E6

Self-similarity and panorama of self-affinity

✦ **Abstract.** This long and essential chapter provides this book with two of its multiple alternative introductions. The mathematically ambitious reader who will enter here will simply glance through Section 1, which distinguishes between self-similarity and self-affinity, and Section 2, which is addressed to the reader new to fractals and takes an easy and very brief look at self-similarity. Later sections approach subtle and diverse facets of self-affine scaling from two distinct directions, each with its own significant assets and liabilities.

Section 3 begins with WBM, the Wiener Brownian motion. In strict adherence to the *scaling principle* of *economics* described in Chapter E2, WBM is self-affine in a statistical sense. This is true with respect to an arbitrary reduction ratio r, and there is no underlying grid, hence WBM can be called the *grid-free*. Repeating in more formal terms some material in Sections 6 to 8 of Chapter E1, Section 3 discusses generalizations that share the scaling properties of WBM, namely, Wiener or fractional Brownian motion of fractal or multifractal time.

Section 4 works within grids, hence limits the reduction ratio r to certain particular values. Being *grid-bound* weakens the *scaling principle of economics*, but this is the price to pay in exchange for a significant benefit, namely the availability of a class of self-affine non-random functions whose patterns of variability include and exceed those of Section 3. Yet, those functions fall within a unified overall master structure. They are simplified to such an extent that they can be called "toy models" or "cartoons."

The cartoons are grid-bound because they are constructed by recursive multiplicative interpolation, proceeding in a self-affine grid that is the simplest case prescribed in advance. The value of grid-bound non-random fractality is that it proves for many purposes to be an excellent surrogate for randomness. The properties of the models in Section 3 can be

reproduced with differences that may be viewed as elements of either indeterminacy or increased versatility. Both the close relations and the differences between the cartoons could have been baffling, but they are pinpointed immediately by the enveloping master structure. At some cost, that structure can be randomly shuffled or more deeply randomized. Its overall philosophy also suggests additional implementations, of which some are dead-ends, but others deserve being explored.

Wiener Brownian motion and its cartoons belong to the *mild* state of variability or noisiness, while the variability or noisiness of other functions of Section 3 and cartoons of Section 4 are *wild*. The notions of states of mild and wild randomness, as put forward in Chapter E5, are generalized in Section 5 from independent random variables to dependent random processes and non-random cartoons. Section 5.4 ends by describing an ominous scenario of extraordinary wildness.

Being constrained to scaling functions, this chapter leaves no room for slow variability. ✦

WHEN DISCUSSING THE ORGANIZATION OF THIS BOOK, the Preface mentions several welcoming entrances. This and the preceding chapters are the entrances most suited for those who do not fear mathematics. (This chapter grew to become too long, and may be best viewed as several chapters bound together.

While Chapter E5 restricted itself to independent random variables, this chapter allows dependence, either deterministic or statistical, but restricts itself to self-affine scaling. This allows for mild and wild randomness, but not for slow randomness. That is, this chapter describes dependent functions or processes in continuing time that generalize a special family considered in Chapter E5, namely sequences of L-stable variables, with their Gaussian limit case.

The term *Panorama* in the title is meant to underline that, beyond the specific needs of this book on finance, this chapter also opens vistas that involve many other fields. Indeed, self-affine random variation is by no means restricted to economics. It is also often encountered in physics, for example, in $1/f$ *noises*. Different examples of those noises involve several of the variants in this *Panorama*, but there is no field in which all variants have been fully implemented. Fuller versions of the same text, updated and with different biases, are scheduled for M 1997N and M 1997H, which mention $1/f$ noise in the title. Those versions will be more technical and perhaps more practical.

The text can first be skimmed and later read in increasing detail. Students of finance who do not favor mathematics may be satisfied to examine the illustrations, and to be aware that this chapter helps organize and relate the Bachelier "B 1900 model," and my successive models in finance, M 1963, M 1965, M 1967 and M 1972, as sketched in Sections 6 to 8 of Chapter E1.

The strong term "cartoons" used to describe the "grid-bound" implementations of self-affinity collected in Section 4. A political cartoon's effectiveness hinges on its being highly simplified, yet preserving the essentials of what it refers to. In the same spirit, as known to readers familiar with elementary fractals and sketched below in Section 2 for the sake of other readers, the non-random Koch islands were mathematical curios until I injected them as "cartoons" of realistic fractal models of coastlines; in turn, those random models are cartoons of real coastlines. Some of the non-random self-affine constructions in Section 4 are cartoons of the random self-affine process in Section 3; the latter, in turn, are cartoons of real price records. Cartoons being unavoidable, the user should learn to like them, and the provider must develop ways to make them simple yet instantly recognizable.

Disclaimers. This *Panorama* is by no means the last word on its topic, in part, because the field of fractals has not yet become unified. Some studies grow from the top down: they first set general principles and then proceed to the consequences. To the contrary, fractal geometry grows from the bottom up. It continues to draw new substance from a succession of explorations with focussed ambitions. In parallel, it continues an effort to rethink the available substance in fashions that are increasingly organized, and suggest new explorations.

Use of the cartoons to resolve a widespread confusion between the M 1963 and M 1965 models. Between the L-stable motion behind the M 1963 model, and the fractional Brownian motion behind the M 1963 model, mathematicians see a number of parallelisms often described as "mysterious." Fortunately, Section 4 suggests that self-affinity may be one of those cases for which order and simplicity are restored, and confusion vanishes, when *a*) the standard models are made *more*, rather than *less*, general, and *b*) the resulting wider family of possibilities is presented in very graphic fashion.

In particular, very simple arguments relative to the cartoons suffice to eliminate a confusing complication that concerns the value of the fractal dimension. Depending on which feature is being singled out, the dimension is as follows:

• Either $D_G = 2 - H$ or $D_T = 1/H$ for the graphs or trails of fractional Brownian motions (M 1965 model; see Section 3.3).

• Either $D = 2 - 1/\alpha$ or $D_T = \alpha$ for the graphs or trails of L-stable processes (M 1963 model).

• Moreover the M 1972 model (see Section 3.13) yields two values, D_T and $D_G > D_T$ that are not functionally related to each other. Other dimensions also enter into contention.

A multiplicity of binary splits. It is useful to underline the versatility of self-affine constructions by describing how they split in several overlapping ways. The following list uses terms that will not be defined until later in this chapter, therefore should be viewed as merely suggestive.

• Between *grid-free* and *grid-bound.*

• Between *mildly* and *wildly variable.*

• Between *continuous* and *discontinuous.*

• Between *monotone,* either non-decreasing or non-increasing, and *oscillating* up and down.

• Between *non-intermittent,* that is, allowing no interval of clock time when motion stops, and *intermittent,* with variation concentrated on a fractal trading time. Variation can also be *relatively intermittent,* if it is concentrated on a new construct: a multifractal trading time.

• Between *unifractal,* characterized by a single exponent H, *mesofractal,* which also includes other values of H restricted to be 0 and/or infinity, and *multifractal,* characterized by a distributed exponent H.

• When H is single-valued, between the case $H = 1/2$ and the cases $H \neq 1/2$.

• Finally (but this will not be discussed in this book), between constructions that are *stable* or *unstable* under wild randomization.

1. CONTRAST BETWEEN SELF-SIMILARITY AND SELF-AFFINITY

This chapter concerns scaling behavior in the graph of a function, more precisely, linearly scaling behavior. Before seeking examples, one must know that this scaling has two principal geometric implementations: self-similar fractals, and the more general self-affine fractals. Self-similarity, the narrowest and simplest, is the most standard topic of fractal geometry, and it is good to begin by briefly considering it in Section 2. But the

remainder of this chapter and this book are limited to functions whose graphs are *self-affine fractals*.

This distinction is essential, and it is most unfortunate that many authors use one word, *self-similar*, to denote two concepts. I gave a bad example, but only until M 1977F, when I found it necessary to introduce the term *self-affine*. This term is now accepted by physicists, engineers, and mathematicians who study non-random constructs. Unfortunately, many probabilists persist in using *self-similar* when they really mean *self-affine*; this is the case in the book by Baran 1994 and Somordnitsky & Taqq 1994.

Let us elaborate. Many geometric shapes are approximately isotropic. For example, no single direction plays a special role when coastlines are viewed as curves on a plane. In first-approximation fractal models of a coastline, small pieces are obtained from large pieces by a similarity, that is, an isotropic reduction (homothety) followed by a rotation and a translation. This property defines the fractal notion of self-similarity. Self-similar constructions make free use of angles, and distances can be taken along arbitrary directions in the plane.

But this book deals mostly with geometric shapes of a different kind, namely, financial charts that show the abscissa as the axis of time and the ordinate as the axis of price. The scale of each coordinate can be changed freely with no regard to the other. This freedom does not prevent a distance from being defined along the coordinate axes. But for all other directions, the Pythagorean definition,

$$\text{distance} = \sqrt{(\text{time increment})^2 + (\text{price increment})^2}\,,$$

makes no sense whatsoever. It follows immediately that circles are not defined. Rectangles must have sides parallel to the axes. Squares are not defined, since – even when their sides are meant to be parallel to the axes – there is no sense in saying that time increments = price increment.

There is a linear operation that applies different reduction ratios along the time and price axes. It generalizes similarity, and Leonhard Euler called it an *affinity*. More precisely, it is a *diagonal affinity*, because its matrix is diagonal. It follows that for graphs of functions in time, like price records, the relevant comparison of price charts over different time spans involves the scaling notion of self-affinity. Self-affinity is more complicated and by far less familiar than self-similarity, therefore this chapter begins by surveying the latter. Readers already acquainted with fractals may proceed to Section 3.

On the measurement of texture, irregularity or roughness. The very irregular and rough shapes often encountered in Nature never tire of exciting the layman's imagination, but science long failed to tackle them. Thus, no serious attempt was made to define and measure numerically the irregularity of a coastline or a price record.

Topology provides no answer, even through its name seems to promise one. For example, consider a chart or time record of prices when filled-in to be continuous; this curve can be obtained from the line without a tear, using a one-to-one continuous transformation. Disappointedly, this property defines *all* price charts as being topological straight lines!

Nor does statistics provide a useful answer. For example, examine the perennial and objective problem of measuring the roughness of physical surfaces. Statistics suggests following a procedure familiar in other fields: first fit a trend-like plane (or perhaps a surface of second or third degree), then evaluate the root-mean-square (r.m.s.) of the deviation from this trend. What is unfortunate is that this r.m.s., when evaluated in different portions of a seemingly homogeneous surface, yields conflicting values.

Does the inappropriateness of topology and statistics imply that irregularity and roughness must remain intuitive notions, inaccessible to mathematical description and quantitative measurement? Fractal geometry is a geomety of roughness, and it answers with a resounding no. It shows that in many cases, roughness can be faced and overcome to a useful extent, thanks to scaling exponents that underlie the scaling principles of mathematical and natural geometry.

For example, coastlines are nearly self-similar, and the most obvious aspect of their roughness is measured by a quantity called *fractal dimension*, which is described in Section 2. Many irregular physical surfaces are self-affine, and their roughness is measured reliably by two numbers. One is the exponent H introduced in Section 3; engineers have already come around to call it simply the "roughness exponent," but mathematically, it is a Hurst-Hölder exponent and a fractal co-dimension. The second characteristic number is a scale factor similar to a root-mean-square, but more appropriately defined. (This topic is treated in M 1997H). The study of roughness in terms of self affinity has become a significant topic in physics; see Family & Vicsek 1991.

2. EXAMPLES OF SELF-SIMILAR RECURSIVE CONSTRUCTIONS

2.1 Getting answers without questions to work on questions without answers

Of the five diagrams in Figure 1, the largest and most complicated is the composite of two wondrous and many-sided broken lines. One of them is violently folded upon itself, and gives the impression of attempting a monstrous task for a curve: to fill without self-contact the domain bounded by the less violently folded second curve. This impression was intended, since we witness an advanced stage of the construction of a variant "space-filling curve," an object discovered by Giuseppe Peano (1858-1932).

Actually, "space-filling curve" is an oxymoron. An improved substitute that I proposed is "space-filling (or Peano) *motion*." Thus, Figure 1 illustrates a variant of the original Peano motion, bounded by a less violently folded fractal "wrapping." By construction, both curves are precisely as complicated in the small as in the large. The wrapping, introduced in M 1982F{*FGN*}, Chapter 6, is patterned after one that Helge von Koch used in a celebrated shape called "snowflake curve." The filling was introduced in M 1982t. Never mind that Koch's motivation was purely mathematical: he was seeking a curve without tangent anywhere, meaning that the direction of a cord joining any two points has no limit as these points converge to each other. To achieve this goal, the simplest was to demand that this cord fluctuate exactly as much in the small as in the large.

Fractal geometry preserved this demand, but changed its motivation from purely mathematical to very practical. When irregularity is present at all scales, it is simplest when, whatever the magnification, the fine details seen under the microscope are the same (scale aside) as the gross features seen by the naked eye. Using the vocabulary of geography, the fine details seen on a very precise map are the same as the gross features seen on a rough map. Concrete reinterpretations of Koch's recursive procedure continually inspire me in empirical work. In summary, one can state two guiding principles.

A) *Scaling principle of natural geometry.* Shapes whose small and large features are largely identical except for scale, are useful approximations in many areas of science.

B) *Scaling principle of mathematical geometry.* Sets wherein small and large features are identical except for scale are interesting objects of study in geometry.

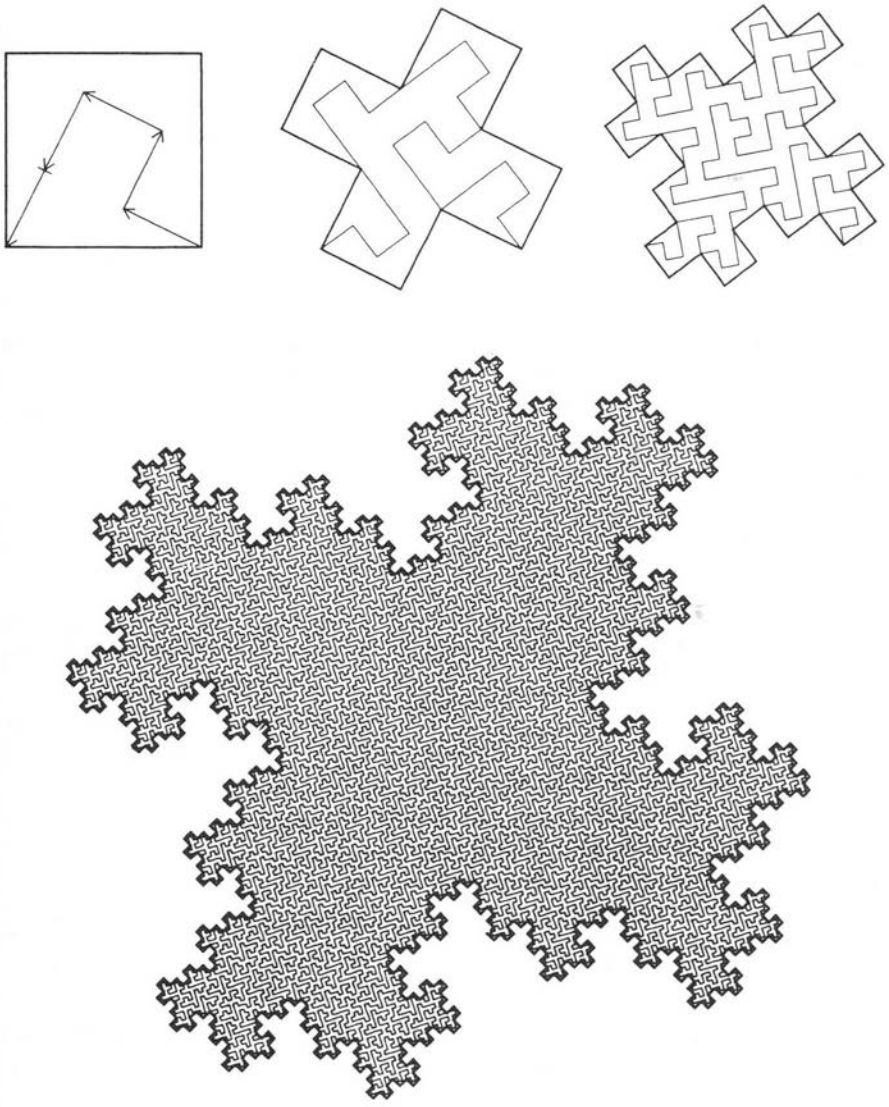

FIGURE E6-1. Construction of a Peano motion "wrapped" is a squared Koch curve that it fills.

Part of my life-work consists in viewing B as providing a collection of answers without question, and setting them to work on the questions without answers summarized under A.

2.2 Examples of self-similar fractal shapes

To implement the goal that Koch stated in his way before I restated it in mine, the easiest is to proceed step by step. Select an *initiator*, often an interval, and a *generator*, also a broken line. The first construction stage replaces each side of the initiator by an appropriately rescaled, translated and rotated version of the generator. Then a second stage repeats the same construction with the more broken line obtained at the first stage, and so on.

The early stages of the constructions shown on Figure 1 are illustrated by four small diagrams to be followed clockwise from left center, in order of increasing complication. The initiators are the four sides of a unit square for each of four repeats of the wrapping of Peano motion and one side of this square for the Peano motion itself. The generator of the motion is an irregular open pentagon that does its best to fill the square, using sides equal to $1/\sqrt{5}$. One perceives an underlying square lattice of lines $1/\sqrt{5}$ apart, and the Peano generator crosses every lattice vertex contained in the original wrapping. The wrapping generator has $N = 3$ sides of length $r = 1/\sqrt{5}$.

In the next stage of the construction, each side of the pentagon is replaced by an image of its whole reduced in the ratio of $1/\sqrt{5}$, and suitably rotated. The result no longer fits within the square, but fills uniformly the cross-like shape obtained by replacing each side of the square by the wrapping generator. The same two constructions are then repeated ad infinitum in parallel. Zooming in as the construction proceeds, one will constantly witness the same density of filling; watching without zooming in, one sees a curve that fills increasingly uniformly a wrapping whose complexity keeps increasing.

The Peano motions which mathematicians designed during the heroic period from 1890 to 1922 filled a square or a triangle, but the present boundaries are more imaginative.

Figure 2 carries the construction of a curve of Figure 1 one step further and the filling is interpreted as the cumulative shoreline of several juxtaposed river networks; the wrapping is the combination of a drainage divide surrounding these networks and of a portion of seashore. To build up the network, one proceeds step by step: (1) Each dead-end square in

the basic underlying lattice – meaning that three sides belong to the filling – is replaced by its fourth side, plus a short "stream" with its source at the center of the dead-end, and its end at the center of the square beyond the newly filled-in side. (2) One proceeds in the same fashion with the polygons left in after the processed dead-ends are deleted. (3) And so on until the filling is changed from a broken line with no self-contact to a collection of "rivers" forming a tree. At this point, the wrapping becomes reinterpreted as the river network's external drainage divide.

To use an old sophomoric line, after you think of it imaginatively, carefully, and at great length, it becomes obvious that a plane-filling motion fails at its assigned task of being a mathematical monster. I proved it to be nothing but a river network's cumulative shore. The converse is also true. Much better-looking river networks are given in my book, M 1982F{*FGN*}, but the basic idea is present here. There is not much else to Peano motions. Thus, the mathematicians who used to tell us that Peano motions are totally nonintuitive had deluded themselves and misinformed the scientists.

2.3 The notion of fractal dimension of a self-similar geometric shape

Each stage of a Koch construction replaces an interval of length 1 by N intervals of length r, therefore multiplies a polygon's length by a fixed factor $Nr > 1$. It follows that the limit curves obtained by pursuing the recursions ad infinitum are of infinite length. Furthermore, it is tempting to say that the filling is "much more infinite" than its wrapping, because its length tends to infinity more rapidly. This intuitive feeling is quantified mathematically by the notion of fractal dimension. The original form was introduced by Hausdorff and perfected by Besicovitch. It is inapplicable to empirical science, and had to be replaced by a variety of alternative definitions.

The explanation of the underlying idea begins with the very simplest shapes: line segments, rectangles in the plane, and the like. Because a straight line's Euclidian dimension is 1, it follows, for every integer $\gamma > 1$, that the "whole" made up of the segment of straight line $0 \le x < X$ may be "paved over" (each point being covered once and only once) by $N = \gamma$ segments of the form $(k-1)X/\gamma \le x < kX/\gamma$, where k goes from 1 to γ. Each of these "parts" can be deduced from the whole by a similarity of ratio $r(N) = 1/N$. Likewise, because a plane's Euclidian dimension is 2, it follows that, whatever the value of γ, the "whole" made up of a rectangle $0 \le x < X; 0 \le \gamma < Y$ can be "paved over" exactly by $N = \gamma^2$ rectangles defined by $(k-1)X/\gamma \le x < kX/\gamma$ and $(h-1)X/\gamma \le y < hY/\gamma$, where k and h

go from 1 to γ. Each part can now be deduced from the whole by a similarity of ratio $r(N) = 1/\gamma = 1/N^{1/2}$. Finally, in a Euclidian space whose dimension is $E > 3$, a D-dimensional parallelepiped can be defined for any $D \leq E$. All those classical cases satisfy the identity

$$D = \frac{-\log N}{\log r(N)} = \frac{\log N}{\log(1/r)} .$$

This expression is the self-similarity dimension. Its value lies in the ease with which it can be generalized. Indeed, the fact that it was first used for a segment or a square is not essential for its definition. The critical requirement is scaling, meaning that the whole can be split into N parts deducible from it by a self-similarity of ratio r (followed by translation, rotation, or symmetry). Such is precisely the case in Figure 1. For the wrapping, $N = 3$ and $r = 1/\sqrt{5}$, hence

$$D = \log 3/\log\sqrt{5} = \log 9/\log 5 = 1.3652.$$

For the filling, $N = 5$ and $r = 1/\sqrt{5}$, hence

$$D = \log 5/\log\sqrt{5} = 2.$$

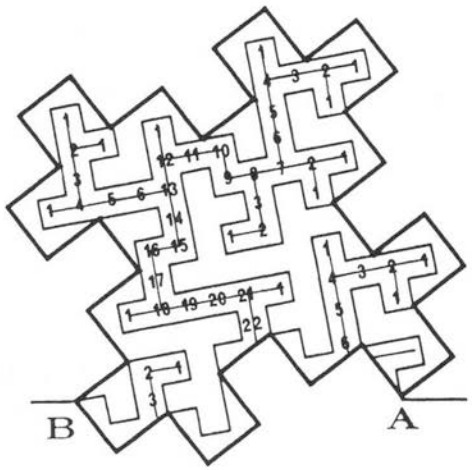

FIGURE E6-2. The third diagram of Figure 1, reinterpreted in terms of a river network. This interpretation led M 1982F{*FGN*} to boast of having "harnessed the Peano Monster Curves."

Thus, the impression that the filling is "more infinite" than its wrapping is both confirmed, and quantified by the inequality between their dimensions. The impression that the filling really fills a plane domain is confirmed and quantified by its dimension being $D = 2$.

The preceding argument may seem overly specialized, so it may be comforting to know (a) that fractal dimension can be defined using alternative methods of greater generality and full rigor and (b) that the result behaves in many other ways like the old-fashioned integer-valued dimension. For example, consider the notion of *measure*. If a set is self-similar and measure is taken properly, then the portion of this set that is contained in a sphere of radius R is of measure proportional to R^D.

3. SELF-AFFINE FRACTAL MODELS IN FINANCE

Joined by readers who knew about fractals and skipped section 2, we now turn from self-similarity to self-affinity and to a collection of possible models of price variation that follow the scaling principle of economics. Sections 3 and 4 cover roughly the same material in two very different ways. There is enough overlap to allow Sections 3 and 4 to be read in either sequence.

The bare facts were already sketched in Sections 6 to 8 of Chapter E1, but *one need not* read these sketches before this Section. Furthermore, this chapter has no room for a full treatment of L-stable motion, fractional Brownian motion and multifractals; L-stable motion is the topic of much of the second half of this book; fractional Brownian motion is the topic of M 1997H, and multifractals are the topic of M 1997N.

3.1 The 1900 model of Bachelier, Brownian motion

$B(t)$ is defined as being a random process with Gaussian increments that satisfies the following "Fickian" diffusion rule:

$$\text{for all } t \text{ and } T,\ E\{B(t+T) - B(t)\} = 0 \text{ and } E\{B(t+T) - B(t)\}^2 = T.$$

A Fickian variance is an automatic consequence if the increments are assumed independent. Conversely, Fickian variance guarantees the orthogonality of the increments. Adding the Gaussian assumption, it guarantees independence.

3.2 Tail-driven variability: the M 1963 model and the L-stable processes

Reference. The concept of L-stability is discussed throughout the second half of this book, and it would be pointless to repeat here the definition due to Paul Lévy. It is enough to say that L-stability means that the sum of N *independent* L-stable variables is itself L-stable. The Gaussian shares this property and indeed it is a limit case of the L-stable variables, when the parameters α tends to 2. Moreover, consider a weighted index of *independent* variables $\Sigma W_g X_g$. When the weights are *not* random, the variables X and the weighted index are L-stable.

The "ruin problem" for the L-stable processes. Suppose a speculator is called ruined if his holdings fall beyond a prescribed level called "threshold." What is the probability that ruin occurs before a time t_{max}? Questions of this type are thoroughly explored for Wiener Brownian motion. For L-stable processes, the literature is limited, but includes Darling 1956 and Ray 1958.

Invariance under non-randomly and randomly weighted forms of addition. This digression is addressed to readers who know the concept of fixed point of a (semi-)group of transformations. L-stable variables are *fixed points* in the operation that consists in transforming independent random variables by taking a *non-randomly weighted average*. A distribution invariant under addition of *independent* addends used to be thought as necessarily Gaussian until M 1960i{E10} injected L-stable addends in a down-to-earth concrete situation. The M 1972 model, to be presented in Section 3.8, involves an actual generalization of L-stability, and it is good to mention how this generalization relates to Lévy stability. The considerations in M 1974f{N15} and M 1974c{N16} also involve a weighted index $\Sigma W_g X_g$; but there is the important innovation that the weights W are not constants, but independent values of the same random variable W.

Given W and N, I investigated the variable Y such that $\Sigma W_g Y_g$ has, up to scale, the same distribution as Y. Using the terminology already applied to L-stable variables, my variables Y are *fixed points* in the operation that consists in taking *randomly weighted averages of independent* random variables. The variables Y range from being close to L-stable (a limit case) to being very different indeed.

3.3 Dependence-driven variability: the 1965 model and fractional Brownian motion

The fractional Brownian motion (FBM) $B_H(t)$ is the random process with Gaussian increments that satisfies the following diffusion rule

$$\text{for all } t \text{ and } T,\ E[B_H(t+T) - B_H(t)] = 0 \text{ and } E[B_H(t+T) - B_H(t)]^2 = T^{2H}.$$

The value $H = 1/2$ yields the Wiener Brownian motion, whose diffusion is called "Fickian." However, the exponent H is only constrained to $0 < H < 1$. For $H \neq 1/2$, the diffusion of FBM is widely called "non-Fickian." In a different terminology, a mysterious but widely used one, H is called "strength of singularity" at time t.

This process was introduced in M 1965h{H} and fully described in M & Van Ness 1968{H} as a model of diverse phenomena that exhibit cyclic non-periodic variability at all time scales. The oldest recorded example concerned the annual discharge of the Nile River and is associated with the Biblical story of Joseph, the son of Jacob. Therefore, I refer to non-periodic cyclicity as the *Joseph Effect*. The use of FBM is economics was pioneered in M 1970e, M 1971n, M 1971q, M 1972c and M 1973j. Recent mathematical references are Baran 1994 and Samorodnitsky & Taqqu 1994 (Section 7.2). Unfortunately, as already mentioned, both books use the word self-similarity where the correct concept, hence the correct term, is self-affinity). A recent book for engineers is Bras & Rodriguez-Iturbe 1993 (pages 210-261).

The property of uniscaling. The above definition implies that the scale factors based on moments satisfy

$$\{E[\,|B_H(t+T) - B_H(t)|^q\}^{1/q} = \text{(a constant) } T^H$$

for all powers $q > -1$. (For $q \leq -1$, this expression becomes infinite.) That is, q-th order scale factor defined by the left hand side, is independent of q. (For $q < -1$, the left hand side is infinite, therefore the equality holds trivially, with an infinite constant). This obvious corollary is said to express *uniscaling*. It will become important in Section 3.8 and 3.9, and the cases when the q-th scale factor depends on q will be called *multiscaling*. By contrast, $B_H(t)$ will be called *uniscaling*, and no complication can arise from writing $\Delta B_H \sim \Delta t^H$.

Hurst, Hölder, and an exponent that bridges mathematics and concrete needs. With a suitable definition of the symbol $\sim$, the functions $B_H(t)$, including $B(t) = B_{1/2}(t)$, satisfy

$$\frac{\log |\Delta B_H|}{\log \Delta t} \sim H.$$

Observe that, in this definition, Δt and ΔB_H are increments over a non-vanishing interval, *not* infinitesimal quantities. One says that, as defined here, H is a *coarse* quantity, not a *fine* or *local* one. In addition, H is defined for all values of t.

The idea behind the exponent H has two thoroughly disparate historic roots. When introduced in M 1965h{H}, $B_H(t)$ was motivated by a difficult problem from civil engineering, and referred to the initial letter of the hydrologist H. E. Hurst (1880-1978), briefly mentioned in Chapter E. But H also has a second set of deep roots in pure mathematics, namely, in the work of L. O. Hölder (1859-1937). Serendipitously, the names of Hurst and Hölder shared the same initial letter. However, Hölder's original definition had to be very much generalized. In Section 4, the underlying idea will split further.

$B_H(t)$ and the phenomenon of long-run statistical dependence. The most striking single property of $B_u(t)$ concerns the quantities $[B_H(0) - B_H(-T)]/T$, called *past* average and $[B_H(T) - B_H(0)/T$, called *future* average. Both are Gaussian random variables, and their correlation is easily seen to be

$$C = \frac{1}{2} \frac{(2T)^{2H} - T^{2H} - T^{2H}}{(T^H)^2} = 2^{2H-1} - 1.$$

That is, C is independent of T. This fact could be called "intuitive" because it follows from self-affine scaling. But an older form of "intuition" of the nature of randomness is more demanding, and insists that a distant past and a distant future "should" become statistically independent. This second intuition is correct in the Wiener Brownian case, where $C = 0$, and in other cases of mild randomness. But it must be "unlearned" in all other cases.

More precisely, $C > 0$ in the "persistent" case $1/2 < H < 1$ and $C < 0$ in the "anti-persistent" case $0 < H < 1/2$. In both cases, $B_H(t)$ *is* definitely *neither* a martingale *nor* a Markov process.

Spectral properties: the fractional Gaussian noise $B'_H(t)$ as a "continuing" or "humming" form of " $1/f$ noise." $B_H(t)$ is continuous but not differentiable. However, one can define for it a "generalized derivative" $B'_H(t)$. The spectral density of $B'_H(t)$ is $\propto f^{-B}$, with the exponent $B = 2H - 1$ ranging between 1 and -1. Physicists denote such phenomena by the curious term of "$1/f$ noises." When $1/2 < H < 1$, the spectral density diverges at $f = 0$. This is one token of long-run statistical dependence. When $0 < H < 1/2$, the integral of the spectral density is 0, which is a different (and far less "robust") token of long-run statistical dependence.

The multiplicity of co-existing fractal dimensions for $B_H(t)$, including the value $D_G = 2 - H$, and the larger value $D_T = 1/H$. Section 2.3 describes how the irregularity of a self-similar fractal curve is in large part measured by a number called its "fractal dimension" D. Self-affine curves are significantly more complicated, as M 1997H will show from several distinct viewpoints.

A first complication is this. While self-similar fractals have a unique fractal dimension, I showed that self-affine fractals demand several, depending on which aspect is being considered. In the case of $B_H(t)$, some careful authors only quote the value $D_G = 2 - H$, while other careful authors only quote $D_T = 1/H > 2 - H$.

Those two sets of authors report different answers because the values $2 - H$ and $1/H$ refer to different geometric objects.

The value $2 - H$ can be shown to be the box dimension of the graph of $X(t)$, hence the suffix G.

The value $1/H$ can be shown to be the box dimension of a different but related geometric object, namely a "trail," hence the suffix T.

The distinction between graph and trail is developed in M 1982F{*FGN*}, but the main facts can be summarized here. First consider a Wiener Brownian motion in the plane. Its coordinates $X(t)$ and $Y(t)$ are independent Brownian motions. Therefore, if a 1-dimensional Brownian motion $X(t)$ is combined with another independent 1-dimensional Brownian motion $Y(t)$, the process $X(t)$ becomes "embedded" into a 2-dimensional Brownian motion $\{X(t), Y(t)\}$. The value $D_T = 2 = 1/H$ is the fractal dimension of the three dimensional graph of coordinates t, $X(t)$ and $Y(t)$, and the projected "trail" of coordinates $X(t)$ and $Y(t)$. However, the dimension $D_G = 2 - H$ applies to the projected graphs of coordinates t and $X(t)$ or t and $Y(t)$. A heuristic derivation of this value is best postponed to Section 3.13, where it will be generalized.

An FBM with $H \neq 1/2$ can only be embedded in a space of dimension $E \geq \max(2, 1/H)$.

We are done now with explaining how two values of the dimension coexist peacefully in the *unifractal* case of the FBM $B_H(t)$. Thinking ahead, Section 3.13 tackles the next case, to be called *multifractal*, and shows that D_T and D_G cease to be related functionally.

There is a second complication (but it is beyond the scope of this book): in the self-affine case, the notion of fractal dimension splits into local and global forms. The above-mentioned values are local, and the global values are different.

3.4 Trading time, compound processes, and a fundamental fact: preservation of the trail dimension D_T under compounding

The variation of most prices is neither tail- nor dependence-dominated, but ruled by both contributions in combination. To model such combinations, one must go beyond the M 1963 and M 1965 models. This is a task I first attacked piecemeal, by seeking suitable random functions and later attacked systematically, by introducing a flexible general family of random functions. (Actually, several options were considered, but the present discussion will be limited to one.)

Trading time and compound processes. The processes in this family are "compound," "decomposable," or "separable" in the following sense: by construction, their variation is "separated" into the combination of two distinct contributions. The first is a trading time θ, a random non-decreasing function of clock time t. In the terminology in Feller 1950 (Vol. II, p. 347), $\theta(t)$ is called *directing function*. The second, which yields price as function of trading time, $X(\theta)$, will be called *compounding function*.

In the absence of further restrictions, the notion of compounding is useless. Indeed, given a function $P(t)$, an arbitrary choice of $\theta(t)$ automatically defines also a function $X(\theta)$ such that $X[\theta(t)] = P(t)$. Our attention will be restricted to the case when the two components are statistically independent.

Furthermore, we wish to insure that the compound process is self-affine, that is, follows the scaling principle of economics. The easiest is to demand that *both* $\theta(t)$ and $X(\theta)$ be self-affine functions. In addition, the directed functions will be WBM and FBM, thus preserving something of the Bachelier model and the M 1965 model. The hope, of course, is that the outcome provides a sensible approximation to interesting data that are driven by a combination of tail and dependence.

Preservation of the trail dimension $1/H_T$ *under continuous compounding.* Section 3.3 distinguishes between the graph of $X(t)$ and an embedded trail of coordinates $X(t)$ and $Y(t)$. Compounding can be continuous or discontinuous, as will be seen momentarily. When it is continuous, it modifies the graph of $X(t)$, but leaves unchanged the trail of coordinates $X(t)$ and $Y(t)$. In particular, the trail dimension remains $1/H$. When compounding is discontinuous, it modifies both the graph and the trail.

Comment. In this section, trading time is a notion that is borrowed from our historical, and therefore intuitive, knowledge of how markets operate. In Section 4, trading time will enter in a far more intrinsic fashion.

3.5 A major but unrecognized "blind spot" of spectral analysis: spectral whiteness is insensitive to change of trading time, therefore misleading

To engineers, successive increments ΔB of Wiener Brownian motion define a *white noise.* They are independent, therefore uncorrelated ("orthogonal"), and their spectral density is a constant, defining a white spectrum. Now, let us follow Brownian motion in a trading time chosen at will (self-affine, or not). The increments of the compound motion are very strongly dependant. However, most remarkably, they are uncorrelated, therefore *they remain spectrally white.* In other words, spectra as applied to a compound process are only sensitive to the whiteness of the directed function, and completely blind to the properties of the directing function.

Indeed, given two non-overlapping time increments $d't$ and $d''t$, the corresponding increments $d'B(t)$ and $d''B(t)$ are, by definition, independent. It is obvious that this property continues to hold when B is followed in a trading time θ that is in a non-linear non-decreasing function of t, and $B(t)$ is replaced by $B^*(\theta) = B[t(\theta)]$. The increments of B^* exhibit very strong dependence, yet they are white, that is, uncorrelated.

Remark concerning statistical method. When interpreting spectra in a non-Gaussian and non-Brownian context, this dangerous possibility must be kept in mind. This serious "blind spot" was noted, but not developed, in my papers on noise of the 1960s, to be collected in M 1997H. It constitutes a fundamental limitation of spectral analysis that statistics must face.

Remark concerning the spectral whiteness of financial data. During the 1960s spectral analysis was introduced into economics with fanfare, but never lived to its promise. The "blind spot" of spectra suffices to account for many puzzling observations reported in the literature. Indeed, Voss 1992 and the contributors to Olsen 1996 are neither the first nor the only

authors to report on such whiteness. Both parties also examined records of absolute price change, or of price change squared. The spectrum is no longer white but instead takes the "$1/f$" form characteristic of FBM. It will be shown at the end of Section 3.9 that this apparent contradiction is characteristic of the M 1972 model, namely, of the Brownian motion in multifractal time.

Remark concerning R and R/S analysis. This form of analysis is mentioned and referenced in Section 7.4 of Chapter E1 and discussed in M 1997H. Changes in tradi g time leaves the range unchanged, but removal of the trend (as it is practiced in R/S) does modify the range. This topic must be withheld for consideration in M 1997H.

3.6 A special form of discontinuous compounding, "subordination;" the notion of fractal time

Definitions. The simplest directing function $\theta(t)$ are functions with non-negative statistically independent increments. This form of compounding is denoted by the term *subordination*, which is due to S. Bochner. The most general implementation is a non-decreasing random function with infinitely divisible increments. The topic is discussed in Feller 1950 (Vol. II, p. 347). When the compounding function is Markovian, so is the compounded function.

Self-affine subordination and the fractal devil staircases. When the directing function is self-affine, it must be an L-stable non-decreasing function, sometimes called "stable subordinator." This is a non-decreasing function of trading time whose graph is an inverse Lévy devil staircase, the latter being a Cantor devil staircase made random.

M 1982F{*FGN*} discusses Lévy staircases in Chapter 31 and illustrates them on Plate 286 and 287. It discusses Cantor staircases and illustrates one in Plate 83 of Chapter 8. The term "staircase" is motivated by the presence of flat steps. The steps are infinitely numerous, and most are infinitesimally small. Between its steps, a fractal staircase moves up by infinitesimal amounts. The values of θ where steps end form a "Cantor dust" or a "Lévy dust." The latter is fully characterized by a single exponent α which is a fractal dimension. Conversely, trading time followed as function of physical time, reduces to a series of jumps of widely varying size. The idea of subordination is that, a fleeting instant of clock time allows trading time to change by a positive amount, generating the price jumps to be considered in Section 3.7.

M 1977F proposed that a trading time ruled by a devil staircase be called a *fractal time*.

Subordination came to play an important role in many aspects of fractal geometry, therefore is discussed in detail in Chapter 32 of M 1982F{*FGN*}, where it is illustrated and interpreted in a variety of contexts.

3.7 Fractal compounding: LSM is identical to WBM, as followed in a trading time defined by a fractal devil staircase

A representation of L-stable motion. The original and simplest form of subordination was used in M & Taylor 1967{Sections 1 and 2 of E21}, to which the reader is referred. It takes price to be a Wiener Brownian motion of fractal trading time. The interesting fact is that the procedure happens to reproduce exactly the L-stable process that M 1963b{E14} proposed for the Noah Effect. The exponent α is "fed in" by the Lévy staircase.

A generalization that calls for detailed exploration: fractional Brownian motion of fractal time. This obvious generalization has two parameters: the α exponent of the Lévy staircase, which is a fractal dimension, and the exponent of the compounding function. $B_H(t)$, the Hölder exponent of the observed process, depends on α and H as we shall see in Section 3.9.

As mentioned in Section 6 of Chapter E1 and Section 3 and Annotations in Chapter E21, Clark 1973 proposed to preserve subordination, while replacing fractal time by a lognormal time, which is non-fractal. M 1973c{E21, Section 3} argued against Clark's substitute. But I never implied that the M 1963 model, as restated in M & Taylor 1967{E21}, said the last word, quite to the contrary. However, instead of "patching up" the subordinator, I propose to replace subordination itself by a suitable more general form of compounding.

3.8 A form of continuous compounding, called multifractal, and a form of variability driven by tail and serial dependence acting together

A direct introduction of dependence into LSM had proven difficult, but compounding beyond subordination opened the gates to diverse possibilities, to which we now proceed. Observe that LSM, FBM and subordination were part of the mathematical literature, but what follows is new, even from the mathematical viewpoint.

Multifractality. The key step in moving beyond subordination consists in changing trading time from fractal to a more richly structured (and more complicated) form called *multifractal.* This step is explained in Chapter ix of M 1975o and in a section on "relative intermittency" on p.

375 of M 1982F{*FGN*}: both argue that many patterns that seem fractal in a first approximation prove on a second look to be multifractal. This step is now taken near-automatically in many fields. It was first taken in M 1969b, a paper concerned with turbulence, and my first full publication in that field, M 1972j{N14} ends (p. 345 of the original) as follows:

"The interplay ... between multiplicative perturbations and the lognormal and [scaling] distributions has incidental applications in other fields of science where very skew probability distributions are encountered, notably in economics. Having mentioned the fact, I shall leave its elaboration to a more appropriate occasion."

Multifractal measures and functions. The concept introduced in M 1972j{N14} and developed in 1974f{N15} and M 1974c{N16} involves non-decreasing multifractal random functions with an infinite number of parameters. Their increments are called *multifractal measures*. The original example introduced in M 1972j{N14} is the "limit lognormal multifractal measure;" it remains after all those years the main example that is "homogeneous" in time. Most explicitly constructed multifractals are grid-bound "cartoons;" they are defined and studied in Section 4. (In the same vein, the main example of fractal trading time with strong homogeneity remains the Lévy staircase used in Section 3.6. Figure 4 of Chapter E1 is a plot of the measures contained within successive intervals of the abscissa, and was originally simulated on a computer in order to model the gustiness of the wind and other aspects of the intermittency of turbulence. But the resulting pattern reminded me instantly of something entirely different, namely, Figure 1 of M 1967j{E15} which represents the variance of cotton price increments over successive time spans.

The limit lognormal multifractal measures are *singular*, and the same is true of all the examples invoked in the early literature – but not of some more recent ones. Being "singular", the integral $M(t)$ of the plot in Figure 5 of Chapter E1 is monotone increasing and continuous, yet non-differentiable anywhere. There is no trace of the step-like intervals corresponding to vanishing variation that characterize the Cantor and Lévy devil staircases. My immediate thought in 1972 was to use this function $M(t)$ as graph of a multifractal trading time $\theta(t)$. In the simplest cases, the inverse function $t(\theta)$ is also multifractal. This thought was not elaborated until recently and is published for the first time here and in three papers by M, Fisher & Calvet in different permutations. Tests delayed for twenty-five years suggest that my 1972 hunch led to a surprisingly good approximation, as will be seen in Section 3.15. I heard rumors of other

investigations of multifractals in finance; after all, once again, this is the next obvious step after fractals. But it is also an extremely delicate one.

A remarkable novelty: Multifractals allow concentration to occur with or without actual discontinuity. The fact that the typical early functions $M(t)$ are continuous is linked to the subtitle of this book and the topic of Section 1.3 of Chapter E2. Indeed, WBM and FBM of multifractal time are capable of achieving an arbitrarily high level of concentration *without* the actual discontinuity that is characteristic of LSM.

As a matter of fact, LSM can be viewed as a limit case. If one looks very closely, this limit is atypical and the convergence to it is singular. But this book need not look close enough to be concerned.

3.9 Characterization of multiscaling: "tau" functions that describe the moments' behavior for the directing and the compound functions

Except for scale, FBM is characterized by one parameter, LSM by two, and the major properties of a self-similar fractal follow from one parameter, its fractal dimension. Multifractals are more complicated: the closer one investigates them, the larger the number of parameters. This is because multifractals are characterized by a plethora of scaling relations, with correspondingly many exponents. The list of principal exponents defines a function "tau" which will now be described in two forms. (While this function is fundamental, it does *not* uniquely describe a multifractal.)

The moment exponent function $\tau_D(q)$ of the directing function. In a multifractal measure, as first shown in M 1974f{N15} and M 1974c{N16}, the moments of ΔM typically take the form

$$E[(\Delta M)^q] = \Delta t^{\tau_D(q)+1}.$$

(*Digression.* Some readers may be surprised by the equality sign, because other writers define a function τ as a limit. The technical reason is that the original method I used to define multifractals focuses on "fixed points" for which equality prevails.)

Moment-based scaling exponents. The q-th root of the q-th moment is a scale factor. For multifractals,

$$\{E(\Delta M)^q\}^{1/q} = \Delta t^{\sigma_D(q)}, \text{ where } \sigma_D(q) = \frac{1+\tau_D(q)}{q}.$$

(*A warning.* The literature also uses the notation $D(q) = \tau(q)/(q-1)$; except in the unifractal case, $D(q) \neq \tau(q)$).

The uniscaling cases. When multifractal time reduces to clock time, $\tau_D(q) + 1 = q$, implying uniscaling, since $\sigma_D(q)$ is independent of q.

The multiscaling cases. In the cases to be considered in this chapter, $\tau_D(q)$ satisfies two conditions: a) $E[(\Delta M)^0] = 1$, that is, $\tau_D(0) = -1$, and b) $E(\Delta M) = \Delta t$, that is, $\tau_D(1) = 0$. However, the graph of $\tau_D(q) + 1$ is not a straight line. It follows that $\sigma_D(q)$ decreases as $q \to \infty$.

Interpretation of the quantity $\tau'_D(1) = D_1$. This quantity has a very important concrete interpretation, as the fractal dimension of the set of values of $\theta(t)$ where the bulk of the variation of θ occurs. It is often denoted as D_1, and will be needed momentarily.

The power exponent function $\tau_C(q) = \tau_D(qH)$ *of the compound function.* Since $\Delta X = G(\Delta\theta)^H$, where G is a reduced Gaussian,

$$E[|\Delta X|^q] = E[|G|^q]E[(\Delta\theta)^{qH}] = (\text{a numerical constant})(\Delta t)^{1+\tau_D(qH)}.$$

This important new result defines an additional "tau" function, namely,

$$\tau_C(q) = \tau_D(qH).$$

"Multifractal formalism." This is the accepted term for the study of the functions $\tau(q)$ and associated functions customarily devoted by $f(\alpha)$. The latter are often called "singularity spectra," but they are best understood by generalizing to oscillating function, the original approach pioneered in M 1974c{N16}: they are limits of probability densities of ΔX, but plotted in a special way. The general idea can be inferred from the discussion in Section 8.4 of Chapter E1, where it is pointed out that linear transformation *cannot* collapse the densities, but *can* collapse the quantities $\rho_\sigma(u)$. Calvet, Fisher & M 1997 sketches the role of the function $f(\alpha)$ in the context of economics, and numerous chapters of M 1997N will fully describe my approach to multifractal measures and functions, and compare it to alternative approaches.

3.10 The FBM of multifractal time accounts for two facts about the tails that constitute "anomalies" with respect to the M 1963 model

It was mentioned repeatedly that reports came out very early that some price records disagree with the M 1963 model. Some authors report tails that follow the scaling distribution but with an exponent α that exceeds

Lévy's upper bound $\alpha = 2$. Other authors report distributions that fail to collapse when superposed with the proper scaling exponent.

We shall now show that multifractals provide a framework compatible with either or both observations.

The critical tail exponent q_{crit}. The equation $\tau_D(q) = 0$ has always the root $q = 1$. In addition, the function τ_D being cap convex, the equation $\tau_D(q) = 0$ may also have a finite second root; when it exists, it is denoted by q_{crit}. The limit lognormal case always yields $q_{crit} < \infty$. An important and suprising discovery is reported in M 1972j{N14}, M 1974{N15} and M 1974c{N16}: when a second root q_{crit} exists, the distribution of ΔM has an asymptotically scaling tail of the form

$$\Pr\{M > u\} \sim u^{-q_{crit}}.$$

Thus, from the viewpoint of the tail, q_{crit} is a "critical tail exponent." It plays the same role as the Lévy exponent α, namely, $E(\Delta M)^q < \infty$ if, and only if, $q < q_{crit}$. The essential novelty is that the range of q_{crit} is *no longer* $0 < q_{crit} < 2$; instead it becomes $1 < q_{crit} < \infty$.

A way to obtain a tail exponent of price change that exceeds the upper bound 2 that is characteristic of L-stability. Now return to compounding, namely to a fractional Brownian function $B_H(t)$ of a limit lognormal trading time. Its increments will satisfy $E(\Delta B_H)^q < \infty$ if, and only if, $q < q_{crit}/H = \alpha$. In the Brownian case $H = 1/2$, q_{crit} can range over $[1, \infty]$, hence α can range over $[2, \infty]$, which conveniently extends the L-stable range $[1, 2]$ of α. Furthermore, choosing H in the range $[1/2, 1]$ extends the range of α to $[1, \infty]$, which is the maximum conceivable in the case where expectations are finite.

Nevertheless, q_{crit} need not exist, that is, a multifractal ΔM *need not* have a scaling tail. This may sound confusing, but only means that not every property of every multifractal is scaling.

It is nice that multifractal trading time makes it possible to extend the range of the asymptotic exponent $\alpha > 2$, but this result is not achieved without major changes. Indeed, the multifractal increments ΔM are *not* scaling in the sense that applies to the L-stable variables. They have more than one characteristic exponent, hence a structure, called *multiscaling*, that is far richer and has many distinct aspects.

Multiscaling implies that the tails of the compound process become increasingly shorter as T increases. This is because the scale factors $[E[\Delta X]^q]^{1/q}$ are scaling and their exponent $\sigma_D(q) = [1 + q_D(\tau H)]/q$ decreases as $q \to \infty$. To

illustrate the importance of this fact, consider the renormalized increments $\Delta X/[E\Delta X^2]^{1/2}$. Contrary to the L-stable increments of the M 1963 model, those multiscaling increments do *not* collapse; instead, their distributions' tails become shorter and shorter as T increases. In other words, the multifractality of trading time is a sufficient explanation of the two anomalies described in Section 3.8.

3.11 Fourier spectral properties before and after rectification

The spectral exponent B_C of the increments of the compound process. The behavior of $E[\Delta X^2]$ for $\Delta t \to 0$ determines the behavior for $f \to \infty$ of the spectral density of the increments of X. That density takes the "$1/f$" form:

$$\text{spectral density} \sim f^{-B_C}, \quad \text{where} \quad B_C = \tau_C(2) = \tau_D(2H).$$

The WBM case, $H = 1/2$, yields $B_C = \tau_D(1) = 0$, as we already know from Section 3.5. When $H \neq 1/2$ but is close to $1/2$, we have

$$B_C = \tau(2H) \sim \tau_D(1) + (2H-1)\tau'_D(1) = (2H-1)\tau'(1) = (2H-1)D_1.$$

Conclusion. In the white case $H = 1/2$, we encounter once again the very important blind spot of spectral analysis noted in Section 3.5. For $H \neq 1/2$, compounding changes the spectral exponent. However, the nearly white cases exhibit an extraordinary and very welcome simplification: the exponent B_C of the compound process "separates" into a product. In the case $D_1 = 1$, which corresponds to FBM in clock time, it is confirmed that the spectral exponent and sole parameter of the increments of the compounding function is $(2H-1)$. As to the directing function, it is not represented by its full function $\tau_D(q)$, only by a single parameter, the dimension D_1. Additional structural details of the directing function, which may be complicated, do not matter.

Value of the spectral exponent, after the increments of the compound process have been "rectified", in the sense of having their absolute values raised to the power $1/H$. Electrical engineers and applied physicists know (more accurately perhaps, used to know) that to understand a "noise," it is good to study it in two steps at least: first in its natural scale, then after it has been "rectified," which usually means taking the absolute value or squaring. This approach motivated the tests carried out in Voss 1992 and mentioned at the end of Section 3.5, and perhaps also the tests in Olsen 1996. In the present context, let us show that a particularly appropriate

rectification consists in "taking the power $1/H$." When $H = 1/2$, this reduces to squaring.

Indeed, take two non-overlapping intervals of duration Δt separated by a time span T, and form the "covariance" of the compounding increments $(\Delta' X)^{1/H}$ and $(\Delta'' t)^{1/H}$. We know that G' and G'' are independent Gaussian variables $\Delta' X = G'(\Delta' \theta)^H$ and $\Delta'' X = G''(\Delta'' \theta)^H$. Hence,

$$E\{(\Delta' X)^{1/H}(\Delta''(X))^{1/H}\} = E(G')^{1/H}E(G'')^{1/H}E(\Delta' \theta \Delta'' \theta).$$

The numerical prefactor $E(G')^{1/H}E(G'')^{1/H}$ depends on H, but otherwise this last expression solely reflects the properties of the directing function. The covariance and the spectral density of ΔX can be shown to be proportional, respectively, to $s^{-\tau_D(2)}$ and $f^{-1-\tau_D(2)}$. For this reason, $\tau_D(2)$ acquired the strange name of "correlation dimension."

Summary. In the WBM case $H = 1/2$, the appropriate rectification boils down to $(\Delta X)^2$. In the FBM case where $H \neq 1/2$ but H is close to $1/2$, one needs corrective factors, but reporting them here would delay us too much.

The spectrum reflects the form of dependence, but only in a limited fashion; it is distinct from, and only distantly related to, the features of $\tau_D(q)$ that affect the shape of the tails. A striking feature of the multifractals is this: scaling may, but need not, be present in the tails, but is always present in the dependence. A Brownian or fractional Brownian function of a multifractal trading time follows the same scaling rule of long-run statistical dependence as found in fractional Brownian motion.

3.12 The notions of partition function, or q-variation, for the directing multifractal time and the compound process.

Take the length of the available sample as time unit, divide it into non-overlapping intervals of lengths $\Delta_i t$, and consider the expression

$$\chi_D(q) = \sum |\Delta_i X|^q.$$

To statisticians, this is a non-normalized "sample estimate" of the moment $\Sigma|\Delta_i X|^q$. To physicists who follow a thermodynamical analogy, $\chi_D(q)$ is a "partition function." To mathematicians who follow N. Wiener, $\chi_D(q)$ is a "q-variation." Extraneous difficulties are avoided by choosing the unit of X so that $\Delta_i X < 1$ for all i.

Unequal $\Delta_i t$. For some purposes, as when we compare the q-variations taken along several alternative "times," it is important to allow the $\Delta_i t$ to be unequal. One takes the infimum of $\chi_D(q)$ for all subdivisions such that $\Delta_i t < \Delta t$, then one lets $\Delta t \to 0$. The values of q such that $\chi_D(q) \to 0$ and $\chi_D(q) \to \infty$, respectively, are separated by a critical value that will be denoted as $1/H_T$.

A very important observation concerning the contribution of the discontinuities to the value of χ. The case of direct interest is when $H_T < 1$. If so, χ_D can be divided into the contribution of the discontinuities and the rest. In the limit, the discontinuities contribute 0 if $q > 1$, therefore if $q > 1/H_T$. As a result, it makes no difference whether or not the discontinuities are included.

Equal $\Delta_i t = \Delta t$. For other purposes, however, one assumes that the $\Delta_i t$ are equal to Δt. This makes it possible to follow $\chi(q)$ as function of Δt, and one finds

$$\chi_D(q, \Delta t) = (\Delta t)^{\tau_D(q)}, \text{ with } \tau_D(q) = \log \Delta\chi(q, dt) / \log \Delta t.$$

The same argument can be carried out when the increments of trading time are replaced by the increments of the compounded process. It yields a new partition function

$$\chi_C(q, \Delta t) = \Delta t^{\tau_C(q)}.$$

3.13 A record's trail and graph have different fractal dimensions

This topic is best approached by a roundabout path.

The special case of FBM in clock time. The function $\tau(q)$ is associated with multifractals, but can also be evaluated for $B_H(t)$. Its value is found to be yielding $X_d(q, \Delta t) = (\Delta t)^{Hq-1}$, hence $\tau(q) = Hq - 1$.

We know from Section 3.4 that the trail dimension is $D_T = 1/H$ with or without compounding. Now let us sketch a standard argument **BUG** that begins with the fact that $\tau(1) = H - 1$, and concludes for the graph dimension with the value $D_G = 1 - \tau(1) = 2 - H$. This argument consists in covering the graph with square boxes of side Δt. Each Δt and the corresponding Δx contribute a stack of $|\Delta x| / \Delta t$ boxes. (Actually, one needs the smallest integer greater than the ratio $|\Delta x| / \Delta t$, but this ratio is $\sim (\Delta t)^{-1/2}$, hence is large when Δt is small.) Denote by $N(\Delta t)$ the total number of boxes in all the stacks and by D_G the box dimension. One has

$$N(\Delta t) = \sum |\Delta x| / \Delta t = (\Delta t)^{\tau(1)-1}, \text{ hence } D_G = \frac{\log N(\Delta q)}{\log(1/\Delta q)} = 1 - \tau(1) = 2 - H.$$

The general case of FBM of a multifractal trading time. The values obtained for $\tau(1)$ and D_G are specific to FBM, but the bulk of the preceding argument is of wider applicability. The total number of boxes of side Δt needed to cover the graph is $\sim \sum |\Delta x| / \Delta t = (\Delta t)^{\tau_C(1)-1}$. Taking a ratio of logarithms, this heuristic argument yields for the dimension of the graph of $X(t)$ the value

$$D_G = 1 - \tau_C(1) = 1 - \tau_D(qH).$$

From $\tau_C(1) < 0$, it follows that $D_G \geq 1$, as is the case for every curve, hence for every graph of a function.

Under multifractal compounding, there is no *functional relation between* D_T and D_G. The unifractal functions FBM are specified by a single parameter H, hence the values of D_T and D_G are necessarily functionally related. Indeed,

$$H = \frac{1}{D_T} = 2 - D_G.$$

A compound process is more complicated, since its specification includes both H and the function $\tau_D(q)$. Hence the values of D_T and D_G *cease to be* functionally related. The best one can say is that an inequality established in M 1974f{N14} implies $D_G < D_T = 1/H$; in fact, $D_G \leq 2 - H$, which we know to be the value relative to FBM.

3.14 Statistical estimation for multifractals, beginning with H, and continuing with the $\tau_C(q)$ function of the multifractal time

The preceding title includes two statistical problems. The good news is that they can be faced separately. This is so because the asymptotic behavior of $\chi(q)$ has the remarkable property of separating the properties of the compounding function $X(\theta)$ from those of the directing function $\theta(t)$.

The estimation of H. It suffices to identify the value of q for which $\tau_C(qH) = 0$. Actually, H can be defined without injecting equal Δt's and the resulting function $\tau(q)$. Indeed, $1/H$ is a "critical value" of the exponent > 1 such that $\chi(q) \to 0$ for $q > 1/H$ and $\chi(q) \to \infty$ for $q < 1/H$.

The estimation of the directing function, once H is known. It suffices to plug H into τ_C to obtain τ_D. When one only wishes to obtain D_1, one can estimate the spectral exponent B and write $D_1 = B/(2H-1)$. Unfortunately, this is the ratio of two factors that may be small simultaneously, therefore, is not very reliable.

The special case of WBM or FBM in clock time. The critical value is $1/H$. Consequently, the behavior of $\chi_C(q)$ suggests a new method of estimating H, to be added to the standard correlation or spectral analysis and the less standard R (range) or R/S methods (see M 1997H).

The case when the trading time θ is multifractal and a continuous function of the clock time t. Once again, the test of whether $\chi(q) \to 0$ or $\chi(q) \to \infty$ does not require the $\Delta_i t$ to be identical, only that they all tend to 0. When $\theta(t)$ is a continuous function, the same critical value φ is obtained by using uniform intervals of θ and uniform intervals of t. Uniform intervals of θ bring us back to the compounding FBM function $B_H(t)$, but trading time is not observable directly, and investigation of actual samples imposes uniform intervals of t.

The discordant case of $B_\sigma\{\theta(t)\}$, when the trading time θ is a discontinuous function of the clock time. This case occurs in the M 1967 representation of the M 1963 model, when compounding reduces to subordination. In some way, it is the limit of the case of continuous directing functions. However, this limit is extremely atypical, the reason being that the Δt can be made increasingly small, but not the $\Delta\theta$. The illuminating behavior of $\chi(q)$ when the $\Delta\theta$ are equal and tend to 0 is inaccessible and not reflected in the behavior of $\chi(q)$ when the Δt are equal and tend to 0.

In particular, recall that the L-stable process of exponent α is the WBM of a fractal time and is twice the exponent of the Lévy devil staircase. In this case, the correct value $H = 1/2$ is *not* revealed by the critical exponent of $\chi(q)$ evaluated with constant Δt. (Digression: this subtle point can be better understood be examining Plate 298 of M 1982F{*FGN*}.)

The converse problems. Now suppose the preceding statistical analysis is carried out on a process that is *not* a FBM of a multifractal trading. The q-variation exponent is defined for *every* function, therefore the algorithm to estimate H_C yields a value in every case.

3.15 The experimental evidence

As mentioned in the Preface, empirical testing of the M 1972 model was slow and could not be as broad and complete as I wished. But we studied the changes in the dollar/deutschmark and other foreign exchange rates

obtained from Olsen Associates in Zürich; the results, which are extremely promising, will be published in three papers by M, Fisher and Calvet in different permutations. The principal figure of Fisher, Calvet & M 1997 is reproduced here as Figure 3. It is a log-log plot of the variation of $\chi_D(q, \Delta t)$ as function of Δt, for several values of q close to 2. Two distinct datasets were matched, namely, daily and high frequency data.

The first observation is that the diagrams are remarkably straight, as postulated by multifactality.

The scaling range is very broad, three and a half decades wide, from Δt of the order of the hour to Δt of more than a hundred days (at least.)

The second observation concerns the value of q for which this graph is horizontal, meaning that $\tau_D(q) = 0$. This value of q defines the trail dimension D_T, and the data show that it is close to the Wiener Brownian value $D_T = 2$. This value was implied when Voss 1992 and Olsen 1996 described the spectrum of the rate changes as being white.

At closer look, however, D_T seems a bit smaller than 2, suggesting $H_T > 1/2$. If confirmed, this inequality would be a token of persistent fractional Borwnian motion in multifractal time.

Increments Δt below one hour seem to exhibit a different scaling, with D_T clearly different from 2. Once again, full detail is to be found in Fisher, Calvet & M 1997.

3.16 Possible directions for future work

A major limitation of the fractional Brownian motion of time was acknowledged in Section 9 of Chapter E1: the resulting marginal distributions are symmetric. A possible way out was also referenced, namely, the "fractal sums of pulses."

This section's context instantly suggests an alternative way out: to replace $B_H(t)$ by an asymetric form of the L-stable process that underlies the M 1963 model. The presence of two parameters (an exponent α and a skewness parameter β) can only help improve the fit of the data. But the resulting process remains unexplored and may prove to be unmanageable.

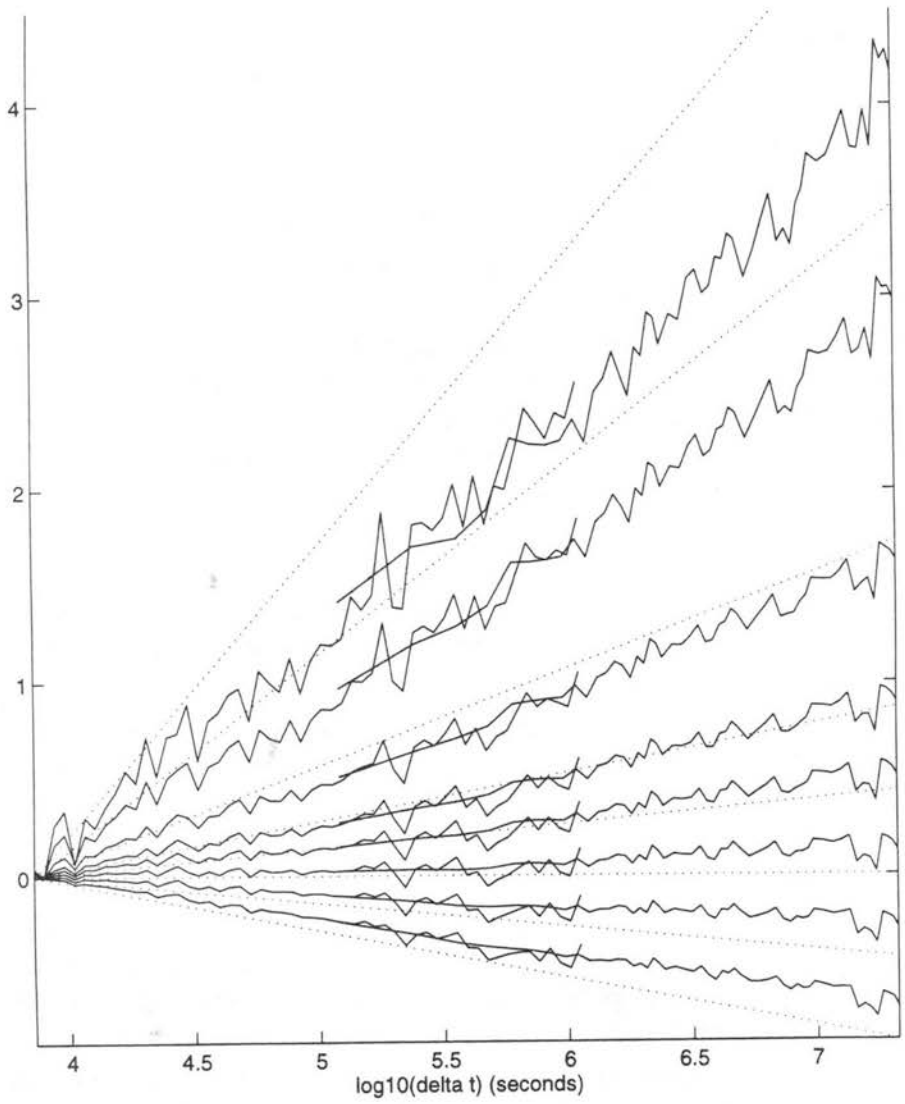

FIGURE E6-3. Doubly logarithmic plot of $\chi_D(q, \Delta t)$, as function of Δt in the case of the Olsen data for the US dollar/Deutschmark exchange rate. The main observations are a) the fact that the plots are straight from Δt of the order of one hour to the end of the data, which corresponds to Δt of more than a hundred days; the slopes of the plots define the function $\tau_D(q)$; b) the fact that the value of $q = D_T$ for which $\tau_D(q) = 0$ is close to 2.

Observation a) is a symptom of multifractality and observation b) is a symptom that the process is close to being a Wiener Brownian motion that is followed in multifractal time. The true value of D_T is a bit smaller than 2, suggesting the inequality $H_T > 1/2$. If confirmed, this would be a token of persistent fractal Brownian motion in multifractal time.

4. DIAGONAL-AXIAL SELF-AFFINE CARTOON SURROGATES

This section covers roughly the same material as Section 3, but in entirely different style. Section 3 concerned constructions that first arose in limit theorems of probability theory (some of them classical and/or difficult). Those constructions are of great independent interest and of "universal" rather than arbitrary character. But their origins are not coordinated, so that, they did not fit comfortably together. To the contrary, this Section proceeds in tightly coordinated fashion and the parts fit well together. The new feature is that those parts are non-universal and to a large extent arbitrary. Their baroque wealth of structure is a loss from the viewpoint of simplicity and esthetics; but it may be a gain from the viewpoint of apprehending the baroque wealth of structure found in nature.

Let me elaborate. By abundantly illustrating self-similarity, M 1982F{*FGN*}, demonstrated that the principle of recursive construction exemplified in Section 2 is very versatile. That is, it is not sharply restrictive but leaves room for many varied implementations. To an even larger extent, self-affinity is versatile almost to excess, hence insufficient by itself for any concrete purpose in science. The goal of this section is to transform the 1900, M 1963, M 1967 and M 1972 models of price variation into constructions that fit together as special examples in a broader, well-organized but diverse collection. The implementation of this goal is distantly inspired by a construction due to Bernard Bolzano (1781-1848). In a terminology that may be familiar to some readers, this implementation is "multiplicative." The more familiar "additive" constructions (patterned on the non-differentiable functions due to K. Wierstrass) proved to be of insufficient versatility.

4.1 Grid-bound versus grid-free, and fractal versus random constructions

The role of grids in providing simplified surrogates. Fractal construction are simplest when they proceed within a grid. Grids are not part of either physics or economics. But suitable grid-based constructs can act as "surrogates" to the grid-free random process, like the 1900, M 1963, M 1965, M 1967, and M 1972 models. When this is possible, the study is easier when carried out on the grid-based cartoons. Besides, the cartoons in this chapter fit as special cases of an overall "master structure" which relates them to one another, is enlightening and is "creative" in that it suggests a stream of additional variants. I came to rely increasingly on this master structure in the search for additional models to be tried out for new or old problems. Striking parallelisms were mysterious when first

observed in the grid-free context, but became natural and obvious in this master structure.

Fractality versus randomness from the viewpoint of variability. The main cartoon constructions in this chapter are non-random. To the contrary, the 1900, M 1963, M 1965, M 1967 and M 1972 models in Section 3 are random, for reasons explained in Section 1 of Chapter E1.

However, an important lesson emerges from near every study of random fractals. At the stage when intuition is being trained, and even beyond that stage, *being random is for many purposes less significant than being fractal.* It helps if the rules of construction are not overly conspicuous, for example, if no two intervals in the generator are of equal length. That is, the non-random counterparts of random fractals exhibit analogous features, and also have the following useful virtue: they avoid, postpone, or otherwise mitigate some of the notorious difficulties inherent to randomness. Those non-random fractals for which acceptable randomizations are absent or limited, are also of high educational value.

Distinction between the contributions of Wiener and Khinchin. In the spirit of the preceding remarks, the readers acquainted with the Wiener-Khinchin theory of covariance and spectrum may recall that a single theory arose simultaneously from two sources: Wiener studied non-random but harmonizable functions, and Khinchin studied second-order stationary random functions. The two approaches yield identical formulas.

A drawback: grid-bound constructions tend to "look "creased" or "artificial." This drawback decreases at small cost in added complication, when the grid is preserved, but the construction is randomized to the limited extent of choosing the generator among several variants. This will be done in Figures 4 and 5. Nevertheless, the underlying grid is never totally erased. To a trained eye, it leaves continuing traces even after several stages of recursion. (A particularly visible "creasing" effect is present in computer-generalized fractal landscapes, when the algorithm is grid-based. Around 1984, this issue was a serious one in the back-offices of Hollywood involved in computer graphics.)

Addition versus multiplication. Specialists know that fractals are usually introduced through additive operations, and multifractals, through multiplicative operations. The reason for selecting the fractal examples that follow is that they can be viewed as either additive or multiplicative, making it unnecessary to change gears in the middle of the section.

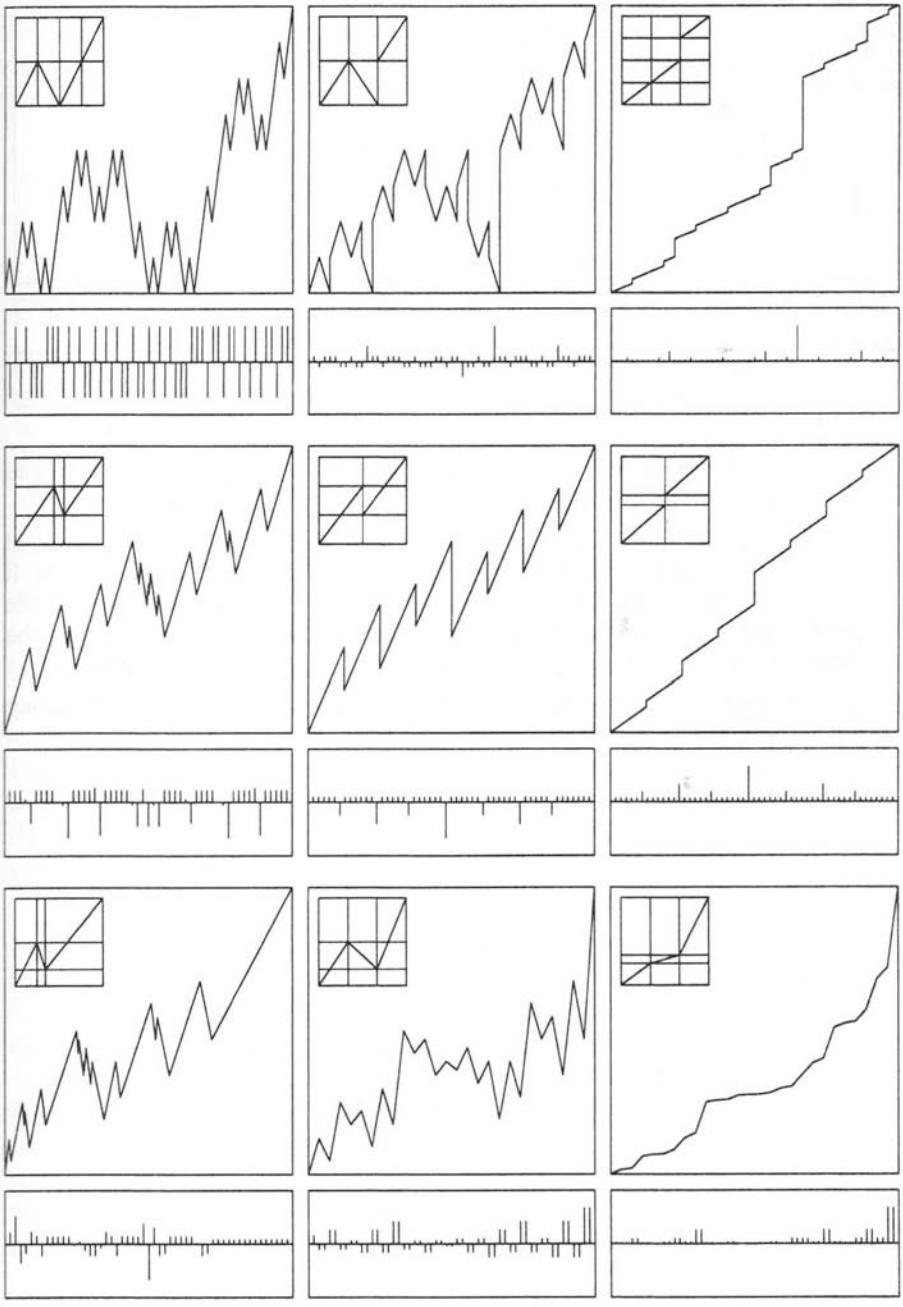

FIGURE E6-4. Six alternative cartoon constructions explained on the next page.

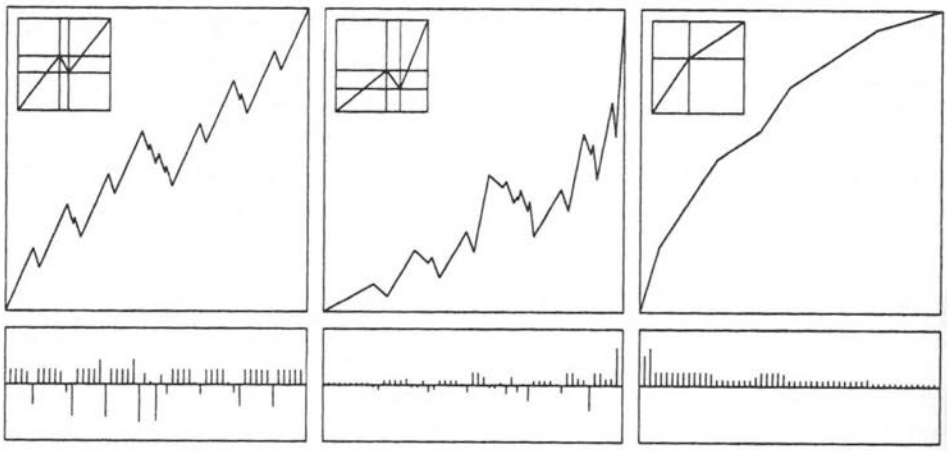

FIGURE E6-5. Three additional cartoons. The following explanation applies to this figure and the preceding one.

Each construction is illustrated by a square diagram and a longitudinal one. The generator is shown in a small window within the square diagram: it is either diagonal or diagonal-and-axial. The square diagram shows the level-2 approximation and the corresponding longitudinal diagram shows the increments of the level-2 approximation, taken over equal time increments.

The diagrams juxtaposed on a horizontal row in Figure 4 are intimately related, as described in Section 4.9.

The power and limitations of the eye. The eye discriminates better between records of *changes* of a function than between records of the function *itself*. Therefore, the results of many constructions that follow will be illustrated in both ways.

Recursiveness also allows many other possibilities that lie beyond the scope of this book: they include the "fractal sums of pulses" (M 1995n and several papers for which the co-author listed first is Cioczek-Georges).

4.2 Description of the grid-bound recursive constructions with prescribed initiator and generator

To start from scratch, each diagram in Figures 4 and 5 is drawn within a *level-0 box*, whose sides are parallel to the coordinate axes of t and x. The "initiator," is an ordered interval (an "arrow") acting as a "hidden string" that crosses the level-0 box from bottom left to top right. Therefore, it is useful to think of this and other boxes as "beads." The width and height of the level-0 box are chosen as units of t and x, making the box a square. (The question of what is meant by a *square* in the affine plane is a subtle issue, to be tackled below, after the definition of H.)

In addition, each diagram contains a *string generator* that joins the bottom left of the initiator to its top right. Alternative descriptions for it are "string of arrows," "broken line," and "continuous piecewise linear curve." The number of intervals in the generator, b, is called *"generator base"*. When the generator is increasing, $b \geq 2$; when the generator is oscillating, the lower bound becomes $b \geq 3$. The larger b becomes, the greater the arbitrariness of the construction. Hence, the illustrations in the chapters use the smallest acceptable values of b.

Axial and diagonal generator intervals. To insure that the recursive construction generates the graph of a function of time, the string generator must be the "filled-in graph" of a function $x = G(t)$, to be called *generator function*. To each t, the ordinary graph attaches a single value of x. To each t where $G(t)$ is discontinuous, the filled-in graph attaches a vertical oriented interval of values of x that spans the discontinuity. The resulting interval in the generator is called *axial*. (The general case, mentioned later but not used in this book, also allows for horizontal intervals.) A non-axial interval is called *diagonal*, and the rectangle that it crosses diagonally from left to right defines a *level-1 box*. In some cases the level-1 boxes can be superposed by translation or symmetry, in other cases they cannot.

Recursive construction of a self-affine curve joining bottom left to top right, using successive refinements within a prescribed self-affine grid. Step 0 is to

draw the diagonal of the initiator. Step 1 is to replace the diagonal of the initiator by the filled-in graph of the generator function $G(t)$. Step 2 is to create a line broken twice, as follows. Each diagonal interval within the generator is broken by being replaced by a downsized form of the whole generator. To "downsize" means to reduce linearly in the horizontal and vertical directions. In the self-affine case, the two ratios are distinct. In some cases, one must also apply symmetries with respect to a coordinate axis. As to the generator's axial intervals, they are left alone. One may also say that they are downsized in the sense that the ratio of linear reduction in one direction is 0, collapsing the generator into an interval.

The "prefractal" approximations of self-affine graphs can take one of two forms. They may consist of increasingly broken lines. Figures 6 and 7 take up important generators and draw corresponding approximations as boundaries between two domains, white and black. This graphic device brings out finer detail, and helps the eye learn to discriminate between the various possibilities. Alternatively, each diagonal interval may be replaced by a rectangular axial box of which it is the diagonal. If so, the prefractal approximation consists in nested "necklaces" made of increasingly fine boxes, linked by axial pieces of string.

As the recursive construction of an oscillating cartoon proceeds, its increments Δu over increasingly small intervals Δt tend to become symmetrically distributed. That is, the ratio of the numbers of positive and negative increments tends to 1. {Proof: After k stages, each increment is the product of k factors, each of the form sign $(\Delta_i x)$. But Π sign $(\Delta_i x)$ is > 0 if Σ sign $(\Delta_i x)$ is even, and is < 0 if Σ is odd. The distribution of Σ sign $(\Delta_i x)$ is binominal and smoothly varying, therefore even and odd values are equally frequent asymptotically.}

4.3 The H exponents of the boxes of the generator

This and the next sections show how the fundamental scaling exponent H of fractional Brownian motion splits into a number of significantly different aspects.

Diagonal boxes and their finite and positive H exponents. Given a diagonal box β_i of sides $\Delta_i t$ and $\Delta_i x$, an essential characteristic is

$$H_i = \frac{\log \Delta_i x}{\log \Delta_i t} = \frac{\log \text{ of the absolute height of the box } \beta_i}{\log \text{ of the width of the box } \beta_i}.$$

In other words,

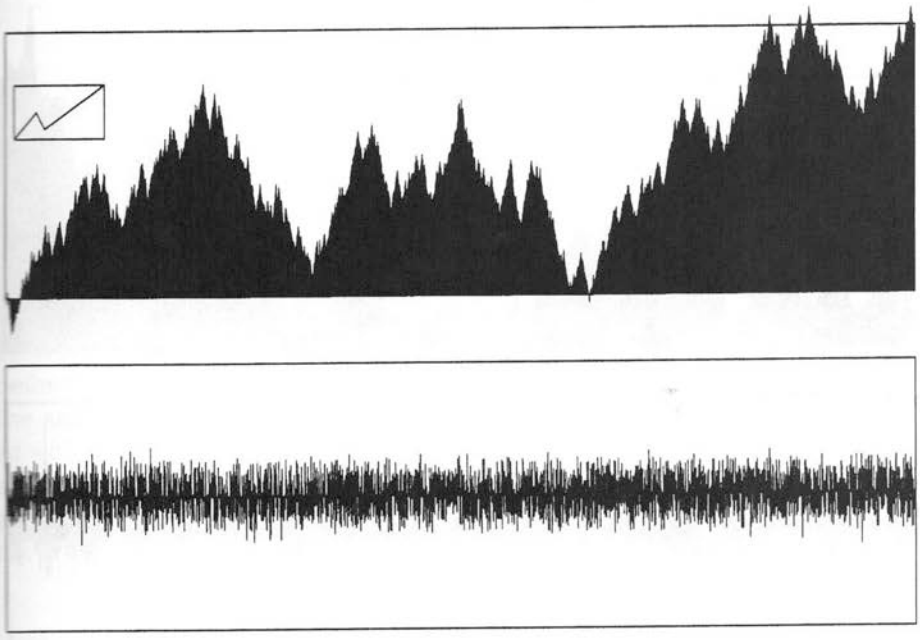

FIGURE E6-6. The top line illustrates a cartoon of Wiener Brownian motion carried to many recursion steps. The generator, shown in a small window, is identical to the generator A2 of Figure 2. At each step, the three intervals of the generator are shuffled at random; it follows that, after a few stages, no trace of a grid remains visible to the naked eye.

The second line shows the corresponding increments over successive small intervals of time. This is for all practical purposes a diagram of Gaussian "white noise" as shown in Figure 3 of Chapter E1.

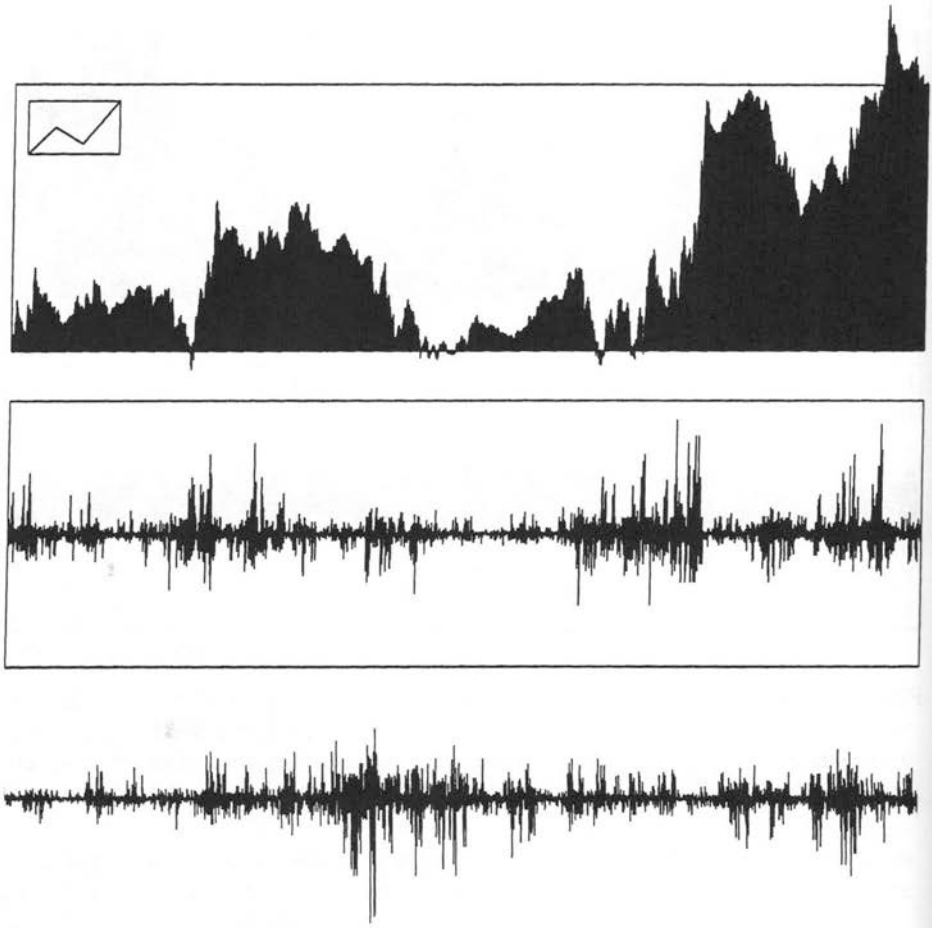

FIGURE E6-7. This figure reveals – at long last – the construction of Figure 2 of Chapter E1. The top line illustrates a cartoon of Wiener Brownian motion followed in a multifractal trading time. Starting with the three-box generator used in Figure 6, the box heights are preserved, so that D_T is left unchanged at $D_T = 2$ (a signature of Brownian motion), but the box widths are modified. (Unfortunately, the seed is not the same as in Figure 6.)

The middle line shows the corresponding increments. Very surprisingly, this sequence is a "white noise," but it is extremely far from being Gaussian. In fact, serial dependence is conspicuously high. The bottom line repeats the middle one, but with a different "pseudo-random" seed. The goal is to demonstrate once again the very high level of sample variability that is characteristic of wildly varying functions.

The resemblence to actual records exemplified by Figure 1 of Chapter E1 can be improved by "fine-tuning" the generator.

$$\Delta_i x = (\Delta_i t)^{H_i}.$$

This identity concerns non-infinitesimal boxes, therefore H_i is a *coarse* coefficient. Now proceed to the limit $\Delta t \to 0$. If $x(t)$ had a well-defined derivative x', one would have $\Delta x \sim x' \Delta t$. Therefore, differentiable functions yield $\alpha = 1$ (but the converse is not true).

Hurst, Hölder, and a way to conciliate mathematical and concrete needs. Section 3.2 mentions that H has roots in the works of both the hydrologist H. E. Hurst (1880-1978), and the mathematician L. O. Hölder (1859-1937). However, these concrete and mathematical contexts require a special effort before they fit comfortably together. For example, assume that all boxes of the same level are equal, with $\Delta t = b^{-1}$ and $\Delta x = b'^{-1}$ for level-1, therefore $\Delta t = b^{-k}$ and $\Delta x = b'^{-k}$ for level-k. It follows that $H_i = \log b' / \log b = H$ for all boxes at all levels; level 0 yields $\log 1/\log 1 = 0/0$, which can be interpreted as equal to H. However, if the level-0 box had sides 1 and B, all the level-k boxes would yield

$$H = \frac{\log b' + \log B/k}{\log b}.$$

In the pure mathematical interpretation due to Hölder, H is a local concept that concerns the limit $k \to \infty$. Its value is not affected by B. By contrast, the concrete interpretation of H that I pioneered do not concern local asymptotics but concrete facts, therefore applies uniformly to all sizes. If the resulting "coarse" H is to serve a purpose, it must be independent of all units of length; this is achieved by setting $B = 1$.

Axial intervals and the values H=0 and H = ∞. One may say that horizontal intervals yield $H = \infty$, and discontinuities yield $H = 0$. The value $H = \infty$ can occur almost everywhere, and the value $H = 0$ can occur at most on a denumerable set, therefore, on a set of dimension 0.

Comments on the examples on Figure 4. The columns are denoted by letters (A, B, and C) from the left, and the rows by numbers (1, 2, 3, 4) from the top. Each example will first be listed by itself, then different examples will be shown to be related to one another. Figure 5 carries further the constructions based on generations A3 and B3, adding randomization for increased realism.

4.4 Unifractality for cartoons, and selected major examples

This section proceeds beyond general statements on grid-bound cartoons. It introduces (by definitions and examples) some key distinctions that explain how those cartoons can serve as "surrogates" of the grid-free self-affine processes described in Section 3. When a process is denoted by XYZ, its cartoon surrogate will be denoted by C(XYZ).

Unifractality. Each box β_i of the generator has its own exponent H_i. If all the H_i are identical, > 0 and $< \infty$, the cartoon construction will be called *unifractal*. The conditions $H > 0$ and $H < \infty$ exclude axial generator intervals. This case is very special, but of fundamental importance, because it includes cartoons of WBM and FBM.

Unibox versus multibox constructions. In the unifractal case, the generator boxes can be either identical, defining the *unibox* case or not, defining the *multibox* case. All unibox constructions are unifractal. Many of their properties depend only on H, but other properties depend on the boxes' two sides, and some properties also depend on the details of the arrangement of the boxes. Multibox constructions depend on a larger number of parameters; they are less regular, hence less "artificial-looking," therefore their fractality is a better surrogate for randomness.

• **C_1(WBM).** (Generator A1). This cartoon of base $b = 4$ is a unibox (hence unifractal) surrogate for Wiener Brownian Motion. It has become widely used in physics (M 1986l, M 1986t, Family & Vicsek 1991). To find where it comes from, consider the Peano-Cesaro motion illustrated (in approximation) on Plate 65 of M 1982F{*FGN*}. Follow this motion as it proceeds from the lower left to the upper right corner, moving through the upper left corner. The projection of this motion on the x-axis will have A1 as its generator. Variants are described in M 1986t{H}, and in various articles collected in M 1997N and M 1997H.

• **C_2(WBM).** (Generator A2). The generator of this multibox cartoon contains $b = 3$ intervals, which is the smallest value that allows oscillations. Denoting the side intervals by x, the middle interval is of height $2x - 1$. Since $H = 1/2$ for WBM, the generating identity becomes $2x^2 + (2x - 1)^2 = 1$, yielding $x = 2/3$.

• **C_3(WBM).** (Generator A3). The novelty is that all three intervals were made unequal, to add realism to the construction. A form of it is carried over many stages in Figure 6.

• **C(FBM).** (Generator A4). Cartoon C_2 (WBM) is readily generalized to $H \neq 1/2$. It suffices to take for x the positive root x_0 of the equation $2x/H + (2x - 1)/H = 1$. For $H > 1/2$, $x_0 < 2/3$; for $H < 1/2$, $x_0 > 2/3$.

Two steps beyond unifractality . The usual contrast to unifractality is provided immediately by *multifractality*, but the present context makes it necessary to single out an intermediate case that did not, until now, warrant a special name. Since "in between" is denoted by the Greek root *meso*, this case will be called *mesofractal*, a word that is used here for the first time.

4.5 Mesofractality for cartoons, and selected major examples

Mesofractality. This term will denote cartoons whose generator includes vertical intervals with $H = 0$, in addition to diagonal intervals sharing a unique H satisfying $0 < H < \infty$. (The definition extends painlessly to allow horizontal intervals with $H = \infty$; such cases are not needed in this book, but will be discussed in detail in the introductory material of M 1997N.)

Like unifractality, mesofractality is a special case, but it too is of fundamental importance because it characterizes the several important cartoons that follow. The first and the second can be described as surrogates of LSM, even though the first only includes negative jumps. The third is a surrogate of fractal trading time, a notion to be defined in Section 4.5.

- **C_1(LSM).** (Generator B2). Begin with C_2(WBM) and modify the generator's boxes by the following transformation: keep the heights constant, expand the first and third box to be of width 1/2, and reduce the second box to be of width 0, hence $H_2 = 0$. All the jumps are negative.

- **C_2(LSM).** (Generator B1). A more realistic surrogate of LSM must have both positive and negative jumps. To achieve this goal, it is necessary to use a generator containing at least $b = 4$ intervals. Begin with C_1(WBM), and modify the generator's boxes by the following transformation: keep the heights constant, expand the first, second and fourth box to be of width 1/3, and reduce the third box to be of width 0. The new H values are $H_1 = H_2 = \log_3 2$, $H_3 = 0$, and $H_4 = \log_3 2$.

- **C_1(FTT) and C_2(FTT).** (Generators C1 and C2). These are inverse functions of variants of the classical devil staircase.

4.6 Multifractality for cartoons, and selected major examples

Multifractality. The most general category of cartoons allows the generator to include diagonal boxes with different values of H_i, ranging from $H_{min} > 0$ to a maximum satisfying $0 < H_{max} < \infty$. Those cartoons are necessarily multibox. Boxes created at the k-th stage of recursion are characterized by H distributed over the interval $[H_{min}, H_{max}]$ in increasingly tight

fashion. The limit of this distribution is described by the multifracatal formalism $f(\alpha)$ mentioned at the end of Section 3.9.

As defined above, multifractality allows some generator intervals to be axial, hence includes cartoons that combine continuous variation with jumps. However, jumps are absent, both from the bulk of the abundant literature on multifractals, and from the single grid-free multifractals of Section 3, like those in M 1972j{N14}, M 1974f{N15} and M 1974c{N16}. They will be discussed in M 1997N and M 1997H, but not in this book.

• **C_1(MFM).** (Generator B3). This example of oscillatory multifractal motion is a much simplified version of a construction due to Bernard Bolzano (1781-1848). Begin with C_3(WBM thru FBM), and modify the generator's boxes by the following transformations: keep the heights constant and change the width to 1/3. (These linear transformations are invertible, therefore called *affinities*. The linear transformations used to define C_1(LSM) and C_2(LSM) cannot be inverted.)

• **C_2(MFM).** (Generator B4). A three interval oscillating generator was chosen haphazardly. A form of this case is carried over to many stages in Figure 7.

• **C_1(MTT).** (Generator C4). A point was chosen in the unit square, and joined to the lower left and upper right corners by a non-decreasing broken line.

• **C_2(MTT).** (Generator C3). This is a three-interval generator yielding a multifractal trading time.

4.7 "First step towards a compound cartoon" representation of a general cartoon: definition of the trail dimension D_T and the trail exponent $H_T = 1/D_T$; spectral density of the form f^{-B} where $B = 1 - 2/D_T$

The major examples in Sections 4.4 to 4.6 include cartoons of WBM, LSM and FBM, which concern M 1900, M 1963 and M 1965 models, and other grid-bound self-affine functions that combine long tails and long memory. The generating functions are of great diversity, and innumerable additional examples immediately come to mind. Their very multiplicity might have been a source of disorder and confusion.

Fortunately, it is not, thanks to a very strong result that will be established in Section 4: every oscillating cartoon construction can be rephrased as a compound function, namely in the form of C(WBM) or C(FBM) as followed in suitable "multifractal" trading time that is a monotone non-decreasing function of clock time.

Recall that the grid-free constructions sketched in Section 3 consist in unifractal Wiener or fractional Brownian motions in a trading time that is either linear, or fractal, or multifractal. That collection of models grew step-by-step from 1963 to 1972, and lacked intrinsic cohesion or legitimacy, except as a rather abstract companion of the functions $\tau_C(q)$ and $\tau_D(q)$.

The grid-based cartoons are much simpler in that respect. In their case, most remarkably, the notion of trading time is intrinsic, compelling and inevitable. We proceed by steps, and begin by defining D_T and $H_T = 1/D_T$.

A simple identity that characterizes unifractal generators. In all cartoons, the box widths satisfy $\sum \Delta_i t = 1$. In the unifractal case where $H_i = H$ for all i, *define* D_T as $1/H$. It follows that the box heights satisfy

$$\sum |\Delta_i x|^{D_T} = 1, \text{ with } D_T = 1/H.$$

A simple identity that characterizes monotone multifractal generators. When $\Delta_i x > 0$ for all i, the equality $\sum \Delta_i t = 1$ trivially implies

$$\sum |\Delta_i x|^{D_T} = 1, \text{ with } D_T = 1.$$

The dimension-generating equation of a cartoon construction. This term will denote the equation $\sum |\Delta_i x|^{\sigma} = 1$, the unknown being σ. We know two cases already: in the unifractal case the only positive root is $\sigma = D_T = 1/H$, and in the monotone case the only positive root is $\sigma = 1$. We now proceed to the remaining possibility.

A new and highly significant concept: generalized values of D_T and $H_T = 1/D_T$, as defined for oscillating multifractal cartoons. In the oscillating multifractal case, the quantities H_i cease to be identical. The generating equation ceases to be a restatement of a mildly relevant identity. However, it remains meaningful and becomes highly significant. Its only positive root D_T satisfies $D_T > 1$ and is a *fundamentally* important characteristic of the construction.

Geometric interpretation of D_T by embedding, as the trail dimension of a closely related vectorial process. As we know from Section 2, a set is a self-similar fractal, when the whole is made of parts that are obtained from the whole by reductions. The generating equation for σ is formally identical to the classical Moran equation that gives the dimension of such a set,

where the reduction ratios, $r_i = |\Delta_i x|$, are not equal, therefore, the simplest formula $D = \log N / \log(1/r)$, is inapplicable.

The procedure that gives substance to this analogy is embedding. It was already used when Sections 3.3 and 3.4 interpreted a scalar FBM $X(t)$ as a projection of a vectorial FBM of at least $1/H$ coordinates. The argument had to refer to a known theorem, that $1/H$ is the dimension of that vectorial FBM. In the present context, embedding is even simpler and requires no delicate reference.

The argument is simplest when $H = 1/2$ and $b = \max i \geq 4$. Consider, in a b-dimensional space, the point P of coordinates $\sqrt{b}\,\Delta_i x$. The squared distance from O to P is $\sum |\Delta_i x|^2 = 1$. Now consider projections on the main diagonal of our b-dimension space. The vector OP projects on an interval of length $1/\sqrt{b}$, and the vector of length $\Delta_i x$ along the i-th coordinate axis projects on an interval of length $\Delta_i x/\sqrt{b}$.

We are now ready to construct a self-similar curve in b-dimensional space, by taking OP as the initiator and the sequence of coordinate vectors of length $\Delta_i x$ as the generator. A classical theorem due to Moran tells us that the fractal dimension of that curve is the root D_T of the dimension-generating equation. This interprets D_T as a fractal dimension, with no reference to the values of the $\Delta_i t$. (The reason for postulating $b \geq 4$ is to some extent esthetic: in a space of $b > 3$ dimensions, one can obtain a spatial curve without double points.)

The preceding construction relies on the Pythagoras theorem, which is why the case $D_T = 2$ is the simplest possible, but the same result can be obtained for all $D_T > 1$.

Spectral density of the embedding vectorial motion. It can be shown to be of the " $1/f$ " form f^{-B}, with $B = 1 + 2H_T$ for the motion itself and $B = -1 + 2H_T$ for the "derivative" of the motion, which is a white noise when $H_T = 1/2$.

4.8 The graph dimension D_G of a cartoon; it is not functionally related to the trail dimension D_T

Having generalized D_T beyond the value $1/H$ relative to FBM, the next step is to generalize D_G beyond the corresponding value $2 - H$.

The special case where $\Delta_i t = 1/b$ for all i. The derivation of D_G for FBM was sketched in Section 3.13. The idea is to cover the graph with stacks of square boxes of side Δt. When $\Delta_i t = 1/b$ for all i , take square boxes of side b^{-k}. One defines $\tau_C(1)$ by writing

$$\chi(1, b^{-k}) = \sum |\Delta X| = (\sum b^{-H_i})^k = (b^{-k})^{\tau_C(1)}.$$

One can define H_G by writing

$$\chi(1, b^{-k}) = (b^{-k})^{1-H_G}, \text{ that is }, b^{-H_G} = \frac{1}{b} \sum b^{-H_i}.$$

For oscillating functions, $b \geq 3$, and the two exponents H_G and H_T are distinct functions of $b - 1 \geq 2$ independent parameters H_i. Therefore, they are *not* linked by a functional relation.

The general case. Multifractal formalism. Section 3.9 defines the functions $\tau_C(q)$ and $\tau_D(q)$. The same definitions apply to the cartoon constructions. The details will be described in M 1997N.

4.9 Constructions of an intrinsic "compound cartoon" representation of a general cartoon; trading time is fractal for mesofractal cartoons and multifractal for multifractal cartoons

Before we complete the task of demonstrating that the cartoon constructions are surrogates of the compound functions examined in Section 3, one last step is needed. Examine the three sets of cartoons that are shown in Rows 1, 2 and 3 of Figure 2, and mark the coordinates as follows: θ and X in Column A, t and X in Column B and t and θ in Column C. Seen in this light, each cartoon in Column B is reinterpreted as a compound cartoon involving its neighbors in the same row. It is obtained from its neighbor in Column A, by replacing the clock time by the fractal or multifractal time defined by its neighbor in Column C. Let us now show here that such a representation can be achieved for every cartoon.

The intrinsic duration of an interval in the generator. We start with a recursive construction of fractal dimension as defined early in this section. To make it over into a cartoon of FBM with the exponent $H_T = 1/D_T$. we must apply the inverse of the linear transformation that led from C_1(WBM) to C_1(LSM) and from C(FBM) to C(MFM). Starting from an arbitrary generator box, the recipe is in two steps:

- keep the height $\Delta_i x$ constant,
- by definition of $\Delta_i \theta$, change the width from $\Delta_i t$ to $|\Delta_i x|^{D_T} = \Delta_i \theta$.

Intrinsic definition of a cartoon's trading time. We are now ready to take a last and basic step. We shall show that an oscillating, but otherwise

arbitrary cartoon can be represented as a unifractal oscillating cartoon of exponent $H_T = 1/D_T$ of a multifractal (possibly fractal) trading time. In Section 3, the notion of trading time entered as a model of our historical and intuitive knowledge of competitive markets. Now, it enters through an inevitable mathematical representation.

The idea is to construct trading time using a cartoon generator defined by the quantities $\Delta_i t$ and $\Delta_i \theta$. Each recursion stage ends with an "approximate trading time that becomes increasingly "wrinkled" as the interpolation proceeds. The limit trading time may, but need not, involve discontinuity, but in all cases, its variation becomes increasingly concentrated in increasingly at each stage of interpolation, and in the limit manifests a high degree of concentration in very short periods of time. How this process builds up is a very delicate topic that cannot be discussed here in detail, but constitutes a core topic of M 1997N.

4.10 The experimental evidence

The visual resemblence between Figure 7 and Figure 1 of Chapter E1 deserves to be viewed as impressive, because of the extraordinary (in fact, seemingly "silly") simplicity of the underlying algorithm. However, multifractal analysis, using $\tau_c(q)$, as in Figure 1, or using the equivalent technique of $f(\alpha)$, shows that the resemblence is not complete. Indeed, the simulated data of Figure 5 yield slopes $\tau_C(q)$ that disagree with Figure 1. This is as it sould be: indeed, in order to simplify the construction to the maximum, Figure 5 uses a single generator, except that the three intervals are randomly shuffled at each iteration. A closer agreement requires the fully random algorithm of M 1972J{N14}, or at least a "canonical" algorithm of in the sence M 1974f{N15}, with lognormal weights. These topics are delicate and must be postponed to M 1997N.

4.11 Why should price variation be multifractal, and would multifractality have significant consequences?

Possibly explanatory power of multiplicative effects. This book is eager to study the consequences of scaling, but reluctant to look for its roots; in particular, Chapter E8 expresses doubts about explanations that involve "proportional effects".

Nevertheless, such an argument underlies multifractals, and is worth sketching.

The structure is especially clear in a generating method that is an alternative to the cartoons described in this section. It concerns the vari-

ance of price movements and proceeds as follows. The originator is a uniform intensity, and it is pertubed by pulses independent of one another and random in every way. Metaphorically, the variability of the variance is attributed to an infinity of "causes," the effects of one cause being described by one pulse. The main feature is that the distribution of pulse lengths must be scaling; this means that the effects are very short-lived for most causes and very long-lived for some.

There is little doubt that, ex-post, such "pulses" could be read into the data, but this is not the proper place to discuss the pulses' reality.

Extrapolation of multifractals and an ominous and possibly inconceivable implication. This brief paragraph simply draws attention to Section 5.4.

4.12 Possible directions for future work

A sketch of directions cannot be comprehensive and comprehensible, without detailed acquaintance will further developments of the theory which are necessarily postponed to M 1997N,H.

5. THE DISTINCTION BETWEEN MILD AND WILD VARIABILITY EXTENDS FROM RANDOM VARIABLES TO RANDOM OR NON-RANDOM SELF-AFFINE FUNCTIONS

As applied to discrete-time sequences of independent random variables, the notions of "mild" and "wild" were discussed in Chapter E5. This section moves on, to establish the same distinction in two additional classes of functions: the continuous time grid-free *random* self-affine functions discussed in Section 3 and the grid-bound *non-random* self-affine cartoon functions discussed in Section 4. Both classes involve numerous "either-or" criteria that sort out diverse possibilities: continuous or not, unifractal or not, and, in the cartoon case, unibox or not. But none of these "either-or" criteria is more important than the distinction between mild and wild.

The fact that one can extrapolate those notions to a non-random context is a special case of the general and important fact, already mentioned in Section 4.1, that fractality is often an excellent surrogate for randomness. The fact that this section is restricted to self-affine functions brings a major simplification, as compared to Chapter E5: there will be no counterpart as slow randomness.

Here is a summary of this section's conclusions. True WBM and its cartoons C(WBM) exhibit *mild variability*. The remaining processes described in Section 3 and other cartoons described in Section 4 contradict mildness in diverse ways, alone or in combination. Those contradictions exemplify different possible forms of *wild variability*. Thus, "Noah wild" recursive functions are cartoons of discontinuous wildly random processes whose jumps are scaling with $\alpha < 2$. "Joseph wild" recursive functions are cartoons of (continuous) Gaussian processes called fractional Brownian motions. The "sporadic wild" recursive functions are cartoons of wildly random processes that I called sporadic because they are constant almost everywhere and supported by Lévy dusts (random versions of the Cantor sets.)

5.1 The notions of mild and wild in the case of random functions

The basic limit theorems. Given that self-affinity forbids slow randomness, it suffices to show that the limit theorems that define mildness remain meaningful beyond random sequences of independent identically distributed variables. For many purposes – including the present one – those theorems are best split into three parts, each concerned with the existence of a renormalizing sequence $A(T)$, such that $\sum_{t=1}^{T} X(t)/A(T) - B(t)$ has a non-degenerate limit as $T \rightarrow \infty$.

LLN. When such a limit exists for $B = 0$ and $A(T) = T$, X satisfies the *law of large numbers.*

GLT. When there exists two functions, $A(T)$ and $B(T)$, such that $X\sum(T)A(T) - B(T)$ converges to the Gaussian, X satisfies the *central limit theorem with Gaussian limit.*

FLD. When $A(T) = T^{1/2}$, $X(t)$ satisfies the *Fickian law of diffusion,* which says that diffusion is proportional to $\sqrt{T}$.

Those theorems hold for independent Gaussian random variables such as the Gaussian, which deserved in Chapter E5 to be called *mildly* random. But all three theorems fail for independent Cauchy variables. And wild variables are those for which one or more of those three theorems fail.

5.2 Extrapolation of recursive cartoon constructions as an "echo" of interpolation

To show that LLN, GLT and FLD have exact counterparts for the extrapolated non-random "cartoons" in Figure 1, and that those counterparts statements may be true or false, we must first extrapolate the recursive construction of our cartoons. The conclusion will be that mild variability

is found only in the cartoon of Wiener Brownian motion. Every other cartoon exemplifies a form of wild behavior, with one added possibility to be mentioned momentarily.

A complete self-affine fractal shape is not only infinitely detailed, but also infinitely large. This infinitely fine detail is absent in the case of sequences of random variables in discrete time and has no significance in either physics or finance. The sole reason why Section 4 did not use extrapolation is because interpolation is far easier to describe, study, and graph. The k-th level of extrapolation can also be called level-$(-k)$ of the construction.

First examine the case where each level-1 box is obtained from the original level-0 box by reduction of ratio b_t horizontally and b_x vertically. A straightforward procedure achieves both k levels of interpolation and k levels of extrapolation: it suffices to start with the prefractal interpolate pushed to the $2k$-th level of interpolation and dilate it in the ratios b_t^k horizontally and b_x^k vertically. This dilation transforms each level-$(-k)$ box into a unit square.

While interpolation is a uniquely specified procedure, extrapolation is not. In order to specify it, it is necessary to first select the fixed point of the dilation. The simplest is the origin 0 of the axes of t and x. However, this is not the only possible choice: the fixed point can be the bottom left or upper right corner of any box in the generator (or the limit of a sequence of such points). The most natural procedure is to select the fixed point at random; in this sense, *all* extrapolated self-affine shapes are intrinsically random. (Interpolated sets do not become random until all selects an origin, but this is an optional step that one may not need to face.)

Observe, however, that the fixed point *cannot* be located on a vertical interval of the generator. If extrapolation is attempted around such a point, the interval it contains will lengthen without bound, into an infinitely large discontinuity characterized by $H = 0$. (It is also impossible to select the fixed point on horizontal interval of the generator. The extrapolation will lengthen this interval without bound, into an infinitely long gap characterized by $H = \infty$.)

5.3 The notions of mild and wild in the case of extrapolated cartoons

We shall examine the law of large numbers and the Fickian law of diffusion.

Counterparts of the law of large numbers (LLN). Several cases must be distinguished.

• *The unifractal case.* The function varies in clock time, H is unique with $H \leq 1$, and $\Delta X \sim (\Delta t)^H$ for every t. If so, the sample average is $\Delta X/\Delta t \sim (\Delta t)^{H-1}$. It follows that the LLN holds and the limit is 0.

• *The mesofractal case.* The function varies in fractal time, and H takes a unique non-degenerate value $H < 1$, except that in interpolation, $H = 0$ for points in a discontinuity. In extrapolation, the fixed point is never in a discontinuity; therefore, the LLN holds.

• *The multifractal case.* The function varies in multifractal time, and in interpolation replaces the single H by a collection of H_i, some of them > 1 while others < 1. One can show that if the fixed point is chosen at random, LLN *fails* with probability 1. The set of fixed points for which LLN holds is very small (of zero measure). We shall return to this issue in Section 5.3.

Absence of counterpart of the central limit theorem (CLT). No choice of $A(T)$ makes $X(T)/A(T)$ converge to a non-trivial limit. This is part of the price one has to pay for the replacement of randomness by non-random fractality.

Counterparts of the Fickian law of diffusion (FLD). One tends to view the Fickian form $A(T) = \sqrt{T}$ as a simple corollary of the Gaussianity of the limit. But it is not. When $H_i = H$ for all i, only $A(T) = T^H$ makes $X(T)/A(T)$ oscillate without end, rather than collapse to 0 or ∞. The Fickian law of diffusion is satisfied when $H = 1/2$. This requirement allows the cartoon of WBM, of course. Mesofractal cartoons also allow $H = 1/2$, but a careful study (which must be postponed to M 1997N and 1997H) suggests that this case is of limited interest.

5.4 An ominous and possibly inconceivable implication of the extrapolation of multifractals

Thus far, the passage from fractals to multifractals deliberately avoided conceptual roadblocks and proceeded in as low a key as possible, but for the next topic a low-key tone would be hard to adopt. In the Gauss-Markov universe and related processes, the effects of large excursions are well-known to be short-lived and to regress exponentially towards the mean. The multifractal world is altogether different and the following inference gives a fresh meaning to the term, "wildness".

Consider a function $x(t)$ drawn as a multifractal cartoon of the kind examined in Section 4, with both t and x varying from 0 to 1. Now hold Δt constant and extrapolate to a time interval of length 1 placed at a distance T away from the original [0, 1]. A striking result concerns the incre-

ments Δx of $x(t)$ over that increasingly distant interval. As $T \rightarrow \infty$, those increments *will not regress at all.* For example (but we cannot stop here for a proof), the average of $|\Delta x|^{1/H}$ will *increase* without bound, like T to the power $-\tau_D(-1) > 0$, if that power is finite, and even faster otherwise.

In words, exponential regression to the mean is replaced by a power law "explosion." We already know that tail lengths explode as one interpolates, and now find that the same is true as one extrapolates.

To understand intuitively the explosion that accompanies extrapolation, the easiest is to reinterpret Figure 5, by imagining that it relates to time span much longer than 1 and that the unit time interval from which one will wish to extrapolate is chosen at random. We know that it is in the nature of interpolation for multifractals that a randomly chosen short interval will with high likelihood fall within a region of *low* variation (and a very short interval will fall within a region of *very low* variation.) A corollary is that the variation is likely to be wilder outside the unit interval than it is inside.

The most likely response to this wildly "unstable" scenario is the usual one: to argue that, well before any explosion occurs, the process is bound to "cross over" to another process obeying different rules. Be that as it may, the consequences of this scenario are fascinating, and will be explored in a more suitable context.