Fractal Measures (Their Infinite Moment Sequences and Dimensions) and Multiplicative Chaos: Early Works and Open Problems

B.B. Mandelbrot

Physics Department, IBM Research Center, Mathematics Department, Harvard University, Cambridge, MA 02138, USA

An infinite sequence of moments is needed to describe a fractal measure. This fact is widely known today, largely thanks to several speakers at this conference, who either refer to it, or push well beyond. Here, I propose to sketch the extensive early background in my work (before 1968) on the theory of turbulent intermittency. This old story matters, because my general procedure also brings forward a number of topics that have not been duplicated, and calls attention to interesting open issues.

1. TWO MAIN TRUNKS OF DEVELOPMENT AND BRANCHES: AN OUTLINE

Having discovered the need for an infinite sequence of moments shortly after the 1966 Kyoto Turbulence Conference, I reported it at the 1968 Brooklyn Symposium [1]. Recently, the telling term "multiplicative chaos" has been attached to the procedures that generate the fractal measures I studied, as well as variants, old or new. This explains the term "M-measure" to be used here.

Two "trunks" separated immediately. The first [2] involves discrete cascades, and fractals that are exactly renormalizable, because of an underlying hierarchical grid. The moments of orders 2/3, 2 and 4 were stressed in Orsay [3], and everything was summarized in Haifa [4]. The second trunk, involving continuous cascades, started at La Jolla [5].

A mathematical branch of the first trunk started in 1974 [6]. Some of my conjectures and theorems were proven or extended by J. PEYRIERE and J. P. KAHANE [7], which triggered other mathematics. Recently, KAHANE [8] proved corresponding conjectures in the second trunk.

The next major event was the rediscovery of results on M-measures by HENTSCHEL and PROCACCIA in 1982 [9], and the many rich developments that followed and are mostly beyond our scope here. Suffices to say that the growth of the main trunk has resumed [10,11]. PARISI and his coworkers [11] call the M-measures "multifractals", but _multi_ is redundant, since all fractals involve a multitude of dimensions, with the exception of the strictly self-similar sets.

Dimensions and Entropies in Dynamical Systems (Pecos River NM, 1985).
Edited by Gottfried Mayer-Kress, New York: Springer, 19-27.

2. ONE PARAMETER MODELS AND WOULD-BE CLASSES OF UNIVERSALITY

The models of intermittency available in 1968 seemed to manage with only one parameter, and to fall into two classes of universality: "all-or-nothing" and "lognormal".

The first models, independent of each other, were by KOLMOGOROV [12] and by BERGER and MANDELBROT [13]. My work concerned noise, but was soon modified to concern turbulence [14]. Then came NOVIKOV and STEWART [15]. The latter performed a recursive interpolation in a hierarchical cubic grid, hence involved self-similarities restricted to ratios the form b^k, with b an integer base b. The parameter b is not of immediate importance. Kolmogorov and I required no grid and allowed self-similarity of arbitrary $r>0$.

The parameter I featured was the fractal dimension D of the support of dissipation in fractally homogeneous turbulence. Novikov-Stewart featured the correlation exponent Q of the turbulent dissipation; their model being fractally homogeneous, this is the fractal co-dimension of the support of dissipation. Kolmogorov used one parameter μ, which specifies a log-normal distribution. In my "Kolmogorov-related" models, $\mu/2$ was to become the fractal co-dimension of the set on which dissipation concentrates. An excellent expository paper [16], which had the great merit of bringing my work to a wide public, stresses a parameter β, which again is not of immediate importance, but led to the term "β-model" often attached to fractally homogeneous turbulence.

Kolmogorov´s model was enormously influential. Unfortunately, I found lognormality to be untenable as he stated it. (The words "Possible refinement..." in the title of [5] only reflect the difficulty then facing a negative comment on a parcel of Kolmogorov´s work.) When a very great scholar stumbles in this way, something subtle is involved.

His basic idea is unchanged to this day: the idea of replacing sums of random processes by products that illustrate the notion of cascade. A physicist expects sums of random variables to be in the "domain of universality" of the Gaussian. So it seems safe to expect products of well-behaved strictly positive variables to converge to the lognormal, and this was proposed by GURVITCH and YAGLOM [17] to justify Kolmogorov´s lognormality on very small scales. However, a step that seems harmless is incorrect in this instance: when a random variable x tends to a Gaussian, the moments of exp(x) need not tend to the moments of exp(G). This is a clear failure of universality, and its consequences are very interesting.

3. MULTIPLICATIVE CHAOS: MICROCANONICAL AND NONRANDOM

The M-measures are "singular" measures, i.e., continuous measures that fail to have a derivative. Examples of strict conservative M-measures abound in pure mathematics, and the new developments since 1968 resided in their use in science, and in their characterization by moments. I also introduced "mean conservative" M-measures; this concept raised altogether new issues.

A cascade process starts with a uniform measure. When the stages are discrete, the k-th stage multiplies the (k+1)st approximate measure by the k-th perturbation $P_k(\underline{x})$. Therefore, the k-th approximate measure of a domain Δ is $\mu_k(\Delta)=\int_\Delta \Pi_{h=1}{}^k P_h(\underline{s})d\underline{s}$, and one is interested in the limit $\mu(\Delta)=\lim_{k\to\infty}\mu_k(\Delta)$. The case $P_k(\underline{x})\geq 0$ is best understood (which is why - Section 6 - the most interesting new problems arise when $P_k(\underline{x})<0$ is allowed.) When the cascade proceeds in a grid of base b, the perturbations are called strictly conservative if $P_k(\underline{x})$ is constant over grid cells of side b^{-k}, and $\int_\Delta P_k(\underline{s})d\underline{s}=1$, with Δ any cell of side b^{-k}.

The B-measure of Besicovitch. This is my term for the special M-measure on a grid obtained when the perturbations are non-random, and $P_k(b^{k-1}\underline{x})=P_1(\underline{x})$, independently of k. $P_1(\underline{x})$ is the generator (="perturbator"?) of the measure. On the line, the generator is built from b "probabilities" p_β, satisfying $\sum p_\beta=1$, and $P_1(t)$ equals $bp_{t(1)+1}$ if $t=0.t(1)t(2)...t(k)$ in base b. Other perturbations at time t are $P_k(t)=bp_{t(k)+1}$. The integral $F_k(t)=\int_o^t \Pi_{h=1}{}^k P_h(s)ds$ is monotone non-decreasing, and is obtained by recursive interpolation. And $F(t)=\lim_{k\to\infty}F_k(t)$ is a self-affine non-random function of t. That is, the portion of F(t) over the interval $[(\beta-1)/b, \beta/b]$ is obtained from the portion F(t) over [0,1] by changing t in the ratio 1/b, and F in the ratio p_β, then translating. Reductions with unequal ratios are not similarities, but affinities [18], and F(t) is fully determined by the collection of affinities under which it is invariant. A generator for these affinities is a nondecreasing broken line with breaks located at multiples of 1/b. While "self-affine function" is a term used in my books, an explicit study is very recent [18] and it provides the proper framework here.

The Hentschel-Procaccia Measures. For many readers of this book, the first contact with the complexity of fractal measures came through [9], where HENTSCHEL and PROCACCIA introduce self-affine non-random fractal measures more general than the B-measures. In the 1-d case, the novelty is that the generator is a non-decreasing broken line with breaks located at arbitrary values of t, instead of multiples of 1/b.

The infinity of exponents. The averages of the quantities $\mu^h(\Delta)$ over all subcells Δ of given size need not be derived in this section, because the argument is identical for the expectations of $\mu^h(\Delta)$ in the random measures in Section 4. In particular, the Hentschel-Procaccia measures involve nearly the same degree of generality as described in Section 4 for random weights in a hierarchical grid.

4. MULTIPLICATIVE CHAOS; MICROCANONICAL IN A GRID AND RANDOM.

The simplest random M-measure is obtained by randomizing, within each cell of side b^{-k}, the positions of the b^k values of $P_k(\underline{x})$.

"Microcanonical" M-measures [2]. The perturbations are conservative, self-affine and stationary within cells. That is, the values of $P_k(\underline{x})$ within different cells of side b^{-k-1} are identically distributed random variables whose sum is 1. It is easiest to start with a random "weight" W satisfying $W>0$ and $\langle W\rangle=1$, and to impose upon the weights W_β in different cells the condition that they must satisfy $\sum W_\beta b^{-d}=1$, i.e., $\sum W_\beta=b^d$. The resulting conditional weight will be denoted by $W_{(d)}$. The values of $P_k(\underline{x})$ in cells of side b^{-k-1}, taken jointly, are sample values of this $W_{(d)}$. Observe that $W_{(d)}<b^d$ and $\langle W_{(d)}\rangle=1$.

The randomized B-measure is the microcanonical M-measure corresponding to W_d having b^d possible values of the form $b^d p_\beta$, with $\sum p_\beta=1$ and $\mathrm{Prob}(W_d=b^d p_\beta)=b^{-d}$ for all β. (Strictly speaking, the assimilation requires that the relation $\sum i_\beta p_\beta=1$, with i_β integer ≥ 0 must be impossible unless $i_\beta=1$ for all β.)

The infinity of exponents. Pick a cell of side b^{-k} at random. For all $h>k$, the measure $\mu_h(\Delta)$ satisfies $\langle\mu_h(\Delta)\rangle=b^{-tk}=|\Delta|$, where $|\Delta|$ is the measure of Δ. Not unexpectedly, all the other moments $\langle\mu_k{}^h(\Delta)\rangle$ are powers of $|\Delta|$. Their exponents, which I evaluated, are $m(h)= -\log_b\langle W^h\rangle + dh$

Their being highly non-universal is well known today, but was a surprise in 1967. To evaluate the fractal dimension of the support of this measure, I introduced a procedure that was new at that time. I observed that a proportion of the measure between 1 and $1-\epsilon$ becomes, after sufficiently many stages $k(\epsilon)$, carried by a self-similar fractal set of codimension arbitrarily close to a quantity independent of ϵ, namely $c(1)=\langle W\log_b W\rangle$.

This may be called the "ϵ-box dimension", the term "box dimension" itself denoting the classical form of fractal dimensions that part of our profession confusingly calls "capacity".

For the randomized B-measure,

$$\langle W \log_b W \rangle = \sum(1/b^d) b^d p_j \log_b b^d p_j = d + \sum p_j \log_b p_j = d - I_1.$$

Hence, the ϵ-box dimension of this measure is I_1, which is the entropy-information of the p_j. It was already well known, however [19], that I_1 is also the Hausdorff-Besicovitch dimension of the set of t´s for which the frequency of the digit β is $p_{\beta+1}$. This set is, loosely speaking, the support of most of Besicovitch measure. This made me conjecture that $\langle W\log_b W\rangle$ is a Hausdorff-Besicovitch codimension for every M-measure, and indeed it is [7].

5. <u>MULTIPLICATIVE CHAOS</u>: <u>CANONICAL</u>. <u>THE LITTLE KNOWN ROLE OF C(h) AS A CRITICAL CODIMENSION</u>. <u>CONTINUOUSLY PERTURBED MULTIPLICATIVE CHAOS</u>

The relations of conservation, $\sum W = b^d$, make a further detailed study of microcanonical cascades very cumbersome. Assuming that conservation only holds on the average makes everything simpler mathematically, and we shall see it yields a richer topic, worth of study on its merits. Anyhow, a low-dimensional cut through a microcanonical M-measure is characterized by partial, not strict, conservation. The reason is that overall conservation expresses that $\sum W_{(d)} = b^d$, the sum being carried over b^d variables, but a cut picks only $b^{d'}$ among these b^d variables. Call these new conditioned variables $W_{(d')}$. When $d'<d$, the $W_{(d')}$ are much less strongly correlated than the $W_{(d)}$. Thus, the model that picks uncorrelated weights and allows the W to be unconditioned and unbounded illustrates a cut through a microcanonical measure of extremely high dimension.

When $W>0$ and $\langle W\rangle=1$ is all that is assumed about W, the measures $\mu_k(\Delta)$ are no longer constructed by recursive interpolation. I showed that strange things may happen. For every domain Δ and $k<\infty$, the k-the approximate measure $\mu_k(\Delta)$ satisfies $\langle\mu_k(\Delta)\rangle=|\Delta|$. However, the seemingly obvious inference that $\langle\lim_{k\to\infty}\mu_k(\Delta)\rangle=|\Delta|$ <u>need not</u> hold. It <u>does</u> hold when $\langle W\log_b W\rangle <d$, but <u>does not</u> hold when $\langle W\log_b W\rangle >d$, and also [7] does not hold when $\langle W\log_b W\rangle=d$. In fact, $\langle W\log_b W\rangle \geq d$ is the necessary and sufficient condition for the cut to be empty almost surely. This result means that a question that seemed a contrived case of mathematical hairsplitting can sometimes become practical. After concrete application has retrained intuition, "hair-splitting" changes to "obvious". In the present case, it suffices to argue as if the measure reduced exactly to being supported by a fractal set of co-dimension $\langle W\log_b W\rangle$ in some high-dimensional space. There is a well-known rule about the effect of intersection upon dimension. Here, this rule shows that $d=\langle W\log_b W\rangle$ is a "critical" dimension: it separates the dimensions of spaces that almost surely miss our fractal, from the dimensions of spaces that hit it with positive probability.

What about the moments of $\mu_k(\Delta)$ when it is non degenerate? I discovered that they may $\to\infty$ as $k\to\infty$. For each space dimension, there is a "critical moment", and for each moment there is a critical space dimension,

$$C(h)=(h-1)^{-1}\log_b\langle W^h\rangle,$$

such that moments are finite for $C(h)>d$ and infinite for $C(h)<d$.

Generalization. Once strict conservation has been abandoned in favor of mean conservation, the perturbation function $P_k(\underline{x})$ need no longer be constant over cells, hence need not be discontinuous. It can be any random function whose correlation range is b^{-k}. Moreover, the base b itself need no longer be an integer. For example, $P_k(\underline{x})$ may be the convolution of a white noise with a kernel having a typical radius of b^{-k}. The effect of this function upon the "texture" of a M-measure very much deserves to be investigated.

The limit lognormal processes of La Jolla [5]. Finally, mean conservation allows the perturbation index k to be made continuous. This was the point of the second trunk of early development mentioned in Section 1. I made $\log P_k(\underline{x})$ a lognormal process, as near as logic allows to Kolmogorov's original idea. There is a sketch in my 1982 book [p. 379]. I showed that $\mu/2$ is the ϵ-box codimension. Recently [8], it has been shown that the Hausdorff-Besicovitch codimension is also $\mu/2$.

The term "Schutzenberger-Renyi Informations." In the special cases of the Besicovitch measure and of related nonrandom fractal measures,

$(h-1)^{-1}\log_b\langle W^h\rangle$ becomes $d-I_h$, where $I_h = (h-1)^{-1} \log_b\sum p_j{}^h$.

Doyne Farmer noticed - after re-deriving I_h - that A. Renyi had called it a "generalized information". A precursor was M. P. Schutzenberger. There is a book that shows rigorously that I_h satisfies axioms that justify calling it "information". However, I happen to subscribe to Lebesgue's wariness of notions that serve no purpose besides being defined. Claude Shannon was not the first to write I_1, but the first to encounter I_1 in unexpected inequalities that inject entropy into the study of communication. In the study of fractal measures, I_1 was first encountered as a Hausdorff-Besicovitch dimension by Besicovitch and his students [19]. But there was no early counterpart for other I_h's.

On the scope of the term "fractal dimension". "Fractal dimension" should now be a generic notion, special cases of which are the box dimension ("capacity"), Frostman´s capacity dimension, the ϵ-box dimension, the similarity dimension, the gap dimension, the Hausdorff-Besicovitch dimension, etc... However, some papers on M-measures follow a usage that restricts the generic term to the fractal dust that supports the M-measures. I feel the usage is misleading.

6. MULTIPLICATIVE CHAOS WITH WEIGHTS OF EITHER SIGN, AND A SURROGATE FOR BROWNIAN MOTION.

Open problems concerning multiplicative chaos are most numerous and obvious in the case when the weight W can take either sign. One new example [18] gives the flavor. On the line, one needs, in addition to the base b, a second base b">0 such that b-b">0 and is even; we shall write $H=\log_b b"$ so that $0<H<1$. The weight W will be two-valued: $W=\pm b/b"$. Strict conservation (of something like electric charge rather than mass!) is achieved by setting $W=+b/b"$ over $(b+b")/2$ cells of length b^{-1} and $W=-b/b"$ over the remaining ones. The sequence of + and - forms the generator. It may be fixed, yielding a non-random M-measure, or chosen each time at random under the above constraint, yielding a microcanonical M-measure. The functions $F_k(t)$ are no longer nondecreasing, and $F(t)=\lim_{k\to\infty}F_k(t)$ is shown in [18] to be a self-affine function, whose increment over an interval b^{-k} in the grid is $|\Delta F| = \pm|\Delta t|^H$, exactly. Similarly, fractional Brownian motion $B_H(t)$ (Wiener´s Brownian motion if H=.5) satisfies $|\Delta B_H|\sim|\Delta t|^H$. However, the distribution of ΔF is not Gaussian but binomial. This makes F(t) a useful surrogate of $B_H(t)$. The exponent of the h-th absolute moment of ΔF is $m^+(h)=-\log_b<|W|^h> + h = hH$.

It is linear in h, which is the simplest possible behavior. (In the case of positive M-measures, m(h) linear in h corresponds to the M-measure that is homogeneous on a fractal dust). The critical exponent is the value of h for which $m^+(h)=hH=1$ is 1/H. To explore its significance, consider the h-variation of F, defined by $\int|\Delta F|^h = |\Delta t|^{hH-1}$, and let $\Delta t\to 0$.

When h>1/H, $\int|\Delta F|^h\to 0$, but when h<1/H, $\int|\Delta F|^h\to\infty$.

$\int|\Delta F|\to\infty$ expresses that F is not of bounded variation. With respect to $\int|\Delta F|^h$, F(t) behaves like $B_H(t)$. Observe that divergence occurs here below the critical h, and concerns the microcanonical case, while for the positive M-measure we know divergence occurs above the critical h, and is found only in the canonical case.

The corresponding canonical M-measure is obtained when W is binomial, with Pr(W=b/b")=(b+b")/2b and Pr(W=-b/b")=(b-b")/2b. Now, ΔF is no longer binomial. Its h-th moment is finite when h<1/H, but infinite when h>1/H. (For example, moments of order h>2 are infinite when H takes the Brown value 0.5.) On both counts, the canonical version is very different from $B_H(t)$. But it is an exciting object for study, and I expect it to be useful; the little I know of its properties will be reported on elsewhere.

In the space of d>1 dimensions, we write $H=\log_b b''/d$, and we select $W=\pm b^{d}/b''=\pm b^{d(1-H)}$. Strict conservation now requires W>0 over $(b^d+b'')/2$ cells and W<0 over the other cells. Again, microcanonical M-measure of a cell Δ, of side b^{-k} and of content $|\Delta|$, satisfies $|\mu_k(\Delta)| = b^{-Hkd} = |\Delta|^H$, and the critical value for the divergence of the h-variation is h=1/H.

REFERENCES

1. B. B. Mandelbrot, in Proceedings of the Symposium on Turbulence of Fluids and Plasmas (Brooklyn Poly, New York, 1968) p. 483 (Interscience, New York,1969).

2. B. B. Mandelbrot, J. Fluid Mech. 62:331 (1974).

3. B. B. Mandelbrot, in Turbulence and Navier Stokes Equation (Orsay, 1975). Lecture Notes in Mathematics. Vol. 565, p. 121 (Springer, New York, 1976).

4. B. B. Mandelbrot, in Statistical Physics Conference (Haifa, 1977) p. 225 (Bristol, Adam Hilger 1978).

5. B. B. Mandelbrot, in Statistical Models and Turbulence (La Jolla, 1972) Lecture Notes in Physics: Vol. 12, p. 333 (Springer, New York, 1972).

6. B. B. Mandelbrot, C. R. Acad. Sci. (Paris) 278A: 289 and 355 (1974).

7. J. Peyriere, C. R. Acad. Sci. (Paris) 278A:567 (1974). J. P. Kahane, C. R. Acad. Sci. (Paris) 278A: 621(1974). J.P. Kahane and J. Peyriere, Adv. Math. 22:131 (1976).

8. J. P. Kahane, C. R. Acad. Sc. (Paris) 301A (1985).

9. H. G. E. Hentschel and I. Procaccia, Physica 8D:435 (1983).

10. B. B. Mandelbrot, J. Stat. Phys. 34: 895 (1984).

11. R. Benzi, G. Paladin, G. Parisi and A. Vulpiani, J. Phys. 17A: 3521 (1984).

12. A. N. Kolmogorov, J. Fluid Mech. 13:82 (1962). Also A. M. Oboukhov, J. Fluid Mech. 13: 77 (1962).

13. J. M. Berger and B. B. Mandelbrot, IBM J. Res. Dev. 7: 224(1963).

14. B. B. Mandelbrot, IEEE Trans. Comm. Techn. 13: 71 (1965). Also Proc. Fifth Berkeley Symp. Math. Stat. and Probability 3:155 (1967). Also IEEE Trans. Inf. Theory 13: 289 (1967).

15. E. A. Novikov and R. W. Stewart, Isv. Akad. Nauk SSSR, Seria Geofiz. 3: 408 (1964).

16. U. Frisch, M. Nelkin and J. P. Sulem, J. Fluid Mech. 87:719 (1978).

17. A. S. Gurvitch and A. M. Yaglom, Physics of Fluids 10: 559(1967).

18. B. B. Mandelbrot, in Fractals in Physics (Trieste 1985) (Amsterdam, North-Holland, 1986).

19. P. Billingsley, Ergodic Theory and Information. (J. Wiley, New York, 1967).