How long is the coast of Britain?
Statistical self-similarity and fractional dimension

Science: 156, 1967, 636-638

B. B. Mandelbrot

Geographical curves are so involved in their detail that their lengths are often infinite or more accurately, undefinable. However, many are statistically ''self-similar,'' meaning that each portion can be considered a reduced-scale image of the whole. In that case, the degree of complication can be described by a quantity **D** that has many properties of a ''dimension,'' though it is fractional. In particular, it exceeds the value unity associated with ordinary curves.

1. Introduction

Seacoast shapes are examples of highly involved curves with the property that — in a statistical sense — each portion can be considered a reduced-scale image of the whole. This property will be referred to as ''statistical self-similarity.'' The concept of ''length'' is usually meaningless for geographical curves. They can be considered superpositions of features of widely scattered characteristic sizes; as even finer features are taken into account, the total measured length increases, and there is usually no clear-cut gap or crossover, between the realm of geography and details with which geography need not be concerned.

Quantities other than length are therefore needed to discriminate between various degrees of complication for a geographical curve. When a curve is self-similar, it is characterized by an exponent of similarity, D, which possesses many properties of a dimension, though it is usually a fraction greater that the dimension 1 commonly attributed to curves. I propose to reexamine in this light, some empirical observations in Richardson 1961 and interpret them as implying, for example, that the dimension of the west coast of Great Britain is $D = 1.25$. Thus, the so far esoteric concept of a ''random figure of fractional dimension'' is shown to have simple and concrete applications of great usefulness.

Figure 1 :fig width=page frame=none depth='4.25i' place=bottom. :figcap. :figdesc. Data from Richardson 1961, Fig. 17, reporting on measurements of lengths of geographical curves by way of polygons which have equal sides and have their corners on the curve. For the circle, the total length tends to a limit as the side goes to zero. In all other cases, it increases as the side becomes shorter, the slope of the doubly logarithmic graph having an absolute value equal to $D - 1$. (Reproduced by permission.) :efig.

Self-similarity methods are a potent tool in the study of chance phenomena, wherever they appear, from geostatistics to economics (M 1963b{E }), and physics (M 1967i{N9}). Very similar considerations apply in the study of turbulence, where the characteristic sizes of the ''features'' (which are the eddies) are also very widely scattered, a fact first pointed out by Richardson himself in the 1920's. In fact, many noises have dimensions Dbetween 0 and 1; therefore, scientists ought to consider dimension as a continuous quantity ranging from 0 to infinity.

2. Review of the methods used when attempting to measure the length of a seacoast

Since geography is unconcerned with minute details, one may choose a positive scale G as a lower limit to the length of geographically meaningful features. Then, to evaluate the length of a coast between two of its points A and B , one may draw the shortest inland curve joining A and B while staying within a distance G of the sea. Alternatively, one may draw the shortest line made of straight segments of length at most G, whose vertices are points of the coast which include A and B . There are many other possible definitions. In practice, of course, the shortest paths must be approximated. We shall suppose that measurements are made by walking a pair of dividers along a map so as to count the number of equal sides of length G of an open polygon whose corners lie on the curve. If G is small enough, it does not matter whether one starts from A or B. Thus, one obtains an estimate of the length to be called $L(G)$.

Unfortunately, geographers will disagree about the value of G, and $L(G)$ depends greatly upon G. Consequently, it is necessary to know $L(G)$ for several values of G. It would be better still to have an analytic formula linking $L(G)$ with G. Such a formula was proposed in Richardson 1961, but unfortunately it has attracted no attention. The formula is $L(G) = F\ G^{1\ -\ D}$, where F is a positive constant prefactor and D is a second constant, at least equal to unity. In Richardson's own words, this D is a ''characteristic'' of a frontier and may be expected to have some positive correlation with one's immediate visual perception of the irregularity of the frontier. At one extreme, $D = 1.00$ for a frontier that looks straight on the map. At the other extreme, the west coast of Britain was selected because it looks like one of the most irregular in the world; this case yields $D = 1.25$. Three other frontiers which, judging by their appearance on the map, were closer to world average in irregularity, yielded the following values: $D = 1.15$ for the land frontier of Germany in about 1899, $D = 1.14$ for the land frontier between Spain and Portugal, and $D = 1.13$ for the Australian coast. South Africa's coast, which appeared to be one of the smoothest in the atlas, had a $D = 1.02$.

Richardson's empirical finding is in marked contrast with the ordinary behavior of smooth curves, which are endowed with a well-defined length and are said to be ''rectifiable.''

3. Self-similarity, dimension and fractional dimension

I will now interpret Richardson's empirical relation as meaning, not only that nonrectifiable curves are natural, but that curves of dimension greater than one are not exclusively an invention of mathematicians.

Let us first review an elementary feature of the concept of dimension, and to show how it naturally leads to the consideration of fractional dimensions.

Figure 2 :fig width=page frame=none depth='4.5i' place=bottom. :figcap. :figdesc. Nonrectifiable self-similar curves can be obtained by the following ''cascade''. Step 1: Choose any of the above drawings. Step 2: Replace each of its N legs by a curve that is obtained from the whole drawing through a similarity of ratio 1/4. One is left with a curve made of N^2 legs of length $(1/4)^2$. Step 3: Replace each leg by a curve obtained from the whole drawing through a similarity of ratio $(1/4)^2$. The desired self-similar curve is approached by an infinite sequence of these steps. :efig.

To begin, a straight line has dimension one. Hence, for every positive integer N , the segment *(0 le x lt X)* can be exactly decomposed into N .hy off nonoverlapping segments of the form .hy on [*(n - 1) X / N le x lt n X / N]* , where n ranges from 1 to N . Each of these parts is obtained from the whole by a similarity of ratio *r (N) = 1 / N* . Similarly, a plane has dimension two. Hence, for every perfect square N , the rectangle *(0 le x lt X ; 0 le y lt Y)* can be decomposed exactly into N nonoverlapping rectangles of the form *[(k - 1) X / Ý √(N le x lt k X) / √(N) ; > < (h - 1) Y / √(N) le y lt h Y / (N)]* , where k and h range from 1 to *√(N)* . Each of these parts is obtained from the whole by a similarity of ratio *r(N) = 1 / √(N)* . More generally, whenever $N^{1/D}$ is a positive integer, a D-dimensional rectangular parallelepiped can be decomposed into N parallelepipeds, each obtained from the whole by a similarity of ratio $r(N) = 1 / N^{1/D}$. Thus, the dimension D is characterized by the relation $D = -\log N / \log r(N)$.

This last property of the quantity D means that it can also be evaluated for any figure that can be exactly decomposed into N parts such that each of the parts is deducible from the whole by a similarity of ratio $r(N)$, by a similarity followed by rotation and even symmetry should also be allowed. If such figures exist, their dimension may be said to be $D = -\log N/\log r(N)$.

(The concept of ''dimension'' is elusive and very complex, and is far from exhausted by simple considerations of the kind used in this paper. Different definitions frequently yield different results, and the field abounds in paradoxes. However, two concepts, the Hausdorff-Besicovitch and the capacity dimensions, when computed for random self-similar figures, have so far yielded the same value as the similarity dimension.

To show that such figures exist, it suffices to exhibit a few obvious variants of von Koch's continuous non-differentiable curve. Each of these curves is constructed as a limit. Step 1 is to draw out the kinked curves of Figure 2, each made up of N intervals superposable upon the segment $[0, 1/4]$. Step 2 is to replace each of the N segments used in step 1 by a kinked curve of step 1 reduced in the ratio $r(N) = 1/4$. This gives a total of N^2 segments of length 1/16. Each repetition of the same process adds further detail. As the number of steps grows to infinity, our kinky curves tend toward continuous limits, and it is obvious that these limits are self-similar, since they are exactly decomposable into N parts deducible from the whole by a similarity of ratio $r(N) = 1/4$ followed by translation. Thus, given N , the limit curve can be said to be of a dimension $D = -log\ N/log\ r(N) = log\ N\ /\ log\ (1/4)$. Since N is greater than 4 in our examples, the corresponding dimensions all exceed unity.

As to measured length, at step number s , our approximation is made of N^s segments of length $G = (1/4)^s$, so that $L = (N/4)^s = G^{1-D}$. Thus, the length of the limit curve is infinite, even though it is a ''line.'' (Note a plane curve may have a dimension equal to 2. An example is Peano's curve, which fills up a square.)

Practical application of this notion of dimension requires further consideration, because self-similar figures are seldom encountered in nature (crystals are one exception). What is often encountered is a statistical form of self-similarity, and the concept of dimension may be further generalized. To say that a (closed) plane figure is chosen at random implies several definitions. First, one must select a family of possible figures, usually denoted by ?.. When this family contains a finite number of members, the rule of random choice is specified by attributing to each possible figure a well-defined probability of being chosen. However, ?. is in general infinite, hence each figure has a zero probability of being chosen. But nonzero probabilities can be attached to appropriately defined ''events'' (such as the event that the chosen figure differs little —. in some specified sense —. from some specified figure).

To be self-similar, two conditions must hold for the family ?, together with the definition of events and their probabilities. First, each of the possible figures must be constructible by somehow stringing together N figures, each of which is deduced from a possible figure by a similarity of ratio r . Second, the probabilities must be so specified that the same value is obtained whether one selects the overall figure at one swoop or as a string. (The value of N may either be arbitrary, or chosen from some specific sequence, such as the perfect squares relative to nonrandom rectangles, or the integral powers of 4, 5, 6, or 7 encountered in the curves built in Figure 2.) When the values of r and N are specified, $D = -log\ N\ /\ log\ r$ can be called a similarity dimension. More commonly, however, given r , N will take different values for different figures of ?.. As one considers points ''sufficiently far'' from each other, the details on a ''sufficiently fine'' scale may become asymptotically independent in such a way that $-log\ N\ /\ log\ r$ almost surely tends to some limit as r tends to zero. In that case, this limit may be considered a similarity dimension. Under wide conditions, the length of approximating polygons will asymptotically behave like $L(G)\ similar\ G^{1-D}$.

To specify the conditions for the existence of a similarity dimension is not a fully solved mathematical problem. In fact, a number of conceptual problems familiar in other uses of randomness in science are also raised by the idea that a geographical curve is random. Therefore, returning to Richardson's empirical law, the most that can be said with perfect safety is that it is compatible with the idea that geographical curves are random, self-similar figures of fractional dimension D . Since empirical scientists have to be content with less than perfect truths, I favor the more positive interpretation stated at the beginning of this report.

ANNOTATIONS

Dates and editorial changes

This text was submitted on November 14, 1966; a revision was received March 27, 1967. Section titles were added in this reprint.

Why this text came to be written? It was intended to be a ''Trojan'' horse allowing a bit of mathematical esoterica to ''infiltrate'' surreptitiously hence near-painlessly, the investigation of the messiness of raw nature.

Today, it means that everyone knows how to answer the question raised in this paper's title. And the notion of fractal dimension is very widely known and used. It is noteworthy that both developments can be credited in part to this very short paper. The story transcends the other topics in this book and deserves to be told here.

By the mid-1960s my record of publications was substantial but presented a serious flaw. Those publications' topics ranged all too widely and was perceived as an aimless juxtaposition of studies of noise, turbulence, galaxy clustering, prices and river discharges. Few persons realized that, to the contrary, I did not deserve to be criticized for immature aimlessness but for increasingly acute single-mindedness. As early as 1956 (as explained in M 1999N, Chapter NX) then increasingly and more seriously in my works on finance and on noise, I had somehow latched on the process of renormalization and found it useful in very diverse contexts. Unfortunately, the nature and worth of that concept was not appreciated until much later when it was rediscovered quite independently in the statistical physics of critical phenomena that arose in 1972.

More specifically, nearly all my works were linked by the ubiquity of ''power-law'' relations, each endowed with an important exponent. Superficially, those exponents seemed both formal and mutually unrelated. But in fact I knew how to interpret them geometrically as ''the'' fractal dimension of suitable sets. Furthermore, this interpretation gave to my work a profound unity that promised further growth. But I soon found out that mention of a fractal dimension in a paper or a talk led all referees and editors to their pencils, and some audiences to audible signs of disapproval. Practitioners accused me of hiding behind formulas that were purposefully incomprehensible. Few mathematicians knew any of the flavors of fractal dimension; if asked, they were worse than useless in explaining this notion to those I was trying to convert. Indeed, I was told that, as ''everyone knew,'' the Hausdorff-Besicovitch dimension was a purely technical device, not a concrete notion. It was believed that the local detail with which it is concerned could not possibly matter in science. It proved to be very difficult to convince them of the contrary: that dimension does matter concretely when an object is self-similar.

Fortunately, I stumbled one day upon Richardson's empirical data on coastline lengths, and recognized instantly that a study of coastlines might lend itself to a ''Trojan horse'' manoeuver. Indeed, everyone has a knowledge of geography, but no one I knew professionally had a vested professional interest in facts and theories concerning coastlines and relief. The manoeuvre succeeded. Everyone was wonderfully objective and receptive to the seemingly wild idea contained in this paper, and as a result, became more receptive to the use of fractal dimension in fields that really matter to me.

How this paper came to be published.

The weekly magazine *Science* accepted this text for publication with marvelous promptness. The referee liked very much my idea that fractal dimension is indeed important in this context, and identified himself as Hugo Steinhaus (1887-1972). But who told *Science* editors that Steinhaus was their man and alive and well in Wroclaw, Poland? I bet it was his former student Mark Kac (1914-1984) whom I knew well. I had many occasions but — to my regret — never checked.

It is chasteining to mention here a 1975 episode that is described in the next chapter's *Annotations*. A later paper of mine went on to study coastlines via the Earth's relief. It too was submitted to *Science* but was resolutely rejected. The irony is that, promptly though indirectly, it became a classic, because it included the recipe for constructing life-like fractal mountains. Clearly, Steinhaus had died and had not been replaced.

How this paper moved non-rectifiability out of mathematical esoterica and into the realm of notions that are so obvious, simple and compelling that one can discuss them with inquisitive children.

My sole *Science* paper became wonderfully effective because, once again, it propelled fractal dimension out of bondage among mathematical esoterica to its proper place among the working scientists' everyday toolboxes. But untold numbers of non-scientists are aware of this paper for the wrong reason: as having proven that a coastline's length increases under increasingly close examination.

The irony is that my paper did not view non-rectifiability in nature was not a highlight only the point of departure, one so familiar as not deserving any specific credit. Was I wrong, and does the final observation of non-rectifiability in nature deserve to be identified and credited? Such a search for origins would have to go very far back, if it is true that Athenian sailors reporting to their Admiralty mentioned the difficulties they encountered in seeking how long is the coast of Sardinia. The search for origins would also be largely idle. To explain why I think so, let us return to Hugo Steinhaus, who welcomed my paper. He identified himself to me by pointing out some papers of his, which is why my final text included the following quotes, which this reprint moved out of the text and into this *Annotation*. (Steinhaus 1954) wrote ''the left bank of the Vistula, when measured with increased precision, would furnish lengths ten, a hundred or even a thousand times as great as the length read off the school map.''

Thus, to quote Steinhaus 1954 again, ''a statement nearly adequate to describe reality would be to call most arcs encountered in nature not rectifiable. This statement is contrary to the belief that non rectifiable arcs are an invention of mathematicians and that natural arcs are rectifiable: it is the opposite that is true.''

Lacking an extraordinary quirk in the refereeing process these papers by Steinhaus would have remained unknown to me. It can safely be said that they had become lost. The question is, ''why''? After all, Steinhaus was a famous scholar and the author of the beautiful and often reprinted *Mathematical Snapshots,* Steinhaus 19XX.

The fact that no one reads Polish scientific journals — even those written in an international language — is not convincing either, in my opinion. After this episode, reading other works Steinhaus wrote in the 1950s, I found it indecent they were written primarily in response to official pressure (even on a 67-year old celebrity) to ''publish or perish'' and ''be relevant.'' Steinhaus was a citizen of the world. Had he viewed this work of his as worthy of wide attention, he would have found a way to publish it where it could not be missed.

Why do I place these thoughts in the mind of a long-dead person? Perhaps because I myself view the nonrectifiability of coastlines as being nothing beyond an idle remark devoid of interest, *unless it is followed up* and placed in an interesting wider setup. History cannot be rewritten, but this wider setup might have been entirely different from that of fractal geometry. Moreover, cute casual remarks are heard and forgotten without encountering resistance, but wider setups must inevitably fight for acceptance. During the period after World War II, when mathematics was dominated by an abstract Utopia, this fight for recognition was bound to be bruising. I cannot avoid quoting here these lines by Alfred North Whitehead (the co-author with Bertrand Russell of *Principia Mathematicae*), ''To come very near true theory and to grasp its precise application are two very different things as the history of science teaches us. Everything of importance has been said before by somebody who did not discover it.''

There is a French saying, that what is true on one side of the Pyrenées is false on the other.