Symmetry 2000, edited by I. Hargittai
and T.C. Laurent, **I**, 2002, 133-141

Symmetry by dilation/reduction, fractals, and roughness

Benoit Mandelbrot

Roughly speaking, *fractals* are shapes that look the same from close by and far away. Fractal geometry is famed (or notorious, depending on the source) for the number and apparent diversity of its claims. They range from mathematics, finance, and the sciences, all the way to art.

Such diversity is always a source of surprise, but this article will argue that it can be phrased so as to become perfectly natural. Indeed, the ubiquity of fractals is intrinsically related to their intimate connection with a phenomenon that is itself ubiquitous. That phenomenon is *roughness.* It has long resisted analysis, and fractals provide the first widely applicable key to some degree of mastery over some of its mysteries.

A self-explanatory term for looking the same from close by and far away is self-similar• Of course, self-similarity also holds for the straight line and the plane. These familiar shapes are the indispensable starting point in the many sciences that have perfected a mastery of smoothness. The point is that the property of looking the same from close by and far away happens to extend beyond the line and the plane. It also holds for diverse shapes called fractals. I named them, tamed them into primary models of roughness, and made them multiply.

The learned term to denote them is forms invariant under *dilation or reduction.* In the professional jargon of some thoroughly developed fields of science, such invariances are called *symmetries,* which would suffice as an excuse for this paper to be included in this book. But there is another far more direct reason: it will be shown that suitable combinations of very ordinary-looking symmetries hold a surprise for the beholder: they produce self-similarity, meaning that they yield fractals.

The theme having been embedded in the order of the words in this paper's title, and every word in the title having already been mentioned (at least fleetingly) and italicized, the etymologies of two of those words are worth recalling. In ancient Greek, the scope of *summetria* went beyond a reference to self-examination in a mirror. A close synonym was in just proportion, and it was mostly used to describe a work of art or music as being harmonious. In time, most words' meanings multiply and diversify. In developed sciences, groups of invariances-symmetries are abstract and abstruse, but those which underlie fractals are – to the contrary – intuitive and highly visual. This paper will show typical ones to be closely related to the simplest mirror symmetry.

As to the word *fractal*, I coined it on some precisely datable evening in the winter of 1975, from a very concrete Latin adjective, *fractus,* which denoted a stone's shape after it was hit very hard. Lacking time to evolve, *fractal* rarely strayed far from the notion of roughness.

The Ubiquity of Roughness

The following list combines, in an intentionally haphazard fashion, many questions that led to ideas underlying fractals, and other questions that the new ideas found it easy to handle.

• How to measure the volatility of the financial charts, for example to evaluate and compare financial risks realistically?

• How long is the coast of Britain?

• How to distinguish proper music (old or new, good or bad) from plain awful noise?

• How to characterize the course of a river untamed by civil engineers?

• How to define wind speed during a storm?

• How to measure and compare the surface structure of ordinary objects, such as a broken stone, a mountain slope, or a piece of rusted iron?

• What is the shape of a cloud, a flame, or a welding?

• What is the density of galaxies in the Universe?

• How to measure the variation of the load on the Internet?

The word, *rough,* appears *in none* of those questions, but the underlying concept appears *in every one. Irregular* is more polite, but rough is more telling.

An inverse question will provide contrast.

• For which shapes do examples of the simplest smooth shapes of Euclid's geometry provide a sensible approximation?

To Early Man, Nature provided just a few smooth shapes: the path of a stone falling straight down, the full Moon or the Sun hidden by a light haze, small lakes unperturbed by current or wind. In sharp contrast, *homo faber* keeps adding examples beyond counting. For example, Man works hard at eliminating roughness from automobile pistons, flat walls and tops, Roman-Chinese-American street grids, and – last but not least – from most parts of mathematics (contributing to the widespread view that geometry is “cold and dry.”)

Forgetting this last question, the others set a pattern one may extend forever. The simple reason is that roughness is ubiquitous in Nature. In the works of Man, it may not be welcome, but is not always avoided, and may sometimes be unavoidable. Examples are found in some parts of mathematics, where they were at one time described as pathological or monstrous, and, once again, in the above list of questions.

Needless to say, each of these questions belongs to some specific part of knowledge (science or engineering) and the practical attitude is not to waste time studying roughness but instead to get rid of it. Fractal geometry, to the contrary, has embraced roughness in all its forms and studies it for its intrinsic interest.

Roughness Lagged Far Behind Other Senses, for Example, Sounds, in Being Mastered by a Science

By customary count, a human's number of sense receptors is five. This may well be true but the actual number of distinct sense messages is certainly much higher.

Take sound. Even today, concert hall acoustics are mired in controversy, recording of speech or song is comfortable with vowels but not consonants, and drums are filled with mysteries. Altogether, the science of sound remains incomplete. Nevertheless, it boasts great achievements.

Let us learn some lessons from its success. Typical of every science, it went far by exerting healthy opportunism. Side-stepping the hard questions, it first identified the idealized sound of string instruments as an icon that is at the same time reasonably realistic and mathematically manageable even simple. It clarifies even facts it does not characterize or explain. It builds on the harmonic analysis of pendular motions and the sine/cosine functions. It identifies the fundamental and a few harmonics then, in due time, builds a full Fourier series. The latter is periodic, that is, translationally invariant, which, in a broad sense, is a property of symmetry. Newton's spectral analysis of light is a related example, although the structure of incoherent white light took until the 1930s to be clarified.

More generally, the harmonies that Kepler saw in the planets' motions have largely been discredited, yet it remains very broadly the case that the starting point of every science is to identify harmonies in a raw mess of evidence.

Scale Invariance Perceived as Playing for Roughness the Role That Harmonious Sound Played in Acoustics

By contrast to acoustics, the study of roughness could not, until very recently, even begin tackling the elementary questions this paper listed earlier. My contribution to science can be viewed as centered on the notion that, like acoustics, the study of roughness could not seriously become a science without beginning by the following step: it had to first identify a basic invariance/symmetry, that is, a deep source of harmony common to many structures one can call rough.

Until the day before yesterday, roughness and harmony seemed antithetic. An ordering of deep human concerns, from exalted to base, would have surely placed them at opposite ends. But change has a way of scrambling up all rankings of this sort. As candidates for the role of harmonious roughness, I proposed the shapes whose roughness is invariant under dilation/reduction.

In the most glaring irony of my scientific life, this first-ever systematic approach to roughness arose from a thoroughly unexpected source in extreme mathematical esoterica. This may account for the delay science experienced in mastering roughness.

Meek Symmetries

In the form known to everyone, the concept of symmetry tends to provoke love or loathing. My own feelings depend on the context. I dislike symmetric faces and rooms but devote all my scientific life to fractals. Let us now move indirectly, step by step, towards examples of fractals.

Among symmetries, the most widely known and best understood involves the relation between an object and its reflection in a perfectly flat mirror. To obtain an object that is invariant – unchanged – under reflection, it suffices to “symmetrize” a completely arbitrary object by combining it with its mirror reflection. As a result, there is an infinity of such objects.

A fairgrounds carousel helps explain symmetry with respect to a vertical axis. Symmetry with respect to a point is not much harder.

Now, replace the symmetry with respect to one mirror by symmetries with respect to two parallel mirrors. Any object can be symmetrized “dynamically,” by being reflected in the first mirror then the second, then again the first, again the second, and so on ad infinitum. The object grows without end and its limit is doubly invariant by mirror symmetry. In addition, it becomes unchanged under translation, in either direction, whose size is a multiple of twice the distance between the planes. Like for a single mirror, the defining constraint allows an infinite variety of such objects.

Symmetry with respect to a circle is a little harder but will momentarily become crucial. One begins with an object that is symmetric with respect to a line. Then one transforms the whole plane by an operation called geometric inversion, which is a very natural generalization of the transformation of x into its algebraic inverse $1 / x$. One finds that geometric inversion transforms (almost) any line into a circle. When the plane has then inverted in this fashion, an object that used to be symmetric with respect to the original line is said to have become symmetric with respect to the circle inverse of the line.

From Meek to Wild Symmetries And on to Self-Similarity

So far, so good. Very elementary French school geometry used to be filled with examples of this kind. Over several grades, everything grew increasingly complex and harder, but only very gradually. Special professional periodicals made believe that they were working on an endless frontier, but it was clear that this old geometry was actually exhausted, dead.

All too many persons hastened to conclude that *all* of more or less visual geometry was dead. Actually, the germ of very different but very visual developments existed since the 1880s(!) in the work of Henri Poincaré. But it did not develop in any way that mattered outside of esoteric pure mathematics. Outside mathematics, it remained dormant until the advent of computer graphics and my work, as exemplified in what follows. Because of fractal geometry, Poincaré's idea now matters broadly and will provide us with a nice transition from the simplest symmetry to self-similar roughness. In the next section, the meaning of roughness is a bit stretched; in the section after the next, almost natural.

The Self-Similar "Thrice-Invariant Dust D," With Three Part Generator Symmetry

Consider, in the plane, a diagram to be called a three-part generator that combines two parallel lines and a circle half-way between them, as shown to the left of Figure 1. Could a geometric shape be simultaneously symmetric with respect to *each* of the three parts of this generator? If this is true of more than one shape, could one identify the smallest, to be called *D*? Painfully learned intuition tempted great thinkers to propose that one should measure the complexity of a notion by the length of the shortest defining formula or sentence. This line of thought would have implied that when the shortest defining formula doubles in length, the corresponding study becomes twice more complex.

Examined in this light, our present generator seems to involve only a small step beyond its separate parts. It seems innocuous. Fortunately for us, but unfortunately for that wrong-headed definition of complexity, this very simple combination of inversions will momentarily bring a great surprise. It will prove to involve a jump across the colossal chasm that separates consideration of very elementary geometric symmetry from the great complexity of fractals.

Both for those who do and those who do not (yet) know much about fractals, the least contrived method to understand them is to take advantage of the computer graphics technology that made all this possible.

The words “minimal thrice-invariant” in the section title might create the fear that *D* is some needle in a haystack, but the precise contrary is the case. The tactic behind the search for *D* is hardly more complicated than the “dynamics” that we used to create a mirror symmetric set: start with an arbitrary object, then add its mirror image, and so on, over an infinite number of times. The novelty is double: the dynamics that searches for a symmetrized *D* is irresistibly "attracted" to its prey and the prey is sharply specific and extraordinarily complex.

The century-old process that yields *D* as a limit set is now called chaos game. To appreciate what is happening with minimal notation and programming effort, it is helpful to know that *D* is entirely contained in the horizontal axis, which is defined as the line that crosses the generating circle's center of abscissa 0 and is perpendicular to the two generating lines of abscissas 1 and 1. Therefore, the generating circle's radius, which is less than 1 being denoted by *r,* inversion simply transforms x *into* r^2 / x .

Altogether, the points symmetric of x with respect to the generating lines and circle have the abscissas $-x-2$, $-x+2$, *and* r^2 / x , respectively. Those formulas' simplicity is hard to beat, but the object grown from this simple seed and will soon be revealed as extremely complex.

Recall the symmetrization of an object with respect to one or two parallel mirrors. Because mirror symmetry preserves an object's size, the outcome depends on the object one started to symmetrize. That is, this form of symmetrization reflects – no pun is intended – the initial conditions. Because reflection into the inside of a circle makes an object smaller, the search for our minimal thrice-symmetric *D* can be altogether different. It remains true that the orbit's first steps depend on the initial point of abscissa x_0 and the sequence of random moves its beginning. But those arbitrarily chosen inputs can be shown to have no perceptible effect on the orbit's limit; in practice, the limit identifies with the orbit from which the first few points have been erased.

The algorithm simplifies if x_0 is picked outside of the circle and its first step replaces x_0 *by either* $-x-2$ *or* $-x+2$. The second and all following steps are best prevented from backtracking, therefore the throw of a coin (or the choice of 0 or 1 at random by the computer) will suffice to decide between the two possibilities other than repetition.

Seeing is believing and, to be believed, the progression of the orbit of this process has best be followed on the computer screen. Figure 2 illustrates successive blowups of a small piece. A striking observation is that all are close to being identical. That is, very small parts of our limit set C are nearly identical – except for size – to merely small parts. That limit set is invariant by reduction; being self-similar, it is a fractal.

The Self-Similar "Four Times Invariant Curve *C*," With Four-Part Generator Symmetry

To move on from dusts to proper rough curves, it suffices to follow another narrow lane in the same conceptual neighborhood. To the two parallel lines of the previous construction, let us now add, not one but *two* circles of equal radii, each tangent to the other and to one of the straight lines, as shown in the right part of Figure 1. To preserve in this article the ideal of almost complete avoidance of formulas, let the pleasure of drawing the limit set be reserved to those who know how to program a geometric inversion in the plane, as opposed to a line. Contrary to Figure 2, Figure 3 is not contained in a straight line, and it is not a dust with gaps devoid of points. It is a continuous curve that has no tangent at any point. It is loop-free, also called singly connected, meaning that it joins two prescribed points in a single way.

Like the thrice-symmetric dust *D* , this four-times-symmetric object *C* is approximately self-similar in many ways. This is especially clear-cut with respect to the midpoint where the curve is osculated (outlined) by two circles intersecting at an angle.

Concluding Remarks

The path this paper takes from plain symmetry to self-similarity is little-traveled but it is very attractive and worth advertising as is now being done. Two alternatives must be mentioned.

By now, many persons know of an example that is analogous to *D* and is exactly (linearly) self-similar. It was provided by a cascade construction due to G. Cantor; it is simple and by now famous, but completely artificial. Many persons also know of an analog to *C,* which is exactly (linearly) self-similar. It was provided by a cascade construction due to H. von Koch; it too is simple and by now famous, but completely artificial.

Figure 3 makes it obvious that one can construct *C* in analogy to the Koch curve. One proceeds by successive replacements of short arcs of a big circle by longer arcs of a smaller circle, an arc's length being measured in degrees. This construction resolves a query by Poincaré. Astonishingly, this query had been left unanswered until Mandelbrot (1982, 1983).

Upon discovering an object related to C, Poincaré called it a curve, then, in an aside, commented “if you can call *this* a curve.” This question was non-obvious as long as C was viewed as “mathematically pathological.” But fractal geometry recognized shapes of this kind as models of nature; no one will any longer deprive them of the dignity of being called curves.

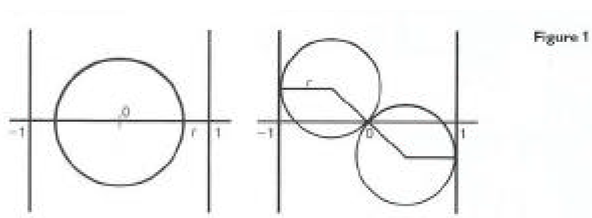

Figure 1. *The generation of Figures 2 and 3.* The figures illustrate two bridges between, on the one hand, the simplest symmetries (with respect to a line or a circle, the latter being a geometric inversion) and, on the other hand, self-similarity, that is, fractality. Both figures are constructed point by point by an orbit that is an infinite random sequence of never repeating symmetries.

The generator of the dust D of Figure 2, shown to the left, is made of two parallel lines and, between them, a circle of radius r. The value taken for $1 - r$ is positive, which implies that there exist genuine empty gaps between the points; hence D is a dust. In Figure 2, $1 - r$ is very small, but, for the sake of clarity, the diagram exaggerates it.

The generator of the curve C of Figure 3, shown to the right, is made of two parallel lines and two circles. Because those circles and lines are – as shown – tangent to one another the orbit merges into a gap-free curve, but does so very slowly.

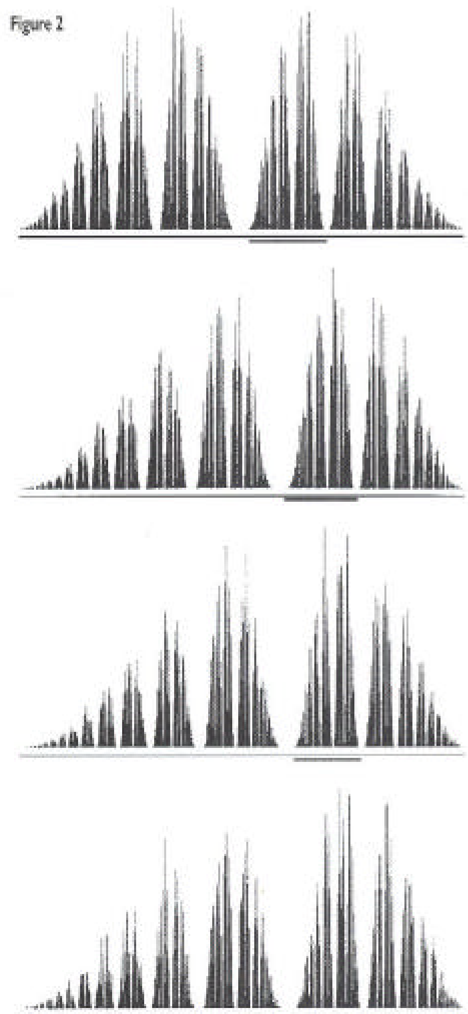

Figure 2. *Histograms of the self-inverse Dust D and their successive “blowups.”* The dust D is entirely located within the horizontal axis and the stack of four parts represents successive blow-ups of a portion of D. To represent the points of a dust exactly is impossible. The best is to divide the axis into small bins and show the number of points in a bin (this is the fractal counterpart of the density) by a vertical bar; this leads to a kind of histogram. The point of this construction is that – except for small deformations – all those blow-up histograms have very much the same form. Each step from the top down represents an orbit whose length (therefore density) increases. At the same time, the overall diagram is spread horizontally, revealing increasing detail, and the bin size is made to decrease.

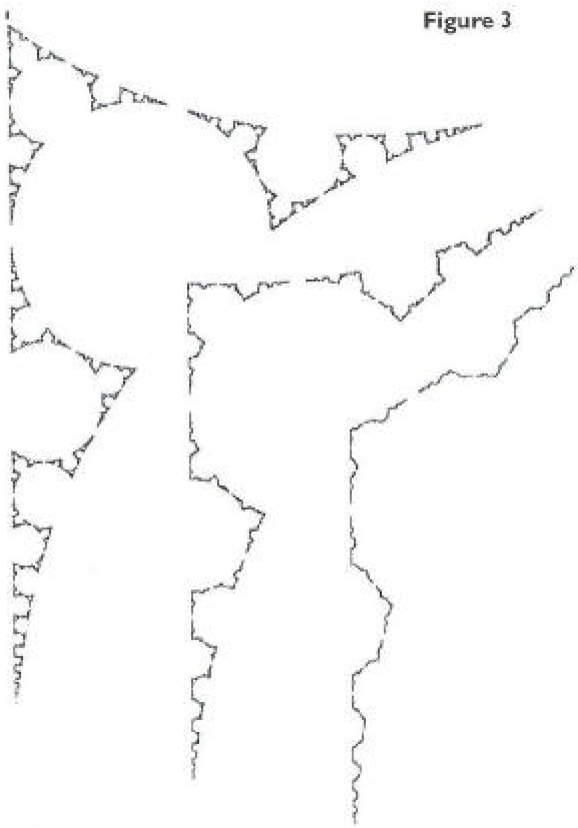

Figure 3. *Three variants of a self-inverse curve C*. Contrary to the dust *D* on Figure 2, these three variants of curve *C* winds up and down and around. The density would have to be drawn along an axis orthogonal to the plane, therefore was omitted. While the gaps of *D* are real, those perceived in *C* are artifacts; in a longer orbit, their lengths decrease to zero.

Bibliography

Mandelbrot B B 1982 *The Fractal Geometry of Nature* Freeman (This is the first and still the most comprehensive book on fractals. Self-inverse constructions are discussed in Chapter 18 and in the next item.)

Mandelbrot B B 1983 *Mathematical Intelligencer:* 5(2), 9-17

Mandelbrot B B 1997 *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk* (Selecta Volume E) Springer-Verlag

Mandelbrot B B 1999 *Multifractals and 1/f Noise: Wild Self-Affinity in Physics* (Selecta Volume N) Springer-Verlag (This book's cover boasts a striking self-inverse construction, a jewel-like necklace made of circles.)

Mandelbrot B B 2001a *Gaussian Self-Affinity and Fractals* (Selecta Volume H) Springer-Verlag.

Peitgen H-O and Richter P *The Beauty of Fractals* New York. Springer-Verlag 1986 (remains the best combination of beauty and solidity)