Self-Affine Fractals and Fractal Dimension

To cite this article: Benoit B Mandelbrot 1985 *Phys. Scr.* **32** 257

Physica Scripta. Vol. 32, 257–260, 1985.

Self-Affine Fractals and Fractal Dimension

Benoit B. Mandelbrot

Physics Department, IBM Research Center, Mathematics Department, Harvard University, Cambridge, MA, U.S.A.

Received June 13, 1985; accepted June 17 1985

Abstract

Evaluating a fractal curve's "approximate length" by walking a compass defines a "compass exponent." Long ago, I showed that for a self-similar curve (e.g., a model of coastline), the compass exponent coincides with all the other forms of the fractal dimension, e.g., the similarity, box or mass dimensions. Now walk a compass along a self-affine curve, such as a scalar Brownian record $B(t)$. It will be shown that a full description in terms of fractal dimension is complex. Each version of dimension has a local and a global value, separated by a crossover. First finding: the basic methods of evaluating the global fractal dimension yield 1: globally, a self-affine fractal behaves as it if were not fractal. Second finding: the box and mass dimensions are 1.5, but the compass dimension is $D = 2$. More generally, for a fractional Bownian record $B_H(t)$, (e.g., a model of vertical cuts of relief), the global fractal dimensions are 1, several local fractal dimensions are 2-H, and the compass dimension is $1/H$. This $1/H$ is the fractal dimension of a self-similar fractal trail, whose definition was already implicit in the definition of the record of $B_H(t)$.

This Note describes an observation that is elementary, yet of both practical and theoretical significance: it explains certain odd or inconsistent results of measurement and adds to our understanding of the notion of fractal dimension.

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

The practical examples in the background are, on the one hand, vertical cuts through either relief or a surface of non-isotropic metal fracture and, on the other hand, records of electric noise. The results are of wide validity, but – to simplify – the arguments are carried out (as described in Section 2) on "records of functions." In this usuage – which seems 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 will be 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 motions that simplify the discussion and are introduced and discussed here for the first time. (They relate to the new "interpolable random walk," which is explored in detail elsewhere). The discussion covers three algorithms that I introduced while setting up fractal geometry, first informally, then increasingly organized fashion.

At first finding (Section 4) concerns the self-affine curve's "box dimension" D_B (in a new streamlined terminology I adopted after [1]). The local value (using small boxes) is 2-H, which coincides with its Hausdorff–Besicovitch dimension, itself a local property.

Two words 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 or ratio r. For these sets, $D_S = \log N/\log(1/r)$. For self-affine sets, D_S simply cannot be evaluated.

The second finding (Section 5) concerns a notion I now call "compass dimension," D_C. It will be shown that for self-affine fractal records, the compass dimension may be theoretically meaningless, but can alaways be evaluated mechanically. The local value (using small compass openings) is $D_C = 1/H$. This value differs from 2-H (except in the excluded limit case $H = 1$, which is degenerate.) On the other hand, $1/H$ coincides with the small box and Hausdorff–Besicovitch dimensions of an important fractal curve other than the record under study: it is the trail of a fractional Brownian motion in E-dimensional Euclidean space, namely, of a self-similar fractal curve that is implicit in the definition of $B_H(t)$, since its E coordinate records ($E > 1/H$) are independent realizations of $B_H(t)$.

The third finding in this Note (Section 6) concerns the quantity I now call "mass dimension" D_M. Again, it may be theoretically meaningless, but can always be evaluated mechanically. The local value (using small radii) is $D_M = 2$-H.

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

Motivation for this study. When studying the fracture surfaces of metals [2], we did not use the compass dimension. I was asked why, which triggered the present investigation.

2. Definition of self-affinity and examples of self-affine sets: the records of Brownian motion, of fractional Brownian motion, and of a new pedagogical variant

Wiener's scalar Brownian motion $B(t)$ is the process within independent and stationary Gaussian increments. It has a well-known invariance property: setting $B(0) = 0$, the random processes $B(t)$ and $b^{-1/2}B(bt)$ are identical in distribution for every ratio $b > 0$. One observes that the rescaling ratios of t and of B are different, hence the transformation from $B(t)$ to $b^{-1/2}$ $B(bt)$ is an "affinity." This is why $B(t)$ was called "statistically self-affine" [1, page 350].

A very important role is played in fractal geometry by the more general fractional Brownian motion $B_H(t)$, where $0 < H < 1$. If $B_H(0) = 0$, the random processes $B_H(t)$ and $b^{-H}B_H(bt)$ are identical in distribution. The value $H = 1/2$ brings $B(t)$ 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 variant of $B_H(t)$ whose rigorous study is elementary; this led me recently to introduce a series of variants of $B_H(t)$ that relate to a random walk one can interpolate.

First, let me describe the essential properties. $M_H(t)$ is defined when H is of the form $H = \log b'/\log b''$, the integer bases b' and b'' being such that $b' - b''$ is positive and even. The idea is that the function $B_H(t)$, whose increments are Gaussian over all Δt, is replaced by a function $M_H(t)$, whose increments over suitable Δt's are binomial with the same mean 0 and the same standard deviation. The requirement is that

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

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

Actual Construction of $M_H(t)$

The details do not really matter here, but are an interesting fresh example of the "multiplicative chaos" procedure I had pioneered in 1972 and in 1974 [1, pages 278 ff.]. (The earlier uses of this procedure generate certain fractal measures of great current importance that have been rediscovered in part by I. Procaccia.) 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 between the sucessive integers of the form p to $p+1$ splits into b' subintervals; in $(1/2)\ (b' + b'')$ subintervals, chosen at random in each interval separately from the other, 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,

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

For example, if $b' = 4$ and $b'' = 2$, $\mu_k(t) < 0$ over one of four subintervals. See fig. 1 for illustration. Now pick statistically independent functions $\mu_k(t)$, and form

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

then integrate to obtain the approximant

$$M^{(k)}(t) = \int_0^t M'^{(k)}(\Delta)\mathrm{d}s$$

and finally take the limit: $M(t) = \lim_{k\to\infty} M^k(t)$. Though this limit depends on b' and b'', the present discussion only involves the value of H, hence the notation $M_H(t)$.

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

It is widely known that for the graph of zeros of $B(t)$, the Haufsdorff–Besicovitch dimension is 1/2, and almost as widely known that for the graph of $B(t)$ itself the Hausdorff–Besicovitch dimension is $1.5 = 1/2 + 1$. 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 non-intuitive notion. While I was delineating and developing the new fractal geometry, I used to pay lip service to it, but I was careful to describe this as a "tactical" and "tentative" step, because this dimension can be no use in empirical work, and is unduly complicated in theoretical work, except for self-similar fractals. Instead, my work introduced (in increasingly formal fashion) several alternative definitions that are useful precisely because they lack generality. For self-similar sets, the values yielded by these dimensions were identical. Now that we move on to self-affine shapes, we shall find that local and global values must be distinguished for each dimension, and that the different local values *cease to be identical*.

The reason is fundamental: "square," "distance," and "circle," are vital notions in "isotropic" geometry, but they are meaningless in affine geometry. More precisely, they are meaningful for relief cross-sections, but are meaningless for noises, because the units along the t axis and along the B-axis are set up independently of each other, hence Δt and ΔB cannot be combined. There being no intrinsic meaning to the notion of equal height and width, a square cannot be defined. Similarly, a circle cannot be defined, because its square radius $R^2 = \Delta t^2 + \Delta B^2$ would have to combine the units along both axes. Furthermore, one cannot "walk a compass' along a self-affine curve, because the distance covered by each step combines a Δt and a ΔB.

On the other hand, a relief cross-section (while self-affine) is a curve in an isotropic plane. And a noise record's purely affine plane is always represented on the same graph paper as an isotropic plane. This causes the above distinction to be elusive, and creates the temptation to draw circles and squares to walk compasses, 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 $M_H(t)$, and its local value is the "correct" 2-*H*. Its global value is 1

After a lattice made of boxes of side $r = 1/b$ is made to cover a set, let $N(b)$ denote the number of boxes in this lattice than intersect the set. "Box dimension" is my present term for a notion that applies to sets for which $N(b)$ behaves like $N(b) \propto b^{D_B}$. What does this mean? It may (as it does for most mathematicians) refer to local behavior, and mean that $\lim_{b\to\infty} \log N(b)/\log b = D_B$. When the set is bounded, one begins by drawing it 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 [1, bottom left of page 237] is heuristic and is not readily made rigorous. For the record of $M_H(t)$, to the contrary, the exact argument is transparent: To cover our fractal 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 a multiplying factor of the order 1, one has $N \sim b'^k(b''^{-k}/b'^{-k}) = (b'^2b''^{-1})^k$. From $H = \log b''/\log b'$, we have $b'' = b'^H$, hence $N = b^{2-H}$. The multiplying factor vanishes when taking $\lim_{b\to\infty} \log N/\log b$, hence $D_B = 2\text{-}H$. Observe that the scales chosen for t and B do *not* matter in this high frequency limit.

The physicist, however, also thinks of the global limit $b \to 0$ or $r \to \infty$, 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 $\lim_{b\to 0} \log N(b)/\log b = 1$. (The detailed argument requires some care, but we shall not dwell on it). Conclusion: two limits that are identical for self-similar fractals are now found to differ!

Thus, a self-affine curve involves a cross-over value of t, call it t_c, defined as being such that $B_H(t + t_c) - B_H(t) \sim t_c$. Stated

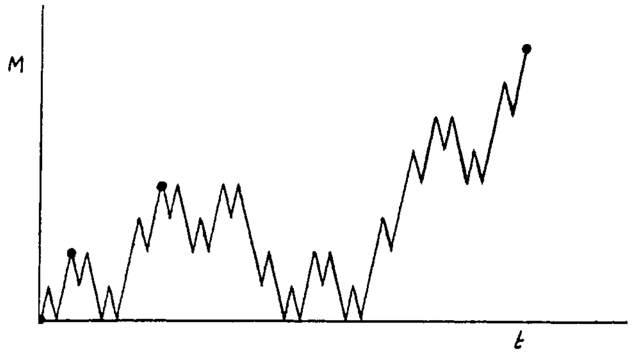

Fig. 1. This broken line of $4^3 = 64$ intervals is the second approximant to a non-random prototype of the random function $M_H(t)$ which Section 2 introduces as a new pedagogical "stand-by" for Wiener's Brownian motion. The bases in this example are $b' = 4$ and $b'' = 2$, hence $H = 1/2$. The constructure is recursive. The leftmost four intervals, which end with dots and proceed UDUU (that is: up, down, up and up), form the generator of this example. The $4^2 = 16$ leftmost intervals (again ending with dots) illustrate the second approximant.

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

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 "similarity dimension," a term I now regret using, because it does not carry over to the self-affine case. Several writers call D_B "capacity," which conflicts with a term Kolmogorov had based on a (non-obvious) analogy with information – theoretical capacity, and with Frostman's dimension, which involves the potential-theoretical capacity and is about to prove central to fractal analysis.

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

"Compass 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 compass of opening η" along the curve behaves like $L(\eta) \propto \eta^{-D_C}$. I developed this notion in order to interpret Richardson's empirical findings about 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 a same relief.

When $\eta \gg t_c$ (e.g., when the unit of B_H is very small), the record is effectively a horizontal line. The compass 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 H.

When, to the contrary, $\eta \ll t_c$ (e.g., when the unit of B_H is huge), the compass walked along the curve remains mostly parallel to the B-axis. The pedagogical variant $M_H(t)$ gives the corresponding D_C with almost no algebra. For example, let $b' = 4$, $b'' = 2$, yielding the Brownian $H = 1/2$. When k is large and $\eta = 2^{-k}$, the quantity $L(\eta)/\eta$, which is the number of compass steps, is seen to be exactly equal to $4^k = \eta^{-2}$, 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 $D_C = 1/H$.

This is an extremely strange value, at first blush. First of all, it can exceed 2 (in fact it can be arbitrarily large), which is impossible for the fractal dimension of a self-avoiding curve in the plane. Secondly, $1/H$ contradicts the value $2\text{-}H$ that definition of the fractal dimension give the other local definitions for the records under study. On the other hand, those familiar with the fractional Brownian motion will identify $1/H$ as being the fractal dimension of the trail (in an E-dimensional Euclidean space 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 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. Indeed, performing a measurement in which the compass is mostly near parallel to the B-axis amounts to collapsing time by nearly flattening the record into a trail along the B-axis. Suppose our scalar B is one coordinate of a vectorial B^* in a space R^E of Euclidean dimension $E \gg 1/H$. If a compass of opening η is walked along the trail of B^*, the steps' projections on any coordinate axis will differ in size, but will mostly be close to $\eta/\sqrt{E}$. Now measure the length of the trail of B^* with steps of length η that are additionally constrained so that their projection on one axis is *exactly* $\eta/\sqrt{E}$. A moment of thought suggests that this last constraint will not much affect the number of steps. Thus (apart from a numerical correction factor dependent on E), walking a compass takes about as many steps along our collapsed record as along the trail in a space of E dimensions. Hence, the compass dimension should indeed be the same in both cases, that is $1/H$.

Remark on fracton/spectral dimension

The box and the Hausdorff–Besicovitch dimension of the zeros of $B_H(t)$ are $1\text{-}H$, and Section 5 deduces a set of dimension $1/H$ from a set of dimension $1\text{-}H$. In a more familiar context, one starts with the trail of $B_H(t)$, of dimension $1/H$, and one deduces $1\text{-}H$ as the dimension of the instants of this motion's recurrence to the origin [3, second paper by B.B.M.]. A further step deduces $2H$ as the fracton/spectral dimension [3, e.g., papers by Rammal and Orbach]. 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 $M_H(t)$. Its value estimated mechanically for small R is $2\text{-}H$. For larger R, the mass dimension is 1.

"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 $2R$.

In affine geometry, the notions of "square" or "circle" are meaningless, but we must tackle the practical question that arises 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 to the contrary, $R \ll t_c$, 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 $M_H(t)$ 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 kth 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 traversal is $\sim (R/\delta)^{1-H}$, and $\delta = b'^{-k} = (b''^{-k})^{1/H} = R^{1/H}$. Hence $M(R) \sim R^{2-H}$, yielding the familiar value $D_M = 2\text{-}H$ in the small R limit.

7. Cross-over pitfalls

To summarize, self-affine fractals *do not* involve exponents such that $N(b) \sim b^{D_B}$, $L(\eta) \sim \eta^{-D_C}$ for all η and $M(R) \sim R^{D_M}$ for all R Different exponents are approached on different sides of 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 of D_C is likely to combine values of η or of R that range on both sides of the cross-over t_c, and the estimate will depend upon where exactly t_c lies in the range of η or R. Such an estimate will be worthless, because it will depend on the units of t and B. Reliability is improved by exaggerating the vertical scale.

8. Discussion

The notion of fractal dimension brings an unavoidable proliferation of distinct quantities, each contributing to a fractal's overall description.

Finally, consider the issue of the foundations of fractal geometry. "Foundations" is a treacherous term, and in particular its figurative meaning in mathematics, in science and in philosophy is profoundly different from its meaning for the architect and the layman. A building's foundations always come first, followed by the basic shelter and later by the decorative work. Before expanding a building, one tests its foundations and one strengthens them if necessary. In successful branches of mathematics and of science, to the contrary, concern with "foundations" tends to come late, after each period of extensive substantive achievement. An endeavor wins little respect as a science if it favors methodology over substance. Ten years ago, laying careful foundations for fractal geometry was not a priority concern. The new discipline had to gain acceptance, and – before that – it had to be built almost from scratch (it includes a few essential parts salvaged from the work of mathematicians engaged in other pursuits, but to this day – as exemplified by this Note – many basic issues had been completely untouched.) Anyhow, events moved fast, and are presently forcing me to write a systematically organized textbook. The present Note is based on excerpts from the book's first draft. A more detailed excerpt will appear in [4].

Acknowledgement

Richard F. Voss, with whom I discussed the material in this Note repeatedly, has made very useful comments and valuable suggestions.

References

1. Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, New York (1982).
2. Mandelbrot, B. B., Passoja, D., and Pallay, A. *Nature* **308**, 721 (1984).
3. *Proceedings of a Symposium on Fractals in the Physical Sciences* U.S. National Bureau of Standards, Gaithesburg, MD, 21–23, November (1983). *J. Stat. Physics.* **36** Numbers 5/6 (1984).
4. *Proceedings of the Sixth Trieste International Symposium, "Fractals in Physics", Ed. L. Pietronero, North-Holland, Amsterdam (to appear).*