

Self-affinity and fractal dimension

- *Chapter foreword.* Mathematicians prefer to construct recursive fractals by interpolation ad infinitum. The historical reason is that fractals arose in the study of local irregularity. A continuing reason is that prefractal finite interpolations converge “strongly” to a fractal limit. Physicists, to the contrary, prefer extrapolation because matter is made of nonvanishing particles so that infinite interpolation has no meaning, and because “practical people” are not worried by flavors (strong or weak) of convergence.

For self-similar fractals, both procedures yield the same fractal dimension. To the contrary – as first observed in M 1985s, the paper reproduced in this chapter – the local (interpolative) and global (extrapolative) fractal dimensions of self-affine fractals can take distinct values.

The original also pioneered in showing that certain difficult new ideas concerning self-affinity are best accepted with the help of recursive “cartoon” constructions of total artificiality. For self-similar fractal curves, the same path was first taken when M1967s faced the question of “How long is the coast of Britain?” by injecting the Koch snowflake curve and more general recursive constructions. Since 1986, I realized that the relation between a fractal model and a recursively constructed artificial surrogate is subtle and permanent, and deserves a stronger term than “variant” or “pedagogical stand-by.” The term selected in M 1997F and M1999N and carried over to this book is “cartoon;” it was retrofitted in this and the subsequent reprints wherever appropriate.

For an excellent exposition of these ideas, see Voss 1985. •

◆ **Abstract.** To define a “divider exponent” for a fractal curve, one walks a divider of decreasing length δ , then evaluates the resulting “approximate length” $L(\delta)$, which is proportional to δ^{1-D} . Long ago, M 1967s showed that for a self-similar curve, the divider exponent D coincides with all other forms of the fractal dimension, e.g., the similarity, box or mass dimensions. But for a self-affine curve, for example, a scalar Wiener Brownian record $B(t)$, a full description in terms of the fractal dimension is complex. Each dimension splits into a *local* and a *global* value, separated by a crossover. Globally, all the basic methods of evaluating the global fractal dimension of $B(t)$ yield 1; that is, a self-affine fractal behaves globally as if it were not fractal. Locally, the box and mass dimensions are 1.5, but the divider dimension is $D=2$. More generally, for a fractional Brownian record $B_H(t)$ (a model of vertical cuts of relief), the global fractal dimensions are 1, several local fractal dimensions are $2-H$ and the local divider dimension is $1/H$. This is the fractal dimension of a self-similar fractal trail in the plane, a curve implicit in the definition of the record of $B_H(t)$. ◆

THIS PAPER DESCRIBES A NEW OBSERVATION that is elementary yet both practically and theoretically significant. It will explain certain odd or inconsistent results of measurement and shall add to our understanding of the notion of fractal dimension.

1. Introduction and summary

It is known that there are several ways of measuring a fractal dimension. This means that several alternative definitions exist, but in the extensively studied case of strictly self-similar fractals all these definitions yield the same value. This paper describes the new and more complex situation that prevails for fractal curves that are *not* self-similar but self-affine.

Two practical examples of such curves are vertical cuts through either a relief or a surface of non-isotropic metal fracture and records of electric noise. The results are of wide validity, but the arguments are carried out (as described in Section 2) on “records of functions.” In this usage – which does not follow M1982F{FGN} and may be new – “record” is contrasted with “trail.” When a point moves in the plane (x,y) , the trail is the set of points (x,y) that have been visited, and the records are the sets of points $(t,x(t))$ and $(t,y(t))$.

The main examples will be records of the scalar Brownian motion $B(t)$, of the more general fractional scalar Brownian motion $B_H(t)$ (whose parameter H satisfies $0 < H < 1$) and of related fractal “cartoons” that simplify the discussion and are introduced and discussed here for the first time. (They relate to the new “interpolable random walk,” which I intend to explore in detail elsewhere.) The discussion covers three algorithms that I introduced while setting up fractal geometry. We shall use for them a new streamlined terminology adopted since *FGN*.

Two sentences suffice to deal with the “(self-)similarity dimension” D_S . This notion applies to self-similar sets, which are made of N parts, each obtained from the whole by a reduction of ratio r . For these sets, $D_S = \log N / \log(1/r)$. For self-affine sets, D_S simply cannot be evaluated.

This paper's *first finding* (Section 4) concerns the self-affine curve's “box dimension” D_B . The local value (using small boxes) is $2 - H$, which coincides with its Hausdorff-Besicovitch dimension, itself a local property.

The *second finding* (Section 5) concerns the “divider dimension” D_C , defined by walking a divider along a curve. Although it may be theoretically meaningless, it can always be evaluated mechanically. It will be shown that for self-affine fractal records, the local value (using small divider openings) is $D_C = 1/H$. This value differs from the box dimension $2 - H$ – except in the degenerate limit case $H = 1$. Conversely, $1/H$ coincides with the small box and Hausdorff-Besicovitch dimensions of an important self-similar fractal curve that is implicit in the definition of $B_H(t)$: it is the curve in E -dimensional Euclidean space whose E coordinate records ($E > 1/H$) are independent realizations of $B_H(t)$.

This paper's *third finding* (Section 6) concerns the “mass dimension” D_M . Again, it may be theoretically meaningless but always can be evaluated mechanically. The local value (using small radii) is $D_M = 2 - H$.

The *fourth and final finding* is that all the above-mentioned local dimensions are high frequency limits. The corresponding low frequency limits are all equal to 1. In the long run, our self-affine fractals are one-dimensional! The point of crossover from 1 to either $2 - H$ or $1/H$ is shown to depend on the ratio of units of t and of B ; in general, this ratio is arbitrary. The biases that may result are investigated.

Motivation for this study. M, Passoja and Paullay 1984 studied fracture surfaces of metals but did not use the divider dimension. I was asked why, and that question triggered the present investigation.

2. Self-affine sets: definition and examples of records of Brownian motion, of fractional Brownian motion, and of cartoons

Wiener's scalar Brownian motion $B(t)$ is the process starting with $B(0) = 0$ and having the following property: for every collection of non-overlapping intervals Δt , the increments of $B(t)$ are independent and stationary Gaussian variables $B(t)$ has the following well-known invariance property:

The random processes $B(t)$ and $b^{-1/2}B(bt)$ are identical in distribution for every ratio $b > 0$.

Since the rescaling ratios of t and of B are different, the transformation from $B(t)$ to $b^{-1/2}B(bt)$ is an "affinity." This is why $B(t)$ was called "statistically self-affine" in M 1982F{FGN} (page 350).

The more general fractional Brownian motion $B_H(t)$, where $0 < H < 1$, plays a very important role in fractal geometry. If $B_H(0) = 0$, the random processes $B_H(t)$ and $b^{-H}B_H(bt)$ also are identical in distribution. For the value $H = 1/2$, one obtains $B(t)$ as a special case of $B_H(t)$.

Unfortunately, a rigorous study of $B_H(t)$ requires difficult arguments. This and related pedagogical needs made it desirable to have a cartoon of $B_H(t)$ whose rigorous study is elementary; this led me recently to introduce a series of cartoons of $B_H(t)$.

First, let me describe the essential properties of the cartoons as exemplified in Figure 1. More generally, a limit process $M_H(t)$ is defined when H is of the form $H = \log b' / \log b''$, where the integer bases b' and b'' are such that $b' - b''$ is positive and even. When the time increments belongs to the construction grid and $\Delta t = b'^{-k}$ the increment of $M_H(t)$, is binomial with mean 0 and standard deviation $(\Delta t)^H$. That is

$$M_H(pb'^{-k}) - M_H[(p+1)b'^{-k}] = \pm (b'')^{-k} = \pm (b'^{-k})^H = \pm (\Delta t)^H$$

holds for all k and p . Clearly, $M_H(pb'^{1-k})$ is a multiple of b''^{-k} . The linear interpolation between these values of $M_H(t)$ is the k -th approximant of $M_H(t)$, to be denoted by $M_H^{(k)}(t)$.

Actual construction of $M_H(t)$. The details are not important here, but one can construct $M_H(t)$ as an interesting fresh example of the "multiplicative chaos" procedure that I pioneered in M1972j{N14} and in M1974f{N15}, (see also M 1982F{FGN} p. 278 ff.). (Recently, this procedure has been rediscovered in part in Hentschel & Procaccia 1983.) The building blocks are "multiplicative effect functions" $\mu_k(t)$ defined as follows. For all k and t ,

$|\mu_k(t)| = b'/b''$, and each interval $[p$ to $p + 1]$, where p is an integer, splits into b' subintervals. In $(1/2)(b' + b'')$ of those subintervals chosen at random, one sets $\mu_k(t) > 0$, and in the remaining $(1/2)(b' - b'')$ subintervals, one sets $\mu_k(t) < 0$. This insures that, for all integers k and p ,

$$\int_p^{p+1} \mu_k(t) = 1.$$

For example, if $b' = 4$ and $b'' = 2$, then $\mu_k(t) > 0$ over three subintervals and < 0 over one. See Figure 1 for illustration. Now pick statistically independent functions $\mu_k(t)$, and form the product

$$M^{(k)}(t) = \prod_{n=-\infty}^k \mu_n(b'^n t).$$

Then integrate to obtain the prefractal approximant

$$M^{(k)}(t) = \int_0^t M^{(k)}(s) ds.$$

Finally take the limit $M(t) = \lim_{k \rightarrow \infty} M^{(k)}(t)$. Although this limit depends on b' and b'' , the present discussion considers only the value of H ; hence, the limit will be denoted by $M_H(t)$.

3. The fractal dimension of the above self-affine cartoons

The graph of zeros of $B(t)$ is widely known to have a Hausdorff-Besicovitch dimension equal to $1/2$. It is almost as widely known that for the graph of $B(t)$ itself the Hausdorff-Besicovitch dimension is equal to $1.5 = 1/2 + 1$. The corresponding dimensions for the records of both $B_H(t)$ and $M_H(t)$ are $1 - H$ and $2 - H$. But the Hausdorff-Besicovitch dimension is a very nonintuitive notion; it can be of no use in empirical work, and is unduly complicated in theoretical work, except in the case of self-similar fractals. As a replacement, my work introduced (in an increasingly formal fashion) several alternative definitions that are useful precisely because they lack generality. For self-similar sets, the values yielded by these various definitions of dimension were identical. But, as we move on to self-affine shapes, we shall find that local and global values must be dis-

tinguished for each dimension and that the various local values *cease to be identical*.

The reason is fundamental: the notions of "square," "distance" and "circle" are vital in the usual "isotropic" geometry but meaningless in affine geometry. More precisely, they are meaningless when the quantities plotted along the t and B -axis are distinct and their increments Δt and ΔB

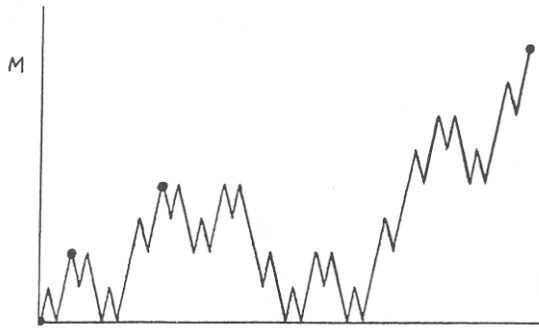


FIGURE C21-1. A prefractal approximant of a non-random self-affine "cartoon" curve. This broken line of $4^3 = 64$ intervals is the second approximant of a nonrandom prototype of the random function $M_H(t)$, which Section 2 introduces as "cartoon" of the Wiener Brownian motion. The bases in this example are $b' = 4$ and $b'' = 2$, hence $H = 1/2$. The construction is recursive. The left most four intervals, which end with dots and proceed $UDUU$ (that is: up, down, up and up), form the generator of this example. The $4^2 = 16$ left-most intervals (again ending with dots) illustrate the second approximant.

In order to randomize this function, each occurrence of the generator is chosen at random from the four possibilities $DUUU$, $UDUU$, $UUDU$ and $UUUD$; for some purposes it is best to narrow the choice to $UDUU$ and $UUDU$.

cannot be compared or combined. When there is no intrinsic meaning to the notion of equal height and width, a square cannot be defined. Similarly, a circle cannot be defined because its squared radius $R^2 = \Delta t^2 + \Delta B^2$ would combine the units along both axes. Furthermore, it becomes impossible to “walk a divider” along a self-affine curve because the distance covered by each step combines a Δt and a ΔB .

A complicating factor is that in the case of vertical cross-sections of a relief, both Δt and ΔB are lengths. Also, the purely affine plane of a noise record is always represented on the same graph paper as an isotropic plane. The above distinction is therefore elusive, and it is impossible to avoid the temptation to draw circles and squares, to walk dividers and to evaluate various “prohibited” dimensions “mechanically.” Sections 5 and 6 describe the results thus obtained.

4. The box dimension is meaningful for the records of $B_H(t)$ and the cartoon $M_H(t)$; its local value is $2 - H$, and its global value is 1

Having covered the unit circle with a lattice of boxes of side $r = 1/b$, let $N(b)$ denote the number of boxes in this lattice that intersect the set. When $N(b)$ behaves like $N(b) \propto b^{D_b}$, the exponent is the box dimension.

In mathematical discussions of dimension, the symbol \propto refers to local behavior and is a shorthand for $\lim_{b \rightarrow \infty} \log N(b) / \log b = D_B$. When the set is bounded, one begins by drawing the whole set within a unit square of the plane. When the set is unbounded, one considers bounded portions obtained as intersections with squares.

The box argument for the records of $B(t)$ or $B_H(t)$, as given in M 1982F{FGN} (bottom left of p. 237), is heuristic and is not readily made rigorous. For the record of $M_H(t)$, on the contrary, the exact argument is transparent. To cover this fractal curve from $t = 0$ to $t = 1$ with boxes of side $1/b = b'^{-k}$, one needs $b^k = b$ stacks of boxes, each with a height between b''^{-k} and $b''^{-k}[(1/2)(1 + b'/b'')]$.

Thus, apart from multiplication by a factor of order 1, one has $N \sim b^k(b''^{-k})/b'^{-k} = (b^2 b''^{-1})^k$. From $H = \log b'' / \log b'$, we have $b'' = b'^H$, hence $N = b^{2-H}$. The multiplicative factor vanishes when taking $\lim_{b \rightarrow \infty} \log N / \log b$, and hence $D_B = 2 - H$. Observe that in this high frequency limit the scales chosen for t and B do *not* matter.

The physicist, however, also thinks of the global limit $b \rightarrow 0$ or $r \rightarrow \infty$, which requires an unbounded record. The portion of a self-affine record from 0 to $t \gg 1$ is covered by a single box. Hence $\lim_{b \rightarrow 0} \log N(b) / \log b = 1$. (The detailed argument requires some care; we

will not dwell on it here). In conclusion, two limits that are identical for self-similar fractals are now found to differ!

Thus, a self-affine curve involves a crossover time increment, call it t_c , such that $B_H(t+t_c) - B_H(t) \propto t_c$. Stated alternatively, the most intrinsic units of t and B_H are such that $B_H(t+1) - B_H(t) \sim 1$.

Terminology. "Box dimension" is a fresh abbreviation for "box-counting dimension." In the case of self-similar fractals, I occasionally called D_B the "similarity dimension," a term I now regret using because it does not carry over to the self-affine case.

Several writers call D_B a "capacity"; this term conflicts with two competing meanings. The first is a term Kolmogorov based on a (non-obvious) analogy with information-theoretical capacity. The second, Frostman dimension, involves the potential-theoretical capacity and is about to prove central to fractal analysis. {P.S. 1999: Unfortunately, this last expectation, which was motivated by the fractal character of the distribution of galaxies, is slow to be implemented. But the good news is that "capacity dimension" has ceased to be used.}

5. The divider dimension cannot always be evaluated for the records of $B_H(t)$ or $M_H(t)$. Its value evaluated mechanically for small η is $1/H$. This is not the fractal dimension of the record of $B_H(t)$ but of a trail related to this record. For large η , the divider dimension is 1

"Divider dimension" is my present term for a notion that applies only to curves having the following property: the length $L(\eta)$ measured by "walking a divider of opening η " along the curve behaves like $L(\eta) \propto \eta^{-D_c}$. I developed this notion to interpret Richardson's empirical findings on geographical coastlines, which are *horizontal* cuts of the relief and are self-similar. The temptation is irresistible to use the same technique for *vertical* cuts of the same relief.

When $\eta \gg t_c$ (e.g., when the intrinsic unit of B_H is very small), the record is effectively a horizontal line. The divider walked along the curve remains mostly parallel to the t -axis, and $L(\eta)$ varies little. This $L(\eta)$ yields $D_c = 1$, irrespective of the value of H .

When $\eta \ll t_c$ (e.g., when the intrinsic unit of B_H is huge), the divider walked along the curve remains mostly parallel to the B -axis. The cartoon $M_H(t)$ yields $D_c = 1/H$ with little algebra. For example, let $b' = 4$ and $b'' = 2$, yielding the Brownian $H = 1/2$. When k is large and $\eta = 2^{-k}$, the quantity $L(\eta)/\eta$, which is the number of divider steps, is seen to be exactly equal to $4^k = \eta^{-2}$, and hence $D_c = 2$. For more general values of $H = \log b'' / \log b'$,

one finds that for small η , $L(\eta)/\eta$ is multiplied by b' when η is divided by b'' . This yields $D_c = 1/H$.

At first blush, this is an extremely strange value. First, it can exceed 2 (in fact, it can be arbitrarily large), which is impossible for the Hausdorff-Besicovitch dimension of a self-avoiding curve in the plane. Second, $1/H$ disagrees with the value $2 - H$ that the other local definitions of the fractal dimension give for the cartoons.

However, readers familiar with the fractional Brownian motion will recognize $1/H$ as being the fractal dimension of the trail (in an E -dimensional Euclidean space \mathbb{R}^E satisfying $E > 1/H$) of a motion whose E coordinates are independent realizations of $B_H(t)$.

In this case, an attempt to measure the fractal dimension for one set actually measures it for a different set. The heuristic argument that follows suggests that this outcome should have been expected. As already mentioned, if $\eta \ll t_c$, our divider remains most of the time nearly parallel to the B -axis; therefore, our experiment nearly collapses time by nearly flattening the record into a trail along the B -axis. Suppose our scalar B is one coordinate of a vectorial \mathbf{B} in a space \mathbb{R}^E of Euclidean dimension $E \gg 1/H$. If a divider of opening η is walked along the trail of \mathbf{B} , the steps' projections on any coordinate axis will differ in size, but most will be close to η/\sqrt{E} . Now measure the length of the trail of \mathbf{B} with steps of length η subjected to the additional constraint so that their projection on one axis is *exactly* η/\sqrt{E} . A moment of thought suggests that this last constraint will not have much effect on the number of steps. Thus (apart from a numerical correction factor dependent on E), walking a divider takes about as many steps along our collapsed record as along the trail in a space of E dimensions. In conclusion, the divider dimension should indeed be the same in both cases, that is, $1/H$.

Remark on the "fracton/spectral" dimension. The box and the Hausdorff-Besicovitch dimensions of the zeros of $B_H(t)$ are $1 - H$, and Section 5 deduces a set of dimension $1/H$ from a set of dimension $1 - H$. In a more familiar context, one starts with the trail of $B_H(t)$ of dimension $1/H$ and deduces $1 - H$ as the dimension of the instants of this motion's recurrence to the origin. A further step led physicists to call $2H$ *fracton dimension* (Alexander & Orbach 1982) or *spectral dimension* (Rammal & Toulouse 1982). Therefore, we deal here with a situation that is the converse of the usual one. The question of whether this remark generalizes to processes restricted to a fractal curve (e.g., a percolation cluster) remains to be examined.

6. The mass dimension cannot always be evaluated for the records of $B_H(t)$ or $M_H(t)$. When its value is estimated mechanically, it is $2 - H$ for small R and 1 for large R

“Mass dimension” is my present term for a notion I had devised for sets having the following property: the mass $M(R)$ contained in the intersection of the set with a disc or ball of radius R behaves like $M(R) \propto R^{D_M}$. The disc or ball can be replaced by a square or cube whose sides are parallel to the axes and of length $2R$.

As already mentioned, the notions of “square” and “circle” are meaningless in affine geometry. Nevertheless, we must tackle the practical questions that arise after a self-affine fractal has been drawn on ordinary graph paper. The “mass” of the record of $B_H(t)$ between times t' and t'' is set to be $|t' - t''|$.

When $R \gg t_C$, the record of $B_H(t)$ is effectively a horizontal interval. It occupies a very thin horizontal slice of the square of side $2R$, hence, $M(R) \propto R$ and $D_M = 1$.

When $R \ll t_C$, the record of $B_H(t)$ is effectively a collection of vertical intervals, one for each zero of $B_H(t)$. Again, the argument is simplified if we replace $B_H(t)$ by the cartoon $M_H(t)$ and consider a square of side $R = b''^{-k}$, with top and bottom ordinates proportional to b''^{-k} . The mass we seek is the same as for the k -th approximant function $M_H^{(k)}(t)$. Thus, mass is the number of times $M_H^{(k)}(t)$ traverses the ordinate of the center of a square, multiplied by the duration δ of each traversal. The number of traversals is $\propto (R/\delta)^{1-H}$, and $\delta = b''^{-k} = (b''^{-k})^{1/H} = R^{1/H}$. Hence $M(R) \propto R^{2-H}$, yielding the familiar value $D_M = 2 - H$ in the small R limit.

7. Crossover pitfalls

To summarize, self-affine fractals do *not* involve a single set of exponents $D_{B'}$, D_D and $D_{M'}$ such that $N(b) \propto b^{D_b}$, $L(\eta) \propto \eta^{D_D}$ for all η and $M(R) \propto R^{D_M}$ for all R . Different exponents are approached as one moves in opposite directions away from the crossover point t_C . And the value of t_C is not always intrinsic, since in the case of noises it depends on the units chosen along the axes. A truly mechanical estimate of $D_{B'}$, D_M or D_C is likely to combine values of η or of R that range on both sides of the crossover t_C . Such an estimate will depend on exactly where t_C lies in the range of η or R . Therefore it will depend on the units of t and B . It will be worthless. Reliability is improved by exaggerating the vertical scale.

volume were not edited before they were reprinted. The reprints in this *Selecta* book should be easier to read.

How this paper came to be written. The distinction between self-similar and self-affine fractals is very clearly made in the *Mathematical Addenda* of M 1977F (Chapter 12, p. 276) and in M 1982F{FGN} (Chapter 39, p. 350). For a long time it was best not to belabor this distinction but the need to face it became irresistible the spring of 1985, when I was at Harvard in the Mathematics Department, teaching a course called MA195. This paper was written during this time. An excellent survey, Voss 1986, is reprinted in Family & Vicsek 1991.

The summer that followed revealed that, for diverse reasons, several other authors had also been thinking of this topic. Mathematical papers include Kono 1986 and Kamae 1986. Quite independently, the physicist P. Wong (see Family & Vicsek 1991) had become worried by his encounters with shapes that should obviously be called fractals yet whose behavior appeared "anomalous."

Now we see that, while vaguely aware of the fact that a fractal need not be self-similar, all too many scientists were unknowingly drawn to think always in terms of self-similar fractals. I must count myself among those careless scientists, who sometimes spoke of "fractals" when they meant "self-similar fractals." For example, I often wrote that the fractal dimension can have many definitions, but always takes a single value when a set is well-defined. When penning this assertion, there seemed to be no need to say that the discussion concerned self-similar fractals.

Cases of near-simultaneous involvement and discovery are, of course, a cliché in the sciences. But in fractal geometry, the serious advent of self-affinity was the first case of near-simultaneity. The discovery of multifractals involved nothing close to near-simultaneity, as seen in M1999N, Chapter N2. Near-simultaneity characterized the direct continuation of my old work in Frisch & Parisi 1985 (an excerpt appears in M 1999N, Section 5.5.2) and soon later its rediscovery by other writers. However, I had been quite alone in studying multifractals in the late 1960s and the early 1970s.