Journal of Statistical Physics, Vol. 110, Nos. 3–6, March 2003 (© 2003)

Multifractal Power Law Distributions: Negative and Critical Dimensions and Other "Anomalies," Explained by a Simple Example

Benoit B. Mandelbrot[1]

Received March 18, 2002; accepted August 6, 2002

"Divergence of high moments and dimension of the carrier" is the subtitle of Mandelbrot's 1974 seed paper on random multifractals. The key words "divergence" and "dimension" met very different fates. "Dimension" expanded into a multifractal formalism based on an exponent α and a function $f(\alpha)$. An excellent exposition in Halsey *et al.* 1986 helped this formalism flourish. But it does not allow divergent high moments and the related inequalities $f(\alpha) < 0$ and $\alpha < 0$. As a result, those possibilities did not flourish. Now their time has come for diverse reasons. The broad 1974 definitions of α and f allow $\alpha < 0$ and $f(\alpha) < 0$, but the original presentation demanded to be both developed and simplified. This paper shows that both multifractal anomalies occur in a very simple example, which has been crafted for this purpose. This example predicts the power law distribution. It generalizes α and $f(\alpha)$ beyond their usual roles of being a Hölder exponent and a Hausdorff dimension. The effect is to allow either f or both f and α to be negative, and the apparent anomalies are made into sources of new important information. In addition, this paper substantially clarifies the subtle way in which randomness manifests itself in multifractals.

KEY WORDS: Multifractals; power-law distribution; negative dimensions; critical dimensions; anomalies; two-valued canonical measure.

There are at least two reasons why the view of multifractals common among physicists demands to be broadened. A first reason is that Duplantier 1999 obtained $f(\alpha)$ explicitly for the harmonic measure around a number of diverse physical clusters (Brownian, percolation, and Ising.) For small but positive values of α, and for large enough values, Duplantier's function $f(\alpha)$ is negative.

Had the measures in question been non-random, $f(\alpha)$ would have been a fractal dimension. Therefore $f(\alpha) < 0$ would have been "seriously

[1] Sterling Professor of Mathematical Sciences, Yale University New Haven, Connecticut 065020-8283; e-mail: mandelbrot@us.ibm.com

0022-4715/03/0300-0739/0 © 2003 Plenum Publishing Corporation

anomalous," in fact, an impossibility according to the otherwise excellent exposition in Halsey *et al.* 1986, and the considerable literature that followed along the same lines. However, random clusters create random measures, and $f(\alpha) < 0$ is allowed and explained in the earlier theory of random multifractals proposed in Mandelbrot 1974ab and made mathematically rigorous in Kahane and Peyrière 1976.

Unfortunately, that original theory was overly general and the original papers, having appeared "before their time," were difficult. The topic remained little-known and/or understood and begs for a simple account.

A second reason to broaden the common view of multifractals involves the fluctuations of certain physical phenomena as well as of financial prices. My latest model (Mandelbrot 1977, 2001abcd), which has since become accepted in the finance literature, consists in a Brownian motion (Wiener or fractional) that does not proceed in clock time, but rather in a "trading time" that is a multifractal function of "clock time." I had discovered in 1962 that actual price increments follow a power law distribution. In order for the present model to include this feature, the multifractal measure in an interval must itself follow a power law of the form $\Pr\{\text{measure} > \mu\} = \mu^{-q_{\text{crit}}}$, with $1 < q_{\text{crit}} < \infty$. Being concerned with nonrandom measures, Halsey *et al.* had no prediction to make about probability distributions. But not only did Mandelbrot 1974ab allow the power-law, but obtained it as a consequence $f(\alpha) < 0$ and $\alpha < 0$. From this double anomaly, it follows that the equation $\tau(q) = 0$ not only has a root $q = 1$ but also a second root that is q_{crit}.

The Goals of this Paper. This paper introduces and studies the "two-valued canonical measure," to be noted by TVCM. This is a new random multifractal that was specifically constructed to illustrate the subtitle of Mandelbrot 1974 in what seems to be the simplest fashion imaginable. This sharply defined goal explains why the references include few items from the immense literature on nonrandom multifractals or finance.

In the multifractals considered by Duplantier, the supports are random fractal sets, but in price variation, the support is the time axis. To simplify and avoid extraneous complications, this paper purposefully restricts itself to very special multifractals on the interval $[0, 1]$.

The Probability Distribution of the Overall Measure $\mu([0, 1]) = \Omega$. Of course, both $f(\alpha)$ and the random variable Ω are determined by the rule that defines the multifractal measure. However, knowing $f(\alpha)$ *does not suffice* to identify the distribution of Ω, not even to determine whether or not Ω is random. For example, we shall compare the Bernoulli binomial multifractal, which is non-random, and a very simple random variant, to be called "canonical," a special case of TVCM. The function $f(\alpha)$ is the same for both, but Ω is non-random (identically 1) for the original Bernoulli and random for the canonical variant. Insofar as all moments are finite, this Ω is "not very random." However, the only general conclusion from the multifractal formalism is that non-randomness, that is, $\Omega = 1$, requires $f(\alpha) > 0$. A fortiori, the distribution of Ω cannot be obtained without additional assumptions.

The simplest case is the "very random" one, insofar as the anomalies $\alpha < 0$ and $f(\alpha) < 0$ allow the tail probability $\Pr\{\Omega > \Omega\}$ to take the power law form $\Omega^{-q_{\text{crit}}}$. In other words, the apparent generality of Halsey *et al.* proves a drawback, while the more special nature of my original model proves an asset. Since 1974, that model has been generalized but became more complicated (see Mandelbrot 1995). However, the key facts are already present in the deliberately non-general example described in this paper.

Hölder Analysis of the Decomposition of a Measure Among Values of α. In order to echo "Fourier analysis," I describe the multifractal formalism as achieving "Hölder analysis." Fourier analysis decomposes a function into a (denumerably infinite) sum of components, each of fixed frequency. Hölder analysis decomposes a measure into a (continuously infinite) sum of measures, each characterized by a value of the Hölder exponent α. The latter notion is borrowed from 1870 mathematical esoterica and measures the strength of local "singularity" or "roughness". Each $\alpha > 0$ is encountered on a set of fractal dimension $f(\alpha) > 0$. The notions of α and f will be generalized by defining $f(\alpha)$ as $f(\alpha) = \rho(\alpha) + 1$, where $\rho(\alpha)$ is a suitably weighted but non-averaged logarithm of a probability.

All that we learn from the Hölder analysis of μ is how the total mass Ω is decomposed among the αs. When $f(\alpha)$ is only defined for $\alpha > 0$, all the moments of Ω are finite and one can write $q_{\text{crit}} = \infty$. When $f(\alpha)$ is defined for $\alpha < 0$, one finds $q_{\text{crit}} < \infty$. All that we have from the critical q_{crit} is how the value of the overall mass is distributed.

Hölder analysis is not directly affected by either the value or the distribution of Ω; it is ruled by a first truncated function $f^{+}(\alpha) =$ maximum of $\{0, f(\alpha)\}$. The randomness of Ω is not directly affected by the Hölder analysis and is largely ruled by a second truncated function $f^{-}(\alpha) =$ minimum of $\{0, f(\alpha)\}$. In the general case, the main constraint on $f(\alpha)$ is that it must be cap-convex. Hence the functions $f^{+}(\alpha)$ and $f^{-}(\alpha)$ are only weakly linked.

Hence, a striking fact is that, to some large extent, the values of Ω and q_{crit} are ruled by two distinct portions of a single generalized function $f(\alpha)$. By and large, Hölder analysis is ruled by the "observable" values of α, defined as those for which $f(\alpha) > 0$. The presence of large excursions in the distribution of Ω is ruled by the "latent" values, defined as those for which $f(\alpha) < 0$. And the power law follows from the presence of values of Ω that are so extraordinarily unlikely that a "sensible" heuristic could (or would) disregard them altogether. This is perhaps why their role remains so little known until now, and the goal of this paper is to provide a presentation that is both "sensible" and correct.

Thermodynamical Formalisms. In every approach to multifractals, the philosophy and the formalism were knowingly borrowed from statistical thermodynamics. For example, the Legendre transform entered into Mandelbrot 1974 indirectly, since the Cramer's theory of the probability of large deviations, on which I relied, is itself rooted in thermodynamical approximations. The formalism in Halsey *et al.* follows (without mentioning it explicitely) the derivation of thermodynamics by the Darwin–Fowler method of steepest descent. A large literature expanded on the thermodynamic connection, but is beyond the scope of this paper.

Beyond Formalisms, a Deep Contrast Between Gases and Multifractals. Analogies can be misleading. Thermodynamics teaches that zero probability events have no observable effect and can be neglected. To the contrary, every theory of multifractals deals with "degrees of thinness" of a set. It deals with diverse properties that one can readily observe—either directly or by close consequences—but that are ruled by events of probability zero. Those properties are known to include sets of dimension < 1 and zero measure corresponding to the "information dimension" and the values of $f(\alpha)$ for $\alpha \neq \alpha_0$. This paper shows that my 1974 theory of multifractals goes one step further. It deals with "degrees of emptiness" of a set that, among others, concern the existence of a finite critical exponent q_{crit}.

Examples of Degrees of Emptiness. The Brownian boundaries studied in Duplantier 1999 provide a new but especially clearcut example of degrees of emptiness. The $4/3$ conjecture in Mandelbrot 1982, now proved, is that the boundary's dimension is $4/3$. The fact that there exist αs for which Duplantier's $f(\alpha) > 0$ expresses that certain properties are only observed on subsets of the boundary of dimension $< 4/3$. Other theorems assert that some properties are almost surely encountered nowhere on the boundary. "Nowhere" means on an empty set. To consider all empty sets as identical used to suffice, but no longer. A new challenge has arisen, to specify "how empty." The answer is the value of a negative probability.

A Fundamental Procedure that Happens to Be Shared by Modern Statistical Physics and the Approach of Mandelbrot 1974. The measure on which this paper focuses is defined on the one-dimensional "time" axis, more precisely, for subintervals of $[0, 1]$. But in order to fully understand negative dimensions and $\alpha < 0$, one must think simultaneously of all measures with a two-valued multiplier that act in d-dimensional spaces with $d > 1$. The dimension d need not be an integer, and one must consider an infinity of distinct critical values of the dimension d.

A technically and intellectually important feature of modern statistical physics falls into three parts. Firstly, a system is not studied in a single space dimension, such as 1 or 3, but in all possible embedding dimensions. Secondly, formal arguments allow the embedding dimension not to be an integer. Thirdly, above some embedding dimension singled out as "critical," the systems' properties follow the so-called "mean-field" theory; below that dimension the properties are very different.

None of these three features belonged to the early theory (due to A. S. Besicovitch) of the Bernoulli binomial. However, quite independently of the developments in physics, all of those three features are present in Mandelbrot 1974. There, the application is turbulence and the first feature—simultaneous multiple embeddings—is intrinsic and unavoidable because a physical phenomenon that occurs in $d = 3$ must be followed through linear cuts with $d = 1$. Secondly, non-integer dimensions are not formal but have a completely rigorous mathematical meaning in terms of fractal dimensions. Thirdly, for each of several distinct "aspects" of TVCM, a certain "normal" behavior prevails when the embedding dimension exceeds the corresponding critical value, while an "anomalous" behavior prevails at low embedding dimensions. Compared to statistical physics, the multifractals' novelty is that the number of different "aspects" is infinite and so is the number of distinct critical dimensions. These points will be discussed elsewhere.

1. BACKGROUND: THE BERNOULLI BINOMIAL MEASURE AND TWO RANDOM VARIANTS: SHUFFLED AND CANONICAL

The prototype of all multifractals is nonrandom: it is a Bernoulli binomial measure. Its well-known properties are recalled in this section, then Section 2 introduces a random "canonical" version. Also, all Bernoulli binomial measures being powers of one another, a broader viewpoint considers them as forming a single "class of equivalence."

1.1. Definition and Construction of the Bernoulli Binomial

A Multiplicative Nonrandom Cascade. A recursive construction of the Bernoulli binomial measures involves an "initiator" and a "generator." The initiator is the interval [0, 1] on which a unit of mass is uniformly spread. This interval will recursively split into halves, yielding dyadic intervals of length 2^{-k}. The generator consists in a single parameter u, variously called *multiplier* or *mass*. The first stage spreads mass over the halves of every dyadic interval, with unequal proportions. Applied to $[0, 1]$, it leaves the mass u in $[0, \frac{1}{2}]$ and the mass v in $[\frac{1}{2}, 1]$. The $(k+1)$th stage begins with dyadic intervals of length 2^{-k}, each split in two subintervals of length 2^{-k-1}. A proportion equal to u goes to the left subinterval and the proportion v, to the right.

After k stages, let φ_0 and $\varphi_1 = 1 - \varphi_0$ denote the relative frequencies of 0's and 1's in the finite binary development $t = 0.\ \beta_1 \beta_2 \cdots \beta_k$. The "pre-binomial" measures in the dyadic interval $[dt] = [t, t + 2^{-k}]$ takes the value

$$\mu_k(dt) = u^{k\varphi_0} v^{k\varphi_1},$$

which will be called "pre-multifractal." This measure is distributed uniformly over the interval. For $k \to \infty$, this sequence of measures $\mu_k(dt)$ has a limit $\mu(dt)$, which is the Bernoulli binomial multifractal.

Shuffled Binomial Measure. The proportion equal to u now goes to either the left or the right subinterval, with equal probabilities, and the remaining proportion v goes to the remaining subinterval. This variant must be mentioned but is not interesting.

1.2. The Concept of Canonical Random Cascade and the Definition of the Canonical Binomial Measure

Mandelbrot 1974ab took a major step beyond the preceding constructions.

The Random Multiplier M. In this generalization, every recursive construction can be described as follows. Given the mass m in a dyadic interval of length 2^{-k}, the two subintervals of length 2^{-k-1} are assigned the masses $M_1 m$ and $M_2 m$, where M_1 and M_2 are *independent* realizations of a random variable M called multiplier. This M is equal to u or v with probabilities $p = 1/2$ and $1 - p = 1/2$.

The Bernoulli and shuffled binomials both impose the constraint that $M_1 + M_2 = 1$. The canonical binomial does not. It follows that the canonical mass in each interval of duration 2^{-k} is multiplied in the next stage by the sum $M_1 + M_2$ of two independent realizations of M. That sum is either $2u$ (with probability p^2), or 1 (with probability $2\ (1-p)\ p$), or $2v$ (with probability $(1-p^2)$.

Writing p instead of $1/2$ in the Bernoulli case and its variants complicates the notation now, but will soon prove advantageous: the step to the TVCM will simply consist in allowing $0 < p < 1$.

1.3. Two Forms of Conservation: Strict and on the Average

Both the Bernoulli and shuffled binomials repeatedly redistribute mass, but within a dyadic interval of duration 2^{-k}, the mass remains exactly conserved in all stages beyond the kth. That is, the limit mass $\mu(t)$ in a dyadic interval satisfies $\mu_k(dt) = \mu(dt)$.

In a canonical binomial, to the contrary, the sum $M_1 + M_2$ is not identically 1, only its expectation is 1. Therefore, canonical binomial construction preserve mass on the average, but not exactly.

The Random Variable Ω. In particular, the mass $\mu([0, 1])$ is no longer nonrandom and equal to 1. It is a basic random variable denoted by Ω and discussed in Section 3.

Within a dyadic interval dt of length 2^{-k}, the cascade is simply a reduced-scale version of the overall cascade. It transforms the mass $\mu_k(dt)$ into a product of the form $\mu(dt) = \mu_k(dt)\,\Omega(dt)$ where all the $\Omega(dt)$ are independent realizations of the same random variable Ω.

1.4. The Term "Canonical" Is Motivated by Statistical Thermodynamics

As is well-known, statistical thermodynamics finds it valuable to approximate large systems as juxtapositions of parts, the "canonical ensembles," whose energy only depends on a common temperature and not on the energies of the other parts. Microcanonical ensembles' energies are constrained to add to a prescribed total energy. In the study of multifractals, the use of this metaphor should not obscure the fact that the multiplication of canonical factors introduces strong dependence among $\mu(dt)$ for different intervals dt.

1.5. In Every Variant of the Binomial Measure, One Can View All Finite (Positive or Negative) Powers Together, as Forming a Single "Class of Equivalence"

To any given real exponent $g \neq 1$ and multipliers u and v corresponds a multiplier M_g that can take either of two values: $u_g = \psi u^g$ with probability p, and $v_g = \psi v^g$ with probability $1-p$. The identity $pu_g + (1-p)\,v_g = 1/2$ demands $\psi[pu^g + (1-p)\,v^g] = 1/2$, that is, $\psi = 1/[2\langle M^g \rangle]$. The expression $2\langle M^g \rangle$ will be generalized and encountered repeatedly, especially through the expression

$$\tau(q) = -\log_2[pu^q + (1-p)\,v^q] - 1 = -\log_2(2\langle M^q \rangle)$$

This is simply a notation at this point, but it will be justified in Section 4. It follows that $\psi = 2^{-\tau(g)}$, hence

$$u_g = u^g 2^{\tau(g)} \qquad \text{and} \qquad v_g = v^g 2^{\tau(g)}.$$

Assume $u > v$. As g ranges from 0 to ∞, u_g ranges from $1/2$ to 1, and v_g ranges from $1/2$ to 0; the inequality $u_g > v_g$ is preserved. To the contrary, as g ranges from 0 to ∞, $v_g < u_g$. For example, $g = -1$ yields

$$u_g = \frac{1/u}{1/u + 1/v} = v \qquad \text{and} \qquad v_g = \frac{1/v}{1/v + 1/v} = u.$$

Thus, inversion leaves both the shuffled and the canonical binomial measures unchanged. For the Bernoulli binomial, it only changes the direction of the time axis.

Altogether, every Bernoulli binomial measure can be obtained from any other as a reduced positive or negative power. If one agrees to consider a measure and its reduced powers as equivalent, *there is only one Bernoulli binomial measure*.

In concrete terms relative to non-infinitesimal dyadic intervals, the sequences representing $\log \mu$ for different values of g are mutually affine. Each is obtained from the special case $g = 1$ by a multiplication by g followed by a vertical translation.

1.6. The Full and Folded Forms of the Address Plane

In anticipation of TVCM, the point of coordinates u and v will be called the *address* of a binomial measure in a *full address space*. In that plane, the locus of the Bernoulli measures is the interval defined by $0 < v$, $0 < u$, and $u + v = 1$.

The *folded address space* will be obtained by identifying the measures (u, v) and (v, u), and representing both by the same point. The locus of the Bernoulli measures becomes the interval defined by the inequalities $0 < v < u$ and $u + v = 1$.

1.7. Alternative Parameters

In its role as parameter added to $p = 1/2$, one can replace u by the ("information-theoretical") fractal dimension $D = -u \log_2 u - v \log_2 v$ which can be chosen at will in this open interval $]0, 1[$. The value of D

characterizes the "set that supports" the measure. It received a new application in the new notion of multifractal concentration described in Mandelbrot 2001c. More generally, the study of all multifractals, including the Bernoulli binomial, is filled with fractal dimensions of many other sets. All are unquestionably positive. One of the newest features of the TVCM will prove to be that they also allow negative dimensions.

2. DEFINITION OF THE TWO-VALUED CANONICAL MULTIFRACTALS

2.1. Construction of the Two-Valued Canonical Multifractal in the Interval [0, 1]

The TVCM are called two-valued because, as with the Bernoulli binomial, the multiplier M can only take 2 possible values u and v. The novelties are that p need not be $1/2$, the multipliers u and v are not bounded by 1, and the inequality $u+v \neq 1$ is acceptable.

For $u+v \neq 1$, the total mass cannot be preserved exactly. Preservation on the average requires $\langle M \rangle = pu+(1-p)\, v = 1/2$, hence $0 < p = (1/2-v)/(u-v) < 1$.

The construction of TVCM is based upon a recursive subdivision of the interval [0, 1] into equal intervals. The point of departure is, once again, a uniformly spread unit mass. The first stage splits [0, 1] into two parts of equal lengths. On each, mass is poured uniformly, with the respective densities M_1 and M_2 that are independent copies of M. The second stage continues similarly with the interval [0, 1/2] and [1/2, 1].

2.2. A Second Special Two-Valued Canonical Multifractal: The Unifractal Measure on the Canonical Cantor Dust

The identity $\langle M \rangle = 1/2$ is also satisfied by $u = 1/2p$ and $v = 0$. In this case, let the lengths and number of non-empty dyadic cells after k stages be denoted by $\Delta t = 2^{-k}$ and N_k. The random variable N_k follows a simple birth and death process leading to the following alternative.

When $p > 1/2$, $\langle N_k \rangle = (\langle N_1 \rangle)^k = (2p)^k = (dt)^{\log(2p)}$. To be able to write $\langle N_k \rangle = (dt)^{-D}$, it suffices to introduce the exponent $D = -\log(2p)$. It satisfies $D > 0$ and defines a fractal dimension.

When $p < 1/2$, to the contrary, the number of non-empty cells almost surely vanishes asymptotically. At the same time, the formal fractal dimension $D = -\log(2p)$ satisfies $D < 0$.

2.3. Generalization of a Useful New Viewpoint: When Considered Together with Their Powers From $-\infty$ to ∞, all the TVCM Parametrized by Either p or $1-p$ Form a Single Class of Equivalence

To take the key case, the multiplier M^{-1} takes the values

$$u_{-1} = \frac{\frac{1}{u}}{2\left(\frac{p}{u}+\frac{1-p}{v}\right)} = \frac{v}{2(v+u)-1} \qquad \text{and} \qquad v_{-1} = \frac{u}{2(v+u)-1}.$$

It follows that $pu_{-1}+(1-p)\,v_{-1}=1/2$ and $u_{-1}/v_{-1}=v/u$. In the full address plane, the relations imply the following: (a) the point (u_{-1}, v_{-1}) lies on the extension beyond $(1/2, 1/2)$ of the interval from (u, v) to $(1/2, 1/2)$ and b) the slopes of the intervals from 0 to (u, v) and from 0 to (u_{-1}, v_{-1}) are inverse of one another. It suffices to fold the full phase diagram along the diagonal to achieve $v>u$. The point (u_{-1}, v_{-1}) will be the intersection of the interval corresponding to the probability $1-p$ and of the interval joining 0 to (u, v).

2.4. The Full and Folded Address Planes

In the full address plane, the locus of all the points (u, v) with fixed p has the equation $pu+(1-p)\,v=1/2$. This is the negatively sloped interval joining the points $(0, 1/2p)$ and $([1/2(1-p)], 0)$. When (u, v) and (v, u) are identified, the locus becomes the same interval plus the negatively sloped interval from $[0, 1/2(1-p)]$ to $(1/2p, 0)$.

In the folded address plane, the locus is made of two shorter intervals from $(1, 1)$ to both $(1/2p, 0)$ and $([1/2(1-p)], 0)$. In the special case $u+v=1$ corresponding to $p=1/2$, the two shorter intervals coincide.

Those two intervals correspond to TVCM in the same class of equivalence. Starting from an arbitrary point on either interval, positive moments correspond to points to the same interval and negative moments, to points of the other. Moments for $g>1$ correspond to points to the left on the same interval; moments for $0<g<1$, to points to the right on the same interval; negative moments to points on the other interval.

For $p \neq 1/2$, the Class of Equivalence of p Includes a Measure that Corresponds to u=1 and v=[$1/2-\min(p, 1-p)$]/[$\max(p, 1-p)$]. This novel and convenient universal point of reference requires $p \neq 1/2$. In terms to be explained below, it corresponds to $\alpha_{\min} = -\log u = 0$.

2.5. Background of the Two-Valued Canonical Measures in the Historical Development of Multifractals

The construction of TVCM is new but takes a well-defined place among the three main approaches to the development of a theory of multifractals.

General mathematical theories came late and have the drawback that they are accessible to few non-mathematicians and many are less general than they seem.

The heuristic presentation in Frisch and Parisi 1985 and Halsey *et al.* 1986 came after Mandelbrot 1974ab but before most of the mathematics. Most importantly for this paper's purpose, those presentations fail to include significantly random constructions, hence cannot yield measures following the power law distribution.

Both the mathematical and the heuristic approaches seek generality and only later consider the special cases. To the contrary, the approach in Mandelbrot 1974ab, the first historically, began with the careful investigation of a variety of special random multiplicative measures. I believe that each feature of the general theory continues to be best understood when introduced through a special case that is as general as needed, but no more. The general theory is understood very easily when it comes last.

In pedagogical terms, the "third way" associates each distinct feature of multifractals with a special construction, often one that consists of generalizing the binomial multifractal in a new direction. TVCM is part of a continuation of that effective approach; it could have been investigated much earlier if a clear need had been perceived.

3. THE LIMIT RANDOM VARIABLE $\Omega = \mu([0, 1])$, ITS DISTRIBUTION AND THE STAR FUNCTIONAL EQUATION

3.1. The identity $EM = 1$ Implies that the Limit Measure Has the "Martingale" Property, Hence the Cascade Defines a Limit Random Variable $\Omega = \mu([0, 1])$

We cannot deal with martingales here, but positive martingales are mathematically attractive because they converge (almost surely) to a limit. But the situation is complicated because the limit depends on the sign of $D = 2[-pu \log_2 u - (1-p)\, v \log_2 v]$.

Under the condition $D > 0$, which is discussed in Section 8, what seemed obvious is confirmed: $\Pr\{\Omega > 0\} > 0$, conservation on the average continues to hold as $k \to \infty$, and Ω is either non-random, or is random and satisfies the identity $\langle \Omega \rangle = 1$.

But if $D < 0$, one finds that $\Omega = 0$ almost surely and conservation on the average holds for finite k but fails as $k \to \infty$. The possibility that $\Omega = 0$

arose in mathematical esoterica and seemed bizarre, but is unavoidably introduced into concrete science.

3.2. Questions

(A) Which feature of the generating process dominates the tail distribution of Ω? It is shown in Section 5 to be the sign of $\max(u, v) - 1$.

(B) Which feature of the generating process allows Ω to have a high probability of being either very large or very small? Section 5 will show that the criterion is that the function $\tau(q)$ becomes negative for large enough q.

(C) Divide $[0, 1]$ into 2^k intervals of length 2^{-k}. Which feature of the generating process determines the relative distribution of the overall Ω among those small intervals? This relative distribution motivated the introduction of the functions $f(\alpha)$ and $\rho(\alpha)$, and is discussed in Section 7.

(D) Are the features discussed under (B) and (C) interdependent? Section 9 will address this issue and show that, even when Ω has a high probability of being large, its value does not affect the distribution under (C).

3.3. Exact Stochastic Renormalizability and the "Star Functional Equation" for Ω

Once again, the masses in $[0, 1/2]$ and $[1/2, 1]$ take, respectively, the forms $M_1\Omega_1$ and $M_2\Omega_2$, where M_1 and M_2 are two independent realizations of the random variable M and Ω_1, and Ω_2 are two independent realizations of the random variable Ω. Adding the two parts yields

$$\Omega \equiv \Omega_1 M_1 + \Omega_2 M_2.$$

This identity in distribution, now called the "star equation," combines with $\langle \Omega \rangle = 1$ to determine Ω. It was introduced in Mandelbrot 1974 and has since then been investigated by several authors, for example by Durrett and Liggett 1983. A large bibliography is found in Liu 2002.

In the special case where M is non-random, the star equation reduces to the equation due to Cauchy whose solutions have become well-known: they are the Cauchy–Lévy stable distributions.

3.4. Metaphor for the Probability of Large Values of Ω, Arising in the Theory of Discrete Time Branching Processes

A growth process begins at $t = 0$ with a single cell. Then, at every integer instant of time, every cell splits into a random non-negative number of N_1 cells. At time k, one deals with a clone of N_k cells. All those random splittings are statistically independent and identically distributed. The normalized clone size, defined as $N_k/\langle N_1\rangle^k$ has an expectation equal to 1. The sequence of normalized sizes is a positive martingale, hence (as already mentioned) converges to a limit random variable.

When $\langle N\rangle > 1$, that limit does not reduce to 0 and is random for a very intuitive reason. As long as clone size is small, its growth very much depends on chance, therefore the normalized clone size is very variable. However, after a small number of splittings, a law of large numbers comes into force, the effects of chances become negligible, and the clone grows near-exponentially. That is, the randomness in the relative number of family members can be very large but acts very early.

3.5. To a Large Extent, the Asymptotic Measure Ω of a TVCM Is Large if, and only if, the Pre-Fractal Measure $\mu_k([0, 1])$ Has Become Large During the Very First Few Stages of the Generating Cascade

Such behavior is suggested by the analogy to a branching process, and analysis shows that such is indeed the case. After the first stage, the measures $\mu_1([0, 1/2])$ and $\mu_1([1/2, 1])$ are both equal to u^2 with probability p^2, uv with probability $2p(1-p)$, and v^2 with probability $(1-p)^2$. Extensive simulations were carried out for large k in "batches," and the largest, medium, and smallest measure was recorded for each batch. Invariably, the largest (resp., smallest) Ω started from a high (resp. low) overall level.

4. THE FUNCTION $\tau(q)$: MOTIVATION AND FORM OF THE GRAPH

So far $\tau(q)$ was nothing but a notation. It is important as it is the special form taken for TVCM by a function that was first defined for an arbitrary multiplier in Mandelbrot 1974ab. (Actually, the little-appreciated Fig. 1 of that original paper did not include $q < 0$ and worked with $-\tau(q)$, but the opposite sign came to be generally adopted.)

4.1. Motivation of $\tau(q)$

After k cascade stages, consider an arbitrary dyadic interval of duration $dt = 2^{-k}$. For the k-approximant TVCM measure $\mu_k(dt)$. The qth power has an expected value equal to $[pu^q + (1-p)\, v^q]^k = \{\langle M^q\rangle]^k\}$. Its logarithm of base 2 is

$$\log_2\{[pu^q + (1-p)\, v^q]^k\} = k \log_2\{pu^q + (1-p)\, v^q\} = \log_2(dt)[\tau(q) + 1].$$

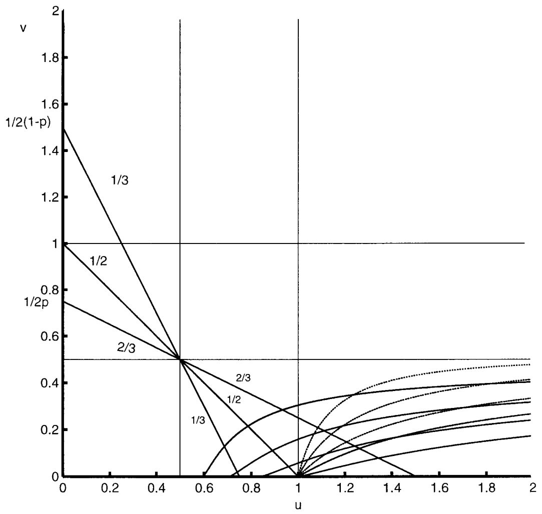

Fig. 1. The full phase diagram of TVCM with coordinates u and v. The isolines of the quantity p are straight intervals from $\{1/[2(1-p)], 0\}$ to $\{0, 1/[2p]\}$. The values p and $1-p$ are equivalent and the corresponding isolines are symmetric with respect to the main bisector $u = v$. The acceptable part of the plane excludes the points (u, v) such that both u and v are $< 1/2$ or both are $> 1/2$. Hence, the relevant part of this diagram is made of two infinite halfstrips. They are reducible to one another by folding along the bisector. The folded phase diagram of TVCM showing isolines of the following quantities: $1-p$ and p—each is made of two straight intervals that start at the point (1, 1) and end at the points $\{1/[2p], 0\}$ and $\{1/[2(1-p)], 0\}$; D—curves that start on the interval $1/2 < u < 1$ of the u-axis and continue to the point $(\infty, 0)$; and q—curves that start at the point (1, 0) and continue to the point $(\infty, 0)$. The Bernoulli binomial corresponds to $p = 1/2$ and the canonical Cantor measure corresponds to the half line $v = 0, u > 1/2$.

Hence

$$\langle \mu_k^q(dt) \rangle = (dt)^{\tau(q)+1}.$$

4.2. A Generalization of the Role of Ω: Middle- and High-Frequency Contributions to Microrandomness

Exactly the same cascade transforms the measure in dt from $\mu_k(dt)$ to $\mu(dt)$ and the measure in [0,1] from 1 to Ω. Hence, one can write

$$\mu(dt) = \mu_k(dt)\, \Omega(dt).$$

In this product, frequencies of wavelength $> dt$, to be described as "low," contribute $\mu_k([0, 1])$, and frequencies of wavelength $< dt$, to be described as "high," contribute Ω.

4.3. The Expected "Partition Function" $\sum \langle \mu^q(d_i t) \rangle$

Section 5 will show that $\langle \Omega^q \rangle$ need not be finite. But if it is, the limit measure $\mu(dt) = \mu_k(dt)\, \Omega(dt)$ satisfies

$$\langle \mu^q(dt) \rangle = (dt)^{\tau(q)+1} \langle \Omega^q \rangle.$$

The interval $[0,1]$ subdivides into $1/dt$ intervals $d_i t$ of common length dt. The sum of the qth moments over those intervals takes the form

$$\langle \chi(dt) \rangle = \sum \langle \mu^q(d_i t) \rangle = (dt)^{\tau(q)} \langle \Omega^q \rangle.$$

Estimation of $\tau(q)$ From a Sample. It is affected by the prefactor Ω insofar as one must estimate both $\tau(q)$ and $\log \langle \Omega^q \rangle$.

4.4. Form of the $\tau(q)$ Graph

Due to conservation on the average, $\langle M \rangle = pu + (1-p)\, v = 1/2$, hence $\tau(1) = -\log_2[1/2] - 1 = 0$. An additional universal value is $\tau(0) = -\log_2(1) - 1 = -1$. For other values of q, $\tau(q)$ is a cap-convex continuous function satisfying $\tau(q) < -1$ for $q < 0$.

For TVCM, a more special property is that $\tau(q)$ is asymptotically linear: assuming $u > v$, and letting $q \to \infty$:

$$\tau(q) \sim -\log_2 p - 1 - q \log u, \quad \text{and} \quad \tau(-q) \sim -\log_2(1-p) - 1 + q \log v.$$

The sign of $u-1$ affects the sign of $\log u$, a fact that will be very important in Section 5.

Moving as little as possible beyond these properties, the very special τ function of the TVCM is simple but Fig. 2 suffices to bring out every one of the delicate possibilities first reported in Mandelbrot 1974a, where $-\tau(q)$ is plotted in that little-appreciated Fig. 1.

Other Features of τ that Deserve to be Mentioned. Direct proofs are tedious and the short proofs require the multifractal formalism that will only be described in Section 10.

The Quantity $D(q)=\tau(q)/(q-1)$. This popular expression is often called a "generalized dimension," a term too vague to mean anything. $D(q)$ is obtained by extending the line from (q, τ) to $(1, 0)$ to its intercept with the line $q = 0$. It plays the role of a critical embedding codimension for the existence of a finite qth moment. This topic cannot be discussed here but is treated in Mandelbrot, 2002c.

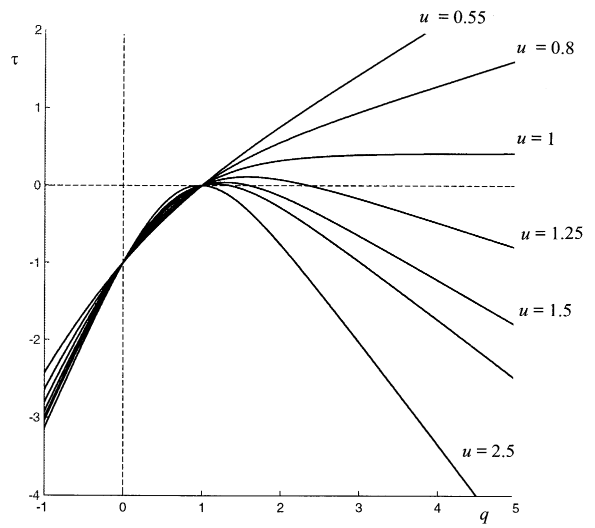

Fig. 2. The functions $\tau(q)$ for $p = 3/4$ and varying g. By arbitrary choice, $g = 1$ is assigned to the case $u = 1$ from which follows that $g = -1$ is assigned to the case $v = 1$. Behavior for $g > 0$: as $q \to -\infty$, the graph of $\tau(q)$ is asymptotically tangent to $\tau = -q \log_2 v$, as $q \to \infty$, the graph of $\tau(q)$ is asymptotically tangent to $\tau = -q \log_2 v$. Those properties are widely believed to describe the main facts about $\tau(q)$. But the TVCM are examples to the contrary. For them, $\tau(q)$ is also tangent to $\tau = q\alpha^*_{min}$ and to $\tau = q\alpha^*_{max}$. Beyond the points of tangency, f becomes < 0. For $g > 1$, one has $u > 1$, hence $\tau(q)$ has a maximum. Values of q beyond this maximum correspond to $\alpha_{min} < 0$. Because of capconvexity, the equation $\tau(q) = 0$ may, in addition to $q = 1$, have a root $q_{crit} > 1$. For $u > 2.5$, one deals with another and very different phenomenon, again first described in Mandelbrot 1974ab (see also Mandelbrot 1990c): the TVCM measure degenerates to zero.

The Ratio $\tau(q)/q$ and the "Accessible" Values of q. Increase q from $-\infty$ to 0 then to $+\infty$. In the Bernoulli case, $\tau(q)/q$ increases from α_{max} to ∞, jumps down to $-\infty$ for $q = 0$, then increases again from $-\infty$ to α_{min}. For TVCM with $p \neq 1/2$, the behavior is very different. For example, let $p < 1/2$. As q increases from 1 to ∞, $\tau(q)$ increases from 0 to a maximum α^*_{max}, then decreases. In a way explored in Section 10, the values of $\alpha > \alpha^*_{max}$ are not "accessible."

4.5. Reducible and Irreducible Canonical Multifractals

Once again, being "canonical" implies conservation on the average. When there exists a microcanonical (conservative) variant having the same function $f(\alpha)$, a canonical measure can be called "reducible." The canonical binomial is reducible because its $f(\alpha)$ is shared by the Bernoulli binomial. Another example introduced in Mandelbrot 1989b is the "Erice" measure, in which the multiplier M is uniformly distributed on $[0, 1]$. But the TVCM with $p \neq 1/2$ is not reducible

In the interval [0,1] subdivided in the base $b = 2$, reducibility demands a multiplier M whose distribution is symmetric with respect to $M = 1/2$. Since $u > 0$, this implies $u < 1$.

5. WHEN $U > 1$, THE MOMENT $\langle \Omega^q \rangle$ DIVERGES IF q EXCEEDS A CRITICAL EXPONENT q_{crit} SATISFYING $\tau(q) = 0$; Ω FOLLOWS A POWER-LAW DISTRIBUTION OF EXPONENT q_{crit}

5.1. Divergent Moments, Power-Law Distributions and Limits to the Ability of Moments to Determine a Distribution

This section injects a concern that might have been voiced in Sections 3 and 4. The canonical binomial and many other examples satisfy the following properties, which everyone takes for granted and no one seems to think about: (a) $\Omega = 1$, $\langle \Omega^q \rangle < \infty$, (b) $\tau(q) > 0$ for all $q > 0$, and (c) $\tau(q)/q$ increases monotonically as $q \to \pm\infty$.

Many presentations of fractals take those properties for granted in all cases. In fact, as this section will show, the TVCM with $u > 1$ lead to the "anomalous" divergence $\langle \Omega^q \rangle = \infty$ and the "inconceivable" inequality $\tau(q) < 0$ for $q_{crit} < q < \infty$. Also, the monotonicity of $\tau(q)/q$ fails for all TVCM with $p \neq 1/2$.

Since Pareto in 1897, infinite moments have been known to characterize the power-law distributions of the form $\Pr\{X > x\} = x^{-q_{crit}}$. But in the case of TVCM and other canonical multifractals, the complicating factor $L(x)$ is absent. One finds that when $u > 1$, the overall measure Ω follows a power law of exponent q_{crit} determined by $\tau(q)$.

5.2. Discussion

The power-law "anomalies" have very concrete consequences deduced in Mandelbrot 1997 and discussed, for example, in Mandelbrot 2001c.

But does all this make sense? After all, $\tau(q)$ and $\langle \Omega^q \rangle$ are given by simple formulas and are finite for all parameters. The fact that those values cannot actually be observed raises a question. Are high moments lost by being unobservable? In fact, they are "latent" but can be made "actual" by a process of "embedding" studied elsewhere.

An additional comment is useful. The fact that high moments are non-observable does not express a deficiency of TVCM but a limitation of the notion of moment. Features ordinarily expressed by moments must be expressed by other means.

5.3. An Important Apparent "Anomaly:" In a TVCM, the qth Moment of Ω May Diverge

Let us elaborate. From long past experience, physicists' and statisticians' natural impulse is to define and manipulate moments without envisioning or voicing the possibility of their being infinite. This lack of concern cannot extend to multifractals. The distribution of the TVCM within a dyadic interval introduces an additional critical exponent q_{crit} that satisfies $q_{\text{crit}} > 1$. When $1 < q_{\text{crit}} < \infty$, which is a stronger requirement that $D > 0$, the qth moment of $\mu(dt)$ diverges for $q > q_{\text{crit}}$.

A stronger result holds: the TVCM cascade generates a measure whose distribution follows the power law of exponent q_{crit}.

Comment. The heuristic approach to non-random multifractals fails to extend to random ones, in particular, it fails to allow $q_{\text{crit}} < \infty$. This makes it incomplete from the viewpoint of finance and several other important applications.

The finite q_{crit} has been around since Mandelbrot 1974 (where it is denoted by α) and triggered a substantial literature in mathematics. But it is linked with events so extraordinarily unlikely as to appear incapable of having any perceptible effect on the generated measure. The applications continue to neglect it, perhaps because it is ill-understood. A central goal of TVCM is to make this concept well-understood and widely adopted.

5.4. An Important Role of $\tau(q)$: If $q > 1$ the qth Moment of Ω Is Finite if, and only if, $\tau(q) > 0$; the Same Holds for $\mu(dt)$ Whenever dt Is a Dyadic Interval

By definition, after k levels of iteration, the following symbolic equality relates independent realizations of M and μ. That is, it does not link random variables but distributions

$$\mu_k([0,1]) = M\mu_{k-1}([0,1]) + M\mu_{k-1}([0,1]).$$

Conservation on the average is expressed by the identity $\langle \mu_{k-1}([0,1]) \rangle = 1$. In addition, we have the following recursion relative to the second moment.

$$\langle \mu^2[(0,1])\rangle = 2\langle M^2\rangle[\langle \mu^2_{k-1}([0,1])\rangle] + 2\langle M^2\rangle\{\langle \mu_{k-1}([0,1])\rangle]^2.$$

The second term to the right reduces to $1/2$. Now let $k \to \infty$. The necessary and sufficient condition for the variance of $\mu_k([0,1])$ to converge to a finite limit is

$$2(EM^2) < 1, \qquad \text{in other words} \quad \tau(2) = -\log_2(EM^2) - 1 > 0.$$

When such is the case, Kahane and Peyrière 1976 gave a mathematically rigorous proof that there exists a limit measure $\mu([0, 1])$ satisfying the formal expression

$$\langle \mu^2([0, 1]) \rangle = \frac{1}{2[(1-2^{\tau(2)}]}.$$

Higher integer moments satisfy analogous recursion relations. That is, knowing that all moments of order up to $q-1$ are finite, the moment of order q is finite if and only if $\tau(q) > 0$.

The moments of non-integer order q are more delicate to handle, but they too are finite if, and only if, $\tau(q) > 0$.

5.5. Definition of q_{crit}; Proof that in the Case of TVCM q_{crit} Is Finite if, and only if, $u > 1$

Section 4.4 noted that the graph of $\tau(q)$ is always cap-convex and for large $q > 0$,

$$\tau(q) \sim -\log_2(pu^q) + -1 \sim -\log_2 p - 1 - q \log_2 u.$$

The dependence of $\tau(q)$ on q is ruled by the sign of $u-1$, as follows.

- *The case when $u < 1$, hence $\alpha_{min} > 0$.* In this case, $\tau(q)$ is monotone increasing and $\tau(q) > 0$ for $q > 1$. This behavior is exemplified by the Bernoulli binomial.
- *The case when $u > 1$, hence $\alpha_{min} < 0$.* In this case, one has $\tau(q) < 0$ for large q. In addition to the root $q = 1$, the equation $\tau(q) = 1$ has a second root that is denoted by q_{crit}.

Comment. In terms of the function $f(\alpha)$ graphed on Fig. 3, the values 1 and q_{crit} are the slopes of the two tangents drawn to $f(\alpha)$ from the origin (0, 0).

Within the class of equivalence of any p and $1-p$; the parameter g can be "tuned" so that q_{crit} begins by being > 1 then converges to 1; if so, it is seen that D converges to 0.

- Therefore, the conditions $q_{crit} = 1$ and $D = 0$ describe the same "anomaly."

In Fig. 1, isolines of q_{crit} are drawn for $q_{crit} = 1, 2, 3,$ and 4. When $q = 1$ is the only root, it is convenient to say that $q_{crit} = \infty$. This isoset $q_{crit} = \infty$ is made of the half-line $\{v = 1/2$ and $u > 1/2\}$ and of the square $\{0 < v < 1/2, 1/2 < u < 1\}$.

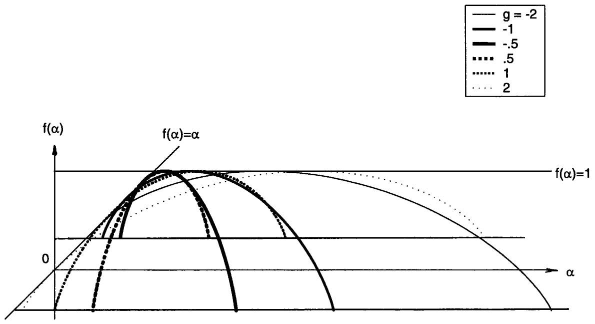

Fig. 3. The functions $f(\alpha)$ for $p = 3/4$ and varying g. All those graphs are linked by horizontal reductions or dilations followed by translation and futher self-affinity. For $g > 0$ (resp., $g < 0$) the left endpoint of $f(\alpha)$ (resp., the right endpoint) satisfies $f(\alpha) < 0$ and the other endpoint, $f(\alpha) > 0$.

5.6. The Exponent q_{crit} Can Be Considered as a Macroscopic Variable of the Generating Process

Any set of two parameters that fully describes a TVCM can be called "microscopic". All the quantities that are directly observable and can be called macroscopic are functions of those two parameters.

For the general canonical multifractal, a full specification requires a far larger number of microscopic quantities but the same number of macroscopic ones. Some of the latter characterize each sample, but others, for example q_{crit}, characterize the population.

6. THE QUANTITY α: THE ORIGINAL HÖLDER EXPONENT AND BEYOND

The multiplicative cascades—common to the Bernoulli and canonical binomials and TVCM—involve successive multiplications. An immediate consequence is that both the basic $\mu(dt)$ and its probability are most intrinsically viewed through their logarithms. A less obvious fact is that a normalizing factor $1/\log(dt)$ is appropriate in each case. An even less obvious fact is that the normalizations $\log \mu/\log dt$ and $\log P/\log dt$ are of far broader usefulness in the study of multifractals. The exact extent of their domain of usefulness is beyond the goal of this paper but we keep to special cases that can be reated fully by elementary arguments.

6.1. The Bernoulli Binomial Case and Two Forms of the Hölder Exponent: Coarse-Grained (or Coarse) and Fine-Grained

Recall that due to conservation, the measure in an interval of length $dt = 2^{-k}$ is the same after k stages and in the limit, namely, $\mu(dt) = \mu_k(dt)$.

As a result, the coarse-grained Hölder exponent can be defined in either of two ways,

$$\alpha(dt) = \frac{\log \mu(dt)}{\log(dt)},$$

and

$$\tilde{\alpha}(dt) = \frac{\log \mu_k(dt)}{\log(dt)}.$$

The distinction is empty in the Bernoulli case but prove prove essential for the TVCM. In terms of the relative frequencies φ_0 and φ_1 defined in Section 1.1,

$$\alpha(dt) = \tilde{\alpha}(dt) = \alpha(\varphi_0, \varphi_1) = -\varphi_0 \log_2 u - \varphi_1 \log_2 v$$
$$= -\varphi_0(\log_2 u - \log_2 v) - \log v.$$

Since $u > v$, one has $0 < \alpha_{\min} = -\log_2 u \leqslant \alpha = \tilde{\alpha} \leqslant \alpha_{\max} = -\log_2 v < \infty$. In particular, $\alpha > 0$, hence $\tilde{\alpha} > 0$. As $dt \to 0$, so does $\mu(dt)$, and a formal inversion of the definition of α yields

$$\mu(dt) = (dt)^{\alpha}.$$

This inversion reveals an old mathematical pedigree. Redefine φ_0 and φ_1 from denoting the finite frequencies of 0 and 1 in an interval, into denoting the limit frequencies at an instant t. The instant t is the limit of an infinite sequence of approximating intervals of duration 2^{-k}. The function $\mu([0, t])$ is non-differentiable because $\lim_{dt \to 0} \mu(dt)/dt$ is not defined and cannot serve to define the local density of μ at the instant dt.

The need for alternative measures of roughness of a singularity expression first arose around 1870 in mathematical esoterica due to L. Hölder. In fractal/multifractal geometry this expression merged with a very concrete exponent due to H. E. Hurst and is continually being generalized. It follows that for the Bernoulli binomial measure, it is legitimate to interpret the coarse αs as finite-difference surrogates of the local (infinitesimal) Hölder exponents.

6.2. In the General TVCM Measure, $\alpha \neq \tilde{\alpha}$, and the Link Between "$\alpha$" and the Hölder Exponent Breaks Down; One Consequence Is that the "Doubly Anomalous" Inequalities $\alpha_{\min} < 0$, Hence $\tilde{\alpha} < 0$, Are Not Excluded

A Hölder (Hurst) exponent is necessarily positive. Hence negative $\tilde{\alpha}$'s cannot be interpreted as Hölder exponents. Let us describe the heuristic argument that leads to this paradox and then show that $\tilde{\alpha} < 0$ is a serious "anomaly:" it shows that the link between "some kind of α" and the Hölder exponent requires a searching look. The resolution of the paradox is very subtle and is associated with the finite q_{crit} introduced in Section 5.4.

Once again, except in the Bernoulli case, $\Omega \neq 1$ and $\mu(dt) = \mu_k(dt)\,\Omega(dt)$, hence

$$\alpha(dt) = \tilde{\alpha}(dt) + \log \Omega(dt)/\log dt.$$

In the limit $dt \to 0$ the factor $\log = \Omega/\log(dt)$ tends to 0, hence it seems that $\alpha = \tilde{\alpha}$. Assume $u > 1$, hence $\alpha_{\min} < 0$ and consider an interval where $\tilde{\alpha}(dt) < 0$. The formal equality

$$\text{“}\mu_k(dt) = (dt)^{\tilde{\alpha}}\text{”}$$

seems to hold and to imply that "the" mass in an interval increases as its interval length $\to 0$. On casual inspection, this is absurd. On careful inspection, it is not—simply because the variable $dt = 2^{-k}$ and the function $\mu_k(dt)$ both depend on k. For example, consider the point t for which $\varphi_0 = 1$. Around this point, one has $\mu_k = u\mu_{k-1} > \mu_{k-1}$. This inequality is nothing paradoxical about it.

Furthermore, Section 7 shows that the theory of the multiplicative measures introduces $\tilde{\alpha}$ intrinsically and inevitably and allows $\tilde{\alpha} < 0$.

Those seemingly contradictory properties will be reexamined in Section 8. Values of $\mu(dt)$ will be seen to have a positive probability but one so minute that they can never be observed in the way $\alpha > 0$ are observed. But they affect the distribution of the variable Ω examined in Section 3, therefore are observed indirectly.

7. THE FULL FUNCTION $F(\alpha)$ AND THE FUNCTION $\rho(\alpha)$

7.1. The Bernoulli Binomial Measure: Definition and Derivation of the Box Dimension Function $f(\alpha)$

The number of intervals of denumerator 2^{-k} leading to φ_0 and φ_1 is $N(k, \varphi_0, \varphi_1) = k!/(k\varphi_0)!\,(k\varphi_1)!$, and dt is the reduction ratio r from $[0, 1]$ to an interval of duration dt. Therefore, the expression

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

is of the form $f(k, \varphi_0, \varphi_1) = -\log N/\log r$. Fractal geometry calls this the "box similarity dimension" of a set. This is one of several forms taken by *fractal dimension*. More precisely, since the boxes belong to a grid, it is a *grid fractal dimension*.

The Dimension Function f(α). For large k, the leading term in the Stirling approximation of the factorial yields

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

7.2. The "Entropy Ogive" Function *f*(*α*); the Role of Statistical Thermodynamics in Multifractals and the Contrast Between Equipartition and Concentration

Eliminate φ_0 and φ_1 between the functions f and $\alpha = -\varphi_0 \log u - \varphi_1 \log v$. This yields in parametric form a function, $f(\alpha)$. Note that $0 \leqslant f(\alpha) \leqslant \min\{\alpha, 1\}$. Equality to the right is achieved when $\varphi_0 = u$. The value α where $f = \alpha$ is very important and will be discussed in Section 8. In terms of the reduced variable $\varphi_0 = (\alpha - \alpha_{\min})/(\alpha_{\max} - \alpha_{\min})$, the function $f(\alpha)$ becomes the "ogive"

$$\tilde{f}(\varphi_0) = -\varphi_0 \log_2 \varphi_0 - (1 - \varphi_0) \log_2 (1 - \varphi_0).$$

This $\tilde{f}(\varphi_0)$ can be called a universal function. The $f(\alpha)$ corresponding to fixed p and varying g are affine transforms of $\tilde{f}(\varphi_0)$, therefore of one another. The ogive function $\tilde{f}$ first arose in thermodynamics as an entropy and in 1948 (with Shannon) entered communication theory as an information. Its occurrence here is the first of several roles played by the formalism of thermodynamics in the theory of multifractals.

An Essential But Paradoxical Feature. Equilibrium thermodynamics is a study of various forms of *near-equality*, for example postulates the equipartition of states on a surface in phase space or of energy among modes. In sharp contrast, multifractals are characterized by extreme *inequality*

between the measures in different intervals of common duration dt. Upon more careful examination, the paradox dissolves by being turned around: the main tools of thermodynamics can handle phenomena well beyond their original scope.

7.3. The Bernoulli Binomial Measure, Continued: Definition and Derivation of a Function $\rho(\alpha) = f(\alpha) - 1$ that Originates as a Rescaled Logarithm of a Probability

The function $f(\alpha)$ never fully specifies the measure. For example, it does not distinguish between the Benoulli, shuffled and canonical binomials. The function $f(\alpha)$ can be generalized by being deduced from a function $\rho(\alpha) = f(\alpha) - 1$ that will now be defined. Instead of dimensions, that deduction relies on probabilities. In the Bernoulli case, the derivation of ρ is a minute variant of the argument in Section 8.1, but, contrary to the definition of f, the definition of ρ easily extends to TVCM and other random multifractals.

In the Bernoulli binomial case, the probability of hitting an interval leading to φ_0 and φ_1 is simply $P(k, \varphi_0, \varphi,) = N(k, \varphi_0, \varphi_1)\, 2^{-k} = k!/(k\varphi_0)!\,(k\varphi_1)!\, 2^{-k}$. Consider the expression

$$\rho(k, \varphi_0, \varphi_1) = -\frac{\log[P(k, \varphi_0, \varphi_1)]}{\log(dt)},$$

which is a rescaled but not averaged form of entropy. For large k, Stirling yields

$$\lim_{k \to \infty} \rho(k, \varphi_0, \varphi_1) = \rho(\varphi_0, \varphi_1) = -\varphi_0 \log_2 \varphi_0 - \varphi_1 \log_2 \varphi_1 - 1 = f(\alpha) - 1.$$

7.4. Generalization of $\rho(\alpha)$ to the Case of TVCM; the Definition of $f(\alpha)$ as $\rho(\alpha) + 1$ Is Indirect but Significant Because It Allows the Generalized f to Be Negative

Comparing the arguments in Sections 7.1 and 7.2 link the concepts of fractal dimension and of minus log (probability). However, when $f(\alpha)$ is reported through $f(\alpha) = \rho(\alpha) + 1$, the latter is not a mysterious "spectrum of singularities." It is simply the peculiar but proper way a probability distribution must be handled in the case of multifractal measures. Moreover, there is a major a priori difference exploited in Section 10. Minus log (probability) is not subjected to any bound. To the contrary, every one of the traditional definitions of fractal dimension (including Hausdorff–Besicovitch or Minkowski–Bouligand) necessarily yields a positive value.

The point is that the dimension argument in Section 7.1 does not carry over to TVCM, but the probability argument does carry over as follows.

The probability of hitting an interval leading to φ_0 and φ_1 now changes to $P(k, \varphi_0, \varphi_1) = p(\varphi_0 k)!/(k\varphi_0)!\,(k\varphi_1)!$ One can now form the expression

$$\rho(k, \varphi_0, \varphi_1) = -\frac{\log[P(k, \varphi_0, \varphi_1)]}{\log(dt)}.$$

Stirling now yields

$$\begin{aligned}\rho(\varphi_0, \varphi_1) &= \lim_{k \to \infty} \rho(k, \varphi_0, \varphi_1)\\ &= \{-\varphi_0 \log_2 varphi_0 - \varphi_1 \log_2 \varphi_1\} + \{\varphi_0 \log_2 p + \varphi_1 \log_2 (1-p)\}.\end{aligned}$$

In this sum of two terms marked by braces, we know that the first one transforms (by horizontal stretching and translation) into the entropy ogive. The second is a linear function of φ, namely $\varphi_0[\log_2 p - \log_2(1-p)] + \log_2(1-p)$. It transforms the entropy ogive by an affinity in which the line joining the two support endpoints changes from horizontal to inclined. The overall affinity solely depends on p, but φ_0 depends explicitly on u and v.

This affinity extends to all values of p. Another property familiar from the binomial extends to all values of p. For all u and v, the graphs of $\rho(\alpha)$, hence of $f(\alpha)$ have a vertical slope for $q = \pm\infty$.

Alternatively, $\rho(\varphi_0, \varphi_1) = -\varphi_0 \log_2 [\varphi_0/p] - \varphi_1 \log_2[\varphi_1/(1-p)]$.

7.5. Comments in Terms of Probability Theory

Roughly speaking, the measure μ is a *product* of random variables, while the limit theorems of probability theory are concerned with *sums*. The definition of α as $\log \mu(dt)/\log(dt)$ replaces a product of random variables M by a weighted sum of random variables of the form $\log M$. Let us now go through this argument step by step in greater rigor and generality. One needs a cumbersome restatement of $\alpha_k(dt)$.

The Low Frequency Factor of $\mu_k(dt)$ and the Random Variable H_{low}. Consider once again a dyadic cell of length 2^{-k} that starts at $t = 0$. $\beta_1 \beta_2 \cdots \beta_k$. The first k stages of the cascade can be called *of low frequency* because they involve multipliers that are constant over dyadic intervals of length $dt = 2^{-k}$ or longer. These stages yield

$$\mu_k(dt) = M(\beta_1)\, M(\beta_1, \beta_2) \cdots M(\beta_1, \dots, \beta_k) = \prod M.$$

We transform $\mu_k(dt)$ into the *low frequency* random variable

$$H_{\text{low}} = \log[\mu_k(dt)\, rb/\log(dt)$$
$$= (1/k)[-\log_2 M(\beta_1) - \log_2 M(\beta_1, \beta_2) - \cdots].$$

We saw in Section 3.5. that the first few values of M largely determine the distribution of Ω. But the last expression involves an operation of averaging in which the first terms contributing to $\mu(dt)$ are asymptotically washed out.

7.6. Distinction Between "Center" and "Tail" Theorems in Probability

The quantity $\tilde{\alpha}_k(dt) = \varphi_0 \log_2 u - \varphi_1 \log_2 v$ is the average of a sum of variables $-\log M$ but its distribution is not Gaussian and the graph of $\rho(\alpha)$ is an entropy ogive rather than a parabola. Why is this so? The law of large numbers tells us that $\tilde{\alpha}_k(dt)$ almost surely converges to its expectation which tells us very little. A tempting heuristic argument continues as follows. The central limit theorem is believed to ensure that for small dt, $H_{\text{low}}(dt)$ becomes Gaussian, therefore the graph of $\log \rho(dt)$ should be expected to be a parabola. This being granted, why is it that the Stirling approximation yields an entropy ogive—not a parabola?

In fact, there is no paradox of any kind. While the central limit theorem is indeed central to probability theory, all it asserts in this context is that, asymptotically, the Gaussian rules the *center* of the distribution, its "bell." Renormalizations reduce this center to the immediate neighborhood of the top of the $\rho(\alpha)$ graph and the central limit theorem is correct in asserting that the top of the entropy ogive is locally parabolic. But in the present context this information is of little significance. We need instead an alternative that is only concerned with the tail behavior which it ought to blow up. For this and many other reasons, it would be an excellent idea to speak of *center,* not *central* limit theorem. The tail limit theorem is due to H. Cramer and asserts that the tail consisting in the bulk of the graph is not a parabola but an entropy ogive.

7.7. The Reason for the Anomalous Inequalities $f(\alpha) < 0$ and $\alpha < 0$ Is that, by the Definition of a Random Variable $\mu(dt)$, the Sample Size Is Bounded and Is Prescribed Intrinsically; the Notion of Supersampling

The inequality $\rho(\alpha) < -1$ characterizes events whose probability is extraordinarily small. The finding that this inequality plays a significant role was not anticipated, remains difficult to understand and appreciate, and demands comment.

The common response is that, even extremely low probability events are captured, if one simply takes a sufficiently long sample of independent values. But this is impossible, even if one forgets that, in the present uncommon context, the values are extremely far from being statistically independent. Indeed, the choice the duration $dt = 2^{-k}$ has two effects. Not only does it fix the distribution of $\mu(dt)$, but it also sets the sample size at the value $N = 1/dt = 2^k$. Roughly speaking, a sample of size N can only reveal values having a probability greater than $1/N$, which means $\rho(\alpha) > -1$.

In summary, it is true that decreasing dt to 2^{-k-1} increases the sample size. But it also changes the distribution and does so in such a way that the bound $\rho = -1$ remains untouched.

This bound excludes all items of information that correspond to $f(\alpha) < 0$ (for example, the value of q_{crit} when finite). Those items remain hidden and latent in the sense that they cannot be inferred from one sample of values of $\mu(dt)$. Ways of revealing those values, supersampling and embedding, are examined in Mandelbrot 1989b, 1995 and the forthcoming Mandelbrot 2003.

Figure 3 shows, for $p = 3/4$, how the graph of $f(\alpha)$ depends on g.

7.8. Excluding the Bernoulli Case $p = 1/2$, TVCM Faces Either One of Two Major "Anomalies:" for $p > -1/2$, One Has $f(\alpha_{min}) = 1 + \log_2 p > 0$ and $f(\alpha_{max}) = 1 + \log_2(1 - p) < 0$; for $p < 1/2$, the Opposite Signs Hold

The fact that the values of $\rho(\alpha_{min}) = f(\alpha_{min}) - 1$ and $\rho(\alpha_{max}) = f(\alpha_{max}) - 1$ are logarithms of probabilities confirms and extends the definition of $p(\alpha) = f(\alpha) - 1$ as a limit rescaled probability. Here, those endpoint values of $f(\alpha)$ are independent of g and the affinity that deduces them from the entropy ogive (with ends on the horizontal axis) characterizes the class of equivalence of p and $1 - p$. If, and only if, $p = 1/2$ and $u + v = 1$, that is, in the familiar Bernoulli binomial case, one has $\rho(\alpha_{min}) = \rho(\alpha_{max}) = \log_2(1/2) = -1$ hence $f(\alpha_{min}) = f(\alpha_{max}) = 0$. When $u + v \neq 1$, one of the endpoints satisfies $f > 0$ and the other satisfies $f < 0$. Sections 7.9 and 9 shall examine the sharply differing consequences of those inequalities.

7.9. The "Minor Anomalies" $f(\alpha_{max}) > 0$ Or $f(\alpha_{min}) > 0$ Lead to Sample Function with a Clear "Ceiling" or "Floor"

Suppose that $f(\alpha_{min}) = 0$ and $f(\alpha_{max}) = 0$, as is the case for $p = 1/2$. Then, using terms often applied to the printed page—but after it has been turned 90°. to the side—the sample functions are "non-justified" or "ragged" for both high and low values. That is, the values tend to be unequal; one is clearly larger than all others, a second is clearly the second largest, etc.

To the contrary, TVCM with $p \neq 1/2$ yield either $f(\alpha_{max}) > 0$ or $f(\alpha_{min}) > 0$. Sample functions have a conspicuous "ceiling" (resp., a "floor"). That is, a largest (resp., smallest) value is attained repeatedly for values of t belonging to a set of positive dimension. To use the printers' vocabulary, when one side is "ragged" the other is "justified." On visual inspection of the data, the ceiling is always visible; the floor merges with the time axis, except when one plots $\log[\mu(dt)]$.

8. THE FRACTAL DIMENSION $D = \tau'(1) = 2[-pu \log_2 u - (1-p)\, v \log_2 v]$ AND MULTIFRACTAL CONCENTRATION

The function $f(\alpha)$ satisfies $f(\alpha) \leqslant \alpha$, with equality $f(\alpha) = \alpha$ when $\alpha = D = \tau'(1)$. From the value of $\alpha = D$ follows one of the most important properties of multifractals. Mandelbrot 2001d proposed to call it "multifractal concentration." This section will first examine its opposite, which is asymptotic negligibility.

8.1. In the Bernoulli Binomial Measure, Weak Asymptotic Negligibility Holds but Strong Asymptotic Negligibility Fails

Recally that during construction, the total binomial measure of [0,1] remains constant and equal to 1. But the first few stages of construction make its distribution become very unequal and a few values that stand out as sharp spikes. After k stages, the maximum measure is u^k, which is far larger than the minimum measure v^k. From the relations

$$2^{-k} = dt, \quad 2^k = N, \quad -\log_2 u = \alpha_{min} < 1, \quad \text{and} \quad -\log_2 v = \alpha_{min} > 1,$$

it follows that

$$u^k = b^{(-\log_b u)(-k)} = (dt)^{\alpha_{min}} = N^{-\alpha_{min}}.$$

In words: even the maximum u^k tends to 0. This is a *weak* form of asymptotic negligibility following a power-law.

The preceding result holds for every multifractal for which there is an $\alpha_{\min} > 0$ that plays the same role as in the binomial case. (In more general multifractals the same role is held by some $\alpha^*_{\min} > \max\{\alpha_{\min}, 0\}$.)

Similarly, the total contribution of any fixed number of largest spikes is asymptotically negligible.

8.2. For the Bernoulli or Canonical Binomials, the Equation $f(\alpha) = \alpha$ Has One and Only One Solution; that Solution Satisfies $D > 0$ and Is the Fractal Dimension of the "Carrier" of the Measure

We now proceed to the total contribution of a number of spikes that is no longer fixed but increases with N. In the simplest of all possible worlds, many spikes would have been more or less equal to the largest, and the sum of all the other spikes would have been negligible. If so, the sum of $N^{\alpha_{\min}}$ spikes would have been of the order of $N^{\alpha_{\min}} N^{-\alpha_{\min}} = 1$.

While the world is actually more complicated there is an element of orderliness. The equality $\varphi_0 = u$ is achieved for $\alpha = f(\alpha) = -u \log u - v \log v = D$. For finite but large k, it follows that

$$\mu(k, \varphi_0, \varphi_1) \sim 2^{-k\alpha} = 2^{-kD} \qquad \text{and} \qquad N(k_1 \varphi_0, \varphi_1) \sim 2^{kf(\alpha)} = 2^{kD}.$$

Hence,

$$\mu(k_1 \varphi_0, \varphi_1)\, N(k_1 \varphi_0 \varphi_1) \qquad \text{is approximately equal to 1.}$$

Actually, this product is necessarily $\leqslant 1$, but the difference tends to 0 as $k \to \infty$. That is, an increasingly overwhelming bulk of the measure tends to "concentrate" in the cells where $\alpha = D$. The remainder is small, but in the theory of multifractals even very small remainders are extremely significant for some purposes.

8.3. The Notion of "Multifractal Concentration"

A key feature of multifractals is a subtle interaction between number and size that is elaborated upon in Mandelbrot 2001d. Section 8.2. showed that the contributions that are large are too few to matter. The small contributions are very numerous, but so extremely small that their total contribution is negligible as well. The bulk of the measure is found in a rather inconspicuous intermediate range one can call "mass carrying." Since $D > \alpha_{\min}$, the N^D spikes of size N^{-D} are far smaller than the largest one. Separately, each is asymptotically negligible. But their number N^D is exactly large enough to insure that their total contribution is nearly equal to the overall measure 1. When a sample is plotted, this range does not stand out but it makes a perfect match between size and frequency.

Practically, the number of visible peaks is so small compared to N^D that a combination of the peaks and the intermediate range is still of the order of N^D. The combined range has the advantage of simplicity, since it includes the N^D largest values. Note that the peaks tend to be located in the midst of stretches of values of intermediate size.

8.4. The Case of TVCM with $p < 1/2$ Allows D to Be Positive, Negative, or Zero

Using the alternative expression for $f(\alpha)$ given in Section 8.4, the identity $f(\alpha) = \alpha$ demands the equality of the two expressions

$$f(\alpha) = -\varphi_0 \log_2[\varphi_0/p] - \varphi_1 \log_2[\varphi_1/(1-p] \qquad \text{and}$$

$$\alpha = -\varphi_0 \log_2 u - \varphi_1 \log_2 v.$$

The solution is, obviously, $\varphi_0 = pu$ and $\varphi_1 = (1-p)\, v$. The sum $\varphi_1 + \varphi_1$ is 1, as it must. Hence, $D = -pu \log_2 u - (1-p)\, v \log_2 v$, as announced. The novelty is that TVCM allow $D > 0$, $D = 0$, and $D < 0$.

Familiar Role of D Under the Inequality $D > 0$. Mandelbrot 1974ab obtained the following criterion, which has become widely known and includes the TVCM case. When positive, D is the fractal dimension of the "set that supports" the measure. Figure 1 shows isolines of D for $D = 0, 1/4, 1/2,$ and $3/4$. The isoline for $D = 1$ is made of the interval $\{u = 1, 0 < v < 1\}$ and the half-line $\{v = 1, u \geqslant 1\}$. The key result is that, contrary to the Bernoulli binomial case, the half line $1 < q < \infty$ subdivides into up to three subranges of values.

Largely Unfamiliar Consequence of the Inequality $D < 0$. For all non-random multifractals, $\tau'(1) > 0$. A casual acquaintance with multifractals takes for granted that this is not changed by randomness. But M74 also allows for an alternative possibility, which has so far remained little known. The example of TVCM shows that, in a canonical case, the formally evaluated D can be negative. In the example of TVCM, D is negative when the point (u, v) falls in a domain to the bottom right of the folded phase diagram in Fig. 1. The consequences of $D < 0$ are drastic: the multifractal reduces to 0 almost surely and is called degenerate.

A Classical "Pathological Limit" as Metaphor. This limit behavior of the distribution of μ seems incompatible with the fact that $E\mu = 1$ by definition. But in fact, no contradiction is observed. A convincing idea of the distribution is provided for each p, by the behavior of the $g \to \infty$ limit of the weights $u^g 2^{\tau(g)}$ and $v^g 2^{\tau(g)}$. This recalls a classical counterexample of analysis, namely, the behavior for $k \to \infty$ of the variable P_k defined as follows: $P_k = k$ with the probability $1/k$ and $P_k = 0$ with the probability $1 - 1/k$. For finite k, one has $EP_k = 1$. But in the limit $k \to \infty$, $P_\infty = 0$, hence $EP_\infty = 0$, so that in the limit the expectation drops discontinuously from 1 to 0. In practice, the preasymptotic measure is extremely small with a high probability and huge with a tiny probability.

The Condition D=0. It defines the threshold of degeneracy.

9. A NOTEWORTHY AND UNEXPECTED SEPARATION OF ROLES, BETWEEN THE "DIMENSION SPECTRUM" AND THE TOTAL MASS Ω; THE FORMER IS RULED BY THE ACCESSIBLE α FOR WHICH $f(\alpha) > 0$, THE LATTER, BY THE INACCESSIBLE α FOR WHICH $F(\alpha) < 0$

Brought together, Sections 3, 6, 7, and 8 imply, in plain words, that what you do not necessarily see may affect you significantly. This section serves to underline that the notion of canonical multifractal is very subtle and deserves to be well-understood and further discussed.

9.1. Definitions of the "Accessible Ranges" of the Variables: qs from q^*_{min} to q^*_{max} and αs from α^*_{min} to α^*_{max}; the Accessible Functions $\tau^*(q)$ and $f^*(\alpha)$

Mandelbrot 1995 was led to introduce the function $ff^*(\alpha) = \max\{0, f(\alpha)\}$. That is,

- In the interval $[\alpha^*_{min}, \alpha^*_{max}]$ where $f(\alpha) > 0$, $f^*(\alpha) = f(\alpha)$,
- When $f(\alpha) \leqslant 0$, $f^*(\alpha) = 0$.

The graph of $f^*(\alpha)$ is identical to that of $f(\alpha)$ except that the "tails" with $f < 0$ are truncated so that $f^* > 0$. In terms of $\tau(q)$, the equality $f(\alpha) = 0$ corresponds to lines that are tangent to the graph of $\tau(q)$ and also go through (0,0). In the most general case, those lines' slopes are α^*_{min} and α^*_{max} and the points of contact are denoted by q^*_{max}(satisfying > 0) and q^*_{min}(satisfying < 0). Therefore, the function $f^*(\alpha)$ corresponds to the following truncated function $\tau^*(q)$.

- When $q < q^*_{min}$, $\tau^*(q) = \alpha^*_{max} q$,
- When $q^*_{min} < q < q^*_{max}$, $\tau^*(q) = \tau(q)$,
- When $q > q^*_{max}$, $\tau^*(q) = \alpha^*_{min} q$.

In other words, the graph of τ^* is identical to that of τ except that beyond $q^*_{\max}$ or $q^*_{\min}$ it follows the tangents that go through the origins. Therefore it is straight.

For the TVCM, one has either $\alpha^*_{\max} = \alpha_{\max}$ with $q^*_{\min} = -\infty$, or $\alpha^*_{\min} = \alpha_{\min}$ with $q^*_{\max} = \infty$.

9.2. A Confrontation

Section 3 noted that the largest values of $\Omega([0, 1])$ are generated when a sample cascade begins with a few large values. Section 6 noted that the value of $\Omega([0, 1])$—irrespective of size—ceases, for $k \to \infty$, to have any impact on α. Section 7 noted that, again for $k \to \infty$, values of α such that $f(\alpha) < 0$ have a vanishing probability of being observed. Section 8.1 followed up by defining the accessible function $f(\alpha)$. Section 8 returned to large values of $\Omega([0, 1])$ and noted their association with $q_{\text{crit}} < \infty$. The values of α they involve satisfy $\alpha < 0$, hence a fortiori $f(\alpha) < 0$. Those values do not occur in multifractal decomposition, yet they are extremely important.

9.3. The Simplest Cases Where $f(\alpha) > 0$ for all α, as Exemplified by the Canonical Binomial

Here, the large values of Ω are ruled by the left-most part of the graph of $f(\alpha)$. That is, the same graph controls those large values and the distribution of $\Omega([0, 1])$ among the $1/dt$ intervals of length dt.

9.4. The Extreme Case Where $f(\alpha) < 0$ and $\alpha < 0$ Both Occur, as Exemplified by TVCM when $u > 1$

Due to the inequality $f(\alpha) < \alpha$, the graph of $f(\alpha)$ never intersects the quadrant where $\alpha < 0$ and $f > 0$. The key unexpected fact is that the portions of $f(\alpha)$ within other quadrants play more or less separate roles. In the TVCM case, those quadrants are parts of one (analytically simple) function. But in general they are nearly independent of each other.

The function $f(\alpha)$ was defined as having a graph that lies in the nonanomalous quadrant $\alpha > 0$ and $f > 0$. This f determines completely the multifractal decomposition of our TVCM measure, in particular, the dimension D and the exponents $q^*_{\min}$, $q^*_{\max}$, $\alpha^*_{\min}$ and $\alpha^*_{\max}$.

To the contrary, q_{crit} is entirely determined by the doubly anomalous left tail located in the quadrant characterized by $f(\alpha) < 0$ and $\alpha < 0$. A priori, it was quite unexpected that this quadrant should exist and play *any* role, least of all a central role, in the theory of multifractals. But in fact, q_{crit} has a major effect on the distribution, hence the value of the total measure in an interval.

9.5. The Intermediate Case Where $\alpha_{min} > 0$ but $f(\alpha) < 0$ for some Values of α

When $p < 1/2$, but $u < 1$ so that $q_{crit} = \infty$ and all moments are finite, large values of μ have a much lower probability than when $u > 1$. As always, however, their probability distribution continues to be determined by the left tail of the probability graph where $f < 0$.

10. A BROAD FORM OF THE MULTIFRACTAL FORMALISM THAT ALLOWS $\alpha < 0$ AND $f(\alpha) < 0$

The collection of rules that relate $\tau(q)$ to $f(\alpha)$ is called "multifractal formalism." TVCM was specifically designed to understand multifractals directly, thus avoiding all formalism. However, general random multifractals more than TVCM demand their own broad multifractal formalism. Once again, the most widely known form of the multifractal formalism does not allow randomness and yields $f(\alpha) > 0$, but the broad formalism first introduced in Mandelbrot 1974 concerns a generalized function for which $f(\alpha) < 0$ is allowed.

10.1. The Broad "Multifractal Formalism" Confirms the form of $f(\alpha)$ and Allows $f(\alpha) < 0$ for Some α

Through a point on the graph of coordinates q and $\tau(q)$, draw the tangent to that graph. Under wide conditions, the tangent's slope is $\alpha(q)$ and its intercept by the ordinate axis is $-f(q)$. Thus

$$\alpha(q) = d\tau(q)/dq \qquad \text{and} \qquad -f(q) = \tau(q) - q\, d\tau(q)/dq.$$

Through the quantities $\alpha(q)$ and $f(q)$, a function $f(\alpha)$ is defined by using q as parameter.

The slope $f'(\alpha)$ is the inverse of the function $\alpha(q)$. The tangent of slope $f'(\alpha)$ intersects the line $\alpha = 0$ at the point of ordinate $-\tau(q)$. The $D(q)$ tangent's equation being $-\tau(q) + q\alpha$, its intersection with the bisector satisfies the condition $-\tau + q = \alpha$, hence $D = \tau(q)/(q-1)$. This is the critical embedding dimension discussed in Section 4.4.

10.2. The Legendre and Inverse Legendre Transforms and the Thermodynamical Analogy

The transforms that replace q and $\tau(q)$ by α and $f(\alpha)$, or conversely, are due to Legendre. They play a central role in thermodynamics, as does already the argument that yielded $f(\alpha)$ and $\rho(\alpha)$ in the original formalism introduced in Mandelbrot 1974ab.

ACKNOWLEDGMENTS

The figures in this paper were prepared by Evgenyi Vilenkin, Yale class of 2003.

REFERENCES

H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Stat.* **23**:493–507 (1952).

P. H. Cootner, ed., *The Random Character of Stock Market Prices* (MIT Press, 1964).

H. E. Daniels, Saddlepoint approximations in statistics, *Ann. Math. Stat.* **25**:631–649 (1954).

R. Durrett and T. M. Liggett, Fixed points of the smoothing transformation, *Z. Wahr.* **64**:275–301 (1983).

U. Frisch and G. Parisi, Fully developed turbulence and intermittency, in *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, M. Ghil, ed. (North-Holland, 1985), pp. 84–86, Excerpted in Mandelbrot 1999a.

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

H. G. E. Hentschel and I. Procaccia, The infinite number of generalized dimensions of fractals and strange attractors, *Physica (Utrecht)* **8D**:435–444 (1983).

J. P. Kahane and J. Peyrière, Sur certaines martingales de B. Mandelbrot, *Adv. in Math* **22**:131–145 (1976). Translated in Mandelbrot 1999a as Chapter N17.

A. N. Kolmogorov, A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number, *J. Fluid Mech.* **13**:82–85 (1962).

Q. S. Liu, An extension of a functional equation of Poincaré and Mandelbrot, *Asian Journal of Mathematics* **6**:145–168 (2002).

B. B. Mandelbrot, The variation of certain speculative prices, *The Journal of Business (Chicago)* **36**:394–419, (1963). Reprinted in Cootner 1964, as Chapter E 14 of Mandelbrot 1997, in Telser 2000, and several other collections of papers on finance.

B. B. Mandelbrot, Une classe de processus stochastiques homothétiques à soi; application à la loi climatologique de H. E. Hurst, *Comptes Rendus (Paris)* **260**:3274–3277 (1965). Translated as Chapter H9 of Mandelbrot (2002).

B. B. Mandelbrot, The variation of some other speculative prices, *Journal of Business (Chicago)* **40**:393–413, (1967). Reprinted as Chapter E14 of Mandelbrot 1997, 419–443, in Telser 2000, and several other collections of papers on finance.

B. B. Mandelbrot, Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence, in *Statistical Models and*

Turbulence, M. Rosenblatt and C. Van Atta, eds. (Springer-Verlag, New York, 1972), pp. 333–351. Reprinted in Mandelbrot 1999a as Chapter N14.

B. B. Mandelbrot, Intermittent turbulence in self similar cascades; divergence of high moments and dimension of the carrier, *J. Fluid Mech.* **62**:331–358 (1974a). Reprinted in Mandelbrot 1999a as Chapter N15.

B. B. Mandelbrot, Multiplications aléatoires itérées et distributions invariantes par moyenne pondérée aléatoire, *Comptes Rendus (Paris)* **278A**:289–292 and 355–358 (1974b). Reprinted in Mandelbrot, 1999a as Chapter N16.

B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, New York, 1982).

B. B. Mandelbrot, Fractals in physics: squig clusters, diffusions, fractal measures and the unicity of fractal dimension, *J. Stat. Phys.* **34**:895–930 (1984).

B. B. Mandelbrot, Multifractal measures, especially for the geophysicist, *Pure and Applied Geophysics* **131**:5–42 (1989a).

B. B. Mandelbrot, A class of multinomial multifractal measures with negative (latent) values for the "dimension" $f(\alpha)$, in *Fractals' Physical Origin and Properties*, L. Pietronero, ed. (Plenum, New York, 1989b), pp. 3–29.

B. B. Mandelbrot, Negative fractal dimensions and multifractals, *Phys. A* **163**:306-315 (1990a).

B. B. Mandelbrot, New "anomalous" multiplicative multifractals: left-sided $f(\alpha)$ and the modeling of DLA, *Physica A* **168**:95–111 (1990b).

B. B. Mandelbrot, Limit lognormal multifractal measures, in *Frontiers of Physics: Landau Memorial Conference*, E. A. Gotsman, Y. Ne'eman, and A. Voronel, eds. (Pergamon, New York, 1990c), pp. 309–340.

B. B. Mandelbrot, Negative dimensions and Hölders, multifractals and their Hölder spectra, and the role of lateral preasymptotics in science, in *J. P. Kahane's meeting* (Paris, 1993). A. Bonami and J. Peyrière, eds., *The Journal of Fourier Analysis and Applications* special issue, 409–432 (1995).

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

B. B. Mandelbrot, *Multifractals and 1/f Noise: Wild Self-Affinity in Physics (Selecta Volume N)* (Springer-Verlag, 1999a).

B. B. Mandelbrot, A multifractal walk through Wall Street, *Scientific American*, February issue, 50–53 (1999b).

B. B. Mandelbrot, Scaling in financial prices, I: Tails and dependence, *Quantitative Finance* **1**:113–124, (2001a). Reprint: *Beyond Efficiency and Equilibrium*, Doyne Farmer and John Geanakoplos, eds. (Oxford UK, The University Press, 2002).

B. B. Mandelbrot, Scaling in financial prices, II: Multifractals and the star equation, *Quantitative Finance* **1**:124–130, (2001b). Reprint: *Beyond Efficiency and Equilibrium*, Doyne Farmer and John Geanakoplos, eds. (Oxford UK, The University Press, 2002).

B. B. Mandelbrot, Scaling in financial prices, III: Cartoon Brownian motions in multifractal time, *Quantitative Finance* **1**:427–440 (2001c).

B. B. Mandelbrot, Scaling in financial prices, IV: Multifractal concentration, *Quantitative Finance* **1**:558–559 (2001d).

B. B. Mandelbrot, Stochastic volatility, power-laws and long memory, *Quantitative Finance* **1**:427–440 (2001e).

B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals (Selecta Volume H)* (Springer-Verlag, 2002).

B. B. Mandelbrot, 2003, Forthcoming.

B. B. Mandelbrot, L. Calvet, and A. Fisher, *The Multifractal Model of Asset Returns. Large Deviations and the Distribution of Price Changes. The Multifractality of the*

Deutschmark/US Dollar Exchange Rate. Discussion Papers numbers 1164, 1165, and 1166 of the Cowles Foundation for Economics at Yale University, New Haven, CT, 1997. Available on the web: http://papers.ssrn.com/sol3/paper.taf? ABSTRACT_ID=78588. http://papers.ssrn.com/sol3/paper.taf? ABSTRACT_ID=78606. http://papers.ssrn.com/sol3/paper.taf? ABSTRACT_ID=78628.

A. M. Obukhov, Some specific features of atmospheric turbulence, *J. Fluid Mech.* **13**:77–81 (1962).

L. Telser, (ed.) *Classic Futures: Lessons from the Past for the Electronic age* (Risk Books, London, 2000).