Physica A 191 (1992) 95–107
North-Holland

PHYSICA A

Plane DLA is not self-similar; is it a fractal that becomes increasingly compact as it grows?

Benoit B. Mandelbrot

Physics Department, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA
and Mathematics Department, Yale University, New Haven, CT 06520, USA

Using two new methods of geometric analysis, this paper establishes that DLA clusters are definitely not self-similar. Compared to small clusters, the morphology of large clusters (of sizes up to 30 million particles) can be characterized, both visually and quantitatively, as being far more "compact" or less "lancunar". Qualitatively, the number of "arms" increases during growth. The evidence does not exclude that the cluster remains fractal, and that its fractal dimension remains constant; however, new pitfalls in the estimation of D are revealed. The gradual change in the morphology of DLA opens the possibility that there is continuity between the standard morphology observed for small to medium computer generated DLA clusters and the compact morphology observed in many actual physical phenomena.

1. Introduction

This paper is an announcement and abstract of recent results from the Yale group that will be reported in detail elsewhere [1, 2]. Being written for specialists, the text dispenses with lengthy preliminaries and is largely a commentary on its figures.

We introduce new methods for the analysis of shape and use these methods to perform direct investigations of the geometric structure of the diffusion limited aggregates (DLA) in the plane [3]. In a first approximation, these aggregates are self-similar fractals [4, 5]. But, applying our tools to off-lattice clusters of up to 30 million particles, we found direct geometric evidence that a more refined model is needed. Two alternative geometric scenarios are advanced and suggested for further investigation: (A) There is a massive transient from one level of "compactness" to a higher level, followed by unchanging compactness. (B) There is a non-terminating drift towards increasing compactness. Qualitatively, compactness increases when the number of branches is perceived to increase. Paradoxically, it is possible (as seen in section 4 and the "Note added in proof") for compactness to increase while the cluster remains a fractal and has an unchanging dimension.

We are well aware that many diverse analytic departures from self-similarity have been reported in the literature; we believe that either of our scenario accounts for these departures; our arguments are described in ref. [2].

Our investigations concern the geometry of the "dead" parts of the cluster, namely, parts that are extraordinarily unlikely to grow further. In that part of the cluster, the fjords have nearly parallel effective walls (they are not at all "fan-shaped", i.e., their width does not increase linearly). This finding was first reported in ref. [6] for medium-sized clusters and is strongly supported in ref. [2] for clusters of up to 30 million particles. It follows that the lowest values of the harmonic measure are extremely low, and our results have a direct bearing on their $f(\alpha)$ distribution. Moreover, it has been demonstrated previously [6] that non-growing portions are scattered throughout the clusters, even near the tips of young branches. Therefore, we expect our results to be pertinent to the study of growth.

2. ε-neighborhoods analysis of DLA. The filling ratio increases as the cluster grows

Our first test uses a new method that we call *ε-neighborhoods analysis*. The results are exhibited in figs. 1, 2 and 3, each containing six subfigures, and concern the dead central part (near the origin) of a single 14.8 million particle cluster, to be denoted by Q. The dead central part is defined conservatively as contained within a circle C_{max} whose radius R_{max} is equal to 3/4 of the radius of gyration of Q. In our cluster, the dead central portion contains 5 million particles. Let us mention immediately that we obtained identical results by investigating in the same way a large number of other clusters, some containing 30 million particles, and by examining the changes in the dead central part of the cluster as it grew.

Reduced truncated clusters. To prepare for ε-neighborhoods analysis, we select a radius $R_{min} < R_{max}$, then define λ by $\lambda^{24} = R_{max}/R_{min}$ and draw for each k the circle C_k of radius $R_k = R_{min}\lambda^k$. In figs. 1 to 3, $R_0 = R_{min} = 78$, $R_{24} = R_{max} = 3746$ and $\lambda = 1.175$. The kth truncated cluster Q_k is defined as the portion of Q that is contained within C_k (i.e., the intersection of Q and the interior of C_k.) Next, each C_k is reduced by a ratio of $1/\pi R_k^2$, hence becomes the circle $\tilde{C}$ of unit area; Q_k is reduced in the same ratio $1/\pi R_k^2$, and becomes a shape to be denoted by $\tilde{Q}_k$. Note that, if $k < 24$, the truncated cluster Q_k has stopped growing long before our cluster reaches its total size Q: thus, Q_k was the dead central part of our cluster at an earlier instant of time.

The notion of an ε-neighborhood. Given a set S, its ε-neighborhood $\varepsilon\{S\}$ is defined (see ref. [7]) as the set of points that are within a distance of ε of a

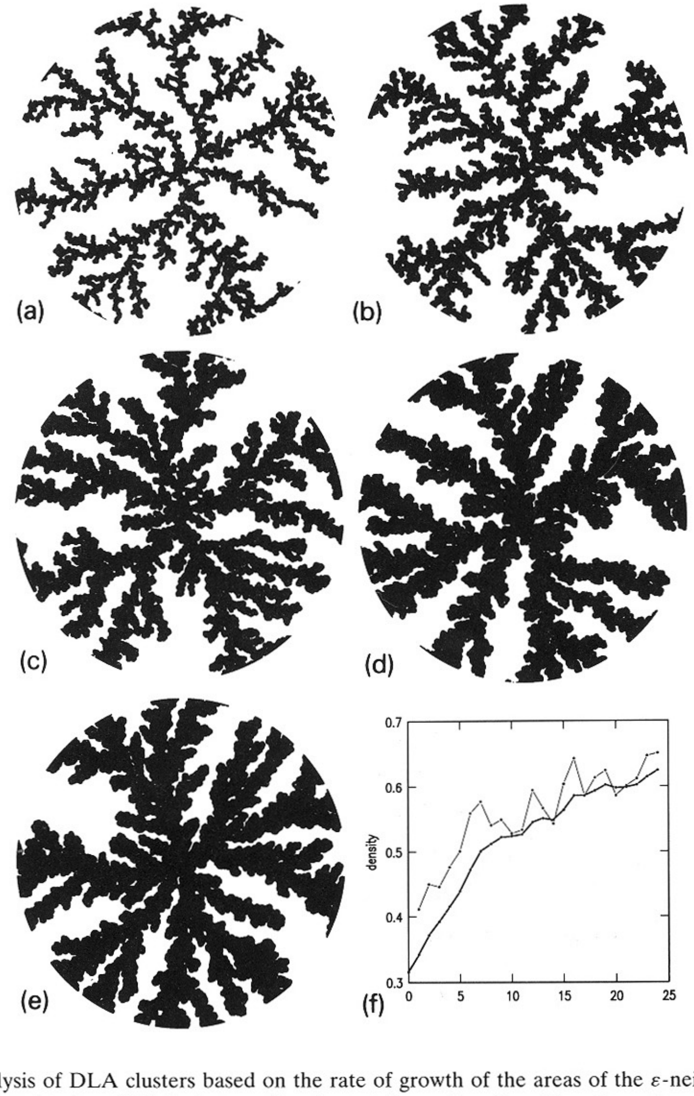

Fig. 1. Analysis of DLA clusters based on the rate of growth of the areas of the ε-neighborhoods of truncated and reduced DLA clusters. See the text for explanation. Here the value of ε is small, very far from saturation for the largest value of k.

point in S. To form $\varepsilon\{\mathrm{S}\}$, draw a circle of radius ε around each point of S; these circles overlap extensively and $\varepsilon\{\mathrm{S}\}$ is their union (sum).

The filling ratio; *explanation of how figs. 1, 2 and 3 are drawn.* Each figure corresponds to one of three values of ε: small, medium and large. For each ε, we evaluated – for the 25 values of k from 0 to 24 – the area of the ε-neighborhood $\varepsilon\{\tilde{\mathrm{Q}}_k\}$. Since the area of $\tilde{\mathrm{C}}$ is 1, those areas are automatically

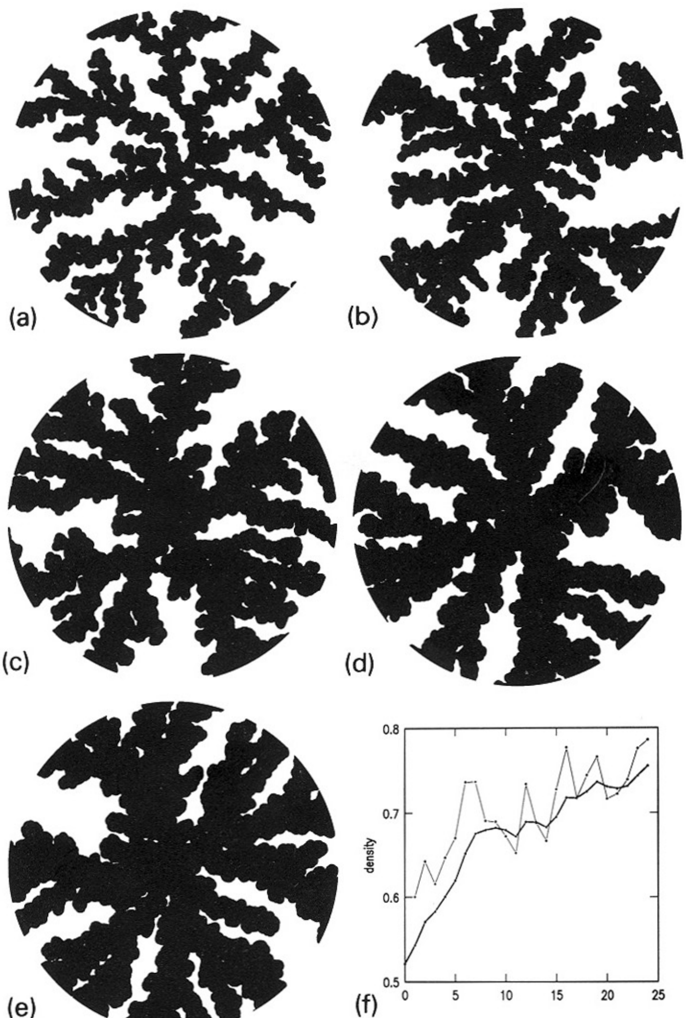

Fig. 2. Analysis of DLA clusters based on the rate of growth of the areas of the ε-neighborhoods of truncated and reduced DLA clusters. See the text for explanation. Here ε is "medium-sized".

renormalized to lie between 0 and 1; they can be called *filling ratios*. Their values are plotted by bold lines in figs. 1f, 2f and 3f. The thin lines represent the renormalized areas of the 24 sets $\varepsilon\{\tilde{Q}_k - \tilde{Q}_{k-1}\}$, each of which is the reduced intersection of Q with the annulus bounded by successive circles C_k and C_{k-1}. Thus, the bold lines refer to integrated data, and the thin lines refer to differential data.

To provide additional evidence of more graphic character, five values $k = 0$, 4, 9, 16 and 24 were singled out (for reasons to be explained in the next

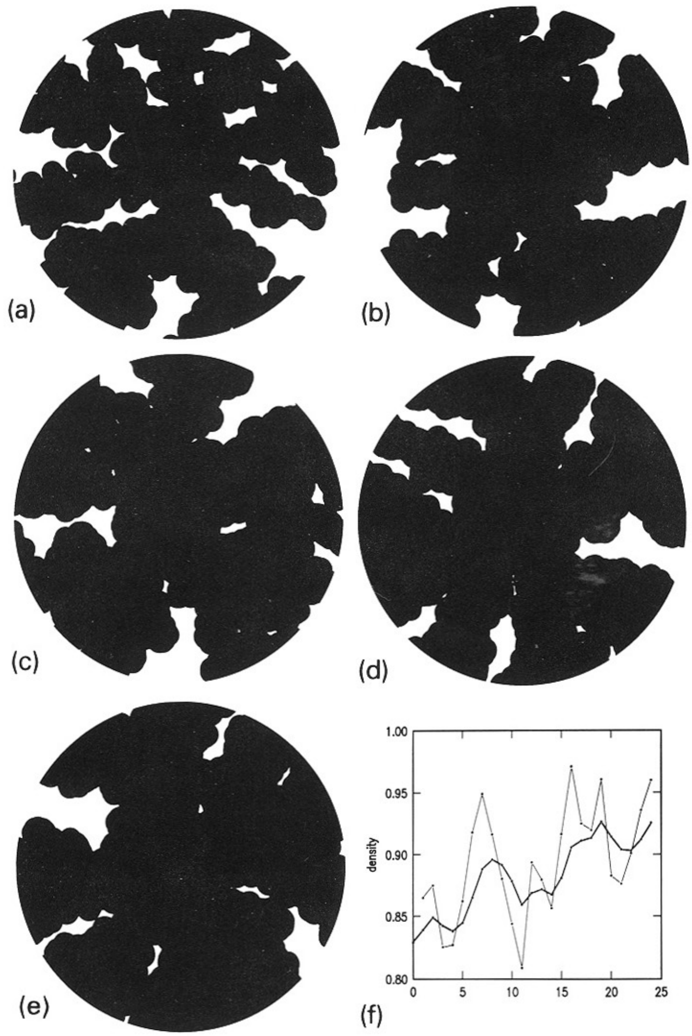

Fig. 3. Analysis of DLA clusters based on the rate of growth of the areas of the ε-neighborhoods of truncated and reduced DLA clusters. See the text for explanation. Here ε is large, close to saturation for the largest value of k.

paragraph) and the corresponding $\varepsilon\{S_t\}$ were plotted as figs. 1a to 1e, 2a to 2e and 3a to 3e, respectively.

Inspection of figs. 1, 2 and 3. If Q had been a Sierpiński gasket, $\varepsilon\{\tilde{Q}_k\}$ would have fluctuated without any systematic trend. Similarly, if Q had been self-similar (in scales above the particle size), the sets $\varepsilon\{\tilde{Q}_k\}$ and $\varepsilon\{\tilde{Q}_k - \tilde{Q}_{k-1}\}$ would have been statistically identical (except for corrections due to particle size). In particular, the curves in figs. 1f, 2f and 3f would have been horizontal, except for up and down statistical fluctuations.

In fact, figs. 1f, 2f and 3f show that these various areas increase with k; as expected, there is less noise in the integrated data. And figs. 1a to 1e, 2a to 2e and 3a to 3c show that the ε-neighborhoods $\varepsilon\{\tilde{Q}_k\}$ fill the circle $\tilde{C}$ increasingly tightly. The growth of the filling ratio with k is slower than linear, and very roughly proportional to $\log k \sim \log\log R$. This explains why the five values of k that we selected for our illustrations are not spaced uniformly.

Mathematical formalization of the preceding findings, using the notion of lacunarity. To express the above findings mathematically, the proper fractal notion is *lacunarity*. The theory of lacunarity was sketched in ref. [7], and extensive recent developments have quantified it by introducing actual numerical rates. When a cluster Q is perceived as becoming increasingly "compact", various measures of its lacunarity are found to decrease. However, this notion is too delicate to be meaningfully discussed here; for details the reader must be referred to ref. [1].

Two geometric scenarios for the preceding findings: "unexpectedly massive transient" and "limitless drift". Our findings can be interpreted in either of the following two ways. (A) They may reveal an unexpectedly massive (that is, long) *transient* from one level of lacunarity to a lower level that holds for all cluster sizes above a certain finite threshold. (B) They may manifest a non-terminating *drift* toward vanishing asymptotic lacunarity.

While examining earlier versions of figs. 1f, 2f or 3f, we have been constantly alert for indications that the plots of the filling ratios had reached a maximum value, implying a transient that has run it course. In some of our early tests, the R_k were the radii of gyration of successive partly grown stages of one cluster; hence the right most portions of our plots of the filling ratios were significantly noisier, and one could not exclude the possibility that had they already crossed over to a horizontal line within the sample that was being examined. So far, however, the crossover has not been confirmed. Hence the "transients" scenario was not yet observed to have run its course, and this scenario remains unproven.

The alternative geometric scenario is that the drift toward increasing compactness continues without end while the cluster remains fractal. Such a scenario cannot be proven by finite size experiments. In addition, it does not become meaningful, and does not cease to appear paradoxical, until the notion of fractal has been generalized beyond the morphology of the familiar Sierpiński gasket and of small or medium sized DLA. This indispensable generalization is one of the main points of the theory of lacunarity [1]; it is better to describe it separately, in section 4 (continuing in the "Note added in proof").

3. Orbital gaps analysis of DLA

The ε-neighborhoods investigated in section 2 are two-dimensional, but a geometric analysis of DLA can also be performed on related one-dimensional shapes. We have analyzed DLA via *orbital* cross sections, namely, via the rescaled cross sections of Q by circles C_k centered at the origin. These are also the rescaled cross sections of the sets $\tilde{Q}_k$ by the circle $\tilde{C}$ in which they are contained. As described, for example. in ref. [7] (p. 135), a "generic" linear section of a fractal of dimension D is a fractal dust of dimension $D-1$. For DLA, a widely accepted value is $D = 1.71$ (it will be discussed in section 5). Thus, we are led to the heuristic expectation that an orbital section of a DLA is a dust of dimension 0.71 on a circle. We shall parameterize each orbital circle by an angle $2\pi x$, which rescales the circle into the interval [0,1].

Second new method: *analysis of the cross sections of $\hat{Q}_k$ by $\tilde{C}$, using the distribution of the gap lengths*. Recall that a gap is defined as an (open) interval whose endpoints belong to the set S but whose interior points do not. Denote by Γ the gap length, by γ a possible value of this length and by $\mathrm{Nr}\{\Gamma > \gamma\}$ the number of gaps of length $>\gamma$. Within the scaling range of S, one expects that $\mathrm{Nr}\{\Gamma > \gamma\}$ will be ruled by the following scaling relation, called the *gap-number rule*: $\mathrm{Nr}\{\Gamma > \gamma\} \sim \gamma^{-(D-1)}/\Lambda_G$, where Λ_G is a gap prefactor ([7], p. 78). Thus, scaling can be tested in experimental situations by checking whether or not the graph of $\log \mathrm{Nr}\{\Gamma > \gamma\}$ versus $\log \gamma$ is straight; if it is, its slope provides an estimate of D.

In the case of DLA, the rough gap-number data for the cluster examined in section 2 are unfortunately very noisy. To eliminate noise, we always average the data, either across several clusters of the same size, or across several neighboring values of R within the same cluster. Fig. 4 was obtained by taking 30 clusters of 10 million particles each, and averaging the 30 quantities Nr for each γ.

The cutoffs. After noise reduction, the doubly logarithmic graph of $\log \mathrm{Nr}\{\Gamma > \gamma\}$ versus $\log \gamma$ is expected to be a straight line of slope -0.71. More precisely, one must expect a straight interval between two cutoffs. Indeed, a gap in an orbital section is bounded by $x = 1$; hence an outer cutoff equal to at most $1/5$ (5 being the agreed number of "main branches" of a DLA cluster). At the other end (choosing the particle's size as the unit of length) thee is an inner cutoff of about $1/2\pi R$.

Inspection of fig. 4. The graphs relative to small R are not straight at all, suggesting that the inner and outer cutoffs are too close to each other. (Had this test been applied to the earlier DLA clusters, it would have failed to reveal their being fractal!) As R increases, a straight portion of slope very nearly

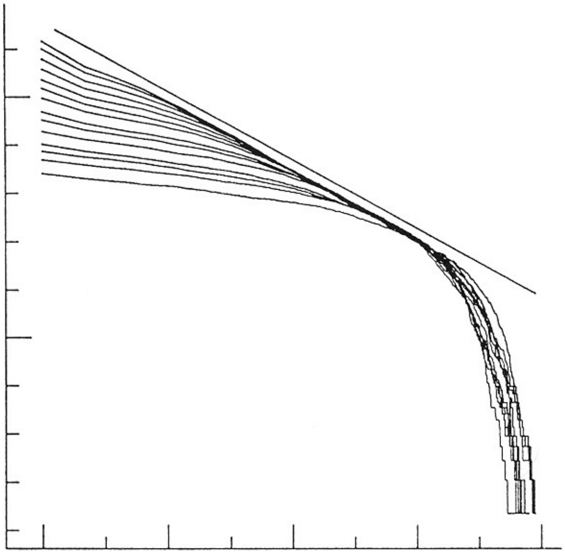

Fig. 4. Analysis of DLA clusters, based on the distribution of orbital gap lengths γ, as measured in fractions of a whole circle. See the text for explanation. These are plots of $\log \mathrm{Nr}\{\Gamma > \gamma\}$ versus $\log \gamma$. Unfortunately, different units were used for the two coordinate axes, but the straight line above the graphs is of slope 0.7.

-0.71 appears in the middle of the graph; it is the anticipated manifestation of an incipient scaling range. As R continues to increase, this range becomes wider. The overall distribution of the gaps splits into a "tail" for small values of γ and a "head" for large values of γ, separated by a scaling range. The overall plot of $\log \mathrm{Nr}\{\Gamma > \gamma\}$ versus $\log \gamma$ is convex ("cap-convex").

Compared to the values one would by extrapolating the scaling range, the few largest gaps (starting with $\gamma_{\max}$) are extremely small, and in fig. 4 they become smaller as the cluster grows. These values correspond to the few widest fjords, located between the major "arms" of the cluster. (This decrease in fjord width has been observed by other authors, for example in ref. [8].)

A first approximation to fig. 4, and the completed transients scenario. To pursue the discussion, it helps to denote by γ_{cross} the (ill-defined) gap where the scaling range crosses over into the head. Now we can return to the analysis of fig. 4, and argue that, in a first approximation, the tail simply "unwraps" monotonically towards a fixed straight line of slope -0.71. Let us show that this first approximation would imply the scenario of a transient that has almost ran its course.

If the tail unfolds on a fixed straight line, the distribution of small gaps $<\gamma_{\text{cross}}$ tends to a limit, and the tail probability $\Pr\{\Gamma < \gamma_{\text{cross}}\}$ must increase, up to the maximum represented by the fixed straight asymptote. This requires the head probability $\Pr\{\Gamma > \gamma_{\text{cross}}\}$ to decrease, and the head itself to "shrink" away from the straight asymptote. That is, the largest gaps must decrease in size, and may also increase in numbers. One can specify this first approximation further, as stating that the number and sizes of the widest gaps eventually cease to change as the cluster grows further. That is, the broadest fjords would begin by widening less than linearly with distance, but would eventually resume a linear (fan-shaped) morphology.

A second approximation to fig. 4, and its relation to the scenario of limitless drift toward an increasingly compact limit. Our second scenario would manifest itself by a limitless shrinking of the distribution's head. Now let us announce an important and unexpected observation that will be justified in section 5. During growth, the value of D need not change, if – while the head shrinks with no limit – the straight (scaling range) portion in fig. 4 drifts up but keeps an invariant slope. Such a drift would express a limitless decrease in the prefactor Λ_G. This phenomenon is as it should be, because ref. [1] shows that Λ_G is a numerical measure of lacunarity.

If the straight scaling portion moves up as it lengthens, plots corresponding to different k will cross increasingly far to the left (and γ_{cross} would decrease). This leads us to consider the envelope of all the plots (i.e., the locus of points accessible from the top without crossing any plot). The limitless drift scenario predicts that the plots' envelope will, in its middle portion, become concave ("cup-convex"), in contrast to the individual plots, which are convex (cap-convex).

Perceived concavity in part of fig. 4. We do indeed perceive such a concavity in fig. 4, as well as in every other figure we have constructed on the same principles. To test our perceptions, we went on to replot fig. 4 to enhance concavity, and its presence was indeed confirmed.

We do not dispute that this concavity is slight. The evidence it provides is of borderline quality. It is no better than the evidence in section 2 in deciding between the scenarios of a transient that has not yet ran its course and the scenario of limitless drift.

4. Example of limitless increase of compactness (decrease in lacunarity) with no charge of dimension

Pick D to satisfy $1 < D < 2$. We know that on the interval [0,1] a Cantor dust with this D is obtained by using generators made of N equal intervals of length $r = N^{-1/D}$. These intervals can be spread uniformly over [0,1], thus yielding an infinity of distinct Cantor dusts $\mathscr{C}_N$ with the same dimension D. (Further Cantor dusts of dimension D are obtained by other procedures.)

The length of the ε-neighborhood of $\mathscr{C}_N$ corresponds to the filling ratio of section 2. It is easy to see (and is shown in detail in ref. [1]) that one has

$$\text{length of}\,[\varepsilon\{\mathscr{C}_N\}] = \varepsilon^{1-D}/\Lambda_{\mathrm{F}}\,,$$

and that Λ_{F} (called the *filling rate of lacunarity*) satisfies $\Lambda_{\mathrm{F}} \to 0$ as $N \to \infty$.

Why are those special Cantor dusts relevant to DLA? Simply because DLA has a hierarchical structure (a tree structure); therefore the cross sections of DLA by a line are also fractals with a hierarchical structure – like the Cantor dusts. An increasing N in the preceding example corresponds to a tree with increasingly heavy branching structure. Therefore, increasing filling ratios (section 2) and an increasingly heavy branching structure are different symptoms of the phenomenon of deceasing lacunarity.

Note that, as $N \to \infty$, our Cantor dust comes increasingly close to being translationally invariant; but it *never* reaches this limit. It "looks" increasingly like an interval; but it *never* becomes an interval. Nevertheless, if N is large, no one would think of evaluating its fractal dimension.

Similarly, it seems taken for granted that the morphology of compact growth is of dimension 2. This belief would explain why, to our knowledge, this dimension is not being measured. The new fact that a dimension <2 is compatible with very low lacunarity, immediately suggests that the experimentalists should actually check D in various examples of compact growth.

Let us end this section with a question. Suppose it is confirmed that the lacunarity of DLA decreases without bound, or that there is a massive transient that has not yet ran its course. If so, is it right to call DLA a fractal? Long ago (after some vacillations) we concluded against a dogmatic definition of the concept of fractal, and the present study confirms that it is best to use loose definitions that make it easy to generalize when need arises. It has just become useful to call a set S fractal, if it is inhomogeneous except that the fractal dimension (defined locally) is the same throughout (see the "Note added in proof").

5. The mass–radius relation and the estimation of D

So far, we have not mentioned the relation $M(R) = \Lambda_M R^D$ for the number of particles, $M(R)$, in a circle of radius R. This is the best-known and best-established property of fractals, and the basis for the standard estimate of D in the case of DLA. Naturally, we drew the diagram of $\log M(R)$ versus $\log R$, and found it to be so nearly straight in every sample that it is not worth reproducing. Its slope for N up to $N = 3 \times 10^7$ takes the value $D \sim 1.71$ [3–5] inferred from the clusters that we view as very small.

A first reaction is that the straightness of $\log M(R)$ simply confirms well-established facts. But the results reported in our figures suggest an altogether different and more complicated view. Indeed, the prefactor Λ_M in the relation $M(R) = \Lambda_M R^D$ is not a numerical constant (like the prefactor for the area in a circle, which is π). Instead, Λ_M is a characteristic of a fractal, to be added to D. In addition, the theory of lacunarity [1] shows that – like Λ_F and Λ_G – the prefactor Λ_M is a numerical rate that measures lacunarity. Therefore, it too should be expected to *decrease* as a DLA cluster grows and becomes increasingly compact.

Assume that the cluster has an underlying true fractal dimension D_{true}, and consider the standard plot of $\log M$ versus $\log R$. Occasional changes in the prefactor would have yielded straight segments of slope D_{true}, lying increasingly low as R increases. But if Λ_M changes gradually, there is only one way for $\log M(R)$ to be represented by a straight line: one must have $\Lambda_M \sim R^{-\Lambda D}$, from which it can be shown to follow (see ref. [2]) that the slope of $\log M(R)$ is $D = D_{\mathrm{true}} - \Delta D$. Therefore, when D is estimated from the mass–radius relation without correcting for the bias ΔD, the estimate should be viewed as merely a lower bound.

There is a long tradition in physics that welcomes relations (of the form) $\Lambda_M \sim R^{-\Delta D}$. In the present instance, this tradition should welcome our scenario of limitless drift. Alternatively, if the decrease of Λ_M stops after a massive transient, one would, *in general*, expect a bend in the graph of $\log M(R)$ versus $\log R$. It is obvious that this prediction can be tested experimentally and deserves much further study.

6. Conclusion

Once again, this is an advance report on ongoing research. An irritating feature of past work on DLA has been that different methods of measurement yield slightly contradictory estimates of D. We are presently seeking accurate estimates of D by our methods, searching whether the previously observed discrepancies are explained by either of our two scenarios, and seeking new criteria to discriminate between those scenarios. A tool needed in this context will be a careful comparison of the distinct measurement of lacunarity provided by the rates Λ_F, Λ_G and Λ_M. Some useful results are given in ref. [1], but little is known.

Acknowledgements

The coauthors of the research summarized in this paper are my associates in the Yale Project for Research on Fractals. Carl J.G. Evertsz and Lionel Woog coauthored ref. [2] and Jürgen Klenk and David Fracchia coauthored ref. [1]. We were assisted by François Normant and Daryl Hepting, and had very interesting discussions with Richard F. Voss. On August 10–13, 1992, this paper's contents were presented in a series of lectures at the University of Oslo, in the Cooperative Phenomena Group of the Physics Department. Extremely useful comments were received from Professors Amnon Aharony, Jens Feder, Torstein Jøssang and Paul Meakin, and from Dr. Rudolf Hilfer. Professor Aharony also read earlier versions of refs. [1,2] and made specific and extremely valuable suggestions that were followed when preparing refs. [1,2] and in revising the present text.

Last but not least, the investigations carried out at Yale University were supported by the Office of Naval Research under grant N00014-90-J-1026, and were performed on two IBM RS/6000 RISC computers, made available at Yale by IBM's Technical Computing Systems Group (Mathematical Sciences Area).

Note added in proof

While correcting the proofs, we noted that two points at least might lead to objections that are easily preempted.

Meaning of the scenario of "non-terminating (limitless) drift." Take a sequence of increasingly large clusters downsized to a fixed size. In our second scenario, the clusters' central portion would converge to a filled-in part of the plane. This limit could be of dimension 2, hence not fractal. But all approximations – however close, but short of the limit – would be fractal. The study of multifractals (in the form I proposed in 1974) has always been rife with instances of limits that differ qualitatively from all their approximations. (This is emphasized in my paper in the Kolmogorov volume of: Proc. R. Soc. (London) A 434 (1991) 79–88.) Now analogous effects may be showing up in the study of fractals themselves.

One cannot explain this paper's observations by familiar transients found in developing self-similar models. In order to test the results reported in our figures, we have repeated the same experiments on the exactly self-similar "M.-V. toy model" of DLA, as described in B.B. Mandelbrot and T. Vicsek, J. Phys. A 22 (1989) L377–L383. The initiator is the interval [0, 1]. The generator is straightened letter λ, made of 3 sticks of length 1/2, two of them linear continuations of each other. The kth step of the construction replaces each of 3^{k-1} sticks of length $2^{-(k-1)}$ by 3 sticks of length 2^{-k} forming a reduced-scale λ. The construction proceeds by interpolation; but in fact the kth interpolation stage is meant to illustrate a reduced kth cluster; therefore it is immediately suited to the tests in this paper.

As k increases, this reduced scale cluster fills in, becomes increasingly "bushy." Hence the area of its ε-neighborhood increases to a limit corresponding to the fully developed tree. Could it be that the effect described in figs. 1, 2, and 3 simply reflects the growth of the ε-neighborhood during the "natural" filling-in of a strictly self-similar structure? For the M.-V. toy model, the answer is, "No". Filling ratios begin by growing, but their growth stops after a short and clearly terminated transient.

The gap distribution in the M.-V. toy model was plotted as in fig. 4. Again, there is no trace of the effects fig. 4 reveals in the case of DLA.

The programs used for the above tests were written by L.N. Siegel.

References

[1] B.B. Mandelbrot, J. Klenk and D. Fracchia, Measures of lacunarity for some fractal dusts on the line, preprint (1992).
[2] B.B. Mandelbrot, C.J.G. Evertsz and L. Woog, Growth with unchanged fractal dimension but decreasing inhomogenity (tentative title) preprint (1992).
[3] T.A. Witten and L.M. Sander, Phys. Rev. Lett. 47 (1981) 1400.
[4] T. Vicsek, Fractal Growth Phenomena (World Scientific, Singapore, 1989).
[5] P. Meakin, in: Phase Transitions and Critical Phenomena, vol. 12, C. Domb and J. Lebowitz, eds. (Academic Press, New York, 1988) p. 335.
[6] B.B. Mandelbrot and C.J.G. Evertsz, Nature 348 (1990) 143.
[7] B.B. Mandelbrot, The Fractal Geometry of Nature (Freeman, New York, 1982).
[8] P. Ossadnik, Physica 176 (1991) 454.