The Journal of Fourier Analysis and Applications

Kahane Special Issue, 1995

Negative Dimensions and Hölders, Multifractals and Their Hölder Spectra, and the Role of Lateral Preasymptotics in Science

B. B. Mandelbrot

ABSTRACT. *Besicovitch and his school discovered cases where the Hausdorff–Besicovitch dimension takes the "entropy-like" form* $-\sum p_j \log p_j$, *where the* p_j *are probabilitylike. The center of the present paper is occupied by related expressions of the form* $-\log pj$, *suitably normalized. Being entropylike but nonaveraged, they will be called "ELNA codimensions." Other new concepts discussed here are those of "ELNA dimensional sequences," and "ELNA Hölder." While all the existing fractal dimensions involve sets or measures the ELNA dimensional sequences, ELNA dimensions, and ELNA Hölders yield a richer structure because they involve sequences of sets and measures, as well as their limits, which can be nonempty or empty. "Survival" occurs when the limit is nonempty; if so an ELNA dimension and ELNA Hölder are positive and difficult proofs show that the ELNA dimension's value is typically identical to those of a Hausdorff–Besicovitch dimension and of other fractal dimensions. But the ELNA dimensional sequence brings important additional information. For example, in the case of multifractals characterized by a Hölder spectrum* $f(\alpha)$, *it defines useful approximating functions* $f_\varepsilon(\alpha)$. *"Extinction" means that the limit of the sequence is empty; the ELNA dimension is then negative, and it is shown that it fulfills a surprising and novel role: It manages to give straightforward interpretations and a numerical value to the so-far empty notion of "the degree of emptiness of an empty set." One interpretation is geometric, in terms of the actual or formal embedding of the generating procedure in a higher dimensional space. The second interpretation is statistical, in terms of a novel procedure called "supersampling," which is motivated by a novel "lateral" passage to the limit. The practical usefulness of negative-valued ELNA dimensions shows that the need may exist in physics for characteristics that become lost in asymptotic results (often described as "fine-grained thermodynamics") but are present in preasymptotic results ("coarse-grained thermodynamics") and can be usefully combined with lateral preasymptotics.*

Math Subject Classification. 28A80, 60G57, 60F10, 60D05, 82A42

Keywords and Phrases. fractals, multifractals, multifractal spectrum, Hölder exponent, Hölder spectrum, negative dimension, degree of emptiness of an empty set, embedding, supersampling, Hausdorff dimension

Acknowledgements and Notes. In the 1960s, in addition to J. P. Kahane and J. Peyrière, I had challenging and beneficial discussions with U. Frisch. Other earlier debts are listed in earlier publications and are too numerous to be repeated here. Around 1990, I had challenging and very beneficial contacts with K. Sreenivasan, C. J. G. Evertsz, and C. Meneveau. Most recently, I was helped by discussions with R. Riedi.

ISSN 1069-5869

Dedication

Several fractal concepts first arose in "pure" mathematics and predate the concrete needs that led to fractal geometry. But concrete needs continually raise new questions and require mathematical tools that are either completely new or modifications of old tools. "ELNA dimension" is an example. It first entered in the study of singular measures called multifractals and as a nameless expression in extinction-or-survival theorems that I conjectured in 1968, 1972, and 1974 and that my old friend Jean-Pierre Kahane, together with his then-student Jacques Peyrière, proved and generalized. This is ample reason for surveying and expanding the new topic of negative ELNA dimensions on this festive occasion. This survey will, I hope, stimulate more general considerations. A new theorem on multifractals is followed by the conjecture that the same result holds under conditions weaker than those I managed to tackle.

No scientist should neglect the advice of William of Ockham that "entities should not be multiplied beyond necessity," but both mathematics and physics meet necessity with great regularity and in many guises. Thus, the study of random fractal sets or measures is enriched by introducing yet another version of the already overloaded concept of dimension.

The need for negative dimensions first impressed itself upon me in several separate ways that recall the introduction of $\sqrt{-1}$. A known procedure was applied, either knowingly or unwittingly, outside its domain of validity, and the outcome was a negative dimension, therefore first seemed meaningless. But it eventually turned out to have a well-defined and novel role: It manages to give a numerical value to the so-far empty notion of "the degree of emptiness of an empty set." This confirms that, under the pressure of applications, fractal geometry continues to be in a state of growth and not yet of conceptual consolidation.

In the space $\mathbb{R}^E$, Hausdorf–Besicovitch dimension also satisfies $D \leq E$. This second bound also turned out to be excessively constraining in physics, and I showed that $D > E$ can be used to give a meaning and numerical value to the so-far empty notion of "the degree of overfilling of space by an overfilling set." The topic is discussed elsewhere.

Section 1 is devoted to motivations and the definitions of the ELNA dimension and the ELNA Hölder. Sections 2 and 3 study ELNA dimensions and Hölders for birth-and-extinction fractals and trema fractals, respectively. These are random sets obtained via birth-and-extinction processes or the cutting out of "tremas." Section 4 studies ELNA dimension in a nonfractal random context: First as a formal device, then concretely. Section 5 outlines my 1974 theory of multifractal singular measures [10, 11, 14]. In this approach, a sequence of functions $f_k(\alpha)$ and a limit function $f(\alpha)$ I called Hölder spectrum are obtained from the probability limit theorems relative to large deviations ("Cramèr theory"). Those theorems introduce two quantities: f is now reinterpreted as an ELNA dimension, and α is now reinterpreted as an ELNA Hölder; $f(\alpha)$ enters as the limit of an ELNA dimensional sequence. An astonishing number of separate fine geometric characteristics of the graph of $f(\alpha)$ turn out to have a direct physical role. Section 6 moves from distribution properties to almost sure properties; it restates some old theorems, draws recent corollaries and states new conjectures. A brief §7 sketches the relation of ELNA dimension to "thermodynamical" thinking and that of ELNA dimensional sequences to a proposed "lateral" extension of thermodynamical thinking.

1. The Notions of Coarse-Graining, ELNA Dimension, Coarse Hölder, Fine Hölder, and ELNA Hölder

For Hausdorff and Besicovitch, fractal dimension concerned well-determined *nonrandom sets*, the empty set being unique and having dimension 0. By contrast, ELNA dimensions concern recursive constructive procedures (indexed by a positive integer or a real) that generate sequences of sets. ELNA stands for "entropylike but nonaveraged." An ELNA dimensional sequence is attached to an indexed sequence of sets, and an ELNA dimension arises as the index $\to \infty$. To simplify notation, this paper uses D without subscript to denote ELNA dimension.

I long resisted attempts to define a dimension for a random set, that is, for a population of sets, but multifractals changed my mind. In addition, I now believe that the simplest fractal dimension would have been more fully understood by nonmathematicians if it had not been attached to a set but rather to a procedure that defines a set. For example, the dimension of the snowflake Koch curve is $D = \log 4/\log 3$; viewed as a similarity dimension, D is a property of the set, but D is better understood if viewed as a property of the construction procedure.

Hölder and Lipschitz introduced an exponent that involves a lim inf and describes the local behavior of nondifferentiable functions. It will be denoted as H or α and extends to describe the local behavior of a measure. It will be called fine Hölder. However, in addition to the usual limit expression, we need a version that is analogous to an ELNA dimension. Although the words "entropylike and nonaveraged" are not applicable, we shall speak of ELNA Hölder. This notion does not apply to one given measure but to an indexed sequence of measures. It can be negative. It first forced itself upon me in the theory of multiplicative multifractals and then proved to be physically essential in physics. In particular, many directly observable properties of multifractals are determined by a Hölder spectrum $f(\alpha)$ (discussed in §5), in which α is an ELNA Hölder and f an ELNA dimension, and the existence of points where both $\alpha < 0$ and $f < 0$ proves to be of mathematical interest and direct physical importance.

1.1. Informal Statement of the Extinction-or-Survival Alternative

Let S denote a random set and s denote a nonrandom sample. The Hausdorff–Besicovitch dimension $D_{\mathrm{HB}}(s)$ is a property of the sample set s, and the ELNA dimension $D(S)$ will be defined in §1.2 as a property of a random event, namely, as a renormalized logarithm of a probability. In other words, D is a function of an underlying probability distribution linked to a generating procedure.

This being granted, $D_{\mathrm{HB}}(s)$ and $D(S)$ are related by extinction-or-survival theorems whose prototypes are found in earlier studies that this paper surveys and extends [10, 11, 7]. Those theorems take either of two distinct forms.

(i) A *surviving* procedure is one that has a positive probability of generating a nonempty limit set. The ELNA dimension of a surviving procedure is positive and can be viewed as another form of fractal dimension, valid for all (or almost all) the sets generated by the procedure. A "survival" theorem describes conditions from which it follows that, with positive probability, s is nonempty and

$$D_{\mathrm{HB}}(s) = D(S), \text{ where } D(S) \geq 0.$$

(ii) An *extinguishing* procedure generates the empty set with probability one. An "extinction" theorem describes conditions from which it follows that a.s., s is empty, hence $D_{\text{HB}}(s) = 0$; on the other hand, $D(S) < 0$.

Thus, all extinguishing procedures can be viewed as equivalent from the viewpoint of D_{HB}. However, the values of their ELNA dimensions $D(S)$ may differ, with the only constraint that they are negative. When an extinction theorem holds, this paper proposes to show that the negative dimension $D(S)$ must not be discarded; it continues to be important and to have a clear geometric meaning and physical significance. Indeed, it makes it possible to split the notion of "a set that is almost surely empty" into a multitude of newly distinguished possibilities. Among those possibilities, the values of the ELNA dimension provide a numerical meaning to the seemingly absurd notion of "degree of emptiness of an empty set."

1.2. Coarse-Graining of Sets and Measures

The self-explanatory term "*coarse graining*" is used in thermodynamics for a procedure that mathematicians introduced independently in many forms. We deal with random sets S in the Euclidean space $\mathbb{R}^E$.

- Each sample set s is replaced by its intersection with a cube of side L, called an L-box, that is subdivided into cubic ε-boxes of side ε.
- ε-coarse-graining replaces s by the set s_ε made of the ε-boxes intersected by s_ε. An ε-box is called nonempty when it intersects s_ε.
- S is assumed to be stationary or translation invariant in the sense that the expression

$$\Pr\{\text{a given } \varepsilon\text{-box } \in S_\varepsilon\} = \Pr\{\text{a given } \varepsilon\text{-box is nonempty}\}$$

is the same for all ε-boxes in the large L-box.

- ε-coarse-graining replaces a measure defined in $\mathbb{R}^E$ by a measure defined for ε-boxes in $\mathbb{R}^E$.

1.3. ELNA Dimensional Sequences and ELNA Dimensions and Codimensions, for which it is Possible that $D < 0$

We define the *sequence* of ELNA codimensions as

$$C(\varepsilon) = -\frac{\log \Pr\{\text{a given } \varepsilon\text{-box is nonempty}\}}{\log \varepsilon}.$$

Once again, this quantity is *entropylike* because it is of the form " – logarithm," but it is *not* averaged. However, it is normalized.

When the *limit* exists, we define the ELNA codimension as

$$C = \lim_{\varepsilon \to 0} C(\varepsilon).$$

The expression $\Pr\{\text{a given } \varepsilon\text{-box is nonempty}\}$ can be arbitrarily small; therefore, the expressions $C(\varepsilon)$ and C can be arbitrarily large, and the corresponding ELNA dimensions $D(\varepsilon) = E - C(\varepsilon)$ and $D = E - C$ can be negative.

1.4. Coarse Graining and the Coarse Hölder Exponent; Fine Graining and the Fine Hölder Exponent H (Pointwise Dimension) That Satisfies $H > 0$

Given space "coarse-grained" into cubic boxes of side ε and a measure μ, we define the *coarse Hölder exponent* in a box as

$$H = \frac{\log(\text{measure } \mu \text{ in an } \varepsilon\text{-box})}{\log \varepsilon}.$$

When the box is chosen at random, H is a random variable. To follow an ingrained notation, the value of H will not be denoted by h, as probabilists would have preferred, but by α.

We define the fine-grained Hölder exponent at a point P as

$$H = \liminf_{\varepsilon \to 0} \frac{\log(\text{measure } \mu \text{ in an } \varepsilon\text{-box containing } p)}{\log \varepsilon}.$$

This last definition is classical, but the term *fine* must be attached to it for the sake of contrast with coarse Hölder. H is sometimes called a *pointwise dimension.* In a more perfect world, the choice of words would not matter, but in our world, "dimensions of measures" are a continuing source of confusion, and I opt to avoid this use.

1.4.1. *Effect of the Choice of Units and the Inequality $H \geq 0$.*

In general, the units of ε, μ can be selected independently, and their choice does not affect the value of the fine Hölder. When a point carries an atom of positive measure, the fine Hölder at that point is $= 0$. When a measure is continuous at a point, its fine Hölder is > 0. On the contrary, coarse Hölders depend on the units of ε and μ, but the dependence disappears in the interesting range where both ε and μ are small.

1.5. The ELNA Hölder, a Generalized Hölder Exponent That Can Satisfy $H < 0$

Multifractals introduce inevitably an expression that is "like a Hölder," except that negative values play an *essential* role. The main point is that, instead of a single measure μ, one must consider a sequence of measures μ_ε that depend upon ε and may fail to converge to a nondegenerate limit μ. In the example of greatest interest, those measures a.s. converge to a limit that vanishes in every interval. To achieve a nicer looking notation, we let ε and μ depend on a parameter k and consider for each point P the sequence of ratios

$$H_k = \frac{\log(\text{measure } \mu_k \text{ in an } \varepsilon_k\text{-box containing } P)}{\log \varepsilon_k}.$$

To denote this sequence and its timing as $\varepsilon_k \to 0$, we keep the letters ELNA.

One can prove that the ELNA dimensions and Hölders satisfy the inequality $H \leq D$.

Section 1 has been entirely formal. The sections that follow discuss how the ELNA dimensions and Hölders can be interpreted in concrete and measurable ways: Either geometrically via embedding or statistically via lateral supersampling.

2. Birth-and-Extinction Fractal Sets; How a Negative-Valued ELNA Dimension Is Interpreted Geometrically via Embedding and Statistically via Lateral Supersampling

2.1. The Basic "Extinction or Survival" Alternative

The birth-and-extinction cascade on the interval [0, 1] is a recursive construction that generates a randomized Cantor dust. The kth cascade stage begins with equal boxes of length $\varepsilon = 2^{-k}$. Each box is further subdivided into two dyadic halves, and each half can behave in one of two ways: It "survives" with the probability p satisfying $0 < p < 1$, or it is "extinguished" with the probability $1 - p$. One defines a "birth-and-extinction process" [*FGN*, Chapter 23] by thinking of mother boxes of length 2^{-k} as being replaced by N daughter boxes of length 2^{-k-1}, using the following rule: $N = 0$ with the probability $(1-p)^2$, $N = 1$ with the probability $2p(1-p)$, and $N = 2$ with the probability p^2. Therefore, $EN = 2p$, and §2.6 will show that the ELNA codimension and dimension defined in §1.2. are

$$D = \log_2 EN = 1 + \log_2 p, \quad \text{and} \quad C = 1 - D = -\log_2 p.$$

The ELNA dimension $D = \log_2 EN$ is a straightforward formal generalization of the basic similarity dimension $\log_2 N$.

The number of boxes after k generations is a random variable N_k, whose behavior is classical.

When $EN > 1$, so that $D > 0$, the birth-and-extinction process has a positive probability of generating a nonempty set called a *birth-and-extinction fractal dust*. For this set, D is the value of the Hausdorff–Besicovitch dimension and also of every other useful form of fractal dimension.

When $EN \leq 1$, so that $D \leq 0$, the "bloodline" extinguishes with probability one. In that case, the "lifetime" of a bloodline is determined by the unique parameter p, hence also by any one-to-one function of p. Let us now explain why the best choice is D.

2.2. Geometric Embedding of an Almost Surely Empty Set Can "Reveal" for This Set a "Latent" Dimension D

The first of two strong reasons for drawing attention to $\log_2 EN$ as a negative dimension follows from the fact that in this instance a negative dimension can be made positive by embedding. That is, an "extinguished" birth-and-extinction fractal set can be viewed as a one-dimensional cut through a "surviving" birth-and-extinction set embedded in a suitably larger space.

Indeed, one can start with the E-dimensional cube $[0, 1]^E$ and apply the same birth-and-extinction procedure to each of its 2^E subcubes. The construction can also be carried out for noninteger E as long as 2^E is an integer. In this embedding space, $EN^E = 2^E p$. Obviously, making E sufficiently large will ensure $2^E p > 1$, hence $D(E) > 0$. Therefore, every extinguished birth-and-extinction set on the line can be interpreted as a cut through a surviving birth-and-extinction set. One can view a negative dimension as *latent* until embedding *reveals* it. Note also that $C = -\log_2 p$ is the upper bound to the dimensions in which the birth-and-extinction construction of probability p generates a set that is almost surely empty.

2.3. Geometric Embedding is Eminently Physical, Whenever It Is the Inverse of Intentional or Unavoidable Dimension Reduction

Examples of dimension reduction include Poincaré sections in theoretical dynamics and velocity-measuring devices in the experimental study of turbulence, which yield a one-dimensional record. In those cases, a negative dimension in a cut may well reflect a positive dimension in the full set. Thus, I first used embedding in the study of turbulence, showing that "anomalies" present in a one-dimensional cut through a flow may vanish if this cut is placed back in a full three-dimensional turbulence.

The conclusion that a physical process is best understood by being investigated simultaneously in spaces of every acceptable dimension was reached independently in several contexts. When applying this idea to turbulence in the late 1960s and early 1970s, I was not aware of its wide and successful use in statistical physics, especially in the context of renormalization groups.

Even in cases where embedding is an artificial mathematical device, it may remain very useful.

2.4. Use of "Supersampling" To Estimate a Latent D

In some cases, embedding is inconceivable, impractical, or unavoidably imperfect. For example, the study of turbulence can use arrays of velocity-measuring devices; but the resulting measurements are not spaced isotropically.

As a substitute, I suggested [15] an "effective embedding" by *supersampling*. The procedure will be described in this section, then its principle and validity will be discussed in §2.5.

Given a birth-and-extinction fractal that eventually will extinguish, the kth stage prefractal approximation is made up of 2^k boxes, some extinguished and other surviving. An embedding in $\mathbb{R}^E$ would involve 2^{Ek} boxes, but would not be reachable. My idea's first step was to think of this unreachable embedding as a bundle of one-dimensional strips, each containing 2^k boxes. The number of such strips would be $2^{(E-1)k}$, and they would, of course, be statistically dependent. My idea's second step was to replace those strips by a "supersample" made of $2^{(E-1)k}$ statistically *independent* samples of the original one-dimensional construction. To obtain a nonempty E-dimensional embedding, one must choose $E - 1 > -D$. The minimum number of independent samples, being 2^k, increases with k and $-D$. Assuming 2^k measurements per sample, the minimum number of measurements is of the order of $2^{(1-D)k}$.

2.4.1. *Asymptotics Are Relevant Only in the Case $D > 0$.*

When $D > 0$, one can estimate D from a single sample and the estimation precision increases with k.

2.4.2. *Pre-asymptotics Are Needed in the Case $D < 0$.*

When $D < 0$, the asymptotic limit $k \to \infty$ is empty. Taking this limit destroys all the information about D and makes it impossible to estimate D. Thus, asymptotics do not matter for $D < 0$. When one knows that one deals with a birth-and-extinction fractal, one improves the estimate of the negative value of D by taking a large number of independent samples for $k = 1$. In more general cases where $D < 0$, we must increase *both* k and the number of sample sequences. This fact is sufficiently significant to be mentioned in this paper's title and will be explored in §7.

2.5. Supersampling Is Analogous in Spirit to a Mixture of Time and Ensemble Averaging: It Is Ergodic When $D > 0$ and Nonergodic When $D < 0$

Time and ensemble averaging are familiar concepts in physics and mathematical ergodic theory. Supersampling also involves an ensemble averaging, and parallelism with ergodic arguments is made complete by viewing the parameter k as a counterpart of time. When $D > 0$, an increase in k leads to an increasingly finely defined fractal; the ergodic theorem's counterpart is the fact that increasingly precise estimation of D can be achieved in either of two ways: by asymptotic "pseudo-time" averaging or by ensemble averaging. When $D < 0$, to the contrary, pseudo-time averaging destroys all information, while ensemble averaging does not. The contrast between ensemble and pseudo-time averaging is necessary when nothing is known beyond $D < 0$ as, for example, in the study of multifractals.

Needless to say, this form of "pseudo-ergodicity" begins as a hypothesis. Its validity, hence the applicability of lateral methods, is not guaranteed, and must be proven separately in each new case.

The simple but powerful idea behind lateral limits has already been used in earlier works of mine and seems to be of wide generality. The underlying thinking begins with the known fact that, when an expression has more than one parameter that can go to infinity, the limit usually depends on the sequence of limits. Some limits may preserve information that is lost in others. Many obvious parameters arising in concrete problems are like the parameter k in a birth-and-extinction cascade. When $D < \infty$ the $k \to \infty$ limit is degenerate, and taking this limit destroys valuable information. The basic insight was that one may be able to identify one or more physically meaningful additional parameters that enter into limit theorems in which this valuable information is preserved and enhanced.

My first use of a lateral limit occurred in 1969 [8]. The notion occurs extensively in my current work on fractal sums of pulses (FSP) processes [20]. The FSP are self-affine, by design. Therefore, they are unchanged by the customary passage to the limit that leads to convergence to the Wiener or fractional Brownian motion. The reason why that passage to the limit is not useful for FSP's is not the same as in the present context: it is not that the usual passage to the limit destroys useful information, but rather that it preserves too much information. However, the definition of a FSP also includes a second parameter; it seems at first sight to be an innocuous density, but it turns out to lend itself to a useful lateral passage to the limit that reveals useful latent information.

2.6. Proof That the Expression $D = \log_2 EN$ Is a Special Case of the Expression Introduced in §1.3

When $D > 1$, the theory of birth-and-extinction processes tells us that, after $k \gg 1$ stages of construction, the number of nonempty boxes is either 0 or takes the form $N_k = \Phi E N^k$, where Φ is a random variable dependent upon p. Hence,

$$\Pr\{\text{a given } \varepsilon\text{-box is nonempty}\} = N_k 2^{-k} = \Phi 2^{-k} E N^k.$$

Therefore,

$$C_k = -\frac{\log_2 \Phi + \log_2 2^{-k} + k \log_2 EN}{\log_2 \varepsilon},$$

and it follows that

$$C = 1 - \log_2 EN,$$

as asserted in §2.1.

When $D < 1$, the argument continues to apply, and the simplest proof is based on embedding.

2.7. Behavior of the Hölder and the ELNA Hölder in the Case of Birth-and-Extinction Fractals

When $D > 0$, the cascade yields a natural way of assigning a measure on these sets. At each cascade stage k, define μ_k by assigning the measure 2^{-Dk} to every surviving box. As $k \to \infty$, this μ_k has a weak limit μ such that $\mu([0, 1]) \neq 0$ with a positive probability. When $\mu([0, 1]) \neq 0$, the μ of a given ε-box is either 0 or of the form $\Phi\varepsilon^D$, where Φ is the random variable introduced in §2.6. Therefore, the coarse Hölder is

$$H = \frac{\log \Phi}{\log \varepsilon} + D = \frac{\log \Phi}{\log \varepsilon} + E + \log_2 p.$$

The fine Hölder satisfies $H = D$ at points belonging to the limit dust and $H = \infty$ elsewhere. Its values are graphed by a Hölder spectrum function $f(\alpha)$ that is defined only for $\alpha = D$ and $\alpha = \infty$ and satisfies $f(D) = D$ and $f(\infty) = 1$.

To extend the notion of Hölder to $H < 0$, start with $H = D > 0$ and observe that $f(\alpha)$ has the following property. When supersampling or a change in E (imbedding or intersection) preserves $H = D > 0$, they translate the point where $\alpha = D$ and $f(\alpha) = D$ along the main bisector $f = \alpha$. It is tempting to consider that, when a decrease in E leads to $D < 0$, it also leads to $H = D < 0$.

The inequality $H < 0$ is paradoxical, but it *does not* imply that there is a fixed μ such that the measure in an ε-box around a given point *increases* as $\varepsilon \to 0$. In fact, $H < 0$ concerns a measure that *changes* as $\varepsilon \to 0$. There is an increase in the number of ε-boxes that become empty and lead to $H = \infty$ and an increase in the measure in those boxes that delay their becoming empty.

3. Trema Sets

The cascade used in §2 to generate a dust and a measure involves dyadic boxes. The method is easy to study, both in mathematics and in physics, but is unnatural. As a substitute, [12, Chapters 31, 33 and 35] introduced sets generated by bounded sets in $\mathbb{R}^E$, called *tremas*. Space lacks to describe them here, but a brief allusion must be included. From a generating set, called a template, in which an origin is singled out, each trema is obtained by dilating or reducing in the ratio p, called *radius*, and moving the origin to a point of coordinates $x_1, \ldots, x_E$. Each trema is parameterized by an address point of coordinates $P, x_1, \ldots, x_E$. The number of trema address points in an elementary box of sides $d\rho, dx_1, \ldots, dx_E$ is assumed to be $CP^{-E-1}dpdx_1, \ldots, dx_E$.

If all the tremas satisfying $\varepsilon < \rho < 1$ are cut out of $\mathbb{R}^E$, one is left with a "remainder set" $S(\varepsilon)$. The remainder sets $S(\varepsilon)$ are indexed by ε; if $\varepsilon' > \varepsilon''$, $s(\varepsilon'')$. The set $S(0) = \lim_{\varepsilon \to 0} S(\varepsilon)$ may be a.s. empty or nonempty with a positive probability for a prescribed point P; one finds that

$$\Pr\{P \in S(\varepsilon)\} = \varepsilon^C.$$

Therefore, the ELNA codimensional sequence of $S(\varepsilon)$ reduces to C, and the ELNA codimension of $S(0)$ is C.

When $C < E$ so that $D > 0$, it can be shown that (a) $S(0)$ is nonempty with a positive probability and that (b) if $S(0)$ is nonempty, it has a.s. D as its Hausdorff–Besicovitch dimension.

When $C > E$ so that $D < 0$, $S(0)$ is a.s. empty and D measures its degree of emptiness. It is a latent quality that can be revealed by embedding into a space of dimension $E' > C$.

4. Intersections of Sets, the Addition Rule for ELNA Codimensions

This section describes a different path that helps understand the nature of ELNA dimension. The context is that of the intersection of two randomly placed sets that need not be fractal.

Minkowski defined a set's ε-neighborhoods and showed that in $\mathbb{R}^3$ a surface is best understood as the limit of the "coarse-grained" approximations represented by its ε-neighborhoods. We shall continue beyond the point where Minkowski stopped and argue that certain constructions that define empty limit sets can be characterized by "latent" ε-neighborhoods.

4.1. The Dimension of Two Sets' Intersection: The Generic Rule and the Major Exception to Its Validity

Take two sets S_1 and S_2 (either Euclidean or fractal) in $\mathbb{R}^E$. Denote their co-dimensions by $E - \dim(S_1)$ and $E - \dim(S_2)$. Define $S = S_1 \cap S_2$ and define latent $\dim(S)$ as

$$E - \text{latent dim}\ (S) = E - \dim(S_1) + E - \dim(S_2).$$

4.1.1. *Generic Rule.*

When latent dim $(S) > 0$, it follows that $\dim(S) =$ latent dim (S).

4.1.2. *Major Exception to the General Rule.*

When latent dim $(S) \leq 0$, it follows that $\dim(S) = 0$.

4.2. Examples: Seemingly Thoughtless Algebra Yields an Intuitive Notion of "Degree of Emptiness"

We begin with seemingly thoughtless algebra in $\mathbb{R}^3$. The intersections of (i) two points, (ii) a point and a line, and (iii) two lines are "generically" empty, but an "intuition" tells us that the intersection of two points should be viewed as even "emptier" than the intersection of a point and a line, or the intersection of two lines. One would wish to replace this intuition by a number. The Hausdorff–Besicovitch dimension fails at this task because it vanishes for the intersections (i), (ii) and (iii).

On the other hand, let codimensions be added, disregarding the major exception. This characterizes our three intersections, respectively, by latent dimensions equal to -3, -2, and -1. If this addition of codimensions could be given a meaning, it would yield a very useful measure of the relationship between nonintersecting sets.

4.3. Steps Needed To Justify the Addition of Codimensions When It Seems To Be Thoughtless

Acknowledge that one can only observe a bounded "window" of space, a cube of side L, containing small blobs, thin sticks, and thin shells and, more generally, containing Minkowski ε-neighborhoods of sets, which are sequences of sets that depend on a parameter ε. The study of such sequences may be richer than the study limited to their $\varepsilon \to 0$ limits. When $D > 0$, the two studies are equivalent. When $D < 0$, the limit for $\varepsilon \to 0$ is degenerate, but if the sets are randomly placed, the preasymptotics for small $\varepsilon > 0$ carry useful information.

Randomness is central to allowing this generalized dimension to become negative. D does not describe a specific set, but describes and classifies a generic reason why a limit set happens to be empty.

4.4. Negative Box ELNA Dimension: Example From Euclidean Geometry

To cover a Euclidean or fractal set of box dimension D_B requires $N(b) \sim b^{D_B}$ boxes of side $r = b^{-1}$. The familiar box dimension D_B simply measures the rate of increase of $N(b)$ with b. A generalized negative dimension might describe the rate of *decrease* of "some quantity" like $N(b)$. The number of boxes, being an integer, could not decrease. To avoid the fact that N is an integer, we must refer to a random ensemble or population of constructions. Let us examine EN.

The algebra is simplest for the intersection of a point and a line in the plane, to show that its box ELNA dimension is equal to -1, start with a square window of side L that includes a pointlike blob of side $1/b$ and a linelike strip of width $1/b$. When the strip intersects the blob, $N = 1$; otherwise, $N = 0$. Intersection occurs when the distance between the point and the line is $< b$, which happens with probability $\sim b/L$. Thus, $EN \sim L/b$, and for large L/b we obtain

$$\lim_{L/b\to\infty} D_B = \frac{\log(1/b)}{\log b} = -1.$$

4.5. Negative ε-Neighborhood (Sausage) Dimension; An Example From Euclidean Geometry

The familiar sausage of S is the set of points that lie within a distance ε of a point in S. The sausage of the union of S_1 and S_2 is therefore the union of the sausages of S_1 and S_2, But what about the sausage of the intersection $S_1 \cap S_2$? Where $S_1 \cap S_2$ is nonempty, the intersection of the sausages of S_1 and S_2 and the sausage of the intersection scale with the same exponent as $\varepsilon \to 0$. When the sets S_1 and S_2 fail to intersect, only the intersection of the sausages continues to be defined. In the present example of a point and a line in the plane, the area of the intersection of sausages is $\sim \varepsilon^2$ with a probability $\sim \varepsilon/L$ and otherwise is 0. Hence, the expected area of the intersection is $\sim \varepsilon^3/L$. The exponent in this expression is a generalized sausage codimension. Its value is 3, which confirms $D = -1$.

5. Multiplicative Multifractal Measures: Derivation of the Hölder Spectrum $f(\alpha)$ via Large Deviations and the ELNA Dimension and Hölder

Sections 5.1 and 5.2 survey two elementary cases and interpret them in ELNA dimensional terms. The remaining subsections study random multiplicative multifractals using Cramèr renormalization. The Hölder spectrum $f(\alpha)$ is obtained not as a property of the limit measure but as a property of a multifractal-generating process.

5.1. Binomial Multifractals

The binomial measures depend upon a real parameter m_0, called a *multiplier* or a *mass* satisfying $0 < m_0 < 1$. The "generating step" redistributes mass between the halves of a dyadic interval, with the relative proportions of m_0 to the left and $m_1 = 1 - m_0$ to the right. Thus, the first stage yields the mass m_0 in $[0, \frac{1}{2}]$ and the mass m_1 in $[\frac{1}{2}, 1]$. After k stages, let φ_0 and φ_1 denote the relative frequencies of 0's and 1's in the binary development of $x = 0.\beta_1\beta_2, \ldots, \beta_k$ written in the counting base $b = 2$. The binomial measure attributes to the dyadic interval $[dx] = [x, x + 2^{-k}]$ a mass equal to $\mu(dx) = m_0^{k\varphi_0} m_1^{k\varphi_1}$. Hence, the coarse Hölder exponent of this interval is

$$\alpha = \alpha(\varphi_0, \varphi_1) = \frac{\log d\mu(dx)}{\log dx} = -\varphi_0 \log_2 m_0 - \varphi_1 \log_2 m_1.$$

The number of intervals leading to φ_0 and φ_1 is $N(k, \varphi_0, \varphi_1) = k!/(k\varphi_0)!(k\varphi_1)!$ Hence, in an informal sense that will be specified momentarily, the set where this α is observed can be said to have a dimension equal to

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

The limit $k \to \infty$ defines

$$\delta(\varphi_0, \varphi_1) = \lim_{k\to\infty} \delta(k, \varphi_0, \varphi_1) = -\varphi_0 \log_2 \varphi_0 - \varphi_1 \log_2 \varphi_1.$$

Eliminating φ_0 and $\varphi_1 = 1 - \varphi_0$ between α and δ, we obtain a function $f(\alpha)$. Its graph is a translated and rescaled version of the familiar bell-shaped graph of the entropy function: We shall say that it is "$\cap$-shaped." It satisfies the following inequalities: $f(\alpha) \le \alpha$, with equality for some α_1; $f(\alpha) \le 1$, with equality for some α_1; $f(\alpha) \ge 0$ in a bounded interval $0 < \alpha_{\min} = -\log_2 m_0 \le \alpha \le \alpha_{\max} = -\log_2 m_0 < \infty$, with equality at the interval's endpoints. Outside of this interval, $f(\alpha)$ is not defined.

The logical standing of the informal dimension $f(\alpha)$ comes out if the problem is restated in terms of probability densities. Even for nonrandom measures a random variable appears when a dyadic box of length 2^{-k} is picked at random among the 2^k boxes. Then the coarse Hölder exponent becomes a random variable H. The probability of hitting $H = \alpha$, to be denoted by $\Pr\{H = \alpha\}$, equals $2^{-k}N(k,\varphi_0,\varphi_1)$ and satisfies

$$C(k,\varphi_0,\varphi_1)=\frac{\log\,\Pr\{H=\alpha\}}{\log dx}=\frac{\log 2^{-k}N(k,\varphi_0,\varphi_1)}{\log dx}.$$

The limit $C(\alpha)=\lim_{k\to\infty}C(k,\varphi_0,\varphi_1)$ is a ELNA codimension in the sum of §1.3. It satisfies $C(\alpha)=1-f(\alpha)$.

5.2. Multinomial Multifractal Measures and the Derivation for the Legendre Multifractal Formalism Using Lagrange Multipliers

In the base $b>2$, a multinomial measure is defined by b multipliers m_β $(0\le\beta\le b-1)$, which add to 1. Denote by Φ the point whose coordinates are the frequencies φ_β 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

$$\mu(dt)=\prod m_\beta^{k_\beta}.$$

The coarse Hölder and the informal dimension δ are given by

$$\alpha=-\sum\varphi_\beta\log_b m_\beta\quad\text{and}\quad\delta=-\sum\varphi_\beta\log_b\varphi_\beta.$$

5.2.1. *The Domain of All Possible Points* (α,δ). *Lagrange Multipliers and Legendre Relations of Gibbs Thermodynamics*

As $k\to\infty$, the Φ's that yield a given α define a portion of a hyperplane, on which the function δ varies continuously. Therefore, the possible values of δ for a given α cover an interval. Those δ's are dominated by the highest among them, which is the solution of the folllowing problem:

$$\text{maximize}\ -\sum\varphi_\beta\log_b\varphi_\beta\quad\text{given}\ -\sum\varphi_\beta\log_b m_\beta=\alpha,\quad\text{and}\ \sum\varphi_\beta=1.$$

The classical method of Lagrange multipliers 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 roles that inverse temperature $1/kT$, "partition function," and "free energy" play in thermodynamics are now played by q, $\sum m_\beta^q$, and $\tau(q)=-\log_b\sum m_\beta^q$, respectively.

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

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

and

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

Hence, the Legendre and inverse Legendre relations

$$\alpha = \frac{\partial \tau(q)}{\partial q}$$

and

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

5.2.2. *The Moments of μ and the function $\tau(q)$. The partition function and its expectation.* When a box is selected at random, the multiplier is a random variable M that takes the value m_β with the probability $1/b$. Its qth moment is $EM_q = \sum(1/b)m_\beta^q$. The "cumulant generating function" is $\log_b EM^q = -1 - \tau(q)$. In terms of $\tau(q)$, one finds that

$$E\mu^q(dt) = [EM^q]^k = [b^{-1} \sum m_\beta^q]^k = [b^{-1-\tau(q)}]^k = (dt)^{1+\tau(q)}.$$

To define the partition function, one proceeds in formal analogy with thermodynamics. The domain of definition of μ is subdivided into boxes of side dx, and the $\mu_i(dx)$ are the measures contained in those boxes. Then, the partition function is the expression

$$\chi(q, dx) = \sum \mu_i^q(dx).$$

It satisfies

$$E\chi(q, dt) = (dt)^{-1} E\mu^q(dt) = (dt)^{\tau(q)}.$$

As we see, the vocabulary of thermodynamics cannot be avoided but must be taken with extraordinary caution. Even when it is suggestive, it may be misleading because the moments of the limits of multifractal measures can differ from the limits of the moments. As a consequence, heuristic thermodynamical arguments are often either incorrect or valid under restricted conditions.

5.3. Random Multiplicative Measures: Distinction Between $\tilde{\mu}$ and μ

Next to be examined is the generalized and randomized form of the multinomial measure that was constructed and investigated in [10, 11, 7]. That construction's first stage takes up the base-b box of length $\varepsilon = b^{-1}$ starting at $\alpha = 0.\beta 000, \ldots$, with $0 \le \beta \le b-1$, and assigns to it a random multiplier $M(\beta)$. The same process is repeated recursively. Thus, the first k stages of multiplication assign to the b' -adic box of length b^{-k} starting at $x = 0.\beta_1\beta_2, \ldots, \beta_k$ the mass

$$\tilde{\mu}_k(dx) = M(\beta_1)M(\beta_1, \beta_2), \ldots, M(\beta_1, \ldots, \beta_k) = \prod M.$$

The successive random multipliers M for given x are independent and identically distributed.

To fit the image of "mass," it is natural to assume some form of conservation.

5.3.1. *The Conservative Variant of the Multiplicative Construction in [ ].*

Here, each cascade stage shuffles mass around while preserving it. Thus, the point of coordinates $M(\beta)$ is a random point in the portion of b-dimensional space defined by $M(\beta) \geq 0$ and $\sum M(\beta) = 1$. These inequalities imply $EM = b^{-1}$. They also imply $M \leq 1$, from which it follows that $\tilde{\mu}(dx) \leq 1$, from which it follows that $\alpha_{\min} \geq 0$. In that case, the mass $\mu_k(dx)$ in a b-adic box of length b^{-k} reduces to $\tilde{\mu}_k(dx)$. It will follow that the complications relative to ELNA Hölder exponents do not occur, but those relative to ELNA dimensions may or may not occur.

5.3.2. *Conservation on the Average and the Canonical Variant of the Multiplicative Construction in [ ]; The Random Variable Ω.*

A more general construction allows mass to be created or extinguished, as long as it is conserved on the average, meaning that $EM = b^{-1}$, but M is not otherwise constrained, and may exceed 1. The simplest subcase, called *canonical*, assumes that the $M(\beta)$ are statistically independent. It follows that the mass $\mu_k(dx)$ is a product of two independent terms.

The low-frequency first k stages yield a mass equal to the above $\tilde{\mu}_k$. A significant consequence is that each step for which $M > 1$ creates mass. As a result, it may happen that $\tilde{\mu}_k > 1$, that is, $\tilde{\alpha}_{\min} < 0$.

But this is not all: The "high-frequency" stages beyond the kth contribute a random multiplier Ω identical in distribution to $\mu_0[0, 1]$. Thus, we can write

$$\mu_k(dx) = \tilde{\mu}_k(dx)\Omega.$$

In conservative cascades, $\Omega \equiv 1$. In canonical cascades with $\alpha_1 < 0$, §6.1 will show that $\Omega = 0$ a.s.. In canonical cascades with $\alpha_1 > 0$, $f(\alpha)$ does not fully determine the probability distribution of Ω but some features of $f(\alpha)$ deeply affect the moments $E\Omega^q$, as will be seen in §6.3.

All these complicated distinctions are important, but Durrett and Liggett [4] have shown that the rules of dependence between the multipliers do not change the extinction or survival criterion, to be introduced shortly. The fact that they do not affect the multifractal formalism may even be viewed as demonstrating that measures that are treated by this formalism as being identical may otherwise be clearly different.

5.4. Random Multiplicative Measures: Distinction Between the ELNA Hölder Exponent $\tilde{H}$ and the Hölder Exponent H; Examples of Functions $f(\alpha)$ and of Diverse Anomalies

Once again, a fine distinction is needed. We can define an ELNA Hölder sequence $\tilde{H}_k$ for the sequence of low-frequency measures $\tilde{\mu}_k$, and a Hölder for the full measure μ. Writing $-\log_b M = V$, our box of base b yields the ELNA Hölder sequence

$$\tilde{H}_k = \frac{\log \tilde{\mu}_k(dx)}{\log(dx)} = \frac{1}{k}[-\log_b M(\beta_1) - \log_b M(\beta_1, \beta_2) \cdots] = \frac{1}{k}$$

Thus, $\tilde{H}_k$ is simply the average of k independent identically distributed random variables of the form V. The ELNA Hölder is $\tilde{H} = \lim_{k\to\infty} H_k$.

If $p(v)$ is the probability density of V and $p_k(v)$ the probability density of $\sum V$, the probability density of $\tilde{H}_k = (1/k)\sum V$ is $kp_k(k\alpha)$. This last density serves to define the ELNA dimensional sequence, and the ELNA dimension $f(\alpha)$ is then obtained by taking a limit for $k \to \infty$. §5.5 will discuss the form of $f(\alpha)$ in a fairly general context. As preparation, it is useful to mention a few easy-to-obtain functions $f(\alpha)$ that exemplify the "nonanomalous case" and two basic "anomalies."

5.4.1. *The Very Special Case Corresponding to Birth-or-Extinction Dusts*

The function f defined in §2.5 is a very special case of the $f(\alpha)$ we are now discussing. In that case, $f(\alpha)$ is only defined for $\alpha = D$ or $\alpha = \infty$ and satisfies $f(D) = D$ and $f(\infty) = 1$.

5.4.2. *The Nonanomalous Case*

It corresponds to $f \geq 0$ and $\alpha_{\min} > 0$, hence, $\alpha_1 > 0$. These conditions are satisfied by the binomial multifractal of §5.1 or the multinomial multifractal of §5.2. At one time, the opinion has spread that this is the "normal" form of $f(\alpha)$ and that all other forms, hence the inequalities $\alpha_{\min} < 0$ and $f_{\min} < 0$, are "anomalous." But they are not, as shown by the following two explicit examples.

5.4.3. *An Example of the $f_{\min} < 0$ Anomaly, Combined with a Nonanomalous $\alpha_1 > 0$*

An example is described in [15]. When M is uniformly distributed over [0, 1], exact elementary calculations yield

$$\tau(q) = \log_2(q+1) - 1,$$

and, denoting $\log_2 e$ by λ,

$$f(\alpha) = \lambda + 1 + \log_2(\alpha/\lambda) - \alpha.$$

In this example, $f_{\min} = -\infty$ and $\alpha_{\min} = 0$.

5.4.4. *An Example of the $\alpha_{\min} < 0$ Anomaly, Combined, or Not, With the $\alpha_1 < 0$ Anomaly*

Negative α's between 0 and $\alpha_{\min} < 0$ must be ELNA Hölders; they cannot be fine Hölder exponents. Since $f(\alpha) < \alpha$, the $\alpha_{\min} < 0$ anomaly implies the $f_{\min} < 0$ anomaly, but goes farther. In an example described in [18], V is Gaussian, hence μ is lognormally distributed. Then the expression of $f(\alpha)$ in terms of α_0 is

$$f(\alpha) = 1 - \frac{(\alpha - \alpha_0)^2}{4(\alpha_0 - 1)};$$

the graph of $f(\alpha)$ is a parabola, $f_{\min} = -\infty$, $\alpha_{\min} = -\infty$, and $\alpha_1 = -\alpha_0 + 2$.

In addition to the $\alpha_{\min} < 0$ anomaly, and the $f_{\min} < 0$ anomaly, this measure may examplify the $\alpha_1 > 0$ anomaly. It occurs when the graph of $f(\alpha)$ is a very flat parabola with $1 < \alpha_0 < 2$, i.e., M is a very long-tailed lognormal.

5.5. Random Multiplicative Measures: The ELNA Codimensions, the Functions $f_k(\alpha)$ and the Hölder Spectrum $f(\alpha)$ as the Limit for $k \to \infty$ of the Distribution of the ELNA Hölder $\tilde{H}$

Section 6.4 showed $\tilde{H}_k$ to be average of random variables. In searching for its asymptotic distribution, standard limit theorems of probability are shown to be of limited usefulness. One needs the Cramèr theorem, which yields a population function $\tilde{f}(\alpha)$ relative to the multiplier M. Given that convergence of random variables can be defined in different ways, one requires a multiplicity of different limit theorems, each true on its own terms.

The law of large numbers tells us that if $E\tilde{H} < \infty$, then $\tilde{H}$ converges to $E\tilde{H}$, which implies that $C(\alpha)$ has its maximum for $\alpha_0 = E\tilde{H}$.

The central limit theorem tells us that as long as $E\tilde{H}^2 < \infty$, the graph of $C(\alpha)$ is parabolic in the immediate neighborhood of $\alpha_0 = E\tilde{H}$.

In the multifractal context, these results give too little weight to α's far from α_0. In order to determine $C(\alpha)$, one needs a very different limit.

Harald Cramèr's "large deviations theorem" As $k \to \infty$,

$$-\lim_{k\to\infty} \frac{1}{k} \log_b \text{ (probability density of } \alpha) \text{ exists and defines a function } C(\alpha).$$

The Cramèr theory shows that the functions $f(\alpha) = 1 - C(\alpha)$ and $\tau(q) = -1 - \log_b EM^q$ are linked by the Legendre and inverse Legendre transforms. The best proof of Cramèr's result, in Daniels [ ], uses the steepest descent argument; it has been rediscovered repeatedly by physicists interested in multifractals. I first used [1] and [2] where the Cramèr theory had barely started being developed, and was rather obscure among physicists and even among mathematicians.

5.5.1. *Comment on the Fact That This Approach Defines $\tau(q)$ via the Expectation EM^q, rather than through the partition function*

Readers acquainted with the physicists' writings on multifractals (Section 5.6.3) are accustomed to seeing $\tau(q)$ defined via a "partition function," which is a sample moment. In the simplest cases, the two definitions are equivalent. But the cases that involve negative dimensions are precisely those for which the two definitions differ, as we shall see.

5.6. Three Variants of the Hölder Spectrum $f(\alpha)$ of a Random Multifractal Measure and Three Views of the Legendre Formalism

The "Hölder spectrum" function $f(\alpha)$ is the "signature" of a multifractal measure. Different authors obtain it by three distinct arguments.

In two widely known approaches to multifractality (§§5.6.2 and 5.6.3), α is a classical Hölder exponent, hence $\alpha > 0$, and f is a Hausdorff–Besicovitch dimension, hence $f \geq 0$. But in my original approach of 1974, now being restated more carefully and forcefully, §5.4 introduces α as the value of ELNA Hölder $\tilde{H}$, and allows $\alpha < 0$. Moreover, f is an ELNA dimension that satisfies $f > 0$ for some $\alpha > 0$ and $f < 0$ for other α (which may be of either sign). One can write

$f(\alpha) = f^+(\alpha) + f^-(\alpha)$, where $f^+ = f$ if $f \geq 0$ and $f^+ = 0$ if $f < 0$; this implies that $f^- = f$ if $f < 0$ and $f^- = 0$ if $f < 0$. Using this notation, we can characterize the approach to multifractals via the Hausdorff–Besicovitch dimension as not yielding the whole function $f(\alpha)$ but only its positive part $f^+(\alpha)$.

5.6.1. The Approach to $f(\alpha)$ via Probabilities, ELNA Hölders, and ELNA Dimensions; the Legendre Formalism Obtained By Applying the Cramèr Theory of Large Deviations to the Coarse-Grained Measures

This approach obtains max $f(\alpha) - f(\alpha)$ as the limit of a certain probability density that has been plotted in properly rescaled logarithmic coordinates. This involves the following steps. (i) Coarse-grain μ by averaging it over boxes of side ε. (ii) Form the coarse ELNA Hölder α. (iii) Form the probability density p of α. (iv) Plot $\log p/\log \varepsilon$ as a function of α. One sees that the variable μ and then its probability density are replaced by properly rescaled logarithms. In this approach, both positive and negative α's and f's are needed. The positive f's serve to characterize the variability of the distribution of the fine-grained μ on its support. The negative f's serve to characterize the randomness in the distribution of coarse-grained μ.

5.6.2. The Approach Through the Hausdorff–Besicovitch Dimension

An alternative definition of $f(\alpha)$ involves almost-sure properties of a fine-grained measure defined for continuous variables. Given $\alpha > 0$, denote by S_α the random Hölder isoset, defined as made of the points x where the classical Hölder exponent satisfies $H = \alpha$. As a corollary of the Mandelbrot–Kahane–Peyrière theorem, one can prove that the Hausdorff–Besicovitch dimension $D_{\mathrm{HB}}(s_\alpha)$ takes the same value for almost all the sample measures of μ and that $D_{\mathrm{HB}}(s_\alpha) = f^+(\alpha)$. $f^+(\alpha)$, introduced in this fashion, but not $f(\alpha)$ itself, can be called "spectrum of singularities."

The properties ruled by $f^+(\alpha)$ are not only the almost-sure properties of the fine-grained μ but also the "typical" coarse-grained properties. Therefore, $f^+(\alpha)$ says nothing about the variability of coarse-grained samples.

5.6.3. The Partition Function Heuristic

This is the approach taken in [4] and [5]. For nonrandom measures with $f(\alpha) = f^+(\alpha)$, the partition function heuristic can be made rigorous and yields $f(\alpha)$, but for the random measures in which we are interested, this heuristic is helpless, and a crucial application is thoroughly misleading.

5.7. Properties of the Hölder Spectrum $f(\alpha)$ Obtained as an ELNA Dimension

The ELNA dimensional sequence concerns ELNA "Hölder isosets." These are sets defined by conditions of the form "$\tilde{H} = \alpha$," where the domain of the parameter α is an interval $[\alpha_{\min}, \alpha_{\max}]$. In all cases, $\alpha_{\max} > 0$. The possibility that $\alpha_{\max} = \infty$ is not discussed in this paper but is not excluded; it characterizes the Minkowski multifractal [19]. As for $\alpha_{\min}$, one can have either $\alpha_{\min} > 0$, or $\alpha_{\min} < 0$, not excluding $\alpha_{\min} = -\infty$, a limit that will be encountered in this paper.

The approach involves a sequence of ELNA dimensional functions to be denoted by $f_\varepsilon(\alpha)$. Their limit $f(\alpha)$ is cap convex (like $-x^2$), but there are cases (notably the Minkowski multifractal) for which the convexity of the $f_\varepsilon(\alpha)$ changes from cap to cup (like x^2) for some critical value of α.

The graph of $f(\alpha)$ is bounded by straight lines, three of which were mentioned in 5.1 in a special case.

A. $f(\alpha) \leq E$, where E is the dimension of the set that supports the measure; $f(\alpha) = E$ for $\alpha = \alpha_0$, with $\alpha_0 > 0$ in all cases.

B. $f(\alpha) \le \alpha$. This shows that 1 is the slope of a tangent drawn from the origin to the graph of $f(\alpha)$. The equality $f(\alpha) = \alpha$ holds for $\alpha = \alpha_1$, with either $\alpha_1 > 0$, or $\alpha_1 < 0$.

C. When $\alpha_{\min} < 0$, one has $f(\alpha) \le \alpha Q_{\text{crit}}$ where $1 < Q_{\text{crit}} < \infty$. This Q_{crit} is the slope of the tangent of slope > 1 that can in certain cases be drawn from the origin to the graph of $f(\alpha)$.

Actually, $\alpha_{\min} < 0$ introduces a mathematical complication that will turn out to be physically important: One must distinguish between the Hölder spectrum $f(\alpha)$ and a related but different function $\tilde{f}(\alpha)$. For $\alpha \ge \alpha_{\text{crit}}$, $f(\alpha) = \tilde{f}(\alpha)$, but for $\alpha < \alpha_{\text{crit}}$, $f(\alpha) = \alpha q_{\text{crit}}$ while $\tilde{f}(\alpha) < \alpha q_{\text{crit}}$. One deals with approximating measures $\tilde{\mu}_k$ that tend to a limit μ, while their ELNA dimensional sequence $\tilde{f}_k(\alpha)$ tends to a limit $\tilde{f}(\alpha)$ that differs from the $f(\alpha)$ function of μ. The critical Q_{crit} and the inequality $\tilde{f}(\alpha) < f(\alpha)$ both involve negative ELNA dimensions and ELNA Hölders. It may therefore come as a surprise that the value of q_{crit} is of directly observable physical significance.

D. Finally, a basic alternative is conveniently expressed by $f(\alpha) \ge f_{\min}$, where $f_{\min} = \inf f(\alpha)$ can be either ≥ 0 or < 0. For certain purposes, it is necessary to distinguish between $\inf\{f(\alpha); \alpha < \alpha_0\}$ and $\inf\{f(\alpha); \alpha > \alpha_0\}$.

When $f(\alpha)$ is introduced as an ELNA dimension, it is possible to have $\alpha_1 < 0$, $f_{\min} < 0$; $\alpha_{\min} < 0$, and each of these inequalities has interesting consequences.

Finer results in the case where $f_{\min} < 0$ hinge upon further fine geometric properties of the graph of $f(\alpha)$. Denote by $[\alpha^+_{\min}, \alpha^+_{\max}]$ the interval where $f(\alpha) \ge 0$. Write $q_{d,\max} = f'(\alpha^+_{\min})$ and $q_{d,\min} = f'(\alpha^+_{\max})$. The interval $[q_{d,\min}, q_{d,\max}]$ will be investigated in §6.4. It will be shown in some cases (and conjectured in others) to be the interval of values of q for which the measure μ^q does not extinguish.

Denote by $\alpha_{\infty,\min}$ and $\alpha_{\infty,\max}$ the solutions of the equation

$$\alpha - \frac{f(\alpha)}{f'(\alpha)} = \alpha_{\min} \quad \text{and} \quad \alpha - \frac{f(\alpha)}{f'(\alpha)} = \alpha_{\max},$$

and write

$$f'(\alpha_{\infty,\min}) = q_{\infty,\max} \quad \text{and} \quad f'(\alpha_{\infty,\max}) = q_{\infty,\min}.$$

The interval $[q_{\infty,\min}, q_{\infty,\max}]$ will be seen in §6.5 to be the interval of values of q for which the partition function $\chi(q, \varepsilon)$ has finite moments of all orders and, therefore, can be estimated reliably.

When $\alpha_{\min} = -\infty$, one can show that $\alpha_{\infty,\min} = \alpha_0$; when $\alpha_{\max} = \infty$, one has $\alpha_{\infty,\max} = \alpha_0$. Denote by $Q_{\text{crit}}(q)$ the largest value of Q such that $E\chi^Q(q, \varepsilon) < \infty$. As q varies from $q_{\infty,\max}$ to $q_{d,\max}$ or from $q_{\infty,\min}$ to $q_{d,\min}$, the value of $Q_{\text{crit}}(q)$ will be shown to decrease from ∞ to 1.

5.8. The Measure $\mu^{(q)}$ Constructed Using the Multiplier $M^{(q)} = M^q/bEM^q$, and the Graph of Its Hölder Spectrum $f(a, q)$

When the ELNA Hölder corresponding to M is $\tilde{H}$, the ELNA Hölder corresponding to $M^{(q)}$ is

$$\tilde{H}^{(q)} = q\tilde{H} - \log[bEM^q(dx)] = q\tilde{H} - \tau(q).$$

Therefore,

$$f(\alpha, q) = f\left(\frac{\alpha + \tau(q)}{q}\right).$$

To each q and p corresponds a characteristic exponent $\alpha_p^{(q)}$ that plays for $\mu^{(q)}$ the role that α_p plays for μ. For the purposes of §6, it is important to compare the main characteristics of $f(\alpha)$ and

$f(\alpha, q)$. The value of $f_{\min}$ is unchanged; therefore, a (non)anomalous $\tilde{f}(\alpha)$ yields a (non)anomalous $\tilde{f}(\alpha, q)$ for all q. As to $\alpha_{\min}$, its replacement $\alpha_{\min}(p)$ satisfies $\alpha_{\min}(q) = q\alpha_{\min} - \tau(q)$ for $q > 1$ and $\alpha_{\min}(q) = q\alpha_{\max} - \tau(q)$ for $q < 1$. Both expressions can be shown to be decreasing functions of $|q|$. Suppose $\alpha_{\min} > 0$ with $f_{p,\min} < 0$, and make q increase, starting with $p = 1$. First, $\alpha_{\min}(q)$ and later $\alpha_1(q)$ will change from positive to negative. Suppose $\alpha_{\min} < 0$, which implies $f_{p,\min} < 0$ with $\alpha_1^{(q)} > 0$; then, as q increases from 1, $\alpha_1^{(q)}$ eventually changes from positive to negative. Sections 6.1 and 6.2 were largely a restatement in terms of $f(\alpha)$ of well-known results from [10, 11, 7]. The present § moves on to the partition function $\mathrm{X}(q, d)$ defined in §5.2.2. For $q = 1$, $\mathrm{X}(1, dx) = \sum \mu_j(dx)$, that is $\mathrm{X}(1, dx) = \Omega$, where Ω is the random variable introduced in §5.3. For $q \neq 1$, all we know until now was a result derived in §5.2.2, that $E\mathrm{X}(q, dx) = (dt)^{\tau(q)}$. In the literature of physics [4, 5], $\mathrm{X}(q, dx)$ is casually assumed to behave like $E\mathrm{X}(q, dx)$. Furthermore, this literature bases the definition of $\tau(q)$, not on the behavior of an observable expectation, but of an observable sample value. This § will point out that the validity of this casual assumption depends on the sign of $\alpha_1^{(q)}$.

The case $\alpha_1^{(q)} > 0$. When $f_{\min} > 0$, this inequality holds automatically. When $f_{\min} < 0$, this inequality only holds. When $\alpha_1^{(q)} > 0$, we shall see that the sample Chi and its expectation follow the same scaling rule.

The case $\alpha_1^{(q)} < 0$. This inequality only occurs when $f_{\min} < 0$ and q is outside of the interval that corresponds to αs satisfying $f(\alpha) \geq 0$. When $\alpha_1^{(q)} < 0$, we shall see that the sample X and its expectation follow *distinct* scaling rules.

The argument presented in this § is, unfortunately, limited to the conservative multiplicative multifractals, for which Ω satisfies $\Omega \equiv 1$. In that case, the normalized parition function $\mathrm{X}(q, dx)(dx)^{-\tau(q)}$ is simply the measure of [0, 1] when the multiplier M is replaced by $M^{(q)}$. As a result, the study of $X(q, dx)(dx)^{-\tau(q)}$ that follows momentarily, reduces to applying the results of §§6.1 to 6.3 to the measure $\mu^{(q)}$.

The measure is neither conservative (§5.3.1), nor canonical (§5.3.2). However, [3] showed that the extinction-or-survival criterion continues unchanged. It is based on the sign of $\alpha_1^{(q)}$, and involves a limit variable $\Omega^{(q)}$ that satisfies $E\Omega^{(q)} = 1$. It is conjectured that that same results hold where Ω is nondegenerate; it is hoped that a reader will soon prove this conjecture.

6. Limit Properties of Sample Measures and of Sample Partition Functions: The Extinction-or-Survival Criterion and the Critical Exponents $Q_{\mathrm{crit}}(q)$

Thus far we examined weak properties of $f(\alpha)$, the *probability distribution* of the measure produced by a multiplicative process for high but finite values of k. A harder task concerns strong or almost-sure properties, that is, the role $f(\alpha)$ plays in characterizing the *limit measure* itself.

Once again, the most basic features of a function $f(\alpha)$ are the signs of α_1, $f_{\min}$, and $\alpha_{\min}$. The shape of $f(\alpha)$ allows for three alternatives already considered in the literature; together, they define four cases, which will be sketched in §§6.1 to 6.3. There §6.4 and 6.5 will describe finer new alternatives.

6.1. Alternatives Based on the Sign of α_1 : The Extinction or Survival Criterion

The case of survival. When $\alpha_1 > 0$, the measure $\tilde{\mu}(.)$ survives with positive probability and converges to a limit μ. This case includes the birth-or-extinction dusts for which $D > 0$ means that $f \geq 0$. But this case is compatible with $\alpha_{\min} < 0$ and $f_{\min} < 0$.

The case of extinction. When $\alpha_{\text{I}} < 0$, the measure is almost-surely extinguished. Necessary conditions for $\alpha_{\text{I}} < 0$ are $\alpha_{\min} < 0$ and $f_{\min} < 0$; for example, this case includes the birth-or-extinction dusts for which $D < 0$. The fact that $f(\alpha)$ includes a portion where $f(\alpha) > 0$ does not prevent extinction. An extinguished measure can be revived by either embedding or supersampling with $E - 1 \geq -\alpha_{\text{I}}$.

How does extinction manifest itself concretely? In the birth-and-extinction fractals of §2, extinction means that, a.s., the measure of every interval vanishes after a variable but a.s. finite number of stages. For multifractals that satisfy $\Pr\{M > 0\} = 1$, the measure of an interval tends to 0 but remains positive for all finite k. Probabilists always hope that quantities that tend to 0 or ∞ can be renormalized to tend to a limit that is > 0 and $< \infty$. But in the present case this is impossible.

6.2. Alternative Based on the Sign of $f_{\min}$, Assuming Survival

The case of survival with $f_{\min} \geq 0$. The best known example is that of $\cap$-shaped $f(\alpha)$, which is cap convex over a bounded interval $0 < \alpha_{\min} < \alpha < \alpha_{\max} < \infty$. When $f_{\min} \geq 0$, which implies $\alpha_{\min} \geq 0$, $f(\alpha)$ is almost surely the Hausdorff–Besicovitch dimension D_{HB} of the Hölder isoset of points where the measure μ has a fine Hölder equal to α. Furthermore, the whole function $f(\alpha)$ can also be obtained by the Hausdorff path or the partition function path. The measure may be random, but, if so, its randomness must be viewed as "irrelevant."

The case of survival with $f_{\min} < 0$. Then $f(\alpha) = f^{+}(\alpha) + f^{-}(\alpha)$, with $f^{-}(\alpha) \neq 0$ for some α. The function $f^{+}(\alpha)$ can be interpreted as an a.s. Hausdorff-Besicovitch dimension. The function $f^{-}(\alpha)$ can be shown to describe the rules of randomness. It will be seen in §6.4 that $f^{-}(\alpha)$ disappears asymptotically as one moves from coarse to fine graining; but it can be interpreted by embedding or supersampling.

6.3. Alternative Based on the Sign of $\alpha_{\min}$, when $\alpha_1 > 0$: the Moments of Ω

When $\alpha_{\min} < 0$, part of the function $f(\alpha)$ corresponds to extraordinarily unlikely values of μ. They could only be observed with the help of massive supersampling, and one would expect the value of $\alpha_{\min}$ to be of small importance. To show that this expectation would be ill-advised, it suffices to consider the distribution of the total measure in the interval $[0, 1]$, which §5.3.2 denotes by Ω. By itself, $f(\alpha)$ fails to determine the distribution of Ω, but the sign of $\alpha_{\min}$ suffices to describe an important aspect of its tail behavior.

The case of survival with $f_{\min} < 0$ *but* $\alpha_{\min} > 0$. In this case, $E\Omega^{q} < \infty$ for all q.

The case of survival with $f_{\min} < 0$ *and* $\alpha_{\min} < 0$. The inequality $E\Omega^{q} < \infty$ holds if and only if $q < Q_{\text{crit}}(1)$, where the quantity $Q_{\text{crit}}(1)$ is obtained as follows: When $\alpha_{\min} < 0$, two tangents to the graph of $f(\alpha)$ go through the origin: one is of slope 1, and the other is of slope $Q_{\text{crit}}(1)$. This is very much in the spirit of supersampling.

6.4. In the Case of Survival with $f_{\min} < 0$, the Behavior of the Partition Function $\chi(q, \epsilon)$ Involves a Finer Subdivision Defined Separately for Each q

The case $\alpha_{\text{I}}^{(q)} > 0$. The extinction-or-survival criterion tells us that the sequence of measures $\mu_{k}^{(q)}(dx)$ converges to a nondegenerate random limit, which is the variable $\Omega^{(q)}$ corresponding to the multiplier $M^{(q)}$. Hence, $\chi(q, dx) = b^{k\tau(q)}\Omega^{(q)}$. It follows that the physicists' heuristic is on the right track in this case. The partition function $\sum \chi(q, dx)$ is equal to the population expectation $b^{k}E\mu^{q}(dx)$, multiplied by the random factor $\Omega^{(q)}$, which satisfies $E\Omega^{(q)} = 1$. That is, the sample value and the expectation follow the same analytic scaling rule. It seems safe to conjecture (but

I have not attempted a full proof) that, as $d \to 0$, the function $\tau(q)$ follows the same analytic scaling rule estimated from the sample partition function, will converge a.s. to the function $\tau(q)$ defined in §5.2.2. The issue of convergence is taken up again in §6.5.

The case $\alpha_1^{(q)} < 0$. Now, the extinction-or-survival criterion tells us that the sample sum of $\mu_k^{(q)}$ converges a.s. to zero. *The unexpected consequence is that the geometric scaling property of* μ *fails to imply that the sample sum* $\chi(q, dx)$ *is analytically scaling.*

However, there is more to say. An elementary but lengthy argument (for which we have no space) shows that (i) for $q < q^+_{\min}$ the sample sum is asymptotically scaling for $dx \to 0$, with the scaling exponent $\tilde{\tau}(q) = q\,\alpha^+_{\min}$, while (ii) for $q > q^+_{\max}$, the asymptotic scaling exponent is $\tilde{\tau}(q) = q\,\alpha^+_{\max}$. For $q_{\min} < q < q_{\max}$, let us define $\tilde{\tau}(q)$ as $\tilde{\tau}(q) = \tau(q)$. The Legendre transform of this function $\tilde{\tau}(q)$ happens to be simply the function $f_+(\alpha)$. The function $f_{-(\alpha)}$ plays no role whatsoever in the study of χ. However, this $\tilde{\tau}(q)$ is only defined asymptotically for $dx \to 0$. This may explain the anecdotal reports that values $f(\alpha) < 0$ have been obtained from the partition function.... and could not be accounted for.

We have carried out extensive computer simulations to test this prediction; the results are very striking. The reader is advised to test them again, for example, using the measures described in [15, 16].

6.5. Addendum: Weak Limit Theorems Rulings the Estimation of $\tau(q)$ from Samples

It is convenient to misplace here some comments on a topic that concerns a weak, rather than a strong, limit theorem. Since $E\Omega^{(q)} = 1$, it was conjectured in Section 6.4 that the $\hat{\tau}(q)$ estimated from a sample partition function converges a.s. to the $\tau(q)$ defined in §5.2.2. The speed of convergence surely depends on the moments of $\Omega^{(q)}$, more precisely, on the order of the highest finite moment, which we denote by $Q_{crit}(q)$. The sample fluctuations of the estimate $\hat{\tau}(q)$ are lowest where $Q_{crit}(q) = \infty$. The domain of attraction of the random variable X is Gaussian when $Q_{crit}(q) > 2$, but Levy stable when $Q_{crit}(q) < 2$. The difference must be reflected in the speed of convergence of $\hat{\tau}$ to τ.

How does the variability in the speed of convergence of τ translate into the speed of convergence of $f(\alpha)$? When $f_{\min} < 0$, the estimate of f is very precise near $\alpha = \alpha_0$ and becomes gradually worse as one approaches the values of α where $f(\alpha) = 0$. When supersampling is used to extend the estimated f to negative values, the reliability of the estimates becomes even worse.

The assertions made informally in this § deserve a careful and rigorous examination.

7. Asymptotics, Information-Theoretical and Thermodynamic Formalisms; Their Meaning and Limitations

The unavoidable introduction of ELNA dimension and an ELNA Hölder and the recognition of the special roles of dimension and Hölders that can be negative, marks a turn in fractal geometry. It affects the profound and hard-to-escape connections that link the geometric notion of dimension to the notion of entropy in statistical thermodynamics as well as the related notion of (Shannon) information.

A first link is that every approach to multifractals involves at least some manipulations that are more or less knowingly or directly borrowed from thermodynamics. For example, the probabilists who work with large deviations view the Cramèr theory as being thermodynamical. Also, the presentation of the multinomial multifractals (§5.2) follows the Lagrange multipliers path toward thermodynamics. Finally, the method of steepest descents is used widely in the Darwin–Fowler approach of thermodynamics, in some presentations of the Cramèr theory [ ], and in the partition function heuristics [4, 5].

Connections between multifractal and thermodynamics are equally obvious when, instead of the logical deductions, one examines the final expressions yielded by those deductions. The expression $-\sum p_j \log p_j$ is a Boltzman's entropy or Shannon's information, and it occurs in theorems of Besicovitch, Eggleston, and Volkmann. These theorems also provide a hint concerning the similarity dimension $\log N/\log b = \log_b N$ of a self-similar fractal set decomposable into N parts reduced in the ratio $1/b$. Indeed, $\log N$ is formally a Gibbs's entropy.

What about the nonaveraged expression of the form $-\log p$, which is used in this paper to define ELNA dimension? A nonaveraged Boltzmann entropy is not used in thermodynamics, but a nonaveraged Shannon information enters in some approaches to information theory. Thus, given J letters of respective probabilities p_j, some writers restrict their attention to the (average) information per letter, $-\sum p_j \log p_j$. But other writers also pay attention to the information $-\log p_j$ corresponding to the letter of index j. Since accepting or dismissing the nonaveraged information brings substantial conceptual changes, is it useful to apply the terms "information theoretical" or "thermodynamical" to those cases when $D < 0$?

Asymptotic descriptions in fractal geometry are associated with probability limit theorems, and one might argue that the notion of thermodynamic property should cover everything that deals with asymptotics. This would describe the whole function $f(\alpha)$ as being thermodynamic, including the part in which $f(\alpha) < 0$ and even $\alpha < 0$.

In my opinion, this view would be overly inclusive. The notion of "thermodynamic property" should be reserved to considerations that include *strong* limit theorems, and involve the resulting fluctuation free asymptotic physics. This restricts $f(\alpha)$ to the values $f(\alpha) > 0$. The values $f(\alpha) < 0$ solely concern Cramèr's *weak* limit theorem, hence, the nature of fluctuations as one approaches a degenerate asymptotia. This is why the actual estimation of $f(\alpha) < 0$ must rely upon "supersamples" or other methods that involve preasymptotics rather than asymptotics.

Using different words, a strict view of the thermodynamical interpretation leaves no room for negative dimensions. However, the thermodynamic description based on probability limit theorems may fail to be sufficiently precise, meaning that those limit theorems destroy valuable information. A fuller description necessitates a knowledge of effects that are *preasymptotic*. One ought to reserve the term "thermodynamic" for the fine-grained and partition-functional properties. The coarse-grained properties that go beyond the thermodynamics are not macroscopic but "mesoscopic." Negative dimension is a mesoscopic notion.

References

[1] Chernoff, H., 1952. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat.* **24**, 493–507.

[2] Daniels, H. E., 1954. Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25**, 631–649.

[3] Durrett, R., and Liggett, T. M., 1983. Fixed points of the smoothing transformation. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*. **64**, 275–301.

[4] Frisch, U., and Parisi, G., 1985. Fully developed turbulence and intermittency. Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics (M. Ghil, ed. North-Holland, Amsterdam, 84–86.

[5] Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., and Shraiman, B. I., 1986. Fractal measures and their singularities: The characterization of strange sets. Phys. Rev. A. **33**, 1141–1151.

[6] Guivarch, Y., 1987. Remarques sur les solutions d'une equation fonctionnelle non-lineaire de Benoit Mandelbrot. C. R. *Acad. Sci. Paris Ser. I Math.* **305**, 139–142.

[7] Kahane, J. P., and Peyrière, J., 1976. Sur certaines martingales de B. Mandelbrot. Adv. Math. **22**, 131–145.

[8] Mandelbrot, B. B., 1969. Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International Economic Review:* **10**, 82–111.

[9] ———, 1972. Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. Statistical Models and Turbulence, (M. Rosenblatt and C. Van Atta, eds.), Springer-Verlag, New York, 333–351.

[10] ———, 1974. Intermittent turbulence in self similar cascades; divergence of high moments and dimension of the carrier. J. Fluid Mech. **62**, 331–358.

[11] ———, 1974. Multiplications aléatoires itérées et distributions invariantes par moyenne pondérée aléatoire. C. R. *Acad. Sci. Paris Ser. I Math.* **278a**, 289–292 and 355–358.

[12] ———, 1982. *The Fractal Geometry of Nature*. W. H. Freeman, New York.

[13] ———, 1984. Fractals in physics: squig clusters, diffusions, fractal measures and the unicity of fractal dimension. *J. Stat. Phys.* **34**, 895–930.

[14] ———, 1989. Multifractal measures, especially for the geophysicist. *Pure Appl. Geophysics.* **131**, 5–42.

[15] ———, 1989. A class of multinomial multifractal measures with negative (latent) values for the "dimension" $f(\alpha)$ *Fractals' Physical Origin and Properties*. (L. Pietronero, ed.). Plenum Press, New York, 3–19.

[16] ———, 1990. Negative fractal dimensions and multifractals. *Physica A*. **163**, 306–315.

[17] ———, 1990. New "anomalous" multiplicative multifractals: left-sided $f(\alpha)$ and the modeling of DLA. *Physica A*. **163**, 95–111.

[18] ———, 1990. Limit lognormal multifractal measures, Landau Memorial Conference. *Frontiers of Physics* (E. A. Gotsman, Y. Ne'eman, and A. Voronel, eds.). Pergamon, New York, 309–340.

[19] ———, 1993. The Minkowski measure and multifractal anomalies in invariant measures of parabolic dynamic systems. *Chaos in Australia* (Sydney, 1990). Edited by Gavin Brown & Alex Opie. Singapore: World Publishing, 1993, 83-94. Slightly ediged reprint: *Fractals and Disordered Systems*. Second edition. Edited by Armin Bunde & Shlomo Havlin, New York: Springer 1995.

[20] ———, 1995. Introduction to fractal sums of pulses. *Lévy Flights and related Phenomena in Physics* (Nice, 1994). Edited by George Zaslawsky, Michael F. Shlesinger & Uriel Frisch. New York: Springer.

Received January, 1994

Mathematics Department, Yale University, New Haven, Connecticut 06520-8283