

THE MINKOWSKI MEASURE AND MULTIFRACTAL ANOMALIES IN INVARIANT MEASURES OF PARABOLIC DYNAMIC SYSTEMS

BENOIT B. MANDELBROT

*Physics Department, IBM T. J. Watson Research Center,
Yorktown Heights NY 10598, U.S.A.*

and

*Mathematics Department, Yale University,
New Haven CT 06520, U.S.A.*

ABSTRACT

The author has recently shown that a multifractal measure due to Minkowski has the characteristic that its $f(x)$ distribution is left-sided, i.e., has no descending right-side corresponding to a decreasing $f(x)$; this creates many very interesting complications. Denjoy had observed that this Minkowski measure is the restriction to $[0, 1]$ of the attractor measure for the dynamical system on the line based on the maps $x \rightarrow 1/2 + 1/4(x - 1/2)$, $x \rightarrow -x$ and $x \rightarrow 2 - x$. This paper points out that it follows from Denjoy's old observation that the new "anomalies" due to the left-sidedness of $f(x)$ extend to the invariant measures of certain dynamical systems.

This short paper is primarily meant to describe an experience I lived through, and to address a warning to the specialist in multifractals. There is a widespread expectation that the $f(x)$ distribution of every multifractal measure μ satisfies $f > 0$ and has a graph shaped like the mathematical symbol \cap . From classical results from the 1930s and the 1940s (by Besicovitch, Eggleston et al.), this expectation is fulfilled when μ is the binomial measure, or near binomial.

But for many measures — some theoretical, other obtained from nature — this expectation grossly *fails* to be fulfilled, in one way or another. This creates a variety of so-called "anomalies." Some anomalies are relative to measures in real space; examples include the distributions of turbulent dissipation (the notion of multifractal first arose in that context⁵) and of the harmonic measure around a DLA cluster. These anomalies are beginning to be well recognized; in particular, the references show that I have contributed several papers to their investigation.

I expect that many of the same anomalies will also be encountered for multifractal measures encountered in dynamical systems. The goal of this paper is to make the students of measures on attractors aware of the preceding references. To do so, I shall discuss the Minkowski measure, μ , which is most attractive but extremely anomalous.

Chaos in Australia (Sydney, 1990).
Edited by Gavin Brown & Alex Opie.
Singapore: World Publishing, 83-94.

When applied to this measure, my approach to multifractals (to be called the *method of distributions*) will be seen to involve necessarily two distinct two aspects of $f(x)$. First, there is a theoretical "population" function $f(x)$. For the Minkowski μ , this $f(x)$ is *not* \cap -shaped; instead, it is left-sided, i.e., monotone increasing toward a maximum $f(\infty) = 1$ and without a right (decreasing) side. But let the measure μ be coarse-grained to intervals of length ε , in order to become observable. If so, the method of distributions defines an empirical "sample" $f_\varepsilon(x)$ for each ε . This $f_\varepsilon(x)$ is observable, and for the Minkowski μ its shape happens to be altogether different from that of $f(x)$.

As $\varepsilon \rightarrow 0$, one has $f_\varepsilon(x) \rightarrow f(x)$, which expresses that the theory behind the method of distribution is logically consistent. But the convergence is excruciatingly slow and extremely singular; very great care is needed to extrapolate the shape of $f(x)$ from that of the $f_\varepsilon(x)$.

When thinking of attractor measures, there is good reason to think first of those supported by strange attractors. But an extraordinary confusion continues to characterize much of the literature on multifractals. This shows that the topic is more delicate than many realize, and that it is best to tackle each issue after every extraneous difficulty has been eliminated.

1. The Minkowski measure on the interval $[0, 1]$

The *Minkowski measure* μ is simple to define and work with, but exhibits very interesting and totally unexpected peculiarities. This μ and the *inverse Minkowski measure* $\tilde{\mu}$ are, respectively, the differentials of two increasing singular functions: $M(x)$ and its inverse $X(m)$.

It is easier to start by defining the inverse Minkowski function $X(m)$, which is constructed step by step, as follows. The first step sets $X(0) = 0$ and $X(1/2) = 1/2$. The second step interpolates: $X(1/4)$ is taken to be the Farey mean of $X(0)$ and $X(1/2)$, where the Farey mean of two irreducible ratios (a/c) and (b/d) is defined as $(a+b)/(c+d)$. More generally, the k -th step begins with $X(m)$ defined for $m = p2^{-k}$, where p is an even integer, and uses Farey means to interpolate to $m = p2^{-k}$, where p is an odd integer. Finally, $X(m)$ is extended to the interval $[1/2, 1]$ by writing $X(1-m) = 1 - X(m)$. The resulting function $X(m)$ is continuous, it increases in every interval, and it is singular; that is, it has no finite derivative at any point. It has an inverse function $M(x)$ with the same properties, illustrated by the "slippery staircase" in Figure 1.

The differentials $\tilde{\mu}$ and μ of the functions $X(m)$ and $M(x)$ are singular measures. The two parts of Figure 2 illustrate the measure μ , as evaluated for intervals of length 10^{-5} .

Minkowski (15, Vol 2, p. 50-51), had called $M(x)$ the " $?(x)$ function," which has few redeeming features. Little (if anything) was written about the measure μ until 1932, when Denjoy¹ observed that it has the following

property. It is the restriction to $[0, 1]$ of the attractor measure for the dynamical system on the line based on the maps

$$x \rightarrow \frac{1}{2} + \frac{1}{4(x - 1/2)}, \quad x \rightarrow -x \quad \text{and} \quad x \rightarrow 2 - x.$$

To transform such a collection of functions into a dynamical system, the standard method is, of course, to choose the next operation at random. This method was used in Plates 198 and 199 of my book ⁶, and the current (and recent) term for it is IFS: *iterated function system*.

Thanks to this interpretation, $M(x)$ proves to have deep roots in number theory (modular functions) and in Fuchsian or Kleinian groups. From this paper's viewpoint, however, the main virtue of the above dynamical system lies in its extraordinary simplicity. I surmise that any complication or difficulty encountered in the study of its invariant measures will *a fortiori* appear in more complex systems grounded in physics. Moreover, one must keep in mind that the paper I wrote with Gutzwiller ³ had two motivations. I was concerned with the above maps, but he was concerned with an important Hamiltonian system in which x is the Liouville measure and m a second invariant measure yielding equally interesting information about individual trajectories.

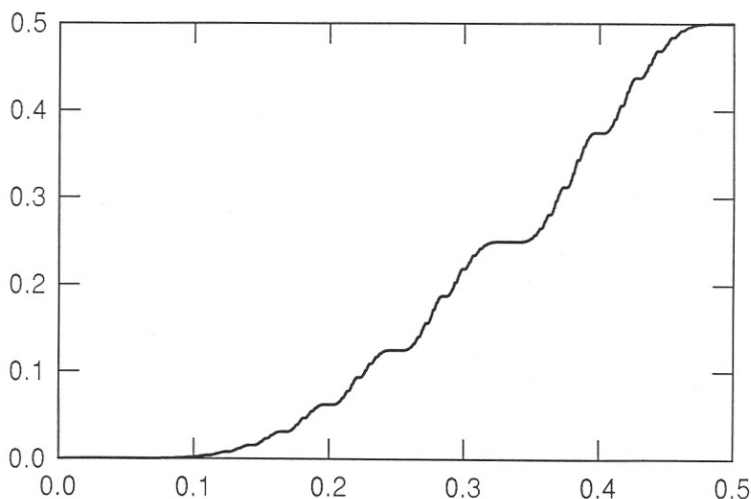


Figure 1. Graph of the Minkowski function $M(x)$ for $0 < x < 1/2$. Contrary to the well-known Cantor devil staircase (my book ⁶, plate 83), the graph of $M(x)$ has no actual steps, only *near steps* that led to its being called a *slippery staircase* in ³.

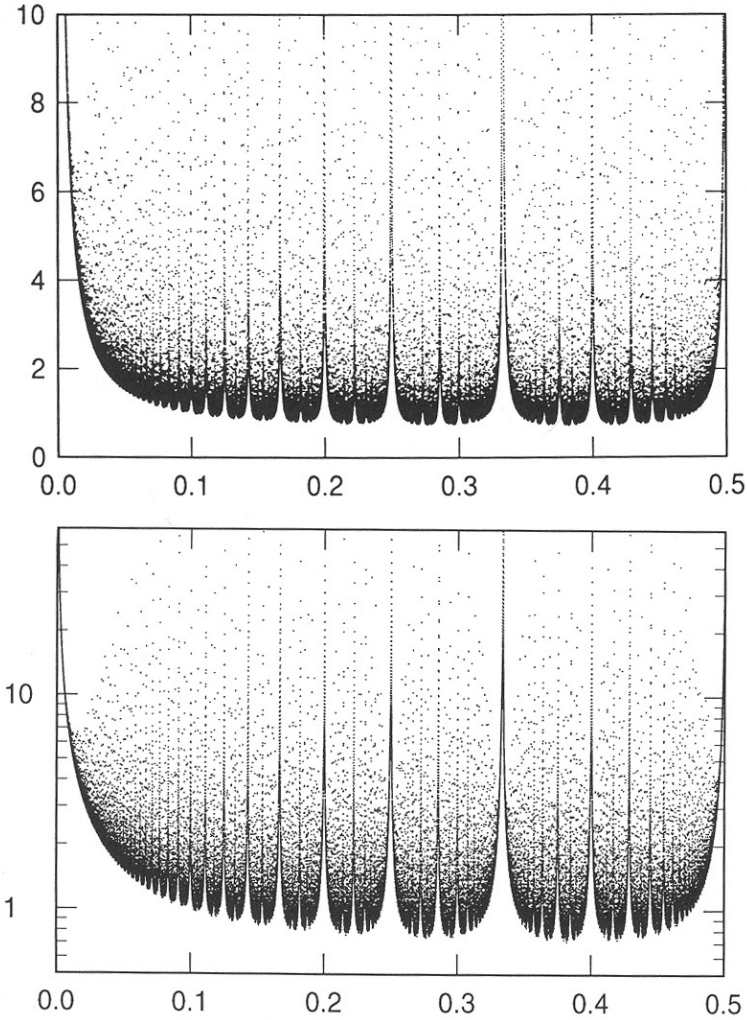


Figure 2. Plots of the Minkowski measure μ in the form of the coarse-grained Hölder exponent and of its logarithm. Taking $\varepsilon = 10^{-5}$, we evaluated the increments $\Delta M = M[(k+1)\varepsilon] - M[k\varepsilon]$. Figure 2a (top) plots the coarse-grained Hölder $\alpha = \log \Delta M / \log \varepsilon$, and Figure 2b (bottom) plots $\log \alpha$. The theory described later in the paper shows that the fine-grained (local) Hölder almost everywhere exceeds any prescribed α . In this figure, this property altogether fails to be reflected.

2. The functions $f(\alpha)$ and $f_c(\alpha)$ of the Minkowski measure

2.1. The theoretical function $f(\alpha)$

For a derivation of the $f(\alpha)$ functions of the Minkowski measure μ and of the inverse Minkowski measure $\tilde{\mu}$, we must refer the reader elsewhere.¹² A first basic fact is that there is a theoretical $f(\alpha)$ for these measures: $f(\alpha)$ is the Hausdorff dimension of the set of points x such that the Hölder exponent $H(x)$ takes the value α . In the case of μ , the graph of $f(\alpha)$ has the following properties:

- $f(\alpha)$ is defined for $\alpha \geq \alpha_{\min} = -1/\log_2 \gamma^2 \sim .7202 \dots$ where γ is the golden mean $\sim .6180 \dots$ (obtained in¹⁶ and — independently — in³)
- $\alpha_1 = [2 \int_0^1 \log_2(1+x) dM(x)]^{-1} = .874 \dots$ (obtained in¹⁷)
- $f(\alpha) \rightarrow 1$ as $\alpha \rightarrow \infty$. This property of $f(\alpha)$ has a strong bearing on the nature of the set of points x where $H(x) > \tilde{\alpha}$: this set is of measure 1.

2.2. The problem of inferring the shape of $f(\alpha)$ from data

After $f(\alpha)$ has been specified analytically, one cannot rest. One must continue by asking whether or not this function can also be inferred when the mechanism of our dynamical system is unknown and only some “empirical data” are available. In this context, “data” may mean one of two things. It may denote the “coarse” (or coarse-grained or quantized) form of the function $M(x)$, as computed effectively for values of x restricted to be multiples of some quantum $\Delta x = \varepsilon$. “Data” may also denote a long orbit of the above dynamical system, that is, a long series of successive values of x ; they too must be recorded in coarse-grained format. For many physical quantities in real space, coarse-graining is physically intrinsic; i.e., they are not defined on a continuous scale, but only for intervals whose length is a multiple of some $\Delta x = \varepsilon$ due to the existence of atoms or quanta; in other physical quantities, there are intrinsic limits to useful interpolation; for example, a turbulent fluid is locally smooth. In the present case, coarse-graining is the result of the necessary finiteness of actual computations and of observed orbits.

Given coarse data, there are at least two ways of seeking to extract or estimate $f(\alpha)$.

2.3. The method of moments as applied to the Minkowski distribution

The better-known way^{2, 4} deserves to be called the *method of moments*. It starts with the coarse-grained measures $\mu_\varepsilon(x)$ contained in successive intervals of length ε and proceeds as follows: a) evaluate the collection of moments embodied in the “partition function” defined by $\chi(\varepsilon, q) = \sum \mu_\varepsilon^q(x)$; b) estimate $\tau(q)$ by fitting a straight line to the data of $\log \chi(\varepsilon, q)$ versus $\log \varepsilon$, and c) obtain $f(\alpha)$ as the Legendre transform of $\tau(q)$.

When applied mechanically to the Minkowski μ , the method of moments either yields nothing or yields nonsense. More precisely, the more prudent mechanical implementations of the method do not fit a slope $\tau(q)$ without also

testing that the data are straight (this can be done by eye). But the Minkowski data for $q < 0$ are not straight at all. Therefore, the prudent conclusion is that "there is no $\tau(q)$." Since no such difficulty arises for $q > 0$, conclusions of this sort are often accompanied by the assertion that the data are "not quite" multifractal. The less prudent mechanical implementations of the method of moments simply forge ahead to fit $\tau(q)$. Depending on a combination of the rule used to fit and of the details of how quantification is performed, those methods may yield an estimated $\tau(q)$ that is not convex. The resulting "Legendre transform" is not a single-valued function, and $f(\alpha)$ is a mystery. In other mechanical methods, the difficulty in estimating $\tau(q)$ is faced by first "stabilizing" the estimate in one way or another; such stabilization may yield some sort of $f(\alpha)$, but one can hardly say what it means and what purpose it serves.

A central feature of the method of moments should be mentioned at this point. The limit process $\varepsilon \rightarrow 0$ is invoked in estimating $\tau(q)$ from the data. But the preasymptotic data corresponding to $\varepsilon > 0$ do *not* define an approximate $f_\varepsilon(\alpha)$.

2.4. The method of distributions

A second way to estimate $f(\alpha)$ is the *method of distributions*, which is used in all my papers listed as references. I have been using it since 1974 and every new development motivates me to recommend it more strongly. The key is simple. While the method of moments rushes to compute the moments of $\mu_\varepsilon(x)$ embodied in the partition function $\chi(\varepsilon, q)$, the method of distributions considers, for every ε , the full frequency distribution of the $\mu_\varepsilon(x)$. These distributions are embodied in graphs statisticians call *histograms*.

First, the range of observed α 's is subdivided into equal "bins." and one records the number of data in each bin. If the number of bins is too small, information is lost, but if there are too many bins, many are empty. In the case of the Minkowski μ , there are many α 's a little above α_{\min} and few α 's strung along up to very high values.

Denote by N_b the number of data in bin b . When N_b is large, $N_b/\Delta\alpha$ serves to estimate a probability density for α . When $N_b = 1$ and the neighboring bins are empty, one estimates probabilities by averaging over a suitably large number of neighboring bins; these probabilities are very small.

Having estimated the probability density $p_\varepsilon(x)$, one forms

$$f_\varepsilon(\alpha) = \frac{\log p_\varepsilon(\alpha)}{\log \varepsilon}.$$

Thus, the method of distribution creates a sequence of functions $f_\varepsilon(\alpha)$. Because $f_\varepsilon(\alpha)$ is the normalized logarithm of a measure, each $f_\varepsilon(\alpha)$ is nothing but a histogram that was replotted in doubly logarithmic coordinates and was

suitably weighted. These histograms should be evaluated for a series of values of ε . When the measure is multifractal, $f_\varepsilon(\alpha)$ converges to a limit $f(\alpha)$. That is, the function $f(\alpha)$ enters the theory as

$$f(\alpha) = \lim_{\varepsilon \rightarrow 0} f_\varepsilon(\alpha).$$

A physicist's typical reaction to histograms is, "Why bother? We all know that the information contained in the histograms is also contained in the moments; besides, moments organize information, and they are familiar and far easier to handle than histograms." Unfortunately, in the context of fractals and multifractals this typical reaction *won't do*.

In the study of fractals, the typical probability distributions are scaling (hyperbolic), and some of their population moments are infinite. The corresponding sample moments — sometimes even the sample average — behave in totally erratic fashion; they bring out no useful information and can be thoroughly misleading.

Now proceed to multifractals. When $f(\alpha)$ is truly \cap -shaped, with $f > 0$, moments raise no major issue, the method of moments works well, and the method of distributions is a less efficient way to obtain $f(\alpha)$. But in all delicate cases, the sample moments embodied in the partition function are treacherous. The method of distributions is the only way to go.

2.5. The method of distributions as applied to the Minkowski measure

In ³, Gutzwiller and I used histograms, and Figure 3 (which is copied from Figure 2 of ³) reproduces the empirical $f_\varepsilon(\alpha)$ we obtained. To obtain this graph, we coarse-grained x , then (in effect) we coarse-grained M . The "quantum" of M was tiny, because it was simply the smallest $M(x + \varepsilon) - M(x)$ our computer allowed in quadruple precision. Thus, the values of the computer could not distinguish from 0 (10% of the whole) were not used.

The resulting data-based curve is utterly different from the theoretical left-sided $f(\alpha)$. It begins with an unquestionably cap-convex left side — as usual. The middle part satisfies $f_\varepsilon(\alpha) > 1$, which cannot be true of $f(\alpha)$, but was expected; this is one of the inevitable biases of the method of distributions, and can be handled. Finally, there is cup-convex right side. This was totally unexpected, because a theoretical $f(\alpha)$ is necessarily cap-convex throughout.

We gave up seeking a better test of this cup-convexity. We did not come close to testing my further hunch, that the estimated $f_\varepsilon(\alpha)$ — if extended far enough — would become < 0 for large enough α . We showed that for $q < 0$ the moment $\chi(q, \varepsilon)$ was not a power law function of ε . But, to our disappointment, we did not succeed in evaluating $f(\alpha)$ analytically. We did conjecture the correct actual form of $f(\alpha)$, (but did not write it down) and were concerned by functions $f_\varepsilon(\alpha)$ and $f(\alpha)$ that differ to such extreme degree.

Recently, I have returned to this problem. Figure 4 was prepared using a method that does not compute $M(x)$ itself, but computes $M(x + dx) - M(x)$ directly. This can be done with arbitrary relative precision, therefore we can reach huge values of x . Figure 4 gives resounding confirmations of the earlier conjectures concerning the existence of a cup-convex right side in the empirical $f(x)$ and of a negative tail.

This sharp mismatch between the theory and even the best experiments spurred me to a rigorous derivation of the theoretical $f(x)$ and of the predicted $f_\varepsilon(x)$. The shape of $f(x)$ has already been mentioned. For $f_\varepsilon(x)$, it suffices to say that, for large x ,

$$f_\varepsilon(x) \sim 1 - (\text{a constant}) \log x / \log \varepsilon.$$

Figure 4 verifies this dependence on the data.

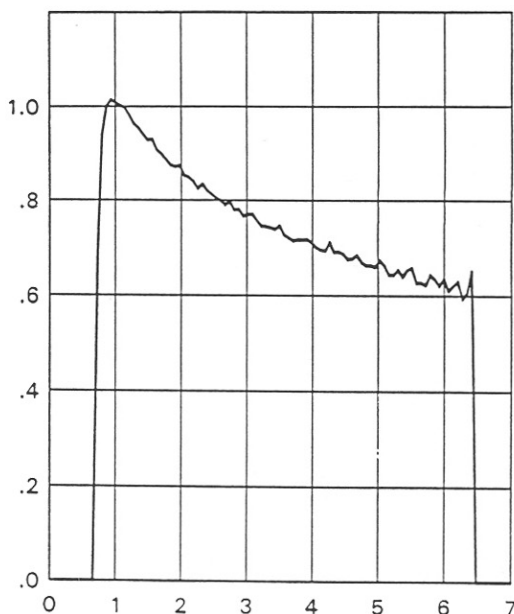


Figure 3. An early plot of the estimated function $f_i(x)$ for the Minkowski measure. Reproduced from³

2.6. Is $f(\alpha)$ a useful notion in the case of the Minkowski μ ?

Once again, our recent evaluations of $f_\varepsilon(\alpha)$ did not come close to reproducing the true shape of the graph of $f(\alpha)$, despite the fact that they involved precision that is totally beyond any conceivable physical measurement. Even the early evaluations³ were well beyond the reach of physics.

Given the difficulties that have been described, should one conclude, in the case of the Minkowski measure, that $f(\alpha)$ is a worthless notion? Certainly this measure confirms that I have been arguing strenuously for a long time: that $f(\alpha)$ is a delicate tool. Its proper context is distributions, that is, probability theory. Moreover, it does not concern the best known and more "robust" parts of that theory, namely, those related to the law of large numbers and the central limit theorem. Instead, it concerns the probability theory of large deviations, which is a delicate topic.

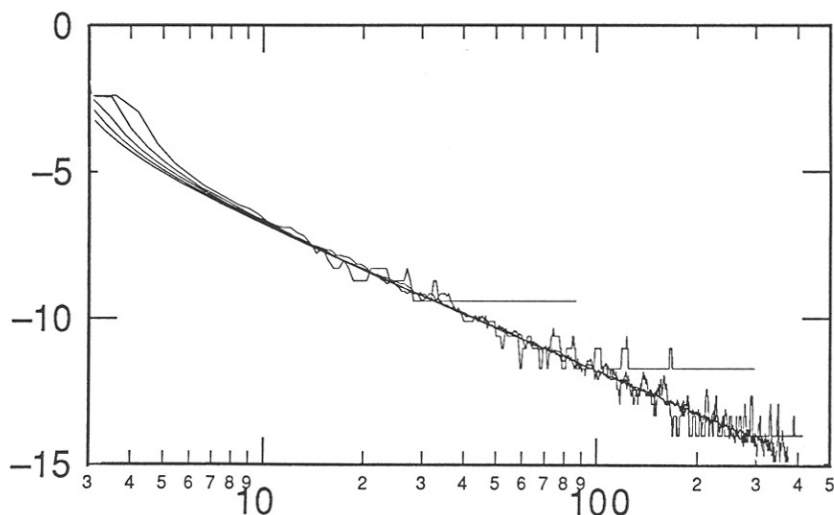


Figure 4. A recent plot of the estimated functions $f_\varepsilon(\alpha)$ for the Minkowski measure restricted to the interval $[1/10, 1/9]$. We started with $f_\varepsilon(\alpha) - 1$ as the vertical coordinate and $-\log \alpha$ in order to straighten $f_\varepsilon(\alpha)$ — we chose $\log \alpha$ as the horizontal coordinate. Then we collapsed the five graphs that correspond to $\varepsilon = 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8}$ and 10^{-9} . This figure strongly confirms the cup-convexity suspected in Figure 3. It shows the occurrence of $f < 0$. Finally, the weighting rule shows that $1 - f_\varepsilon(\alpha)$ decreases as $\varepsilon \rightarrow 0$.

3. Remarks

3.1. On continuous models as approximations, and on "thermodynamics"

Why do physicists study limits that cannot be attained? Simply because it is often easier to describe a limit than to describe a finite structure that can be viewed as an approximation to this limit. In particular, this is why it is often taken for granted in the study of multifractals that a collection of "coarse-grained" approximations can be replaced by a continuous "fine" or "fine-grained" description. The latter involves Hausdorff dimensions and introduces the functions $\tau(q)$ and $f(\alpha)$ directly, not as limits.

For the Minkowski μ , however, the actual transition from coarse to fine-graining (as $\Delta x \rightarrow 0$ and $\Delta m \rightarrow 0$) is extraordinarily slow and many aspects of the limit differ qualitatively from the corresponding aspects of even close approximations. Therefore, the role of limits demands further thoughts, which I propose to describe elsewhere.

The continuous limit approximation has been described as ruled by a "thermodynamical" description. Thus, the fact that the convergence is slow and singular in the case of the Minkowski μ reveals a fundamental practical limitation of the thermodynamic description.

3.2. Parabolic versus hyperbolic systems

To pinpoint the essential ingredient of our special dynamical system, it is important to see how its properties change if the system itself is modified. If one wants $f(\alpha)$ to become \cap -shaped, it suffices to replace the first of our three maps by

$$x \rightarrow \frac{1}{2} + \frac{\rho}{x - 1/2},$$

with $\rho < 1/4$. As $\rho \rightarrow 1/4$, the right side of $f(\alpha)$ lengthens and is pushed away to infinity, and the anomalies disappear asymptotically. Formally, the system changes from being hyperbolic to parabolic. Hence, the anomalies we have investigated are due to the system's parabolic. In terms of the limit $f(\alpha)$, the differences between parabolic and hyperbolic cases increase as $\rho \rightarrow 1/4$. But actual observations lie in a preasymptotic range; for a wide range of ε , $f_\varepsilon(\alpha)$ will be effectively the same for ρ close to $1/4$ as it is for $\rho = 1/4$.

3.3. Multiplicativity

The multifractals known longest and best ⁵ are the multiplicative measures. But the Minkowski measure μ is not a multiplicative multifractal. Nevertheless, several properties of μ were first conjectured on the basis of an approximation of μ by a multiplicative multifractal ⁸ and later proved to be correct.

3.4. In lieu of conclusion

The apparent "strangeness" of the facts described in this paper *must not* discourage the practically minded reader. Repeating once again a pattern that is typical of fractal geometry, it turns out that what had seemed strange should be welcomed and not viewed as strange at all.

4. Acknowledgements

The work that led to this paper began in 1987. Over the years, I had invaluable discussions with M. C. Gutzwiller, C. J. G. Evertsz, T. Bedford and Y. Peres. Figures 1, 2, and 4 were prepared by J. Klenk. Peres wrote his Ph.D. thesis on the Minkowski measure. He had informed me that, while both Gutzwiller and I had independently rediscovered μ before we joined forces to write about it in 1988, we had been anticipated by Minkowski and Denjoy. However, the $f(\alpha)$ function of μ was not discussed (however imperfectly) until our joint paper.

5. References

1. A. Denjoy, *Comptes Rendus* (Paris) **194** (1932) 44-46. *Journal des Mathématiques Pures et Appliquées* **62** (1938) 105-151. Reprinted in A. Denjoy, *Articles et Mémoires*, 1955, **2**, 925-971.
2. U. Frisch and G. Parisi, *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, ed. M. Gil., (Amsterdam: North-Holland, 1985) pp 84-86.
3. M. C. Gutzwiller and B. B. Mandelbrot, *Phys. Rev. Lett.* **60** (1988) 673.
4. T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia and B. I. Shraiman. *Phys. Rev. A* **33** (1986) 1141-1151.
5. B. B. Mandelbrot, *J. Fluid Mech.* **62** (1974) 331-338. Also *Comptes Rendus* (Paris) **278A** (1974) 289-292 and 355-358.
6. B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman, New York, 1982).
7. B. B. Mandelbrot, *Pure and Applied Geophysics* **131** (1989) 5-42. See also in *Fluctuations and Pattern Formation* (Cargèse, 1988), eds. H. E. Stanley and N. Ostrowsky (Kluwer, Dordrecht and Boston, 1988) pp. 345-360.
8. B. B. Mandelbrot, *Physica A* **168** (1990) 95-111.

9. B. B. Mandelbrot, *Physica A* **163** (1990) 306-315. Developed in *Fractals. Physical Origin and Properties (Proceedings of the Erice meeting, 1988)*, ed. L. Pietronero (Plenum, New York, 1989) 3-29.
10. B. B. Mandelbrot, *Proc. Royal Soc. (London)* **A 434** (1991) 79-88.
11. B. B. Mandelbrot, *Multifractals and $1/f$ Noise. 1963-1976.* (*Selecta*, Vol. N) (Springer, New York, forthcoming).
12. B. B. Mandelbrot, forthcoming.
13. B. B. Mandelbrot and C. J. G. Evertsz. *Physica A* **177** (1991), 386-393.
14. B. B. Mandelbrot, C. J. G. Evertsz and Y. Hayakawa, *Phys. Rev. A* **42** (1990) 4528-4536.
15. Minkowski, H. 1911. *Gesammelte Abhandlungen* (Chelsea, New York, reprint).
16. R. Salem, *Trans. Am. Math. Soc.* **53** (1943) 427-439. Reprinted in R. Salem *Oeuvres Mathématiques*, pp. 282-294, (Hermann, Paris, 1967).
17. J. R. Kinney. *Proc. Am. Math. Soc.* **11** (1960) 788-794.