PAGEOPH, Vol. 131, Nos. 1/2 (1989) 0033-4553/89/020005-42$1.50 + 0.20/0

Multifractal Measures, Especially for the Geophysicist*

BENOÎT B. MANDELBROT[1,2]

Abstract—This text is addressed to both the beginner and the seasoned professional, geology being used as the main but not the sole illustration. The goal is to present an alternative approach to multifractals, extending and streamlining the original approach in MANDELBROT (1974). The generalization from fractal *sets* to multifractal *measures* involves the passage from geometric objects that are characterized primarily by one number, namely a fractal dimension, to geometric objects that are characterized primarily by a function. The best is to choose the function $\rho(\alpha)$, which is a limit probability distribution that has been plotted suitably, on double logarithmic scales. The quantity α is called Hölder exponent. In terms of the alternative function $f(\alpha)$ used in the approach of Frisch-Parisi and of Halsey *et al.*, one has $\rho(\alpha) = f(\alpha) - E$ for measures supported by the Euclidean space of dimension E. When $f(\alpha) \geq 0$, $f(\alpha)$ is a fractal dimension. However, one may have $f(\alpha) < 0$, in which case α is called "latent." One may even have $\alpha < 0$, in which case α is called "virtual." These anomalies' implications are explored, and experiments are suggested. Of central concern in this paper is the study of low-dimensional cuts through high-dimensional multifractals. This introduces a quantity D_q, which is shown for $q > 1$ to be a critical dimension for the cuts. An "enhanced multifractal diagram" is drawn, including $f(\alpha)$, a function called $\tau(q)$ and D_q.

Key words: Fractal, multifractal, measure, Hölder, limit theorem.

1. Introduction and Motivation. Reasons Why Multifractals are Indispensable in Geophysics and in Other Sciences

The topic of multifractals is bound to become of increasing importance to geophysics, in particular if the present volume becomes influential.

In one phrase, the generalization from fractal sets to multifractal measures involves the passage from geometric objects characterized primarily by one number, to geometric objects characterized primarily by a function. This function can be a probability distribution that has been renormalized and plotted suitably.

In a different single phrase, the generalization for fractal sets to multifractal measures involves the passage from a finite number of fractal dimensions to an

---

* Note: This text incorporates *and supersedes* MANDELBROT (1988). A more detailed treatment, in preparation, will incorporate MANDELBROT (1989).

[1] Physics Department, IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, U.S.A.
[2] Mathematics Department, Yale University, New Haven, CT 06520, U.S.A.

infinite number of "dimensions." Moreover (and this is a special point of this paper), these "dimensions" can be negative. We shall have gone far from the integer-valued dimensions of Euclid.

1.1. The Example of Copper

Correspondent to the simplest fractals, the basic idea is self-similarity, either exact or approximate. The closely related notions of self-similar fractal or self-similar multifractal can be phrased in many ways, but the geophysicist might best understand them in the context of the distribution of a rare mineral, such as copper. We first consider high-grade copper, then consider gradually lower grades.

High-grade copper is of course distributed nonuniformly: it concentrates in very few regions of the world. If one examines one such region in detail, however, copper continues to be found to be nonuniform: it concentrates in few sub-regions. And so on. It is reasonable, therefore, to suppose that the relative distribution of high-grade copper is the same (in the statistical sense) within each copper-bearing region, whether it is small or large. This being granted, take a large region, and cover it by a grid of equal "cells." As the cell sizes are made smaller, the total area of the cells that contain high-grade copper is found to shrink.

Mathematics has long known a construction that follows this process, and fractal geometry has "tamed" this construction to make it a model of nature. For example see MANDELBROT (1982), *The Fractal Geometry of Nature (FGN).* In the language of fractal geometry, high-grade copper is usefully viewed as "concentrated on," or "supported by," a self-similar fractal set of low fractal dimension.

Next, examine lower grade copper. The fact that it is more widespread in nature is expressed by its being supported by a fractal set of higher fractal dimension.

Overall, in order to give a full representation of the distribution of copper, it is seen that fractals are necessary and that no single fractal set is sufficient. A simple description consists of giving the fractal dimensions corresponding to each of a sequence of grades, as defined by thresholds varying from 0 to a very high value that is rarely exceeded.

The overall idea of the preceding paragraph has been combined with the generalization of the notion of self-similarity from sets to measures, and has thereby led to the notion of self-similar multifractal *measure*. To say that a multifractal *is a measure* and *not a set* is a very important distinction. It will be explained in Section 3.1. Our work on multifractals was initially concerned with the intermittency of turbulence, and was mostly carried out in the period 1968 to 1976, but it had started about 1962. My book *FGN* surveys multifractals on pp. 375–376, but this survey is overly sketchy and is now obsolete.

1.2. An Interesting Old Quote

The simplest of all multifractals, which is nonrandom, is called binomial and is

discussed in Section 5. This construction that has long been known to mathematicians, and has been tamed by fractal geometry, to make it a model of Nature. It happens that the basic circumstances that call for the binomial multifractal measure are very intuitive, and have nearly been rediscovered in the earth sciences context described in Section 1.1. Indeed, the geologist DE WIJS (1951) (quoted in *FGN*, p. 376) has described them as follows:

"Consider a [body of ore] with a tonnage W and an average grade M. With an imaginary cut we slash this body into two halves of equal tonnage $\frac{1}{2}W$, differing in average grade. Accepting for the grade of the richer half $(1+d)M$, the grade of the poorer half has to be $(1-d)M$ to satisfy the condition that the two halves together average again M... A second imaginary cut divides the body into four parts of equal tonnage $\frac{1}{4}W$, averaging $(1+d)^2M$, $(1+d)(1-d)M$, $(1+d)(1-d)M$, and $(1-d)^2M$. A third cut produces $2^3=8$ blocks, namely 1 block with an average grade of $(1+d)^3M$, 3 blocks of $(1+d)^2(1-d)M$, 3 blocks of $(1+d)(1-d)^2M$, and one block of $(1-d)^3M$. One can visualize the continued division into progressively smaller blocks... The coefficient d as a measure of variability adequately replaces the collective intangibles [dear to those who feel that ore estimation is an art rather than a science], and statistical deductions based upon this measure can abolish the maze of empirical and intuitive techniques."

Of course, de Wijs did not even begin to explore the geometric aspects of his first sketch of a model, and neither he nor notable followers (including G. Matheron) had an inkling of fractals or of multifractals, e.g., of the basic notion of fractal dimension. However, assume that the ore density is independent of grade, making tonnage equivalent to volume, and allow the (reinterpreted) scheme of de Wijs to continue *ad infinitum*. We shall see that this leads to the conclusion that the ore "curdles" into a binomial multifractal.

1.3. Relative Intermittency in a Context Broader than that of Metals

To broaden the scope of multifractals, let us quote from the subsection on *Relative Intermittency* of my book *FGN*, p. 375 ss.

"The phenomena to which [multi] fractals are addressed are scattered throughout this Essay, in the sense that many of my case studies of natural fractals negate some unquestionable knowledge about Nature.

"We forget in Chapter 8 that the noise that causes fractal errors weakens between errors but does not desist.

"We neglect in Chapter 9 our knowledge of the existence of interstellar matter. Its distribution is doubtless *at least* as irregular as that of the stars. In fact, the notion that it is impossible to define a density is stronger and more widely accepted for interstellar than stellar matter. To quote deVaucouleurs, 'it seems difficult to believe that, whereas visible matter is conspicuously clumpy and clustered on all

scales, the invisible intergalactic gas is uniform and homogeneous . . . [its] distribution must be closely related to . . . the distribution of galaxies . . . .'

"And in Chapter 10 the pastry-like sheets of turbulent dissipation are an obviously oversimplified view of reality.

"The end of Chapter 9 mentions very briefly the fractal view of the distribution of minerals. Here, the use of closed fractals implies that, between the regions where copper can be mined, the concentration of copper vanishes. In fact, it is very small in most places, but cannot be assumed to vanish everywhere.

"In each case, [portions of space] of less immediate interest were artificially emptied to make it possible to use *closed* fractal sets, but eventually these areas must be filled. This can be done using a fresh hybrid [namely, a] mass distribution in the cosmos such that no portion of space is empty, but, [given two] small thresholds θ and λ, a proportion of mass at least $1 - \lambda$ is concentrated on a portion of space of relative volume at most θ."

1.4. A Feature of Most Direct Importance in Many Sciences: Many Measures are not Observable Directly, only Through "Cuts"

The exploration of the earth cannot be carried out fully in three dimensions. Very often it must follow a straight bore-hole to obtain a straight 1-dimensional cross-cut through a real system that is intrinsically 3-dimensional. Often flat cuts are all that is available for inspection. The same constraint is encountered when turbulence in 3-dimensional space is explored via 1- or perhaps 2-dimensional cuts. Typically, the positions of these cuts bear no relation to the overall turbulence, and can therefore be thought of as having been chosen at random.

Consider also the context of strange attractors. Their full natural space has a very high dimensionality. But they are typically examined via a "Poincaré section" by a plane. The position of the plane, again, often bears no relation to the full attractor, and can be viewed as having been chosen at random. As we shall see by examining typical cases, the measure observed along a random cut has properties that are without counterpart in the measures studied in their natural space, and vice versa. This raises the issue of what can and what cannot be inferred from a cut to the whole measure. This issue is extraordinarily important and has motivated our early work of 1968–1976, especially MANDELBROT (1974).

2. Two Alternative Summaries

The present text begins with introductory material, continues with the binomial measure (Section 5) and then proceeds to step by step generalizations. One must wait until Section 8 to initiate a discussion of the cuts. The result is longer than we would have preferred.

We hope to interest both the reader who is not yet fully familiar with multifractals, and the reader who is already familiar with them via the variant approach of FRISCH and PARISI (1985), which was adopted by HALSEY *et al.* (1986).

As to the reader familiar with our earlier work, MANDELBROT (1988), he will find the present text to be much more detailed; for example, Sections 3.4 and 3.5 of the earlier work have been replaced by Sections 5 to 8.

Our goal is not to teach manipulations, but to present the reader who is new to multifractals with what we believe is the most understandable and simplest form of their theory, and to provide the skilled reader with the surprisingly simple but subtle explanation of the formal manipulations with which he is familiar.

2.1. Summary for the Reader New to Multifractals

The work is best summarized as follows:

The notion of self-similarity extends readily from fractal sets to measures. As we shall see in Section 3 a measure is simply a way of specifying a method of spreading mass, or probability, or other "stuff," over a supporting set. The distribution may be spread on a Euclidean "support", like an interval or a square, or it may be restricted to a fractal support, like a Cantor set. The function is the same, except for detail.

In order to describe quantitatively a self-similar method of spreading stuff around, one standard first characteristic is a sequence of moments, or an exponent in the cumulant generating function, which is denoted in this context by $\tau(q)$. The earliest graph of $-\tau(q)$ is given as Fig. 2 in MANDELBROT (1974).

Another characteristic, equivalent to the first one, is a limit distribution function. This limit is akin to an ordinary probability distribution function. Also, the renormalization must follow a very unusual and ill-known path: one needs a *multiplicative* renormalization that is unfamiliar. The *additive* renormalization, which is $\bar{I}$ and familiar to every student of probability, and which leads to the central limit theorem, is worthless. Because of this renormalization, the limit distribution function is best considered after logarithmic transformations of both axes, that is, as plotted with doubly logarithmic coordinates. We shall denote the resulting function by $\rho(\alpha)$. An alternative quantity $f(\alpha)$, given by $f(\alpha) = \rho(\alpha) + 1$ for measures on the line, has become entrenched in the literature to specify the limit distribution.

One can obtain $f(\alpha)$ from $\tau(q)$, and $\tau(q)$ from $f(\alpha)$, by a direct or inverse Legendre transform. This property follows immediately from the Lagrange multipliers approach to the Gibbs distribution in thermodynamics, an approach long familiar to every physicist. Later, we give a full mathematical justification of the formalism, valid in a broader context in which $f(\alpha)$ can very well be negative. This justification is provided by reference to existing (but little-known) limit theorems of probability due to Harald Cramèr, and concerned with "large deviations". The Legendre transform expresses the intuitive fact that $f(\alpha)$ can be obtained as the cap-convex envelope of an easily drawn family of straight lines.

On a plot of $f(\alpha)$, the quantity—$\tau(q)$ is the ordinate of the intercept of that tangent to $f(\alpha)$, whose slope is q, with the vertical axis. And the intercept of the same tangent with the main bisector of the coordinate axes is the quantity $D_q = \tau(q)/(q-1)$, which is interpreted later as a critical dimension.

MANDELBROT (1974) has introduced two distinct kinds of random self-similar multifractals, respectively called *conservative* (or *microcanonical*) (Section 7) and *canonical* (Section 8). The latter are the more sophisticated, and needed for the study of low-dimensional cuts of multifractals embedded in a high-dimensional space (see Section 1.4). In particular, the canonical 1974 multifractals are necessary to understand the very peculiar standing of the lognormal distribution, which is widely mentioned when applying multifractals to turbulence. Observe that FRISCH and PARISI (1985) explicitly note that, while their approach had originally stemmed from MANDELBROT (1974), it was of lesser generality, because it did not accommodate the lognormal.

2.2. *Summary for the Reader Familiar with the Multifractal Formalism, as Originated by Frisch and Parisi*

To start with a question: Since "multifractal" is not a physical notion, but a probabilistic (and measure-theoretical) concept and tool, where is the probability theory in the many recent papers devoted to this topic? The answer is simple and is given in this work. It is best summarized as follows:

The function $\tau(q)$ with which you are familiar is a standard probabilistic tool to represent measures, called cumulant generating function.

The quantity α with which you are familiar is a standard notion, called Hölder exponent. There is no gain in calling it "strength of singularity." HENTSCHEL and PROCACCIA (1983) call it dimension, but experience shows that many confusions are avoided if this term (in all its multiple and still multiplying forms!) is reserved to *sets* and never used to apply to *measures*.

There is a wide belief among those familiar with the multifractal formalism that the function $f(\alpha)$ with which you are familiar is a *new* concept that is *defined* via a Legendre transform from the function $\tau(q)$, therefore can be labeled at will (for example as the "spectrum of singularity"). In reality, if the measure is supported by a Euclidean interval of dimension 1, $f(\alpha)$ is simply equal to $\rho(\alpha) + 1$, where $\rho(\alpha)$ is a limit probability distribution function plotted on doubly logarithmic coordinates. But the set that supports the measure can be a fractal. This leads to a slightly more complex rule described in Section 6.6.

There are several methods for estimating $\rho(\alpha)$ and $f(\alpha)$. The Legendre transform starting with $\tau(q)$ yields one such method. But the estimation of the $\tau(q)$ involves many diverse smoothing operations. They are not innocuous, yet they have not been subjected to the careful analysis they deserve and demand.

The quantities $D_q = \tau(q)/(q-1)$ are "critical dimensions" for $q > 1$. They were introduced in MANDELBROT (1974) through a theorem asserting that along a cut of dimension D through a measure, one has $\langle \mu^q(dx) \rangle = \infty$ when $q > q_{\text{crit}}(D)$. The function inverse of the function $q_{\text{crit}}(D)$ is D_q.

The attractiveness of multifractals may to some extent be due to an aura of mystery. All mystery is eliminated, however, when one understands the nature of the formal manipulations. But does it *really* matter that $\tau(q)$, α and $f(\alpha)$ are properly labeled? To respond, let us note that our approach obtains the Legendre transform between $\tau(q)$ and $f(\alpha)$ via Lagrange multipliers, as is usual in first courses of thermodynamics. In thermodynamics, there are several standard ways of making the use of these multipliers rigorous after the fact. One way is to follow the Darwin-Fowler method of steepest descents. This used to be a standard topic in textbooks, but is unfortunately no longer familiar to every physicist. Thus, in effect, the Frisch-Parisi method skips Lagrange multipliers, and proceeds to the Darwin-Fowler method immediately. No teacher of thermodynamics would proceed in this way in a first course! Nor should a teacher of multifractals.

The preceding remarks help explain that, in our experience, the Frisch-Parisi approach to multifractals has generated substantial confusion. No one has an intuition of what a "spectrum of singularities" can or cannot be like. Proper foundations flush out confusion. As a foremost example, the Frisch-Parisi interpretation of $f(\alpha)$ as a fractal dimension has led to paradoxes linked to negative $f(\alpha)$. These paradoxes had begun as surprising anecdotal evidence, and have of course ended by being explained away by suitable special developments of the theory. They do not even arise in our properly probabilistic approach.

The formalism originated by Frisch and Parisi has now given rise to a very extensive literature. But this is not the proper place to survey it, even in part. Our goal is to set out an alternative approach on its own terms, and the relations between the two approaches have better be developed elsewhere.

3. *Spatial Variability Beyond Fractal Homogeneity*

3.1. *The Essential Distinction Between a Set and a Measure, and the Notion of "Singular Measure"*

To understand multifractals beyond the crude formalism, a full appreciation of the distinction between the concepts of *set* and of *measure* is essential.

Fortunately, a knowledge of either set theory or measure theory is completely unnecessary. More precisely, the complications of measure theory enter into the theory at a stage beyond the topics discussed in this paper, for example when one wishes to follow the detailed proofs of some of the theorems to which we shall refer.

Basic ideas are always explained best when stated in the simplest context, which in the case of this paper is usually the interval [0, 1], the square $[0, 1]^2$ or the cube $[0, 1]^3$. In a Euclidean space of dimension E, the interval, square or cube are examples of basic sets of the form $[0, 1]^E$.

The notion of a set. In order to define a set S in $[0, 1]^E$, one needs a rule that says for each point $P \in [0, 1]^E$ whether it satisfies $P \in S$ or $P \notin S$. Such a rule is provided by the "indicator function" $I(P)$ defined as $I(P) = 1$ if $P \in S$ and $I(P) = 0$ if $P \notin S$.

Uniform or homogeneous measure. It is well-known that one can associate with the whole $[0, 1]^E$ a uniform measure. The set and its uniform measure are mathematically equivalent. Next, consider a fractal set constructed recurvisely, e.g., a Sierpinski gasket, which is described in Figure 1. The most natural measure, again called "uniform," gives equal weight to each of its thirds. Again, the set and its uniform measure are equivalent. However, the uniform measure is very special, and in many cases very unrealistic, as has already been stressed in Section 1.

The notion of the limit of a set. Multifractals play two main roles, each of which can begin with a complicated set of points. The first role is exemplified by the distribution of copper on earth. Since the mining engineer can see no harm in assuming

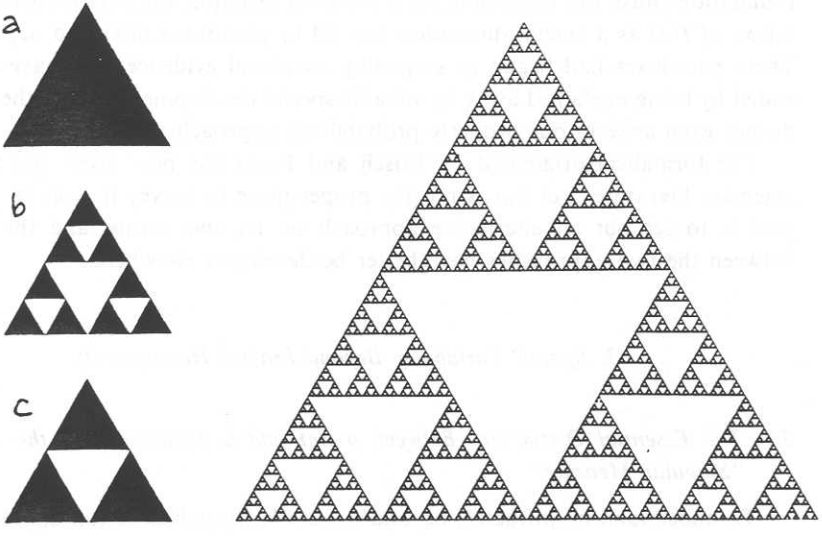

Figure 1

The construction of Sierpinski's gasket proceeds as follows. The initiator, shown in Figure 1a is a triangle. In the stage between Figures 1a and 1b, this triangle is divided into four and the middle fourth is erased, yielding 1b. Between 1b and 1c, each of the three remaining fourths is, divided into four and the middle fourth of each remaining triangle is erased. The same process is repeated *ad infinitum*.

that each copper atom reduces to a point, all the copper atoms on earth form a set of points. The second role is exemplified by the successive "sample values" of a random process examined at discrete instants of time, or the successive "orbit points" of a deterministic process. Again, one deals with a sequence of points or ordered set.

In all the cases of interest, however, these sets are very large, and to specify them completely would be an unmanageable task. It would also be pointless. This leads to a natural impulse, especially in the second example: let time run forever and take the limit of the set of successive positions of our process. Question: is the limit a set? In order to answer usefully, consider a collection of points that have been chosen at random on [0, 1].

First suppose that their probability distribution is continuous, with a density $\mu'(x)$ such that $\mu'(x) > 0$ for all x. If so, the limit of an increasingly large set of sample points is indeed a set, namely the whole interval [0, 1]. The fact that this limit is independent of the function $\mu'(x)$ means of course that the limit contains less information than the finite set that tends to this limit. The answer is obvious: the least that we must do in either case is to specify, not only a limit set, but also a density on this set.

A second familiar case is when our random point can take only certain admissible values x_g, each of which has a known probability μ_g. In this case, the limit is a (denumerable) set $\{x_g\}$, but we must also specify the μ_g which are said to be "supported' by this set.

How to describe a random process that generates x with a density $\mu'(x)$? A first characteristic is $\mu'(x)$ itself, or its integral $\mu(x) = \int_0^x \mu'(s)\, ds$, which is a cumulative probability distribution. An alternative though less complete characteristic is the function

$$x(\mu) = \text{total length of the intervals of } x \text{ within which } \mu'(x) < \mu.$$

The densities $\mu'(x)$ and the probabilities μ_g happen to define the two classical examples of "measures." Multifractals are also measures, but they belong to a third class, which has no density and is not discrete, and is called "singular." The best is to first discuss them in an example.

3.2. An Old but Good Illustration of the Concept of Singular Measure

Since "to see is to believe", Figure 2 reproduces the earliest completed illustration of a random multifractal, and appeared first in our earliest full paper on this topic (MANDELBROT, 1972). Many geologists and geophysicists may find it difficult *not* to find never-ending new interpretations for it.

The horizontal axis shows the "time" t, as "discretized" into small boxes of uniform width Δt, and the vertical axis shows the sequence of the "measures" of

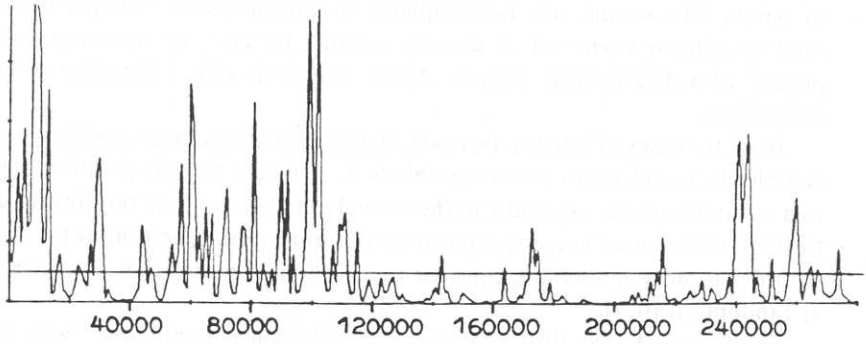

Figure 2
The earliest simulation of a sample from a multifractal measure, namely the limit lognormal measure of MANDELBROT (1972).

these boxes. Let us say again that, if the total integral measure over the time span [0, 1] is set to 1, one can think of *the measure in a box as the probability of hitting this box*. But the measure in a box is, more generally, the amount of any kind of "stuff" contained within it.

The measure of the interval $[0, t]$, call it $\mu([0, t])$, is of course an increasing function of t, and what is plotted here is the sequence of its finite increments $\mu([0, t + dt]) - \mu([0, t])$. The increments' values are joined to form a curve, but this curve is *not* what it seems to be. It is not the graph of an intrinsically interesting function, and it cannot readily be transformed into one.

To provide contrast and familiar background, let us draw the analogous diagram for a measure having a continuous density $\mu'(t)$. In that case, $\mu(\Delta t)/\Delta t$ is an approximation to the graph of the function representing $\mu'(t)$, and a first characterization of our measure would be provided by the distribution of this approximate density $\mu(\Delta t)/\Delta t$ in time. As $\Delta t \to 0$, the approximate density would tend to the true $\mu'(t)$.

In the present instance, however, the situation is extremely different. By design, the measure is approximately self-similar. This notion will be discussed in Section 3.4, but it may already be said that when the measure in Δt is examined on much finer scales $\Delta_2 t \ll \Delta t$, it proves to be about as irregular as the whole Figure 2. For example, if Δt is halved, the sharing of the $\mu(\Delta t)$ between the two halves is usually very unequal.

It follows that this $\mu([\Delta t])$ *grossly fails to have a local density*. Nor is it discrete. Therefore, taking the limit $\Delta t \to \infty$ transforms the wiggly curve on Figure 2 into a curve that is even more wiggly, hence does not become an increasingly close approximation to a function.

Suppose that the measure $\mu([0, t])$, considered as a function of t, is continuous but is not differentiable. Then the measure $\mu(dt)$ has no local density, and is called "singular." (This is a technical term, due to Lebesgue.) For a singular measure,

the notion of asymptotic "distribution" for the values of $\mu([\Delta t])$ has no meaning. Fortunately, a suitable substitute is available.

3.3. *The Limit Probability Distribution $\rho(\alpha)$, and the Corresponding Function $f(\alpha)$*

Begin with the equivalent of Figure 2 for a sequence $\Delta_k t$ of decreasing values of Δt. Then, for each $\Delta_k t$, replace $\mu([\Delta_k t])$ by the quantity $\alpha_k = \log \mu([\Delta_k t])/\log \Delta_k t$, which is called Hölder exponent. Next, form the probability density of α_k, and replace it by the quantity $\rho_k(\alpha) = -\log(\text{probability density of } \alpha_k)/\log \Delta_k t$. Each of these two steps consists of first taking logarithms, and then renormalizing them by dividing by the factor $-\log \Delta_k t$. This renormalizing factor is not familiar, and may seem peculiar and artificial. However, it will be justified in Section 7.2. There, we examine successive examples for which the measure is known to be multifractal, and we find that $\rho_k(\alpha)$ converges as $\Delta t \to 0$ to a limit that does not reduce to either 0 or ∞. Any other renormalizing factor, to the contrary, would yield a degenerate limit.

The limit of $\rho_k(\alpha)$ will be denoted by $\rho(\alpha)$. It will be seen that $\rho(\alpha)$ is negative for all α, except for the value where $\rho(\alpha)$ reaches its maximum, which is the expectation $\langle \alpha \rangle$. But one must realize that the convergence to the limit may be slow. The property that $\rho_k(\alpha) \to \rho(\alpha)$ can be turned around, and used to *define* the notion of multifractal.

As has been first pointed out by FRISCH and PARISI (1985), it is also worth considering the quantity $f(\alpha) = \rho(\alpha) + 1$. In effect, Frisch and Parisi have discussed $f(\alpha)$ in the special case $\rho(\alpha) \geq -1$, so that $f(\alpha) > 0$, for all α. They have pointed out that, in this special case, one can interpret $f(\alpha)$ as being the fractal dimension of a suitable set. The replacement of $\rho(\alpha)$ by $f(\alpha)$ has virtues in some cases, but our feeling is that, fundamentally, it hides the nature of the multifractals.

Between my early papers and the preprint of FRISCH and PARISI (1985) (which was distributed in 1983), multifractals continued to develop only in the sense that the mathematical background was cleaned up and extended (e.g., in KAHANE and PEYRIÈRE, 1976). But they did not receive new applications (nor were they mentioned in *Physical Review Letters*). Their spread is a recent phenomenon, and most readers who have heard of them are likely to know presentations that follow the approach common to Frisch and Parisi and to HALSEY *et al.* (1986). Unfortunately, the algebra of these presentations is needlessly complicated, artificial, and of limited applicability and the terminology of HALSEY *et al.* hides the extremely simple nature of the underlying structure. We shall, therefore, adopt the notation of HALSEY *et al.*, but follow our original approach in the form into which it has lately developed.

3.4. *Self-similar Measures and Beyond*

We need a few more formal definitions concerning positive measures. The

measure of a set S will be written as $\mu(S)$. Again, those not familiar with measure can think of μ as being the probability of hitting the set S. The multifractal measures obtained as a result of multiplicative cascades are closest in spirit, to the exact self-similar fractal sets.

Recall that a fractal set is exactly self-similar, if it can be decomposed into parts, each of which is obtained from the whole by a transformation called isotropic contraction, or contracting similitude, to be denoted by $\mathscr{K}$. A self-similar set is fully determined by a collection of contractions. For example, each third of a basic fractal called a Sierpinski gasket is obtained from the whole by a contracting similitude of ratio $r = 1/2$. Starting with any triangle in a "prefractal collection of triangles", the interpolation of the shape itself continues without regard to the "past" construction steps.

Now suppose that a (positive) measure $\mu(P)$ is defined for each third of the gasket, then is interpolated for each third of a third etc. More generally, when the part P' of a fractal is obtained from the part P by the contracting transformation $\mathscr{K}$, so that $P' = \mathscr{K}(P)$, the conditional measure of P' in P is defined precisely like a conditional probability. It is the ratio $\mu(P')/\mu(P)$ of the measure $\mu(P')$ to the conditioning measure $\mu(P)$.

Now, the idea of self-similarity for a measure expresses that the process of successive interpolations continues without regard to the "past" steps. The rough idea is that, as the parts contract, the measures they carry contract proportionately. To express this idea formally, take a second contracting transformation $\mathscr{L}$, and compare $\mu(P')/\mu(P) = \mu\{\mathscr{K}(P)\}/\mu(P)$ with $\mu\{\mathscr{L}(P')\}/\mu\{\mathscr{L}(P)\}$. If these conditional measures are identical for every choice of $\mathscr{K}$ and $\mathscr{L}$, the measure μ will be called a *strictly self-similar multifractal.*

A random measure is called *statistically* self-similar if, given one or a finite collection of nonoverlapping parts $P_\gamma = \mathscr{K}_\gamma(P)$, the distribution or the joint distribution of the quantities $\mu(P_\gamma)/\mu(P)$ depends only on the contractions $\mathscr{K}_\gamma$.

Side remark. In a more general mathematical fractal set, the parts are obtained from the whole by transformations that are *nonlinear*. Examples where the contractions are in some sense *almost linear* include the Julia sets of polynomial maps. The corresponding multifractals include the harmonic measures on these sets. Other examples of multifractal measures concern the limit sets of groups based upon inversions in circles (*FGN*, Chapters 18 and 20). The limit set itself may be a straight line, as in the example examined by GUTZWILLER and MANDELBROT (1988). Finally, the "fat fractals" (a new term for the fractals investigated in *FGN*, Chapter 15) and the *Mandelbrot* set involve essentially nonlinear transformations. As yet, there is no general agreement about which transformations are acceptable in defining the terms "fractal" and "multifractal."

3.5. Five Basic Concepts Concerning Multifractals

Aside from the distinction between set and measure, multifractals involve five basic notions, each of which will be introduced in the simplest possible context. A)

The Hölder α is introduced in Section 4, in the context of the homogeneous Cantor measure. B) The functions $\rho(\alpha)$ and $f(\alpha)$ are introduced in Section 3, in the context of the binomial measures. C) The Legendre transform "multifractal formalism" is introduced in Section 6, in the context of the multinomial measures. D) The notion of latent α's is introduced in Section 7, together with a first form of randomness, in the context of "conservative 1974 multifractal measures." E) The notions of critical exponent q_{crit}, and of critical dimension function D_q (D_q is the inverse function of q_{crit}), and the notion of virtual α are introduced in Section 8, together with a second form of randomness, in the context of "canonical 1974 multifractal measures."

4. *The Basic "Unifractal" is the Cantor Measure. It is Nonrandom and Self-similar. Introduction of the Hölder α*

4.1. Definitions

It is probably not necessary here, to repeat the definitions of the triadic Cantor dust $\mathscr{C}$, and of its devil staircase function $C(x)$ but it takes little space. To define $\mathscr{C}$, one starts with the interval [0, 1] and one cuts out the middle third open interval]1/3, 2/3[. A closed interval is denoted by [ ], and an open one by] [. Then one cuts out the middle third open interval of [0, 1/3] and [2/3, 1], etc. The staircase function $C(x)$ is a special case of the function $\mu([0, t])$ of Section 2.3. The left and right halves of $\mathscr{C}$ are assigned the equal measures 1/2, so that $C(1/3) = C(2/3) = 1/2$. Then the four quarters of $\mathscr{C}$ are assigned the equal measures 1/4, so that $C(1/9) = C(2/9) = 1/4$ and $C(7/9) = C(8/9) = 3/4$. And one continues to interpolate *ad infinitum*. Thus, the measure of every interval $[x', x'']$ in [0, 1], is obtained from the limit function $C(x)$ as $\mu([x', x'']) = C(x'') - C(x')$.

This measure is called Cantor measure, or *fractally homogeneous measure* on the Cantor dust, for the following reason. If two pieces of Cantor dust contained in $[x_1', x_1'']$ and $[x_2', x_2'']$ can be superposed by translation, they carry identical measures: $\mu([x_1', x_1'']) = \mu([x_2', x_2''])$.

4.2. The Devil Staircase $C(x)$ is a Diagonally Self-affine Fractal Curve

Indeed, it subdivides into the following three parts. The "middle part" is deduced from the whole by a degenerate (noninvertible) affinity of matrix $\begin{pmatrix} 0 & 0 \\ 0 & 3^{-1} \end{pmatrix}$ and fixed point $(2^{-1}, 2^{-1})$. The two "side parts" are deduced from the whole by affinities that share the same matrix $\begin{pmatrix} 2^{-1} & 0 \\ 0 & 3^{-1} \end{pmatrix}$ and have fixed points (0, 0) and (1, 1), respectively. In the present context, degenerate matrices "do not count." Therefore, the unique simplicity of the homogeneous multifractal on $\mathscr{C}$ is due to the fact that the matrix has only one possible nondegenerate form.

More generally, consider a fractal obtained by a recursion such that there are N contractions, all having ratios equal to the same r, and that each recursive stage divides the measure equally between the N parts. Then the measure is self-affine, the matrix being either degenerate of the form $\begin{pmatrix} 0 & 0 \\ 0 & r \end{pmatrix}$ or nondegenerate of the form $\begin{pmatrix} N^{-1} & 0 \\ 0 & r \end{pmatrix} = \begin{pmatrix} r^D & 0 \\ 0 & r \end{pmatrix}$.

4.3. *Application of the Notion of Hölder Exponent to Intervals Instead of Points. α and $f(\alpha)$ for the Unifractal Measures*

Adapting a classical mathematical notion due to Hölder to apply to the dyadic cell $[dt]$, we write

$$\alpha(t) = \frac{\log[\mu(dt)]}{\log(dt)}.$$

For the Cantor dust of dimension D, the interval $[0, 1]$ contains $N^k = 2^k$ nonempty cells of length $dt = b^{-k}$, each containing the same measure $N^{-k} = b^{-kD}$. For each of these cells, one has $\mu(dx) = (dx)^D$. Hence $\alpha = D$.

The remaining cells are empty, and they form a set of fractal dimension equal to 1. In these cells, one can say that $\alpha = \infty$.

Infinitesimals of the form $(dy) = (dx)^\alpha$ with $\alpha \neq 1$ are called *nonstandard*, and Abraham Robinson has constructed around them a field called *nonstandard analysis*. It is not needed in this discussion.

The function $f(\alpha)$. For future reference, it is useful to introduce a function $f(\alpha)$ that is defined for two points only:

$$\text{for} \quad \alpha = D, f(\alpha) = D, \quad \text{and for} \quad \alpha = \infty, f(\alpha) = 1.$$

4.4. *Finite Sample Prefactors*

In general, the Cantor measure on a Cantor set of base b is studied using cells whose base b' is such that $b' \neq b$ and $b' \neq b^h$ for all integer h. In that case, the number of nonempty cells continues to be "of the order" of N^k, but it is no longer equal to N^k. One can write it in the form $F_1 N^k$, where the prefactor F_1 depends on k and remains positive and finite. The measure in a nonempty cell, similarly, can be written as $F_2 N^{-k}$, where the prefactor F_2 depends on k, varies between cells, is finite but can be arbitrarily small. Therefore, the sample values of α are of the form $D + \log F_2/k$. These values may overlap several of the bins in which the values of α are sorted for the purposes of statistics.

5. The Basic Multifractal is the Binomial Measure. It is Nonrandom and Self-similar. Introduction of the Function $\rho(\alpha)$ and $f(\alpha)$, and of the Multifractal Representation

The binomial measure on the interval [0, 1] is the simplest example of a multifractal measure. By studying this example more carefully and in greater detail than is usual, this section hopes to make many of the important features of multifractals stand out clearly. In particular, the function $f(\alpha)$ is best understood when developed in a context in which no formalism is needed.

Under various names, the binomial and multinomial measures of Sections 5 and 6 have been long known in pure mathematics. But the first concrete application (to the distribution of turbulent diffusion, MANDELBROT, 1974), as well as various extensions in a "fractal" context have come with our work. *FGN*, calls them *Besicovitch measures*, to honor the principal early contributor of their study, but this terminology has never taken hold. It has now become obsolete, because it lends itself poorly to the necessity of distinguishing between the binomial and the multinomial cases.

5.1. Construction of the Binomial Measure, Derivation of the Equivalent Functions $f(\alpha)$ and $\rho(\alpha)$

The binomial measures are a family of multifractals that depend upon a single parameter m_0, variously called a *multiplier* or a *mass*. One assumes that m_0 satisfies $0 < m_0 < 1$ and $m_0 \neq 1/2$, and one defines $m_1 = 1 - m_0$. Often it is nice to choose m_0 larger than m_1, so that $1/2 < m_0 < 1$. (Note that, in terms of the de Wijs quote in Section 1.2, $m_0 = (1 + d)/2$.)

The mass in the "initiator" [0, 1] is taken as unity, then the basic "generating step" is to spread mass over the halves of every dyadic interval, with the relative proportions m_0 and m_1. The first stage yields the mass m_0 in [0, 1/2] and the mass m_1 in [1/2, 1], etc. After k stages, suppose that $t = 0, \eta_2\eta_2 \ldots \eta_k$ is the development of t in the counting base $b = 2$ and let φ_0 and φ_1 denote the relative frequencies of 0's and 1's in the binary development of t. Then the binomial measure is defined as the measure that assigns to the dyadic interval $[dt] = [t, t + 2^{-k}]$ of length $dt = 2^{-k}$ the mass

$$\mu(dt) = m_0^{k\varphi_0} m_1^{k\varphi_1}.$$

The binomial measure is self-similar. By construction, $\mu(dt)$ satisfies the property of self-similarity described in Section 3.4.

The notion of "pre-multifractal". One can define a sequence of measures $\mu_k(dt)$ such that $\mu_k(dt) = \mu(dt)$ if $[dt]$ is a dyadic cell of length 2^{-k}, and the mass is distributed uniformly over each cell. The measures $\mu_k(dt)$ can be called "pre-multifractal", because they play the same role relative to the binomial measure, as the broken line "prefractal" approximations play relative to the fractal sets such as the Koch curve.

The function $\mu(t)$, an analog of the Cantor staircase function $C(t)$, is diagonally self-affine. As we know, this means that its variation within a cell of side b^{-k} is a reduced-size replica of its variation over the original cell of side 1. In every cell, the reduction is an affinity whose matrix is diagonal. The horizontal linear reduction is b^{-k} in every cell. The vertical reduction is the measure within the cell, and varies from cell to cell.

The Hölder α. In the present case,

$$\alpha = \alpha(\varphi_0, \varphi_1) = -\varphi_0 \log_2 m_0 - \varphi_1 \log_2 m_1,$$

and $0 < \alpha_{\min} = -\log_2 m_0 \leq \alpha \leq \alpha_{\max} = -\log_2 m_1 < \infty$.

The box dimension. The number of intervals leading to φ_0 and φ_1 is $N(k, \varphi_0, \varphi_1) = k!/(k\varphi_0)!(k\varphi_1)!$. One can use N to form the following expression:

$$\delta(k, \varphi_0, \varphi_1) = -\frac{\log N(k, \varphi_0, \varphi_1)}{\log(dt)} = -\frac{\log[k!/(k\varphi_0)!(k\varphi_1)!]}{\log(dt)}.$$

This quantity δ is of the form $-\log N/\log r$ that characterizes the similarity dimension of a set. Hence one can call δ a *box fractal dimension*. More precisely, since the boxes belong to a grid, it is a grid fractal dimension.

For large k, the replacement of the factorial by the leading term in the Stirling approximation shows that

$$\lim_{k \to \infty} \delta(k, \varphi_0, \varphi_1) = \delta(\varphi_0, \varphi_1),$$

with

$$\delta(\varphi_0, \varphi_1) = -\varphi_0 \log_2 \varphi_0 - \varphi_1 \log_2 \varphi_1.$$

Derivation of a function $f(\alpha)$. Eliminating φ_0 and φ_1 between α and δ, we obtain a function written in parametric form. It will be denoted as $f(\alpha)$, instead of $\delta(\alpha)$, the reason being explained in Section 6.2.

Note that $0 \leq f(\alpha) \leq \min\{\alpha, 1\}$, with equality being achieved in four points: 1) $f(\alpha) = 0$ for $\alpha = \alpha_{\min} = -\log m$, 2) $f(\alpha) = 0$ for $\alpha = \alpha_{\max} = -\log m_1$, 3) $f(\alpha) = 1$ for $\alpha = (1/2)(\alpha_{\min} + \alpha_{\max})$ and 4) $f(\alpha) = \alpha$ for $\alpha = -m_0 \log m_0 - m_1 \log m_1$.

A first reading of this paper may proceed directly to Section 6.

5.2. Entropy, Information, Entropy Dimension and Information Dimension

The notion of entropy, as every physicist knows, is extremely important but far from easy to handle. One of the basis forms of entropy is written as $-\Sigma\, p_\beta \log_b p_\beta$, where the p_β are probabilities, and entropy written in this form has since been reinterpreted by Claude Shannon as an "information". Thus Section 5.1 proves the binomial case of the important result that $f(\alpha)$ is often formally an entropy-information. This role has led many authors to refer to it as "information dimension."

Though I do not like to argue about words, this term has always seemed inappropriate. Even the simplest similarity dimension $\log_b N$ is already formally identical to one of the basic expressions for entropy-information.

5.3. *Restatement of Section 5.1 in Terms of Rescaled Doubly Logarithmic Plots of Probability Densities*

Having obtained $f(\alpha)$, we shall spend several sections in a discussion of what it really does and does not mean. First, let us reinterpret the combinatorics in Section 5.1 in probabilistic terms. This will be easy, and very important because it will show that $f(\alpha)$ is *not* a new notion requiring an entirely new intuition, but a new form of a familiar notion, only requiring the further development of old intuitions.

Let a dyadic cell of length 2^{-k} be picked at random among the 2^k such cells. For the probability of hitting a cell that corresponds to a prescribed α or the equivalent φ_0 and φ_1, we shall use the notation $\Pr\{A = \alpha\}$. Here, A is to be read as "capital α", and we follow the probabilists' custom of denoting a random variable by a capital letter, and a possible value of this variable by the corresponding small letters. We have of course $\Pr\{A = \alpha\} = 2^{-k} N(k, \varphi_0, \varphi_1)$. As $k \to \infty$, $\Pr\{A = \alpha\} \to 0$ for all $\alpha \neq \langle \alpha \rangle$. The next step is to raise the basic question of probability theory. It is possible to "renormalize" $\Pr\{A = \alpha\}$ in such a way that it has for $k \to \infty$ a limit that is neither 0 nor ∞. Clearly, this is possible and even easy: it suffices to define

$$\rho(k, \varphi_0, \varphi_1) = -\frac{\log \Pr\{A = \alpha\}}{\log dt} = -\frac{\log 2^{-k} N(k, \varphi_0, \varphi_1)}{\log dt}.$$

Obviously, $\rho(k, \varphi_0, \varphi_1) = \delta(k, \varphi_0, \varphi_1) - 1$, with $\delta(k, \alpha)$ as in Section 5.1. Hence,

$$\lim_{k \to \infty} \rho(k, \varphi_0, \varphi_1) = \rho(\alpha) = f(\alpha) - 1.$$

5.4. *Restatement of $f(\alpha)$ and $\rho(\alpha)$ in Terms of "Tail" Dimensions and Probabilities*

What we have done is best stated in terms of "tail" quantities. After $k < \infty$ has been specified, the quantities φ_0 and φ_1 are both multiples of $1/k$. The "tail" quantities are defined by

$$N^+(k, \varphi_0, \varphi_1) = \sum_{\psi_0 \geq \varphi_0} N(k, \psi_0, -\psi_0)$$

$$N^-(k, \varphi_0, \varphi_1) = \sum_{\psi_0 \leq \varphi_0} N(k, \psi_0, -\psi_0).$$

By analogy with $-\log N/\log dt$, one can use N^+ and N^- to define the tail box dimensions $\delta^+(k, \varphi_0, \varphi_1)$ and $\delta^-(k, \varphi_0, \varphi_1)$, and then to seek and define their limits $\delta^+(\varphi_0, \varphi_1)$ and $\delta^-(\varphi_0, \varphi_1)$. A bit of combinatorics and Stirling approximations easily yields the following facts:

For $\varphi_0 > 1/2$, $\delta^+(\varphi_0, \varphi_1) = \delta(\varphi_0, \varphi_1)$ and $\delta^-(\varphi_0, \varphi_1) = 1$.

For $\varphi_0 = 1/2$, $\delta^+(\varphi_0, \varphi_1) = \delta^-(\varphi_0, \varphi_1) = 1$.

For $\varphi_0 < 1/2$, $\delta^+(\varphi_0, \varphi_1) = 1$ and $\delta^-(\varphi_0, \varphi_1) = \delta(\varphi_0, \varphi_1)$.

Next, the function δ^* is defined by stringing together the information-carrying portions of δ^+ and δ^-. Thus,

$$\delta^*(k, \varphi_0, \varphi_1) = \min\{\delta^+(k, \varphi_0, \varphi_1), \delta^-(k, \varphi_0, \varphi_1)\}.$$

The next obvious step is to define the positive (resp. negative) "tail probabilities" of α as being $\Pr\{A \geq \alpha\}$ (resp., $\Pr\{A \leq \alpha\}$), and to define "the" tail probability as being

$$\Pr\{A \geq \alpha\} \quad \text{when} \quad \alpha > \langle\alpha\rangle$$

$$\Pr\{A \leq \alpha\} \quad \text{when} \quad \alpha < \langle\alpha\rangle.$$

By a further analogy, we define the quantities $\rho^+(k, \varphi_0, \varphi_1) = \delta^+(k, \varphi_0, \varphi_1) - 1$, and similarly $\rho^+(k, \varphi_0, \varphi_1)$, $\delta^+(k, \varphi_0, \varphi_1)$ and $\rho^-(\varphi_0, \varphi_1)$. It is easy to see that, as $k \to \infty$,

For $\varphi_0 > 1/2$, $\rho^+(\varphi_0, \varphi_1) = \rho(\varphi_0, \varphi_1)$ and $\rho^-(\varphi_0, \varphi_1) = 0$.

For $\varphi_0 = 1/2$, $\rho^+(\varphi_0, \varphi_1) = \rho^-(\varphi_0, \hat{\varphi}_1) = \rho^-(\varphi_0, \varphi_1) = 0$.

For $\varphi_0 < 1/2$, $\rho^+(\varphi_0, \varphi_1) = 0$ and $\rho^-(\varphi_0, \varphi_1) = \rho(\varphi_0, \varphi_1)$.

The quantities $\rho(k, \varphi_0, \varphi_1)$, $\rho^+(k, \varphi_0, \varphi_1)$, $\rho^-(k, \varphi_0, \varphi_1)$ and $\rho(\varphi_0, \varphi_1)$ will now be interpreted in probabilistic terms, and show to be nothing but probability distributions in doubly logarithmic coordinates.

5.5. *The Notion of Multiplicative Measure and a Very Surprising Limit Theorem, a Special Case of a Theorem of Cramer*

It is useful to restate the fact that the measure $\mu(dt)$ in a cell of length 2^{-k} is the *product* of k *multipliers* equal to either m_0 or m_1. We consider the interval [0, 1] globally and construct a cascade as follows. It begins with a uniform density $\mu_0'(t) = 1$. After the first stage, let the density become $\mu_1'(t) = m(2t)$, where $m(t)$ is the following periodic function of period 2:

$$m(t) = m_0 \quad \text{if} \quad 0 < t < 1 \quad \text{and} \quad m(t) = m_1 = 1 - m_0 \quad \text{if} \quad 1 < t < 2.$$

After stage 2, $\mu_2'(t) = m(2t)m(2^2t)$.

After stage k, $\mu_k'(t) = m(2t)m(2^2t) \ldots (2^kt)$.

(The value of $\mu_k'(t)$ at the dyadic points does not matter.) This $\mu_k'(t)$ is the density of the premultifractal $m_k(dt)$.

Hence, the binomial measure is called a *multiplicative* multifractal. It is the simplest example of multiplicative fractal constructed in a grid. This property makes it exactly renormalizable.

The process of multiplication of random quantities is unfamiliar, but it is easy to transform it into a process of addition, by taking logarithms. Thus,

$$\alpha(k) = -(1/k)[\log_2 m(\eta_1) + \log_2 m(\eta_2) + \cdots + \log_2 m(\eta_k)],$$

where $m(\eta_j) = m_0$ if $\eta_j = 0$ and $m(\eta_j) = m_j$ if $\eta_j = 1$.

Now, to select at random the cell $[dt]$, means to select at random the k first digits of t. Hence we see that $\alpha(k)$ is the average of k values of a random variable equal to

$$-\log_2 m_0 \text{ with the probability } 1/2, \text{ and to}$$

$$-\log_2 m_0 \text{ with the probability } 1/2.$$

Comments concerning lognormality. In terms of the representation of $\alpha(k)$ as the average of k random variables. The results of Section 5.1 express a limit theorem of probability theory. The form of this theorem ought to surprise those readers who believe that the classical central limit theorem shows that $\log m(\eta_1) + \log m(\eta_2) + \cdots$ is asymptotically Gaussian in "every case." This would imply that the measure $\Pi M(\eta_k)$ is asymptotically lognormal. These familiar assertions are perfectly true, and they result from the application of a different renormalization. From the central limit theorem one does obtain information about the multiplicative multifractals, but this information only implies that $\rho(\alpha)$ and $f(\alpha)$ are parabolic in the central bell near their maximum. Concerning their form away from the maximum, the central limit theorem does not pretend to say anything. This is a very important issue, to which we shall return in a more general context in Sections 7 and 8.

5.6. The Two Forms of the Multifractal Representation

The term "multifractal," which is due to FRISCH and PARISI (1985), is attractive and has contributed to drawing attention to the multifractals. It is motivated by the following representation that, in turn, is often inverted and made into a definition of what a multifractal is.

"A multifractal measure can be represented as the union of a continuous infinity of addends. Each addend is an infinitesimal 'unifractal measure.' It is characterized by a single value of α, and it is supported by a fractal set having the fractal dimension $f(\alpha)$. The sets corresponding to different α's are intertwined."

The situation after a finite number $k < \infty$ of cascade stages. After $k < \infty$ stages "fractal dimension" is defined for the premultifractal multiplicative measure as in Section 4.2. It is an informal box dimension that involves a single reading of $-\log N/\log r$, and does not involve a limit. The multifractal representation follows; it is mostly useful when k is large enough to take the bite off the critique to be described in Section 5.7.

The limit for $k \to \infty$. Role of the Hausdorff-Besicovitch dimension. For $k \to \infty$, the multifractal representation involves far subtler issues, but in the present case it

was justified in advance by theorems in EGGLESTON (1949) and VOLKMAN (1958) that were available "off-the-shelf," a pleasant surprise. These theorems involve the Hausdorff-Besicovitch dimension, D_{HB}. The quantity D_{HB} had played an important early role in fractal geometry, but this role has been decreasing sharply in recent years.

The theorem in EGGLESTON (1949) covers the t's that are "normal" for the probabilities φ_0 and φ_1. These are the real values of t whose binary development $t = 0.\ \eta_1\eta_2\ldots$ is such that the frequencies of 0 and 1 in the sequence $\eta_1, \eta_2, \ldots$ converge to well-defined limits φ_0 and φ_1. Concerning the "Eggleston set" formed by these t's, the Eggleston theorem tells us that its Hausdorff-Besicovitch dimension is the familiar $\delta = \varphi_0 \log_2 \varphi_0 - \varphi_1 \log_2 \varphi_1$. Furthermore, the Hölder exponent α can simply be defined by taking a sequence of increasingly small nested intervals dt that include t, for example the sequence of dyadic intervals, and one can write

$$\alpha(t) = \lim_{dt \to 0} \frac{\log \mu([dt])}{\log dt}.$$

When t is itself dyadic, the sequence of η_j ends by an infinity of either 0's or 1's. Hence, one must define left and right α exponents, using $\log \mu([t - dt, t])$ and $\log \mu([t, t + dt])$. But this hardly matters, because dyadic t's are of fractal dimension equal to $f(\alpha) = 0$. As a matter of fact, the binomial measure's $\alpha_{\min}$ and $\alpha_{\max}$ are attained to the left and the right sides of the same dyadic values of t (except that the endpoint $t = 0$ yields only $\alpha_{\min}$ and the endpoint $t = 1$ yields only $\alpha_{\max}$).

Now, let us go beyond the Eggleston theorem, and consider the t's for which φ_0 and φ_1 fail to be defined. To account for these t's, the definition of the Hölder exponent takes its original mathematical form: one must replace $\lim_{dt \to 0}$ by

$$\alpha^{\sup}(t) = \limsup_{dt \to 0} \frac{\log \mu([dt])}{\log dt} = \limsup_{dt \to 0} [-\varphi_0(k) \log_2 m_0 - \varphi_1(k) \log_2 m_1],$$

where $\varphi_0(k)$ and $\varphi_1(k)$ are the frequencies of 0 and of 1 in the k first digits of t. To continue, we must replace $\varphi_1 = 1 - \varphi_0$ by two independent quantities $\varphi_0^{\sup}$ and $\varphi_0^{\inf}$, with

$$\varphi_0^{\sup} = \limsup_{k \to \infty} \varphi_0(k) = 1 - \varphi_1^{\inf} = 1 - \liminf_{k \to \infty} \varphi_1(k)$$

$$\varphi_0^{\inf} = \liminf_{k \to \infty} \varphi_0(k) = 1 - \varphi_1^{\sup} = 1 - \limsup_{k \to \infty} \varphi_1(k).$$

Volkmann (1958) proves that $-\Phi_0 \log_2 \Phi_0 - \Phi_1 \log_2 \Phi_1$ is the Hausdorff-Besicovitch dimension of the set of t where $\varphi_0^{\sup} = \Phi_0$, and also of the set of t where $\varphi_0^{\inf} = \Phi_0$. It follows that it is also the Hausdorff-Besicovitch dimension of the set of t where $\alpha^{\sup}(t) = \alpha = -\Phi_0 \log_2 m_0 - \Phi_1 \log_2 m_1$. Thus, Volkmann's theorem extends to *all* values of t the interpretation of $f(\alpha)$ as a dimension.

Interpretation of the two multifractal representations in terms of two different probability limit theorems. In terms of the probabilistic foundation of $f(\alpha)$ via $\rho(\alpha)$,

the above two forms of the multifractal representations are totally different in spirit. The box dimensional representation is the analog of the *weak* law of large numbers, which says that, as the size k of a sample increases,

$$\Pr\{|\text{sample average} - \text{expectation}| > \varepsilon\} \to 0.$$

The Hausdorff dimensional representation is the analog of the *strong* law of large numbers which is indeed far stronger (and harder to prove) and asserts that

$$\Pr\{\lim_{k\to\infty} \text{sample average} = \text{expectation}\} = 1.$$

Space lacks to elaborate upon this analogy.

5.7. *Convergence Towards the Limit $f(\alpha)$ is Slow, and "Sampling Biases" are Ordinarily Very Large. Therefore, $f(\alpha)$ Represents the Generating Mechanism of the Multifractal Very Indirectly*

Recall that the derivation of the function $f(\alpha)$ in Section 5.2 involves a limit argument based on the leading term in the Stirling formula for the factorial. The next question is, how fast do $\delta(k, \varphi_0, \varphi_1)$, δ^+ and δ^- converge to their limits, which we know is either $\delta(\varphi_0, \varphi_1)$ or 1?

Figure 3 compares the limit "Stirling" expression with the quantities $\delta(\alpha, \varphi_0, \varphi_1)$, δ^+ and δ^- based upon the full factorials. The value $k = 56$ corresponds to $dt = 2^{-56} \sim 10^{-17}$, which is minuscule. To study a probability distribution on such fine grids would require sample sizes well beyond any ordinary experiment. Yet, even at this stage, we see that the limit $f(\alpha)$ is far from having been reached.

Particularly visible is the value of α where the graphs of the functions δ^+ and δ^- cross each other. This is where the two tail probabilities are equal to each other, hence to 1/2. Probabilists call this the "median" value of α. Here, $\delta^+ = \delta^- = \log 2^{k-1}/\log 2^k = 1 - 1/k$. This is a special case of the extremely simple result, $\delta^+ = \delta^- = 1 + 1/\log_2(dt)$, which holds for *all* multifractals. Anyhow, whenever the distribution of α is symmetric, as it is in the binomial case, it is obvious that the median α is also the mode α (the most probable α) and the expectation $\langle\alpha\rangle$. For other multiplicative multifractals, the median, the mode and the expectation need not converge for $k \to \infty$, but a sufficient condition for convergence is that the bell of the limit distribution be asymptotically symmetric. This is the case when $\langle \log M \rangle < \infty$.

This bias elicits a subtle conceptual question. There is no question that $f(\alpha)$ does describe the process that generates the binomial measure. But what does $f(\alpha)$ really tell us about actually observable quantities, namely, those relative to the premultifractal measures that correspond to values of $k < \infty$ that are of "sensible size?"

To pinpoint the issue by contrast, recall what happens with self-similar fractals. For them, the concept of fractal dimension is also best defined by a limit. Yet, we argued (convincingly) in our books that it is concretely and practically representative and useful, because it also has a direct relevance to ordinary-sized fractals.

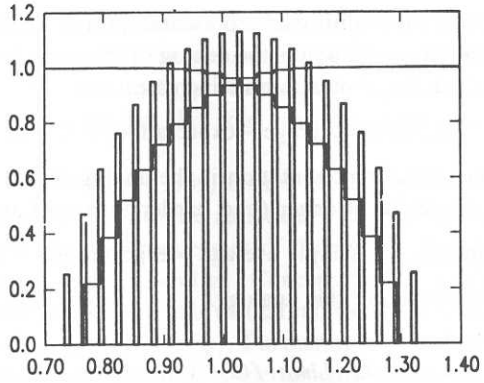

Figure 3

Three forms of the theoretical distribution of the binomial measure $\mu(dt)$. Calculations were carried out for $k = 20$, resulting in $dt = 2^{-20}$, namely, about a millionth. The three alternative distributions are plotted in "rescaled doubly logarithmic" form. That is, the horizontal is not μ itself, but $\alpha = \log \mu(dt)/\log(dt)$. The vertical bars simulate the "density" of α, i.e. (α being discrete) a probability divided by $\Delta\alpha$. The monotone increasing and decreasing step curves represent the tail probabilities $\Pr\{A < \alpha\}$ and $\Pr\{A > \alpha\}$. All three functions are plotted as (log of a function)/$\log(dt)$. For $k \to \infty$, all three curves have the same limit $\rho(\alpha)$. One has $\rho(\alpha) = 0$ and $f(\alpha)\rho(\alpha) + 1 = 1$ for the center of the diagram. Thus, the densities "overshoot" the limit, and the tails "undershoot" it. See Sections 5.7 and 6.6 for a discussion.

In a binomial multifractal (and *a fortiori* in more general ones), the situation is far harder. The $f(\alpha)$ is exclusively an asymptotic notion that is quite distant from representing the distribution of the measures over intervals of ordinary sizes dt.

The issue of the lack of direct significance of $f(\alpha)$ is especially acute in the case of turbulence, or of all other phenomena in which the fractal range has a lower cut-off that is not infinitely far, but reached by ordinary experiments. In those cases, indeed, the passage to the limit that seems to define $f(\alpha)$ has no concrete meaning. Thus, to say that the measure $\mu(dt)$ is *directly* represented by the limit would be a fiction. However, again, $f(\alpha)$ does indeed provide a useful *indirect* representation of the generating mechanism. Section 6.6 will tackle the problem of how $f(\alpha)$ can be inferred from a finite sample of data.

6. The Second Most Basic Multifractal is the Multinomial Measure. Introduction of the Legendre Multifractal formalism

6.1. Basic Background. The Domain of all Possible Points (α, δ)

To construct a multinomial of base $b > 2$ requires b masses $m_\beta (0 \le \beta \le b - 1)$ adding to 1. Denote by Φ the point whose coordinates are the frequencies φ_β

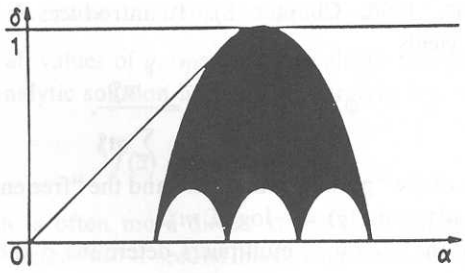

Figure 4
Rough idea of the domain of (α, δ) for a multinomial multifractal with $b = 4$. The domain's upper boundary defines the function $f(\alpha)$. Here, all the m_β are different. Clearly, $\alpha_{\min} = \min(-\log_b m_\beta) > 0$, and $\alpha_{\max} = \max(-\log_b m_\beta) < \infty$.

of the digits β in the base-b development $0, \eta_1 \eta_2 \ldots \eta_k$. Every b-adic interval characterized by Φ yields

$$\mu(dt) = \Pi m_\beta^{k\varphi_\beta}, \text{ hence } \alpha = -\sum \varphi_\beta \log_b m_\beta \quad \text{and} \quad \delta = -\sum \varphi_\beta \log_b \varphi_\beta.$$

In the binomial case, a single $\delta = f(\alpha)$ could be deduced from the value of α, but this possibility is not available here. A given α, indeed, allows a host of possible sets of values of φ_c constrained by $\Sigma \varphi_c = 1$. Each of these sets has its own δ, and the multinomial representation would not be a sum over a sample index α, but a sum over two indexes α and δ. More precisely, the Φ's yielding a given α define a portion of hyperplane, on which the function δ varies continuously. Therefore, the possible values of δ for given α cover an interval. After the points (α, δ) corresponding to all the values of α have been combined, the result is a domain of the plane, as shown in black in Figure 4.

This domain is bounded *to the left* by $\alpha_{\min} = \min\{-\log m_\beta\}$, and *to the right* by $\alpha_{\max} = \max\{-\log m_\beta\}$. It is bounded from *below* by $\delta \geq 0$. In fact, it has as its more precise lower boundary a certain scalloped curve. For example, when no two m_β are identical, (including $\alpha = \alpha_{\min}$ and $\alpha = \alpha_{\max}$), $\delta_{\min} = 0$ when $\alpha = \alpha_\beta$, and between two successive α_β's, $\delta_{\min}$ follows an entropy arch of height $\log_b 2$. When several values of β yield the same α_β, the arches on both sides of α_β are rescaled by an affinity. For example, if $\alpha_{\min}$ is attained for N_β distinct β's, $\delta_{\min}(\alpha_{\min}) = \log_b N_\beta$.

6.2. *Thermodynamical Manipulations. The Lagrange Multipliers Argument, and the Legendre Relation of the Gibbs Theory*

For a given value of α, the δ's are dominated by the term whose dimension is the highest. This is the term that maximizes $-\Sigma \varphi_\beta \log_b \varphi_\beta$, given $-\Sigma \varphi_\beta \log_b m_\beta = \alpha$, and $\Sigma \varphi_\beta = 1$. To solve this problem, we use the classical method of Lagrange

multipliers (HUANG, 1966, Chapter 8). It introduces a multiplier q, with $-\infty < q < \infty$, and yields

$$\varphi_\beta = \frac{b^{q \log_b m_\beta}}{\sum b^{q \log_b m_\beta}} = \frac{m_\beta^q}{\sum m_\beta^q}.$$

The customary roles of the "partition function" and the "free energy" are played here by the quantities Σm_β^q and $\tau(q) = -\log_b \Sigma m_\beta^q$.

In terms of $\tau(q)$, the Lagrange multipliers determine q and $f(\alpha)$ from α by

$$\alpha = -\sum \varphi_\beta \log_b m_\beta = -\frac{\partial}{\partial q} \log_b \sum m_\beta^q \frac{\delta \tau(q)}{\delta q}$$

$$\max \delta = f(\alpha) = -\frac{\sum \left(q \log_b m_\beta - \log_b \sum m_\beta^q \right) m_\beta^q}{\sum m_\beta^q}.$$

That is,

$$\alpha = \frac{\partial \tau(q)}{\partial q} \quad \text{and} \quad f(\alpha) = q \frac{\partial \tau}{\partial q} - \tau = q\alpha - \tau.$$

The black domain in Figure 4 is now replaced by its upper boundary, which is the graph of a function $f(\alpha)$. Since in the multinomial case the black domain of α and δ satisfies $\alpha > 0$ and $\delta \geq 0$, we see that $f(\alpha) \geq 0$, $\alpha_{\min} > 0$, $f(\alpha_{\min}) \geq 0$ and $f'(\alpha_{\min}) = \infty$, and $\alpha_{\max} < \infty$, $f(\alpha_{\max}) \geq 0$ and $f'(\alpha_{\max}) = -\infty$. Multifractals that are not multinomial, yet possess these properties we call "pseudo-multinomial."

Note, incidentally, that

$$\varphi_\beta = -\frac{1}{q} \frac{\partial \tau}{\partial (\log_b m_\beta)} = \log_b \frac{m_\beta}{q} \frac{\partial \tau}{\partial m_\beta}.$$

The moments of μ. When a cell is selected at random, the multiplier is a random variable M that takes the value m_β with the probability $1/b$. Its q-th moment is $\langle M^q \rangle = \Sigma (1/b) m_\beta^q$. The quantity $\log_b \langle M^q \rangle = -1 - \tau(q)$ in a common way of specifying a random variable in mathematical statistics, and it is called "cumulant generating function." (Note: The sign attributed to τ in the usual multifractal notation is unfortunate, and a number of authors, e.g., FEDER (1988), spontaneously select the opposite sign.) Similarly,

$$\langle \mu^q(dt) \rangle = [\langle M^q \rangle]^k = \left[b^{-1} \sum m_\beta^q \right]^k = [b^{-1-\tau(q)}]^k = (dt)^{1+\tau(q)}.$$

Adding the contributions μ_i of the $(dt)^{-1}$ intervals of length dt, we find

$$\left\langle \sum \mu_i^q \right\rangle = (dt)^{-1} \langle \mu^q(dt) \rangle = (dt)^{\tau(q)}.$$

In the approach of HALSEY *et al.* (1986), this is the *definition* of $\tau(q)$.

6.3. *Inverse Legendre Relations, Analytic and Graphical*

Knowing $\tau(q)$ for all values of q, one can trace all the straight lines of equation $\delta_q(\alpha) = q\alpha - \tau$. The analytic solution is that these straight lines define $f(\alpha)$ as their envelope, namely

$$f(\alpha) = \min_q (q\alpha - \tau).$$

A graphical approach is often more direct. If the lines represented by $\delta_q(\alpha)$ are traced in black, they merge into a second black domain in the (α, δ) plane, which "surrounds" the black domain of $\{\alpha, \delta\}$ that we have considered previously.

6.4. *Thermodynamical Analogies*

If the present approach is used, the observations that, $q \sim$inverse temperature, $\tau \sim$Gibbs free energy, and $f \sim$entropy are not after-the-fact "thermodynamic reinterpretations" of multifractals. They are embedded in the multifractals' origin in the Lagrange multipliers. Further use of the thermodynamic interpretation is beyond the scope of this paper.

6.5. *The "Darwin Fowler" Derivation of the "Multifractal Formalism" and Actual Computation*

FRISCH and PARISI (1985) and HALSEY *et al.* (1986) obtain the same equations $\alpha = \tau'$ and $f = q\alpha - \tau$ via a steepest-descent argument. Experts who have learned statistical mechanics in older books recall that Darwin and Fowler have used steepest descents to justify rigorously the Lagrange multipliers procedure. (See also HUANG, 1966, Section 10.1.) However, *no one* teaching thermodynamics will start with the Darwin–Fowler method, without having first presented the Lagrange multipliers. Therefore, the path towards the formalism taken in Section 6.2. involves the least effort and the fullest understanding. Section 7 describes the next simplest generalization.

6.6. *Estimation of $f(\alpha)$. Statistical Corrections for the Small Sample Bias*

To deduce $f(\alpha)$ from a finite sample of data can be done in many different ways. We shall give one "quick and dirty" method, then several examples of inference from an observable fractal range to an inaccessible limit.

A quick and dirty correction from a single histogram. Section 5.7 shows that the convergence to the limit $\delta(\varphi_0, \varphi_1)$ is much faster for the tail functions $\delta^+(k, \varphi_0, \varphi_1)$ and $\delta^-(k, \varphi_0, \varphi_1)$ than it is for the density functions $\delta(k, \varphi_0, \varphi_1)$. Moreover, we already know that the value of the bias at the median where $\delta^+ = \delta^-$ is "universal", in the sense that bias $= 1/\log_2(1/dt)$, independently of $f(\alpha)$. Assuming that the bias is the same for all α, which is the binomial case, suggests the following

"Recipe": Suppose that $\mu(dt)$ is known down to cells of length $2^{-k_{\max}}$. Evaluate the sample tail function $\delta^*(k_{\max})$. An estimate of $\delta(\alpha)$ is provided by $^*(k_{\max}) + 1/k_{\max}$.

The moments and Legendre transform route to the estimation of $f(\alpha)$ *and* $\zeta(\alpha)$. The preceding procedure is extremely crude, but there are several alternatives. Sections 6.2 and 6.3 suggest one alternative, which has become widely used. First, estimate $\tau(q)$ from the data; next, compute $f(\alpha)$ as the Legendre transform of the estimated $\tau(q)$. This path is deeply imbedded in the approach of Frisch and Parisi and of Halsey *et al.*, since their approach *defines* $f(\alpha)$ in this fashion.

However, the path via $\tau(q)$ *is not unique*. Let us now describe the principle underlying this and many other estimation procedures, then list alternative implementations.

How to infer $f(\alpha)$ *from data within an observable fractal range*. Suppose that [0, 1] has been divided into $2^{k_{\max}}$ minimal intervals of length $2^{-k_{\max}}$. Then the measure $\mu(dt)$ is of course known for each of these intervals. But for every $k < k_{\max}$, $\mu(dt)$ is also known for 2^k intervals of length 2^{-k}. The values of $\mu(dt)$ carry considerable amounts of information. This information can be processed by evaluating $\tau(q)$, but have been a vocal advocate of alternative methods.

Two such methods are found in MENEVEAU and SREENIVASAN (1989) and in CHHABRA and JENSEN (1989). There is room for many more.

6.7. *Generalization of the Scope of the Binomial and Multinomial Multifractals Defined on [0, 1]*

The preceding discussion is of wider generality than has been suggested thus far. Many physicists seem to be under the impression that multifractal measures are always supported by fractal sets. The discussion thus far has shown, to the contrary, that the simplest multifractals are those supported by [0, 1]. Now, we proceed to the easy task of demonstrating that, when the bases are matched in a way to be described, a multinomial measure on [0, 1] maps upon one that is supported by a cube or by a fractal set constructed recursively.

Mapping a fractal on [0, 1]. Consider a fractal set that is constructed recursively using a generator made of N "sticks", which are intervals of equal lengths. The best known examples are the generalized Koch curves (FGN, p. 142 shows that this family also includes the Sierpinski gasket). Recursion defines a point P on such a fractal as being on stick number η_z in the first recursion stage, . . . stick number η_k in the k-th stage. Together, the integers η_k define a real number in the base N, namely $t = 0.\eta_1, \ldots, \eta_k, \ldots$, with $0 \le t \le 1$. In this way, every interval $[dt]$ of [0, 1] is mapped on a piece of our fractal, which *FGN* calls a "fractal interval."

Mapping a multifractal. Now distribute on [0, 1] an arbitrary multinomial measure of base N. The above map transforms it automatically into a multinomial measure on our fractal.

Double points. It is known, of course, that some P's are characterized by several

values of t. Fortunately, however, the multifractal measures are continuous (though not differentiable). When the P's for which t is ambiguous are denumerable, these ambiguous t's do not matter. (There are several sources of ambiguity. The first is that if t is a multiple of N^{-k}, it can be represented as ending in an infinity of 0's or of $(N-1)$'s. A second source of ambiguity follows from double points in the Koch fractal; for example, the Sierpinski gasket has asymptotic double points.)

Cubes. A Peano motion (our term to replace the improper term, Peano curve) can, similarly, be used to map a given multinomial measure on [0, 1] upon a cube $[0, 1]^E$. The only condition is that the multinomial's base N be such that $N^{1/E}$ is an integer.

Effect of mapping upon α, $f(\alpha)$ *and* $\rho(\alpha)$. This effect is, most fortunately, very simple and even amusing. To fit a widely used notation, denote our fractal's dimension by D_0. Denote by $|dP|$ the Euclidean distance between the endpoints of the map dP of the linear interval dt. It is easy to see that $|dP| = |dt|^{1/D_0}$.

Therefore for the measure μ' on the map of [0, 1], α is replaced by

$$\alpha' = \log \mu'(dP)/\log |dP| = D_0 \log \mu(dt)/\log(dt) = D_0\alpha.$$

Moreover, our map multiplies the fractal dimension of a set in [0, 1] by D_0. Hence, the f-function is replaced by $f' = fD_0$. Thus,

$$f'(\alpha')/D_0 = f(\alpha'/D_0).$$

When $f'(\alpha')$ is plotted in the "reduced" coordinate states, α'/D_0 and f'/D_0, the shape of its graph is independent of D_0.

7. *The Random 1974 Multifractal Measures, Part I. The Conservative Case and the Introduction of "Latent" α's, with Negative f(α)'s*

Now we proceed to the first of two fully studied cases of the exactly renormalized multifractals. They are the "1974 multiplicative multifractals" introduced in MANDELBROT (1974). The reader who has access to MANDELBROT (1989) will find it fruitful to study that paper at this point.

7.1. *The Conservative 1974 Measures on [0, 1]*

First, let us modify the multinomial measure as follows: The indexes of the masses m_β are shuffled at random before each stage of the cascade that distributes the mass within each of the k-th level intervals into masses within the $(k+1)$st level intervals. (Note that one must *not* shuffle mass between *all* the k-th level intervals!) The resulting "shuffled multinomials," are the next exactly renormalizable stage beyond the multinomials. Observe that shuffling does not modify $\rho(\alpha)$, hence does not modify $f(\alpha)$.

The 1974 multifractals enter logically as the next stage after the shuffled multinomials. They require a base $b \geq 2$ and a multiplier that is not limited to the multinomial's prescribed finite collection of m_β's, but may be a more general random variable. Since it may depend on β, it will be denoted by $M_\beta \geq 0$.

Since mass is conserved while it is spread around within a cell, strong "conservation" rules constrain the multiplier. For example, the b random multipliers $M_\beta (0 \leq \beta \leq b-1)$ must add to 1. Consider the representative point whose coordinates are these b multipliers M_β. This is a random point restricted to an "admissible set," namely, to the portion of b-dimensional space defined by $M_\beta \geq 0$ and $\Sigma M_\beta = 1$. When a cell is chosen at random, the mass is multiplied by one of the M_β, chosen with equal probabilities $1/b$. Therefore, the average multiplier M is the mixture of the b coordinates of the representative point. Clearly, $0 \leq M \leq 1$ and $\langle M \rangle = b^{-1}$. The conservative 1974 measures *can, but need not*, add the further assumption that the coordinates M_β are identically distributed. If they are, M is the same random variable as any of the coordinates M_β.

By a repetition of this scheme, the b-adic cell of length b^{-k}, starting at $t = 0, \eta_1 \eta_2 \ldots \eta_k$ is given the mass

$$M(\eta_1) M(\eta_1, \eta_2) \ldots M(\eta_1, \ldots, \eta_k) = \prod M.$$

Here, the successive random multipliers M for given t are identically distributed and independent. Our b-adic cell, therefore, yields

$$\alpha = -(1/k)[\log_b M(\eta_1) + \log_b M(\eta_1, \eta_2) \ldots].$$

Thus, α is the average of k independent random variables. Let us now show that, for the desired α's, its limit distribution for $k \to \infty$ is far from "intuitive."

7.2. *Five Examples of Sums of Independent Random Variables, for which the Probability Distribution is Known in Analytically Closed Form for Both the Individual Addends, and for their Sum or Average*

To obtain the limit distribution of α for $k \to \infty$, the limit theorems known to every physicist (law of large numbers and central limit theorem) give results without interest. One needs, instead, the general results to be described in Section 7.3. In order to motivate these theorems, let us begin by special random multipliers M that allow an analytically explicit evaluation of the distribution of α.

Originally, these examples were singled out as being among the few known examples with the desired property. After the fact, however, the first three examples have proven to be of intrinsic interest. Example A) illustrates Section 5. Example B) illustrates the new feature that enters in this Section 7, and Example C) illustrates the new feature that enters in Section 8. Our random variables, whether discrete or

continuous, will stand for $-\log_b M$ or for $-\Sigma \log_b M$. To avoid unwieldy formulas, however, they will be denoted by X, and their values by x. In each case, we denote the density by $p(x)$ for each addend, by $p_k(x)$ for the sum of k addends, by $kp_k(k\alpha)$ for the average α of k addends. We also consider the quantity $(1/k)\log_b [kp_k(k\alpha)]$, which (in effect) was already used in Section 5.3.

A) *The binomial distribution.* Section 5.3 implies that $(1/k)\log_2 [kp_k(k\alpha)] \to \rho(\alpha) = f(\alpha) - 1$, where $f(\alpha)$ is the "entropy" function described in Section 4.2.

The striking features of this example are that $\alpha_{\min} > 0$ with $f'(\alpha_{\min}) = \infty$ and $\alpha_{\min} < \infty$ with $f'(\alpha_{\min}) = -\infty$.

B) *The Gamma distribution.* When the parameter is γ, its density is $p(x) = x^{\gamma-1}e^{-x}/\Gamma(\gamma)$. The sum of two Gamma random variables of respective parameters γ' and γ'' is a Gamma random variable of parameter $\gamma' + \gamma''$. Therefore, $p_k(x) = x^{k\gamma-1}e^{-x}/\Gamma(k\gamma)$. By the Stirling approximation, $(1/k)\log_e [kp_k(k\alpha)] \to \gamma \log(\alpha/\gamma) - \alpha + \gamma = \rho_e(\alpha)$.

Let us show that this example is encountered for the conservative 1974 multifractals when $b = 2$, so that the admissible set for the multipliers M_0 and M_1 is the interval [0, 1]. If both multipliers are distributed over this admissible set with uniform density, one has $\Pr\{M > m\} = 1 - m$. Writing $Y = -\log M$ yields $\Pr\{Y > y\} = e^{-y}$. This exponential is simply the Gamma distribution corresponding to $\gamma = 1$. It leads to $\rho(\alpha) = \log_2(\alpha \log_e 2) - \alpha + \log_2 e$.

The striking features of this example are that $\alpha_{\min} = 0$ and $\alpha_{\max} = \infty$, and that $\rho(\alpha_{\min}) = \rho(\alpha_{\max}) = -\infty$. These properties could not happen in the multinomial case, and serve in Section 7.4 to introduce the notion of *latent* α.

The Gamma distribution and related examples are discussed in detail (with several illustrations) in MANDELBROT (1989), which–again–the reader would benefit from at this point.

C) *The Gaussian distribution* $p(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/2\sigma^2)$. Here, $[kp_k(k\alpha)]^{1/k} = (2\pi\sigma^2/k)^{-1/(2k)}\exp(-\alpha^2/2\sigma^2)$. Thus, $(1/k)\log [kp_k(k\alpha)] \to -\alpha^2/2\sigma^2 = \rho(\alpha)$.

Computationally, this example is the simplest of the present five. But it raises novel issues, which is why it has not been listed first. The Gaussian has two unbounded tails, while the conservative model requires $M < 1$. From $M_{\max} > 1$ follows $\alpha_{\min} = -\infty$, which will serve in Section 8 to introduce the notion of *virtual* α.

D) *The Poisson distribution.* When $\langle X \rangle = \gamma$, one deals with the discrete probability $p(x) = e^{-\gamma}\gamma^x/x!$. Here $p_k(x) = e^{-k\gamma}(k\gamma)^x/x!$ By the Stirling approximation,

$$(1/k)\log_e [kp_k(k\alpha)] \to -\gamma + \alpha\log_e(\gamma e/\alpha) = \rho_e(\alpha).$$

E) *The Cauchy distribution.* Its density is, of course, $p(x) = [\pi(1 + x^2)]^{-1}$. By a property called "Lévy stability," one has $p_k(x) = [\pi k(1 + x^2/k^2)]^{-1}$, hence $\rho_e(\alpha) = 0$. This result means the Cauchy case falls quite outside the pattern found in cases A) to D). The renormalization that it requires is altogether different. But the random variable $M = \exp$ (Cauchy variable) satisfies $\langle M^q \rangle = \infty$ for all q. Therefore, it is impossible to use it as a multiplier satisfying conservation on the average.

Features common to examples A) to D). 1) To form the quantity $(1/k)\log p_k(k\alpha)$ is to renormalize the average in a way that greatly accentuates the low probability "tails," and deemphasizes the central "bell." In examples, we have found that the newly renormalized average still converges to a limit, for which $\rho(\alpha) \leq 0$, with $\rho(\alpha) = 0$ for one value of α. Namely, for the expected value $\langle\alpha\rangle$ of α.

In examples B), C) and D), the analytic form of $\exp[\rho(\alpha)]$ is nearly identical to that of $p(\alpha)$, except that it is "wrongly normalized." It does not integrate to 1, and would not be proper probability density. A consequence of nonintegrability is that, before the asymptotic range is reached, each of the above quantities $(1/k)\log p_k(k\alpha)$ is >0 over a range of α's near the maximum of $\rho(\alpha)$. That is, as in the binomial case, direct estimates from data collected for a finite k overestimate $\rho(\alpha)$ near its maximum.

Differences between the five examples. Gross nonuniversality. All the above limits differ from one another. For example, the limits for the non-Gaussian cases differ from the limit relative to the Gaussian case. In Section 7.3, this surprising result will be shown to hold very generally. It is the second most important probabilistic fact underlying the multifractals (the most important fact being the behavior of moments along low dimensional cuts; Section 8).

7.3. *Generalization of the Multifractal Formalism by Application of the Cramèr Limit Theorem. Profound Nonuniversality*

The above examples are very special cases of the "large deviations theorems" of H. Cramèr; see BOOK (1984), CHERNOFF (1952), DANIELS (1954, 1987). That the proper mathematical result is available "off-the-shelf" is a pleasant surprise. The "local form" of the Cramèr theorem asserts the following:

> As $k \to \infty$, the quantity $(1/k)\log_B$(probability density of α) converges to a limit, to be denoted as $\rho(\alpha)$.

The tail quantities $(1/k)\log_B$ (probability of $\alpha > \langle\alpha\rangle$) and $(1/k)\log$ (probability of $\alpha < \langle\alpha\rangle$) converges to the same limit. It is easily shown that

$$f(\alpha) = \rho(\alpha) + 1 = \rho(\alpha) + \text{dimension of the measure's support.}$$

The Legendre formalism. In this multinomial case, $\tau(q)$ could be written either as $-\log_b \Sigma m_\beta^q$ or as $-1 - \log_b\langle M^q\rangle$. This second form remains meaningful for all random variables M. With this definition of $\tau(q)$, the generalized Gibbs formalism resulting from the Cramèr theorem remains valid. That is, the relation between $\tau(q)$ and $f(\alpha)$ follows the Legendre rules we have encountered in Section 5 for the multinomial measures.

Failure of universality. Moreover, there is *no* universality, in the specific narrow sense exemplified in Section 7.2: different multipliers yield different $f(\alpha)$'s, and conversely.

Obviously, Cramèr-type theorems extend to the case when the factors M are weakly dependent or weakly nonidentical. See AZENCOTT *et al.* (1980) and DAWNHA-CASTELLE (1979) for various generalizations.

7.4. Negative $f(\alpha)$'s, and the Distinction Between Manifest and Latent Values of α

Negative $f(\alpha)$. The 1974 multifractals introduce a very important theme. First, suppose that M is a discrete random multiplier and that its largest possible value $M_{\max}$ satisfies $\Pr\{M = M_{\max}\} < b^{-1}$, which is perfectly allowable. It follows that $f(\alpha_{\min}) = \log_b [\Pr\{M_{\max}\}] + 1 < 0$. As a matter of fact, $f(\alpha_{\min})$ can be arbitrarily large in absolute value, and negative. Similarly, $\Pr\{M = M_{\min}\} < b^{-1}$ results in $f(\alpha_{\max}) < 0$.

Secondly, the example of the Gamma distribution suffices to show that the 1974 multifractals allow $\alpha_{\min} = 0$ and $\alpha_{\max} = \infty$, with $f(\alpha_{\min}) = -\infty$ and/or $f(\alpha_{\max}) = -\infty$. To show that this result is intuitive, take a more general M that is continuously distributed, but can be approximated arbitrarily close by a discrete multiplier with $\log M_{\max}$ arbitrarily close to 0 and/or $\log M_{\min}$ arbitrarily close to $-\infty$, with both extremes having correspondingly arbitrary tiny probabilities. Then $\alpha_{\min} = 0$ and/or $\alpha_{\max} = \infty$.

A new and important distinction. The α'*s* such that $f(\alpha) > 0$ will be called *manifest*. The remaining α's will be called *latent*, which means "hidden but present." (For negative latent dimension in a different context, see MANDELBROT (1984).)

In the special case $\Pr\{M_{\max}\} \geq b^{-1}$ and $\Pr\{M_{\min}\} \geq b^{-1}$, all α's are manifest. Such a multifractal can be called "pseudo-multinomial."

Crossover from the manifest to the latent α's. The existence of latent α's is characterized by crossover values $\alpha^*_{\min}$ and $\alpha^*_{\max}$, such that $f(\alpha^*_{\min}) = 0$ and $f'(\alpha^*_{\min}) = q^*_{\max} < \infty$ and/or $f(\alpha^*_{\max}) = 0$ and $f'(\alpha^*_{\max}) = q^*_{\min} > -\infty$.

The effect of latent α's on the moments. Let us call "manifest $f(\alpha)$" the function $f(\alpha)$ restricted to manifest values of α. Its slopes, $q^*_{\max}$ for $\alpha_{\min}$ and $q^*_{\min}$ for $\alpha_{\max}$ fail to be infinite. Therefore, the manifest $f(\alpha)$ differs from the full $f(\alpha)$ function when there are latent α's. How much do the latent α's contribute to $\tau(q)$ and D_q? For $q^*_{\min} < q < q^*_{\max}$, they contribute nothing. But for large positive and/or negative q, their contribution is predominant. Therefore, the estimates of $\tau(q)$ or D_q made on the basis of manifest α's is truncated and biased.

Note that the "full" functions $\tau(q)$ and $D_q = \tau(q)/(q-1)$, evaluated on the basis of all α's are analytic functions. But the "manifest" $\tau(q)$ or D_q evaluated on the basis of manifest α's is incorrect (truncated and biased) for large $|q|$'s.

There is a seeming paradox here. On the one hand, the probability outside of the central bell tends to 0 as $k \to \infty$, meaning that the tails become thoroughly

insignificant. In the limit $k \to \infty$, the most probable value, the expectation and the other usual parameter of location all converge with each other. On the other hand, those "negligibly" few values in the tails are so huge that their contributions to all the moments of order $q \neq 0$, and to $\tau(q)$, are predominant. Moreover, the moments and $\tau(q)$ depend on the exact form of $f(\alpha)$, that is, are nonuniversal.

The meaning of negative values for the "dimension" $f(\alpha)$. When $f(\alpha)$ is viewed as a fractal dimension, the notion that $f(\alpha) < 0$ would be self-contradictory, that is, would be impossible. This self-contradiction does not deny that $\rho(\alpha)$ is useful. It only expresses the fact that $f(\alpha) = \rho(\alpha) + 1$ cannot always be interpreted as a fractal dimension. As a matter of fact, several authors have approached this issue from different angles, and have shown that restricting multifractals to $f(\alpha) \geq 0$ lead to self-contradiction (CATES and DEUTSCH, 1987), or is otherwise not acceptable (FOURCADE *et al.*, FOURCADE and TREMBLAY, 1987).

From the b^k data given by a single sample of measure, one can estimate $f(\alpha)$ if the probability of α is $> b^{-k}$, which means that $f(\alpha) \geq 0$. The range of the α's *is not* expanded by increasing k, because the population from which the sample is drawn also depends upon k. This invalidates the intuitive belief that the sample value necessarily converges with the population expectation as the sample increases.

In order to measure the negative $f(\alpha)$ for latent α's, the only generally valid method is to combine data from very many samples.

A remark linking Sections 7 and 8. Very often, negative $f(\alpha)$'*s* have a different cause, to be described in Section 8. They appear because one considers low-dimensional cuts through higher-dimensional measures. In that case, one extends the ranges of observable α's by investigating the original measure through higher-dimensional cuts. The reason is that "generically" one should expect the overall original measure to include cells of higher or lower density, "hotter" or "colder", than the hottest or coldest cells along the one-dimensional cut.

8. The Random 1974 Multifractal Measures, Part II. The Canonical Case and the Properties of Random Cuts. Introduction of the Critical Exponent q_{crit} and of its Inverse Function D_q

8.1. The Canonical 1974 Measures on [0. 1] and Random Cuts

In this next random generalization of the binomial multifractal, the multipliers $M \geq 0$ are allowed to be statistically independent. This assumption implies, of course, that (almost surely) the cascade *does not conserve mass*. However, it is easy to insure that *mass is conserved on the average*. It suffices that the condition $\langle M \rangle = 1/b$, which used to be a corollary of conservation, be preserved as an autonomous requirement. The assumption of independence is motivated by an issue already discussed in Section 1.4. Many multifactor measures can only be observed

through cuts. To illustrate, we imagine a conservative cascade of cubic cells of dimension E. Each cubic cell contains b^E cubic subcells, but a straight edge of the initial cube intersects only b cells. If $b^E \gg b$, the conservation laws $\Sigma M(\eta_1) = 1$ hardly constrain the multipliers M within these subcells. These M can be assumed independent.

Independence of the multipliers brings essential new features.

8.2. *The Random Variable* Ω

In order to understand the structure of the cuts and the role of the "critical dimension" D_q, it is necessary to introduce a random quantity to be denoted by Ω (MANDELBROT, 1974). Consider a canonical cascade that has started with mass 1 in [0, 1], and has continued over infinitely many stages. Because of the lack of conservation the ultimate mass is not identical to 1, but is a random variable. It will be denoted by Ω.

An alternative way to write Ω is to add the masses in the b subcells of length $1/b$. In the first cascade stage, the β-th interval is given the mass $M(\beta)$. Ultimately, it contains the mass $M(\beta)\Omega(\beta)$, the various quantities $M(\beta)$ and $\Omega(\beta)$ being statistically independent. The sum of the ultimate masses can be written as either $\Sigma M(\beta)\Omega(\beta)$ or Ω. This identity can be written as

$$\sum M_\beta \Omega_\beta \equiv \Omega \quad (\equiv \text{means "identical in distribution"}).$$

Thus Ω is the fixed point of the operation of randomly weighted averaging using as weights the random quantities $M(\beta)$. This Ω, first introduced in MANDELBROT (1974), has been the object of several papers of pure probabilists.

8.3. *The Mass* $\mu(dt)$ *in a Cell of Length* b^{-k}*: Low- and High-Frequency Factors*

Let the cell start at $t = 0$, $\eta_1\eta_2 \ldots \eta_k$. The first k-stages of the cascade can be called *low frequency* stages. They generate precisely the same effects as in the conservative cascade of Section 7.1. That is, they yield the multiplicative term

$$M(\eta_1)M(\eta_1, \eta_2) \ldots M(\eta_1, \ldots \eta_k) = \prod M.$$

This ΠM is to be called a *low frequency* (or *long wavelength*) *factor*, because it is generated by multipliers that are constant over b-adic intervals of length $dt = b^{-k}$ or shorter.

However, in contrast to the conservative case, that is not all! Each cell is also subjected to precisely the same process as has been described for [0, 1]. Therefore, the measure does *not* reduce to ΠM, but instead takes the form

$$\mu(dt) = \Omega(\eta_1, \ldots \eta_k)M(\eta_1)M(\eta_1, \eta_2) \ldots M(\eta_1, \ldots \eta_k) = \Omega \prod M.$$

That is, each cell involves a sample value of Ω, to be called a *high frequency* (or *short wavelength*) *factor*, which results from multipliers that vary over intervals that are shorter than $dt = b^{-k}$. Successive M and Ω are independent, and all the M (or Ω) are identically distributed.

8.4. The Hölder α and the Function $f(\alpha)$. Virtual Values of α

Write $\alpha = \alpha_L + \alpha_H$, with

$$\alpha_L = -(1/k)[\log_b M(\eta_1) + \log_b M(\eta_1, \eta_2) \ldots]; \quad \alpha_H = -(1/k) \log_b \Omega.$$

When $k \to \infty$, $\alpha_H = -(1/k) \log_b \Omega \to 0$, hence α_H does not affect $f(\alpha)$. That is, on the logarithmic scale of the Hölder α, the high frequency factor Ω is asymptotically negligible. Moreover, the term α_L continues to be covered by the theorem of Cramèr, and the "Gibbs" quantities $\tau(q)$ and $f(\alpha)$ and D_q are precisely as for the conservative 1974 measures. However, a canonical 1974 cascade allows *two* possibilities that have not been seen in the preceding Sections, namely virtual α's and misbehaving $\tau(\alpha)$.

Virtual α's. First possibility: a canonical cascade allows $M_{\max} > 1$, hence $(\alpha_L)_{\min} < 0$, with $f[(\alpha_L)_{\min}] < 0$. If we forget Ω, our α_L would be a Hölder exponent, and $\alpha_L < 0$ would be absurd, because it would imply that a subinterval b^{-k-1} includes more mass than its embedding interval b^{-k}. But α_L is *not quite* a Hölder exponent. The paradox of the $\alpha_L < 0$, to be called *virtual*, can be shown to vanish if one considers Ω.

We shall see that, either both factors ΠM and Ω are "well behaved," or both "misbehave" simultaneously, each in its own way.

8.5. The Prefactors $\langle \Omega^q \rangle$ Are Infinite when q is above the Critical Exponent q_{crit}

We have seen that Ω is negligible in the scale of the α. But in the scale of μ, taking account of Ω changes the familiar $\langle \mu^2(dt) \rangle = (dx)^{\tau(q)+1}$ into $\langle \mu^q(dt) \rangle = \langle \Omega^q \rangle (dt)^{\tau(q)+1}$. The new moment prefactors $\langle \Omega^q \rangle$ reflect high frequency effects due to scales $< b^{-k}$. They have far reaching effects, because *they need not be finite.* Indeed, MANDELBROT (1974) shows (and KAHANE and PEYRIÈRE, 1976 then proves rigorously) the following.

The inequality $\langle \Omega^q \rangle < \infty$ holds if, and only if, $q < q_{\text{crit}}$, where q_{crit} is the solution of the equation $\tau(q) = D_q = 0$. When $q > q_{\text{crit}}$, then $\tau(q) < 0$ and $D_q < 0$.

Without the presence of divergent prefactor $\langle \Omega^q \rangle$, the inequality $\tau(q) < 0$ for $q > q_{\text{crit}}$ would have meant that $\langle \mu^h(dt) \rangle$ actually *decreases as dt* increases. This would be an extreme anomaly, but it does *not* occur, because of $\langle \Omega^q \rangle = \infty$.

Furthermore, the paradox of *virtual* α's can be shown to vanish if one considers the moments of Ω.

8.6. *The Tail Behavior of* Ω

MANDELBROT (1974) has conjectured, and GUIVARC'H (1987) has proven, that $\Pr(\Omega > \omega) \sim \omega^{-q_{\text{crit}}}$. This is the simplest tail behavior compatible with the role q_{crit} plays for Ω. Since q_{crit} plays the same role for $\mu(dt)$, the same tail behavior also holds for $\mu(dt)$. More precisely, such is the case if dt is well above the inner cutoff (if there is a cutoff > 0).

This prediction of the model is extremely important to the analysis of experimental data. But it should be stressed that it only holds asymptotically after the cascade has been allowed to proceed down to scales well below dt. In the case of turbulence, this requires an extremely high Reynolds number R, and indeed there are reports of ill-behaved sample moments of high q. But for the lower R reached in laboratory turbulence, Ω is effectively truncated at a level increasing with R, which avoids divergent moments. Experimentalists are urged to compare the dissipation's $\tau(q)$ for different R's.

8.7. *Sequences of Cuts with Different Bases, but Identical W Multipliers*

To renormalize once for all bases, it suffices to consider a cascade that multiplies the densities (instead of the masses) by a fixed *weight* $W = bM$, satisfying $\langle W \rangle = 1$. When 1974 measure in E-dimensions, with weight W and base b, is intersected by a fractal of dimension $D = \log B / \log b < E$, it yields a measure with the same W and the new base B. In fact, by using random cuts, one may view B as continuously variable. One has $-\tau(q) = 1 + \log_b \langle M^q \rangle = \log_b \langle W^q \rangle - (q-1)$. This is precisely the function that MANDELBROT (1974) plots (his Fig. 2) and discusses for different values of b. (The same paper also shows that the critical q_{crit}, which it denotes by α, depends on D.)

Now consider for different B's or D's the enhanced graphs showing $f(\alpha)$, $\tau(q)$ and D_q (as in Figure 5), and translate the origin of each graph to the point of coordinates $\alpha = \langle \alpha \rangle$ and $f(\alpha) = D_0$. This is the proper "renormalization" to use here, because one can show that the translated graphs *coincide* if one includes the manifest, the latent and the virtual α's. As a corollary, increasing B eventually changes latent α's into manifest α's. That is, it changes negative $f(\alpha)$'s into positive $f(\alpha)$'s, which one can estimate through the B^k data from a single sample. Also, increasing B eventually changes virtual α's into latent, then into manifest ones. Decreasing B eventually reaches $D = D_0$, beyond which the measure becomes degenerate (MANDELBROT, 1974; KAHANE and PEYRIÈRE, 1976).

As corollary, the manifest portion of $f(\alpha)$ fails to represent a 1974 multifractal in an intrinsic fashion. But a full description of the effects of the same W in different D_0's is provided in the context of Cramèr's limit theorem by the intrinsic quantities $\alpha - \langle \alpha \rangle$, $\rho(\alpha)$, and $D_q - D_0$.

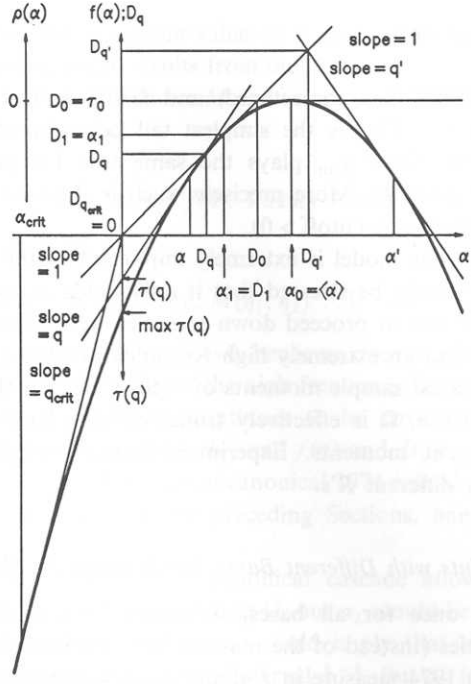

Figure 5

Enhanced version of the multifractal diagram. One vertical scale shows $\rho(\alpha) = \lim_{k \to \infty}(1/k) \log_b$ (probability) versus $\alpha = -(1/k) \log_b$ (measure). A second vertical scale shows $f(\alpha)$. In this particular graph, the weight W is lognormal, but μ itself *is not lognormal at all*. This example combines all three multifractal regimes described in the text: manifest α's ($\alpha < 0$ and $0 \leq f(\alpha) \leq \min(\alpha, D_0)$), latent α's ($\alpha > 0$ but $f(\alpha) < 0$) and virtual α's ($\alpha < 0$ and $f(\alpha) < \vartheta < 0$). As is known, $q = f'(\alpha)$ and $-\tau(q)$ is the intercept of the tangent of slope q by the vertical axis. Let us also observe, in addition, that D_q is either coordinate of this tangent's intercept by the line $f(\alpha) = \alpha$.

8.8. Role of the D_q as Critical Dimensions

The last task is to explain why, for 1974 measures, D_q is a *critical* dimension for the moments of order q, as shown in MANDELBROT (1974). Later, the D_q were written down, in HENTSCHEL and PROCACCIA (1983), without any motivation, and called "generalized dimensions" (to call D_q a dimension can make no geometric sense when $D_q > E$ for low enough q's).

The key to our motivation for D_q, again, resides in the study of low dimensional cuts of higher dimensional multifractals with the same $f(\alpha)$.

The discussion is simplest when $\max D_q < \infty$. The main step is to embed the cascade in a space of dimension $E > \max D_q$. We know that the equation $D_q = 0$ defines the q_{crit} relative to the property that $\langle \mu^q(d\mathbf{x}) \rangle < \infty$ and that the variability of the Ω's (which are independent in a sample) overwhelms the variability of the low frequency component of $\mu(d\mathbf{x})$. Therefore, the law of large numbers implies

that along the cut with $D > D_q$, the q-th sample moment is the sum of many contributing terms, each of which is negligible in relative value. Along a cut with $D < D_q$, to the contrary, $\tau(q) < 0$, which can be shown to imply that, while $\langle \mu^q \rangle = \infty$, all the percentiles of the q-th moment tend to 0 with dt. Hence, the sample moment is very small with a very high probability, and the exceptional instances when it is *not* very small occur because of the exceptional contribution of a single huge sample addend.

Thus, the set of points that contribute to the sample q-th moment of μ is hit by generic cut spaces of dimension $> D_q$, and missed by generic cut spaces of dimension $< D_q$.

9. Final Remarks

We realize that the later part of this paper is overly hasty. MANDELBROT (1974) is (unfortunately but perhaps unavoidably) a difficult paper, and not a good reference for the student. We are preparing a more complete presentation on this topic, which will incorporate MANDELBROT (1989).

Acknowledgements

Many thanks to Jens Feder, Tamas Vicsek, Jacques Peyrière, Carl Evertsz, and David di Vincenzo for pointing out rough and/or obscure spots in earlier versions of this text.

In Memoriam Abraham Robinson (1918–1974). From the mathematical viewpoint, expressions of the form $dy \sim (dx)^\alpha$ are examples of "*nonstandard infinitesimals*". This is a concept from *nonstandard analysis*, a creation of Abraham Robinson. He and I had once hoped to explore this joint field of interest together, but he fell ill and passed away. The memory of this project, together with my recent accession to the Abraham Robinson Professorship of Mathematical Sciences at Yale, bring my thoughts back to this remarkable man.

REFERENCES

AZENCOTT, R., GUIVARC'H, Y., and GUNDY, R. F., *Ecole de Saint-Flour 1978, Lecture Notes in Mathematics* (Saint Flour, 1978) Vol. *774* (Springer, New York 1980).

BILLINGSLEY, P., *Ergodic Theory and Information* (J. Wiley, New York 1967) p. 139.

BOOK, S. A. (1984), *Large deviations and applications*, In *Encyclopedia of Statistical Sciences 4*, 476.

CATES, M. E., and DEUTSCH, J. M. (1987), *Spatial Correlations in Multifractals*, Phys. Rev. *A35*, 4907.

CHERNOFF, H. (1952), *A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations*, Ann. Math. Stat. *23*, 493.

CHHABRA, A., and JENSEN, R. V. (1989), *Direct Determination of the $f(\alpha)$ Singularity Spectrum*, Phys. Rev. Lett. *62*, 1327.

DACUNHA-CASTELLE, D., *Grandes Déviations et Applications Statistiques* (Astérisque 68) (Societé Mathématique de France, Paris 1979).

DANIELS, H. E. (1954), *Saddlepoint Approximations in Statistics*, Ann. Math. Stat. *25*, 631.

DANIELS, H. E. (1987), *Tail Probability Approximations*, International Statistical Review *55*, 37.

ELLIS, R. S. (1984), *Large Deviations for a General Class of Random Vectors*, The Annals of Probability *12*, 1.

FEDER, J., *Fractals* (Plenum, New York 1988).

FOURCADE, B., BRETON, P., and TREMBLAY, A.-M. S. (1987), *Multifractals and Critical Phenomena in Percolating Networks: Fixed Point, Gap Scaling and Universality*, Phys. Rev. *B36*, 8925.

FOURCADE, B., and TREMBLAY, A.-M. S. (1987), *Anomalies in the Multifractal Analysis of Self-similar Resistor Networks*, Phys. Rev. *A36*, 2352.

FRISCH, U., and PARISI, G., *Fully developed turbulence and intermittency*, In *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics* (International School of Physics "Enrico Fermi," Course 88) (eds. Ghil, M.) (North-Holland, Amsterdam 1985) p. 84.

GRASSBERGER, P. (1983), *Generalized Dimensions of Strange Attractors*, Phys. Lett. *97A*, 227.

GUTZWILLER, M. C., and MANDELBROT, B. B. (1988), *Invariant Multifractal Measures in Chaotic Hamiltonian Systems, and Related Structures*, Phys. Rev. Lett. *60*, 673.

GUIVARC'H, Y. (1987), *Remarques sur les Solutions d'une Equation Fonctionnelle Non Linéaire de Benoît Mandelbrot*, Comptes Rendus (Paris) *305I*, 139.

HALSEY, T. C., JENSEN, M. H., KADANOFF, L. P., PROCACCIA, I., and SHRAIMAN, B. I. (1986), *Fractal Measure and their Singularities: The Characterization of Strange Sets*, Phys. Rev. *A33*, 1141.

HENTSCHEL, H. G. E., and PROCACCIA, I. (1983), *The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors*, Physica (Utrecht) *8D*, 435.

HUANG, K., *Statistical Mechanics* (J. Wiley, New York 1966).

KAHANE, J. P., and PEYRIÈRE, J. (1976), *Sur Certaines Martingales de B. Mandelbrot*, Adv. in Math. *22*, 131.

MANDELBROT, B. B., *Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence*, In *Statistical Models and Turbulence* (Lecture Note in Physics, Vol. 12), Proc. Symp. La Jolla, Calif. (eds. Rosenblatt, M, and Van Atta, C.) (Springer-Verlag, New York 1972) p. 333.

MANDELBROT, B. B. (1974), *Intermittent Turbulence in Self Similar Cascades*; *Divergence of High Moments and Dimension of the Carrier*, J. Fluid Mech *62*, 331; also Comptes Rendus *278A*, 289, 355.

MANDELBROT, B. B., *The Fractal Geometry of Nature* (W. H. Freeman, New York 1982) pp. 373–381 discuss α and the "Lipshitz-Hölder heuristics," then the "nonlacunar fractals" = multifractals.

MANDELBROT, B. B. (1984), *Fractals in Physics: Squig Clusters, Diffusions, Fractal Measures and the Unicity of Fractal Dimension* J. Stat. Phys. *34*, 895.

MANDELBROT, B. B. (1986), *Letter to the Editor: Multifractals and Fractals*, Physics Today *11*.

MANDELBROT, B. B., (1988), *An introduction to multifractal distribution functions*, In *Fluctuations and Pattern Formation* (Cargèse, 1988) (eds. Stanley, H. E., and Ostrowsky, N.) (Kluwer, Dordrecht-Boston 1988) pp. 345–360.

MANDELBROT, B. B., *Examples of multinomial multifractal measures that have negative latent values for the dimension $f(\alpha)$*, In *Fractals* ("Etteor Majorama" Centre for Scientific Culture, Special Seminar) (ed. Pietronero, L.) (Plenum, New York 1989).

MANDELBROT, B. B. (forthcoming), *Fractals and Multifractals: Noise, Turbulence and Galaxies* (Selecta, Vol. 1) (Springer, New York).

MEAKIN, P., *The growth of fractal aggregates and their fractal measures*, In *Phase Transitions and Critical Phenomena* (eds. Domb, C., and Lebowitz, J. L.) (Academic Press, London 1988) *12*, 335.

MENEVEAU, C., and SREENIVASAN, K. R. (1987), *Simple Multifractal Cascade Model for Fully Developed Turbulence*, Phys. Rev. Lett. *59*, 1424.

MENEVEAU, C., and SREENIVASAN, K. R. (1989), *Measurement of $f(\alpha)$ from Scaling of Histograms, and Applications to Dynamic Systems and Fully Developed Turbulence*, Phys. Lett. *A137*, 103–112.

PRASAD, R. P., MENEVEAU, C., and SREENIVASAN, K. R. (1988), *Multifractal Nature of the Dissipation Field of Passive Scalars in Fully Turbulent Flaws*, Phys. Rev. Lett. *61*, 74.

VOLKMANN, (1958), *Ober Hausdorffsche Dimensionen Von Mengen, Die Durch Zifferneigenschaften Sind*, Math. Zeitschrift *68*, 439.

(Received/accepted March 1, 1989)