are statistically self similar like Brownian motion, river flood levels, and turbulence.

Fractal sets do not have to be continuous as evidenced by the Cantor set (dimension .6309) and the distribution of matter in the galaxy (dimension about 1.3). However, all are nowhere differentiable.

PREVIOUS METHODS

Three basic approaches to generating fractal sets using a digital computer have been published. They are shear displacement processes, modified Markov processes, and inverse Fourier transforms.

A shear displacement process [6] on a line cuts the line at a uniform random point and displaces the left and right parts vertically in opposite directions a Gaussian random distance. This is repeated until the detail is sufficiently fine. In the case of a plane, the shear lines are determined by uniformly distributed points and angles. While



Triadic Koch Curve

Figure 1.

these methods produce interesting sets, they are computationally very expensive (roughly second degree for the line and fourth degree for the plane).

Mandelbrot [4,5] has developed a modified Markov process which can produce fractal sequences of dimension 1 through 2. The modification entails using a weighted sum of the past values of the sequence plus a random variable to generate the next value.

It exhibits persistence in that the next value is affected by values in the recent past. Mathematically it is quite intricate and difficult to work with. The method is satisfactory for producing time sequences, but it has several shortcomings in a graphics environment. It tends to run away in a positive or negative direction. It cannot be made to pass through two or more specified points. It is at best extremely difficult to reverse, that is, generate the same numbers in the reverse order from some point on the curve. Extension to surfaces is, to the author's knowledge, an unsolved problem. Finally, no guidance is provided in how to refine a sequence if local scale magnification is desired.

The third method depends on the I/f power spectrum property of fractal sets. One constructs a spectrum having amplitude decreasing with the reciprocal square root of frequency and white noise for the phase component. An inverse Fourier transform gives a fractal curve. The inverse Fourier transform requires a substantial amount of complex arithmetic and, in any practical sense, is limited to sequences of length 2**k. The whole sequence is produced at once. It cannot be generated incrementally without losing the N/log2(N) time advantage of the FFT. As with the modified Markov process, the step size is constant.

THE SUBDIVISION METHOD

The fractal subdivision technique is based on a generalization of conventional subdivision methods [3]. Conventional subdivision of a curve or surface element produces two or more like elements that, together, exhibit the same geometric properties of their parent; for example, the same overall shape, derivatives, normals, continuity, etc. The subdivision method introduced here splits a fractal region of finite size in such a way as to maintain its statistical properties.

We represent a fractal element by a set of points and a statistical factor related to its dimension. A fractal curve, for example, is composed of its two endpoints and an offset or roughness factor. Subdivision of a fractal curve consists of determining a "midpoint" by a constrained random process. The type of process employed determines the character of the resulting set. The new point becomes the common endpoint of the two new fractal subcurves. To guarantee self similarity, and thus preserve its statistical properties, it is sufficient that the deviation of the "midpoint" from the actual midpoint be in some way proportional to the distance between the endpoints. As a result the roughness of the final shape will be independent of scale.

A computer graphic representation of a fractal set is subject to inner and outer scale limitations. The graphical inner scale refers to the tolerance or stopping criteria for the subdivision process. At some point (indistinguishable by eye, if carried far enough) the fractal segment is replaced by a simple geometric element like a straight line, triangle or bicubic patch. The outer scale is determined by the starting point set. The fractal character extends from the given starting points down to the geometric primitive subelements used to represent the geometry of the resulting form. This "spatial bandwidth" determines the richness or resolution of the object. The sets produced by application of the principles described above are not true fractal sets, but approximations. Between the inner and outer scale they are self similar, etc. however they do not have fractional dimension since that would require infinitely many subdivisions. The behavior of the geometry inside the inner scale is up to the implementor. In all figures presented here the inner geometry is linear unless otherwise stated.

Mandelbrot I71 uses the terms line-to-line, line-to-plane, plane-to-line and several others in classifying fractal functions. For the purposes of this discussion one can think of these classes as containing functions of m real parameters into a real Euclidean space of n dimensions. The terms line, plane, space, etc. are taken as shorthand for the dimensionality of the function's domain and range.

THREE FRACTAL CURVE CONSTRUCTION PROCEDURES

We construct a fractal line-to-line function [7] as:

Let $h = f(t, o, e, R)$

where t is a real parameter in $[a,b]$
o is a real offset or roughness factor
e is a real tolerance or resolution limit
R is a random real function on $[a, b]$ with zero mean

The construction procedure is recursive:

Let h_a and h_b be given and $t1 = a$, $t_2 = b$

BEGIN

IF $t_2 = t_1 > e$ THEN

DEFINE $t_m = (t_1 + t_2) / 2$

$h_m = (h_1 + h_2) / 2 + o\,(t_2 - t_1)\,R(t_m)$

AND REPEAT WITH

$[t_1, t_m]$ AND $[t_m, t_2]$

END

The effect of this is to offset the midpoint by a random displacement proportional to the product of the offset and the parametric length. See Figure 2.

That the above definition produces a fractal curve may be shown in a variety of ways. Viewed as a time signal, the function is essentially a superposition of triangular pulses with random amplitude and the expected value of the amplitude proportional to the pulse length. It is easy to show that the function has a power spectrum of l/f.

For some $k > 0$ and $[t_1, t_2]$ in $[a, b]$

we have a length $(b = a)\,2^{-k} = t_2 - t_1$

and average amplitude $o\,(t_2 - tl)$ $<|R|> = c\,(t_2 - tl)$

for some $c > 0$.

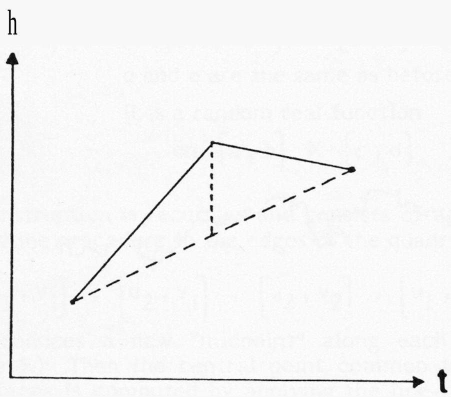
For the total power supplied for the k-th level we have

$$2^k c\,(t_2 - t_1)(t_2 - t_1)$$
$$= c\,2^k (b - a)\,2^{-k} (b - a)\,2^{-k}$$
$$= c\,2^{-k} (b - a)^2$$

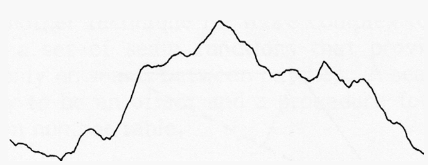
Since $f = 2^k$, the power is proportional to l/f.

Some ways to show that a set is fractal are to show that its similarity length (remember the coastline and the measuring stick) increases by a (statistically) constant ratio, or that it is nowhere differentiable (in the limit), or that its Hausdorff-Besicovitch dimension (again, in the limit) is greater than its topological dimension.

The dimension of a fractal curve created by binary subdivision may be determined from its similarity ratio. For example, if its length increases by an average ratio of 1.3 each time it is subdivided then its similarity ratio is 1.3. The formula for its dimension may be derived from the more general equation found in



Construction Procedure



Typical Resulting Curve

Line-to-Line Function

Figure 2.

$$D = \frac{\ln(2)}{\ln(2 / \text{ratio})}$$

The fractal dimension for a similarity ratio of 1.3 is then 1.61. Note that D is defined only if the similarity ratio is less than 2.

We now present two axis-independent fractal line-to-plane curve methods: one-variable and two-variable.

The one-variable line to plane function is:

Let $p = g(t, o, e, R)$

with t, o, e, R the same as before

We construct it recursively:

Let p_a and p_b be given, with tl = a and t_2 = b

BEGJN

IF $t_2 - t_1 > e$ THEN

DEFINE $t_m = (t_1 + t_2) / 2$

$p_{mx} = (p_{1x} + p_{2x}) / 2$
$\quad o\,(p_{2y} - p_{1y})\,R(t_m)$

$p_{my} = (p_{1y} + p_{2y}) / 2$
$\quad + o\,(p_{2x} - p_{1x})\,R(t_m)$

AND REPEAT WITH

$[t_1, t_m]$ AND $[t_m, t_2]$

END

Geometrically, we are inserting the midpoint on the perpendicular bisector of p_1p_2. (Figure 3).

The two variable line-to-plane function is a simple generalization of the one variable case:

Let $p = h(t, o, e, Rl, R_2)$

with t, o, e the same as before and R_1 and R_2 two independent random real functions on $[a, b]$.

Let p_a and p_b be given and tl = a, $t_2 = b$

BEGIN

IF $t_2 - t_1 > e$ THEN

DEFINE $t_m = (t_1 + t_2) / 2$

$p_{mx} = (p_{1x} + p_{2x}) / 2$
$\quad + o\,((p_{2x} - plx)\,R_2(t_m) -$
$\quad (p_{2y} - p_{1y})\,Rl\,(t_m))$

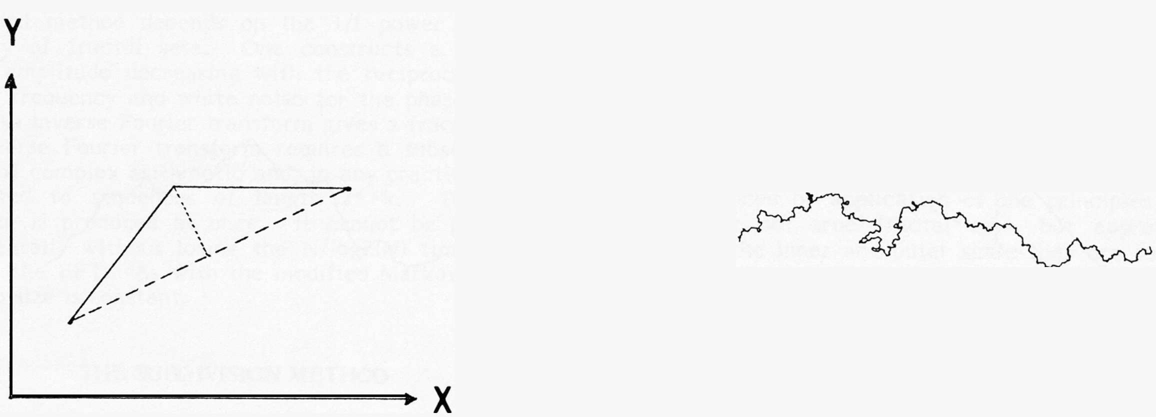
$p_{my} = (p_{1y} + p_{2y}) / 2$
$\quad + o\,((p_{2x} - p_{1x})\,R_1(t_m) -$
$\quad (p_{2y} - p_{1y})\,R_2(t_m))$

AND REPEAT WITH

$[t_1, t_m]$ AND $[t_m, t_2]$

END

In this case the new "midpoint" is located somewhere inside a square box centered on the actual midpoint of the line segment and aligned with it. (Figure 4.)

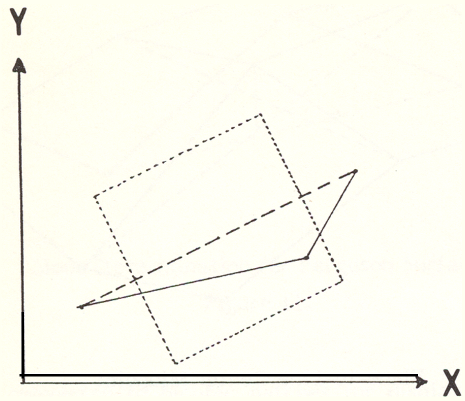


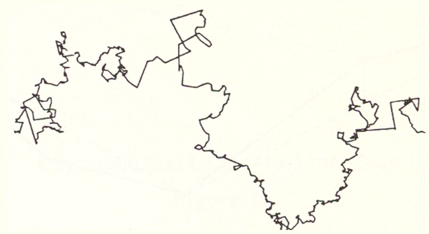
Construction Procedure

Typical Resulting Curve

Line-to-Plane Function # 1

Figure 3.





Construction Procedure

Typical Resulting Curve

Line-to-Plane Function # 2

Figure 4.

Note that we have used $t_2 = t_1 <= e$ to terminate the recursion. This produces a predictable number of points but it may not be visually uniform. A more visually pleasing termination criterion is

$$\|P_2 - P_1\| < e.$$

The infinity norm

$$|P_{1x} - P_{2x}| + |P_{1y} - P_{2y}|$$

is a cheap and sufficient approximation for this procedure.

A FRACTAL PLANE-TO-LINE FUNCTION

We construct a fractal plane-to-line function as:

Let $h = f(u, v, o, e, R)$

where u and v are real parameter values in $[a, b] \times [c, d]$

o and e are the same as before

R is a random real function on $[a, b] \times [c, d]$

The construction is recursive and consists of applying the line-to-line procedure to the edges of the quadrilateral

$$[u_1, v_1], [u_2, v_1], [u_2, v_2], [u_1, v_2].$$

This produces a new "midpoint" along each boundary. (Figure 5.) Then the central point common to the four subsurfaces is computed by applying the line-to-line procedure to the two line segments joining pairs of opposing "midpoints", and adding the result. (Figure 6.) The effect of this process is that of summing two orthogonal independent sets of fractal line-to-line functions.

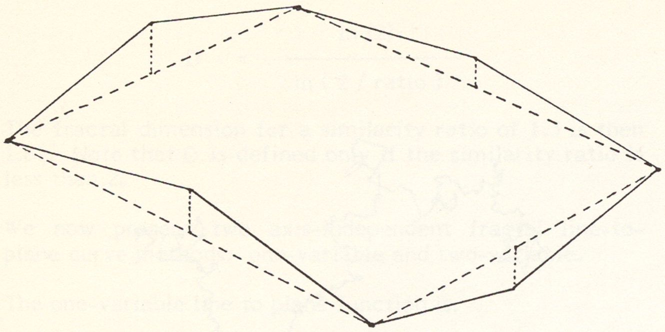
Note that R is a random function and not a random variable or random number generator. This is critically important. When the common edge between two newly created subsurfaces is itself subdivided, the results must be the same for both adjoining surfaces or a gaping hole will be created. (Figure 7.) A simple, fast way to achieve a random function is via table lookup. Random number tables used to create figures for this paper have a Gaussian distribution, zero mean, and variance one. For curves, where a table is not always necessary but is fast, the curve parameter t modulo 10 bits is an adequate table index. Where two completely independent random sequences are needed, they can be obtained simply by circularly offsetting the indexes into the same table. In the case of surfaces with two parameters, the situation is somewhat more complex. One cannot simply add the parameter values since a strong symmetry about the line $u = v$ would result. For example, we would have $R(1,0) = R(0,1)$. The technique used in creating figures 8 to 10 represents u and v as binary fractions. To evaluate $R(u,v)$, v is rotated a few bits, the bytes of u and the rotated v are summed, the result is masked to 10 bits, and used as an index. In reality, almost any hashing method is satisfactory.

For objects constructed of many separate fractal patches the random function must give equal values on common edges between surface elements. The surfaces in Figures 8 to 10 were embedded in a single common parametric range. Another technique for more complex topologies is to define a set of seam functions that provide random numbers only on seams between patches. A seam function needs only to be an offset and a procedure for accessing the random number table.

EXAMPLES OF PLANE-TO-LINE FUNCTIONS

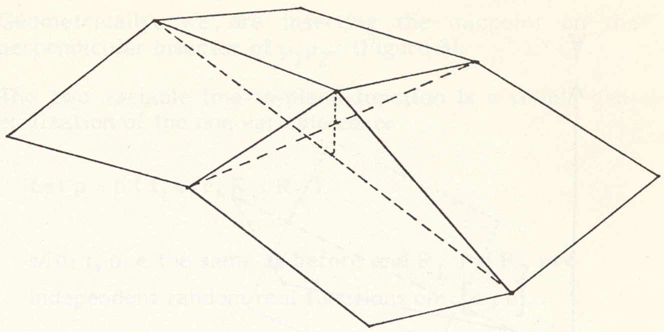
Figures 8, 9 and 10 are examples of fractal plane-to-line functions illustrating different inner geometry. All were generated automatically from the same random seed, offset, and mesh size.

Figure 8 is composed of 1024 bilinear surface patches plus a ground plane polygon. The altitudes of the corners of the bilinear patches were computed by the fractal to-line procedure described above. The bilinear patches



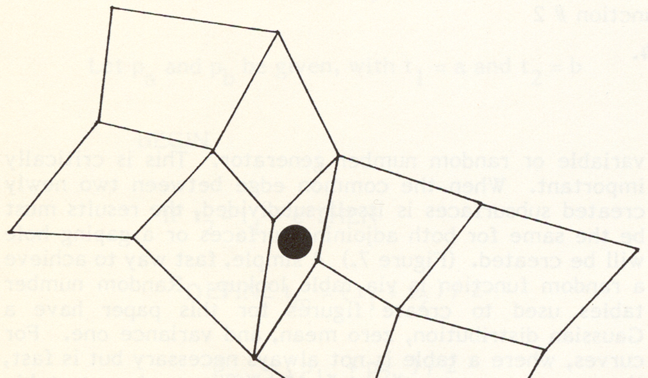
Plane-to-Line Edge Midpoints

Figure 5.



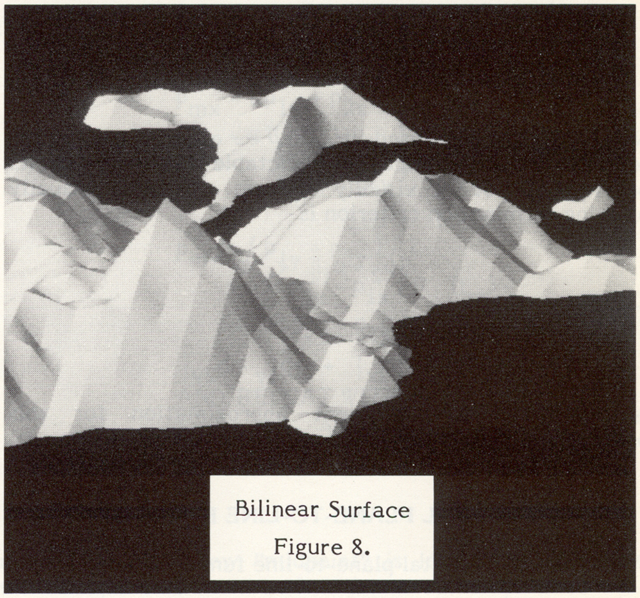
Plane-to-Line Center Midpoint

Figure 6.



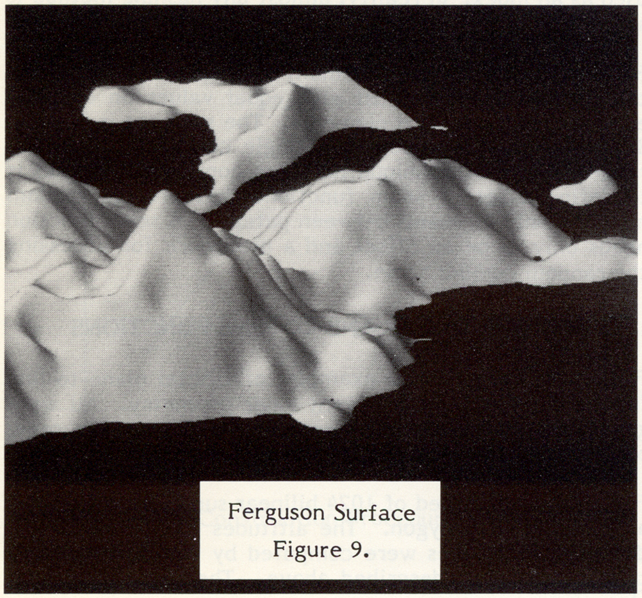
Hole Resulting from Inconsistent Random Numbers

Figure 7.



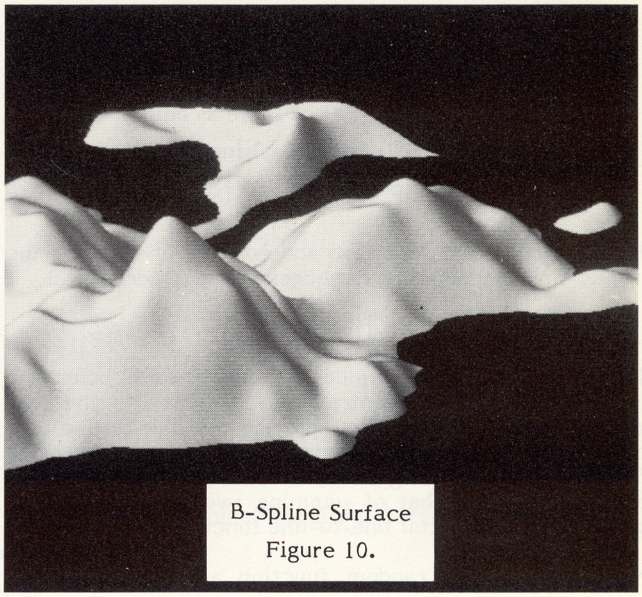
Bilinear Surface

Figure 8.



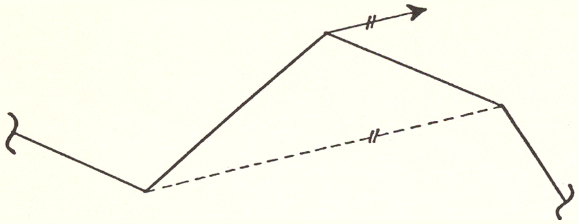
Ferguson Surface

Figure 9.



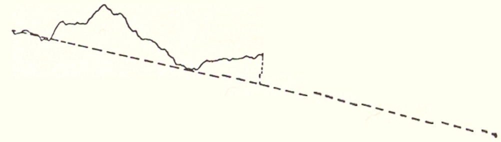
B-Spline Surface

Figure 10.



Slope Determination for Ferguson Surface

Figure 11.



Extrapolation (Line-to-Line Case)

Figure 12.

were converted to bicubic patches for display using the Carpenter/Lane hidden surface algorithm [3].

Figure 9 is composed of 1024 Ferguson I21 patches plus a ground plane polygon. A Ferguson patch surface is made up of Hermite bicubic patches with zero twist and has Cl continuity with its neighbors. The partial derivatives at the corners of individual patches were derived from the differences between the two opposite neighboring corners as in Figure 1 I. This surface is locally flat at the patch corners and exhibits Mach banding.

Figure 10 consists of 1024 bicubic Bezier patches and a ground plane polygon. The patches were derived from a bicubic b-spline surface whose vertices were calculated by a fractal plane-to-line function. The entire surface is C2 continuous everywhere and appears remarkably smooth.

SUMMARY

The example methods in this paper do not generate all fractal curves and surfaces. In fact, the subset they do generate is very small. However, many other subdivision techniques and random functions can be employed to expand the range of realizable forms.

Advantages of the subdivision approach are plentiful. Calculation is cheap, amounting to a few multiplies and adds per point. Curves and surfaces can be made to pass through specified points simply by using the subdivision process to interpolate between the points. Unfortunately, the statistical character, e.g. l/f power spectrum, is lost across boundaries and is the subject of continuing investigation. If the points are not too radically distributed and the offset parameter is not too small, the intermediate interpolation points are not detectible. Curves and surfaces can also be extrapolated. To extrapolate a curve in a particular direction, one determines the location of a new far “endpoint” such that the current far endpoint is a valid “midpoint” of the extended curve. (Figure 12.) Then interpolation between the old endpoint and new “endpoint” fills out the rest of the curve. Using a simple stack, curves and surfaces may be generated all at once, from any boundary, or in any or all directions from somewhere in the middle. The random function guarantees the repeatability of the shape regardless of the direction taken. Many totally different curves and surfaces may be created using the same table by adding offsets to the computed table indices. A fast scanline-like algorithm is possible. And, the step size (in parameter) is 2**(-k) where k is not necessarily constant. This means that the local resolution, or inner scale, can be made dependent on the local magnification, thus more efficiently distributing computer power over the image.

REFERENCES

1. Besicovitch, A. S., Ursell, H. D., Sets of fractional dimensions (V): On dimensional numbers of some continuous curves. Journal of the London Mathematical Society 12, (1937), 18-25.

2. Ferguson, J. C., Multivariable curve interpolation. J. Assoc. Comput. Mach. 2, (1969), 221-228.

3. Lane, J. M., Carpenter, L. C., Whitted, T., Blinn, J. F., Scan line methods for displaying parametrically defined surfaces. Communications of the ACM. 23 (Jan. 1980), 23-34.

4. Mandelbrot, B. B., Computer experiments with fractional Gaussian noises. Water Resources Research. 5, (1969), 228.

5. Mandelbrot, B. B., A fast fractional Gaussian noise generator. Water Resources Research 7, (1971), 543-553.

6. Mandelbrot, B. B., Stochastic models for the Earth’s relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islan [redacted] Proc. Nat. Acad. Sci. USA. 72, (1975), 3825-3828.

7. Mandelbrot, B.B., FRACTALS: Form Chance, and Dimension. W. H. Freeman Company. San Francisco, (1977).

8. Richardson, L. F. The problem of contiguity: an appendix of statistics of deadly quarrels. General Systems Yearbook 6, 134-187.