Annals of the Israel Physical Society Volume 5

Editor-in-Chief

Raoul Weil, Technion — Israel Institute of Technology, Haifa

PERCOLATION STRUCTURES AND PROCESSES

Edited on behalf of the Israel Physical Society by

G. Deutscher
Tel Aviv University

R. Zallen
Xerox

Joan Adler
Technion

Published by

Adam Hilger, Bristol
and
The Israel Physical Society, Jerusalem
in association with
The American Institute of Physics, New York

CHAPTER 3

AN EXPLICIT FRACTAL MODEL OF PERCOLATION CLUSTERS

Benoit B. Mandelbrot

IBM Watson Research Center, Yorktown Heights, NY 10598

Annals of the Israel Physical Society, v. 5

Percolation Structures and Processes
Eds. G. Deutscher, R. Zallen and J. Adler

ABSTRACT

It is now widely agreed that, on scales below ξ, the geometry of percolation clusters involves a number of dimensionality-like exponents. Nevertheless, saying that such an exponent is a dimensionality need not be of any usefulness, and in particular such an exponent need not be a fractal dimensionality. More generally, a reference to fractals remains purely formal until a fully specified implementation has been described. The present article describes a very simple such structure applicable to percolation clusters.

1. INTRODUCTION[1]

1.1. THREE STAGES OF SCALING IN PHYSICS

Soon after scaling entered statistical physics[3], it was applied to the study of percolation near criticality. Diverse *analytic* relationships between different characteristics of a system gave rise to a long list of critical exponents.

A second stage of application of scaling ideas to percolation occurred after fractal geometry was conceived, as first presented in Ref. 4. This stage centers around the postulate that scaling can extend to *geometric* characteristics, instead of remaining a purely analytic notion. In particular[5], several writers were prompt to postulate that percolation clusters at criticality can be characterized by one or more effective dimensionalities, taking "anomalous" non-integer values. These numbers are to be added to the existing list of critical exponents. After a few false starts, the fractal dimensionality was related to standard exponents.

A third stage of application of scaling ideas to percolation should be distinguished. It started with the fully specified fractal model that I devised in 1978 and discussed widely, but first published in Ref. 2. Use of the term "model", instead of "representation", stresses that this construction makes no pretense of being either unique or wholly realistic. However, I claim that a fully described fractal model *does not* reduce to its generating algorithm, or to the numerical parameters that characterize its construction.

1.2. IN PRAISE OF GEOMETRY

More generally, the fractal side of the study of critical phenomena *is not* merely an appendage to the usual scaling analysis, but a genuine geometric counterpart. It is interesting and a bit curious that geometry should need to be defended among physicists. One may observe that students of percolation who wish to present this topic to outsiders seem unavoidably led to stress its geometry, before they veer towards analysis (and, in most cases, a loss of general comprehensibility). Similarly, the early sections of Ref. 8 seem to promise a geometrical picture which the paper itself fails to develop.

Still more generally, I claim that the contention of the mathematical purists, that geometry has been reduced to analysis, deserves nothing but scorn on the part of physicists. No one will deny that such a reduction may be harmless or even fruitful during certain stages when a discipline seems ripe as physics but demands to be developed as mathematics. Furthermore, twentieth century physics includes several branches whose geometry came straight from analysis and appears to have little intuitive content of its own. But in other branches the careful avoidance of geometric intuition is nothing but a stunt. Thus, allowing the above experiences to set a general rule would, in my view, be extraordinarily counterproductive in theoretical—and even in mathematical—physics.

Since it is hard today to define geometry as an intrinsically separate branch of mathematics it is safest to think of geometry as being as an art of thinking intuitively with the help of pictures. It is far harder to teach than the art of algebraic manipulation, and it may be hard to automate. Fortunately, the gift for geometry is quite widespread, and the current emergence of computer graphics is bound to help many to realize that they have this gift. If they do, they must resist any attempt to force them to box like analysts: blindfolded.

Needless to say, I do not propose to proscribe analysis, but hope for a return to the situation that prevailed a hundred years ago, when Henri Poincaré and Felix Klein (among others) wrote eloquently of geometry and analysis as equal partners.

1.3. GEOMETRY OF PERCOLATION

First, let me take the specific example at hand to dispose of the common belief that geometric intuition based on pictures often leads to outright mathematical mistakes. This criticism disregards the more numerous mistakes intuition would have helped avoid. An explicit fractal model of percolation would have prevented the confusion that prevailed in 1977. Since one of the basic analytic roles of an anomalous dimensionality is to replace the Euclidean dimensionality as an exponent, analysis tempts one to view every such exponent as a fractal dimensionality. Geometrically, this is totally unwarranted. Ref. 9 (p. 197) warned against pseudo fractals, but did not go far enough in the deviation of geometry (on this account), since it did not develop an actual model.

In more positive examples of the usefulness of my explicit fractal model of percolation clusters, experimental tests of its validity have now been performed using either Monte-Carlo runs of percolation on geometric lattices[10] or on actual films of gold[11] or of lead[11].

Many further uses of fractal models in the study of percolation are unfortunately beyond the scope of this article. For example, effective construction of fractals and careful inspection of the resulting illustrations circumscribed the problems raised by the fact that a fractal cannot be translationally invariant. To measure the strength of a fractal's departure from translational invariance, I introduced the notion of lacunarity (Chapters 34 and 35 of Ref. 2), and showed how one can adjust lacunarity, often at will, by modifying the model.

One will also observe that this article's fractal model leaves open the precise topological structure of the individual clusters; for example, it may but need not include "dangling bonds". In other words, the model must be narrowed down considerably if it is to agree with additional observations. In this respect, the reader's attention is drawn to Chapter 14 of Ref. 2 (especially page 132), which claims that individual clusters "backbones" can be modeled by a special lattice structure which I retrieved from the "Gallery of mathematical monsters", and to which I gave a name that was widely accepted: "Sierpinski gasket". See also Ref. 13.

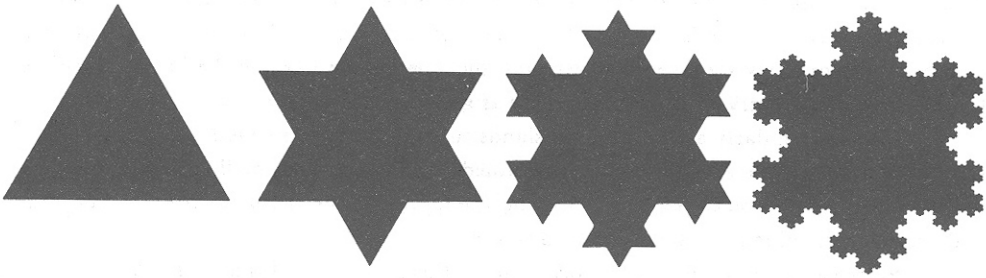
The gasket is also very useful in the study of critical phenomena on fractals[14,15]. It turns out to be a valuable addition to the sadly small collection of "physical" systems whose analysis demands no approximation.

It is known that spaces of "anomalous" dimensionality occur in many other branches of statistical physics. Typically, a formal dimensionality appears as an exponent, without a specified underlying geometry, and is followed by expansions in the dimensionalities $4-\varepsilon$, $1+\varepsilon$, or $2-\varepsilon$. While many requirements are imposed upon the underlying spaces, the first proof that these requirements are actually compatible has not been an analytical proof, but a constructive geometrical proof involving fractals; see Ref. 16.

One reason why explicit drawings can be fruitful in fractal geometry is of course psychological. Experience shows that the properties most useful to application are *not* those which were built in by definition, but other properties that the eye sees while the logical mind misses...and in many cases had missed for decades. Thus, Refs. 4, 9 and 2 provide steadily growing evidence that fractal constructions, once drawn and understood on their geometric terms, have very much mattered in many domains of science. For example, this is how I came to introduce lacunarity. This is also how, 85 years after Peano discovered the plane filling curves, I discovered behind most of them two "conjugate" plane filling trees; this discovery tamed the Peano curves. Similarly, 40 years after Hausdorff described certain infinitesimal limit effects through an unwieldy anomalous dimensionality, I found that in the case of self-similar fractals this dimensionality also rules finite size effects of interest to the physicist.

1.4. THE SNOWFLAKE PROTOTYPE OF FRACTAL CURVES. THE NOTION OF FRACTAL GENERATOR. SIMILARITY DIMENSIONALITY. FRACTAL DIMENSIONALITY.

The fractal model of percolation clusters generalizes a classic construct due to H. von Koch, known as the snowflake curve. The figure that follows shows a sequence of approximations to one third of the snowflake curve. Readers who desire a more detailed discussion are referred to my books.



The construction of each third of the snowflake curve is recursive, and its basic ingredient is a shape I call generator; it contains the points 0 and 1, and links them by N segments of length r. Here, $N=4$ and $r=1/3$. The k-th recursion stage begins with a broken line made of segments of length r^{k-1} and ends with a broken line made of segments of length r^k; it consists in replacing each leg of the input broken line by a suitably reduced version of the generator. The limit curve is called self-similar, because it is made of N replicas of itself, reduced in the ratio r. The similarity dimensionality is defined by $D=\log N/\log(1/r)$.

Calling the snowflake's contour a curve refers to a topological property. Without tear and without self-overlap, this contour can be deformed into a circle. By contrast, the present article is devoted to fractals that decompose into an infinity of (denumerable) disjoint fragments, each of them a connected curve. They are sums (Σ, sigma) of curves. To denote such shapes, I coined the term sigma-curve, or σ-curve, and several more specific terms to be defined below.

Curves and σ-curves can be said to have a topological dimensionality equal to 1. A self-similar curve or σ-curve is a fractal (according to the definition I introduced in my book) if its similarity dimensionality is >1. More generally, a self-similar set is a fractal if its similarity dimensionality D exceeds its topological dimensionality D_T.

1.5. COMMENT ABOUT ORGANIZATION OF THIS ARTICLE

In order to underline the descriptive and explanatory power of fractals, I did not organize my books around subject matter fields but around themes, each of which is useful in many subject matters. This principle is used in Chapter 13 of Ref. 2 and is preserved here. The concrete cases range from the coastlines of islands in an archipelago to percolation. The material in the first few sections was new to Ref. 9, and the bulk of the remainder was new to Ref. 2.

2. *AN EXPLICIT MODEL OF FRACTAL ISLANDS AND CLUSTERS*

2.1. INTRODUCTION TO GEOGRAPHICAL ISLANDS

The first working chapter of each of my books is titled "How Long Is the Coast of Britain"? That question, at first slightly foolish, improves upon acquaintance. The only correct answer is "it depends upon the beholder". As the contributions of increasingly small bays and promontories are taken into account, the length increases and the "true length" is both very large and very ill-determined. It is as good as infinite.

One can similarly ask how many islands surround Britain's coast? Surely, their number is both very large and very ill-determined. As increasingly small rock piles become listed as islands, the overall list lengthens, and the total number of islands is both very large and very ill-determined. It is as good as infinite.

Since earth's relief is finely "corrugated," there is no doubt that, just like a coastline's length, an island's total area is geographically infinite. But the domains surrounded by coastlines have well defined "map areas." And the way in which a total map area is shared among the different islands is an important geographic characteristic. One might even argue that this "area-number relation" contributes more to geographic form than do the shapes of the individual coastlines. For example, it is difficult to think of the Aegean Sea's shores without also including those of the Greek islands. The issue clearly deserves a quantitative study, and this chapter provides one, by generalizing the Koch curves.

Next, this chapter examines diverse other fragmented shapes obtained by generalizing the familiar fractal-generating processes: either the Koch procedure or curdling. The resulting shapes are called *contact clusters* here, and the diameter-number distribution is shown to be the same for them as for islands.

Special interest attaches to the plane-filling contact clusters, in particular those clusters generated by certain Peano curves, whose teragons do not self-intersect but have carefully controlled points of self-contact. The saga of the taming of Peano monsters (Chapter 7 of Ref. 2) is thereby enriched by a new scene!

Last but not least, this chapter includes the first part of a case study of the geometry of a very important physical phenomenon: Bernoulli percolation. This case study continues at the beginning of Chapter 14 of Ref. 2.

2.2. KORČAK EMPIRICAL LAW, GENERALIZED

List all the islands of a region by decreasing size. The total number of islands of size above a is to be written as Nr(A>a) on the same pattern as the notation Pr(A>a) of probability theory. Here, a is a possible value for an island map's area, and A denotes the area when it is of unknown value.

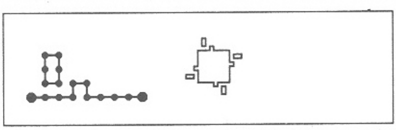
B and F' being two positive constants, to be called exponent and prefactor, one finds the following striking area-number relation:

$$\mathrm{Nr}(A>a)=F'a^{-B}\ .$$

Korčak 1938 (the name is pronounced Kor'chak) comes close to having discovered this rule, except that it claims that B = 1/2, which I found incredible, and which the data showed is unfounded. In fact, B varies between regions and is always >1/2. Let me now show that the above generalized law is the counterpart of the distribution Chapter 8 obtains for the gap lengths in a Cantor dust.

2.3. KOCH CONTINENT AND ISLANDS, AND THEIR DIVERSE DIMENSIONS

To create a Koch counterpart to the Cantor gaps, I let the generator split into disconnected portions. To insure that the limit fractal remains interpretable in terms of coastlines, the generator includes a connected broken line of $N_c<N$ links, joining the end points of the interval [0,1]. This portion will be called the *coastline generator*, because it determines how an initially straight coastline becomes transformed into a fractal coastline. The remaining $N-N_c$ links form a closed loop that "seeds" new islands and will be called *island generator*. Here is an example:



N=16
r=1/8
D=4/3

In later stages, the sub-island always stays to the left of the coastline generator (going from 0 to 1), and of the island generator (going clockwise).

A first novelty is that the limit fractal now involves two distinct dimensionalities. Lumping all the islands' coastlines together, $D = \log N/\log(1/r)$, but for the coastline of each individual island $D_c = \log N_c/\log(1/r)$, with the inequalities

$$1 \leq D_c < D.$$

The cumulative coastline, not being connected, is not itself a curve but an infinite sum (Σ, sigma) of loops. I propose for it the term *sigma-loop*, shortened into *σ-loop*.

Note that modeling of the observed relation between D and D_c in actual islands requires additional assumptions, unless it can be derived from a theory, as in Chapter 29 of Ref. 2.

2.4. THE DIAMETER-NUMBER RELATION

The proof that the generalized Korčak law holds for last section's islands is simplest when the generator involves a single island, and teragons are self-avoiding. (The term *teragons* denotes a fractal's approximating broken lines). Then the first stage of construction creates 1 island; let its "diameter," defined by $\sqrt{a}$, be λ_0. The second stage creates N islands of diameter $r\lambda_0$, and the mth stage creates N^m islands of diameter $\lambda = r^m\lambda_0$. Altogether, as λ is multiplied byr, $Nr(\Lambda > \lambda)$ is multiplied by N. Hence the distribution of Λ (for all values of λ of the form $r^m\lambda_0$) takes the form

$$Nr(\Lambda > \lambda) = F\lambda^{-D},$$

in which the crucial exponent is the coastline's fractal dimensionality! As a corollary

$$Nr(A > a) = F'a^{-B}, \text{ with } B = D/2,$$

we have thus derived the generalized generalized Korčak law. For other values of λ or a, one has the staircase curve familiar from the distribution of Cantor gaps' lengths, Chapter 8 of Ref. 2.

This result is independent of N_c and D_c. It extends to the case when the generator involves two or more islands. Now we can note that the empirical B regarding the whole Earth is of the order of 0.6. This value is very close to one half of D measured from the coastline lengths.

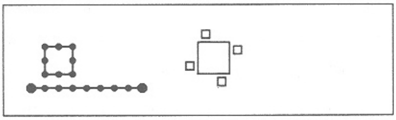
2.5. GENERALIZATION TO E>2

Now let the same construction be extended to an Euclidean space, whose dimensionality will be denoted by E. (This notation, in preference to the physicists' d, helps minimize the possible confusion between different notions of dimensionality.) In this generalization, it continues to be true that the E dimensional diameter, defined as volume$^{1/E}$, is ruled by a hyperbolic expression of the form $Nr(\text{volume}^{1/E}>\lambda) = F\lambda^{-D}$, wherein the crucial exponent is D.

The exponent D also rules the special case of Cantor dusts for $E=1$, but there is a major difference. The length outside the Cantor gaps vanishes, while the area outside the "Koch" islands can be, and in general is, positive. Ref. 2 returns to this topic in Chapter 15.

2.6. FRACTAL DIMENSION MAY BE SOLELY A MEASURE OF FRAGMENTATION

The preceding construction also allows the following generator



$N=16$
$r=1/8$
$D=4/3$

The overall D is unchanged, but the coastline D_c takes the smallest allowable value, $D_c=1$. In the present model, island coastlines are allowed to be rectifiable! When such is the case, the overall D is not a measure of irregularity, as is the case for the snowflake curve, but solely of fragmentation. Instead of the wiggliness of individual curves, D measures the number-area relationship for an infinite family of rectangular islands.

It is still true that, when the length is measured with a yardstick of ε, the result tends to infinity as $\varepsilon \to 0$, but there is a new reason for this. A yardstick of length ε can only measure islands with a diameter of at least ε. However, the number of such islands increases as $\varepsilon \to 0$, and the measured length behaves like ε^{1-D}, exactly as in the absence of islands.

In the general case where $D_c>1$, the value of D_c measures irregularity alone, while the value of D measures irregularity and fragmentation in combination.

A FRAGMENTED FRACTAL CURVE MAY HAVE TANGENTS EVERYWHERE. By rounding off the islands' corners, one may make every coastline have a tangent at every point, while the areas, hence the overall D, are unaffected. Thus, being a fractal σ-curve and being without tangent are *not* identical properties.

2.8. THE INFINITY OF ISLANDS

AN INNOCUOUS DIVERGENCE. As $a \to 0$, $Nr(A>a) = Fa^{-B}$ tends to infinity. Hence, the generalized Korčak law agrees with our initial observation that islands are practically infinite in numbers.

LARGEST ISLAND'S RELATIVE AREA. This last fact is mathematically acceptable because the cumulative *area* of the very small islands is finite and negligible. All islands of area below ε have a total area that behaves like the integral of $a(Ba^{-B-1})=Ba^{-B}$ from 0 to ε. Since $B<1$, this integral converges, and its value $B(1-B)^{-1}\varepsilon^{1-B}$ tends to 0 with ε.

Consequently, the largest island's relative contribution to all the islands' cumulative area tends to a positive limit as the islands increase in numbers. It is *not* asymptotically negligible.

LONGEST COASTLINE'S RELATIVE LENGTH. On the other hand, assuming $D_c=1$, the coastline lengths have a hyperbolic distribution with the exponent $D>1$. Hence the cumulative coastline length of small islands is infinite. And, as the construction advances and the number of islands increases, the coastline length of the largest island becomes relatively negligible.

RELATIVELY NEGLIGIBLE SETS. More generally, the inequality $D_c<D$ expresses that the curve drawn using the coastline generator alone is negligible in comparison to the whole coastline. In the same way, a straight line ($D=1$) is negligible in comparison to a plane ($D=2$). Just as a point chosen at random in the plane almost never falls on the x-axis, a point chosen at random on the coastline of a "core" island surrounded with sub-islands almost never falls on the core island's coastline.

2.9. SEARCH FOR THE INFINITE CONTINENT

In a scaling universe, the distinction between the islands and the continent cannot be based on tradition or "relative size." The only sensible approach is to define the continent as a special island with an *infinite* diameter. Let me now show that for the constructions at the beginning of this chapter, the probability of a continent being generated is zero.

In a sensible search for a continent, we must no longer choose the initiator and the generator separately. From now on, the same generator must be made to serve both for interpolation and for extrapolation. The process runs by successive stages, each subdivided into steps.

The first step upsizes our chosen generator in the ratio of $1/r$. The second step puts a "mark" on *one* of the links of the upsized generator. The third step displaces the upsized generator, to make its marked link coincide with $[0,1]$. The fourth and last step interpolates the upsized generator's remaining links.

The same process is repeated ad infinitum, its progress and outcome being determined by the sequence of positions of the "marked" links. This sequence can take diverse forms.

The first form requires the coastline generator to include a positive number N_c-2 of "nonextreme" links, defined as belonging to the coastline generator but not ending on either 0 or 1. If the mark is consistently put on a nonextreme link, each stage of extrapolation expands the original bit of coastline, and eventually causes it to be incorporated into a fractal coastline of infinite extent in both directions. This proves that it is indeed *possible* to obtain a continental coastline in this setup.

Secondly, always mark an extreme link of the coastline generator, each possibility being chosen an infinite number of times. Then our bit of coastline again expands without end. If we always choose the same link, the coastline expands in only one direction.

Thirdly, always mark a link that belongs to the island generator. Then the biggest island before extrapolation is made to lie off a bigger island's shore, then off-off a still bigger island's, and so on ad infinitum. No continent is *ever* actually reached.

Next suppose that the marks fall according to the throws of an N-sided die. In order for the extrapolation to generate a continent, it is obviously necessary that all the marks beyond a finite (kth) stage be placed upon one of N_c-2 nonextreme links of the coastline generator. Call them "winning" links. To know one has reached a continent after k stages, one must know that thereafter *every* throw of our die, with *not one* exception, will win. Such luck is not impossible, but it is of vanishing probability.

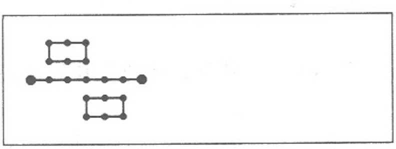
2.10. ISLAND, LAKE AND TREE COMBINATIONS

The Koch islands being mutually similar, their diameter Λ can be redefined as the distance between any two specified points, best chosen on the coastline. Next, we observe that the derivation of the diameter-number relation makes specific use of the assumption that the generator includes a coastline generator. But the assumption that the generator's remaining links form islands, or are self-avoiding, is never actually used. Thus, the relation

$$Nr(\Lambda>\lambda)=F\lambda^{-D}$$

is of very wide validity. One can even release the condition the teragons initiated by two intervals must not intersect. Let us now show by examples how the configuration of the original $N-N_c$ links can affect the resulting fractal's topology.

COMBINATION OF ISLANDS AND LAKES. Relieve the generator from the requirement of being placed to the left, going clockwise. When it is placed to the right, it forms lakes instead of islands. Alternatively, one may include *both* lakes and islands in the same generator. Either way, the final fractal is a σ-loop whose component loops are nested in each other. This is illustrated by the generator



N=18
r=1/6
D~1.6131

THE ELUSIVE CONTINENT. When the generator is as above and the initiator is a square, the initiator's side injects a nonintrinsic outer cutoff. A more consistent approach is to extrapolate it as we did for islands without lakes. Again, it is almost sure that no continent is ever reached, and that the nesting of islands within lakes within islands continues without bound.

AREA-NUMBER RELATION. In order to define the area of an island (or lake), one may at will take either the total area, or the area of land (or water), within its coastline. The two differ by a fixed numerical factor, hence affect $Nr(A>a)$ through its prefactor F', not its exponent $D/2$.

COMBINATION OF INTERVALS AND TREES. Now assume that the $N\text{-}N_c$ links form either a broken line with two free ends, or a tree. In either case, the fractal splits into an infinite number of disconnected pieces, each of them a curve. This σ-curve is no longer a σ-loop; it is either a σ-tree or a σ-interval.

2.11. THE NOTION OF CONTACT CLUSTER

The generator may also combine loops, branches and diverse other topological configurations. If so, the limit fractals' connected portions recall the *clusters* of percolation theory (as seen later in this chapter) and of many other areas of physics. To us, this usage is terribly unfortunate, due to the alternative meaning of *cluster* in the study of galaxies. We need therefore a more specific and cumbersome term. I settled on "contact cluster." Luckily, the term σ-cluster is not ambiguous.

PLANE-FILLING CONTACT CLUSTERS. As D reaches its maximum $D=2$, the arguments in the preceding section remain valid, but additional comments become necessary. Each individual cluster tends to a limit, which may be a straight line, but in most cases is a fractal curve. On the other hand, all the clusters together form a σ-curve, whose strands fill the plane increasingly tightly. The limit of this σ-curve is no longer a σ-curve, but a domain of the plane.

THE ELUSIVE INFINITE CLUSTER. *No actually infinite* cluster is involved in the present approach. It is easy to arrange the generator's topology so that any given bounded domain is almost surely surrounded by a contact cluster. This cluster is in turn almost surely surrounded by a larger cluster, etc. There is no upper bound to cluster size. More generally, when a cluster seems infinite because it spans a very large area, the consideration of an even larger area will almost surely show it to be finite.

2.12. MASS-NUMBER AND WEIGHTED DIAMETER-NUMBER RELATIONS. THE EXPONENTS $D-D_c$ AND D/D_c.

Now let us reformulate the function $Nr(\Lambda>\lambda)$ in two ways: first by replacing a cluster's diameter λ by its mass μ, then by giving increased weight to large contact clusters.

Here, a cluster's mass is simply the number of links of length b^{-k} in the clusters itself (do *not* count the links *within* a looping cluster!). In effect we center a square of side b^{-k} on each vertex, and add half a square at each end-point.

The mass of a cluster of diameter Λ being the area of its modified sausage, $M \propto (\Lambda/b^k)^{D_c}(b^k)^2 = \Lambda^{D_c}/(b^k)^{D_c-2}$. Since $D_c<2$, $M \to 0$ as $k \to \infty$. The mass of all the contact clusters taken together is $\propto (b^k)^{D-2}$; if $D<2$, it too $\to 0$. And the relative mass of any individual contact cluster is $\propto (b^k)^{D_c-D}$; it tends to 0 at a rate that increases with $D-D_c$.

MASS-NUMBER RELATION. Clearly,

$$\mathrm{Nr}(M>\mu) \propto (b^k)^{-D+2D/D_c}\mu^{-D/D_c}.$$

DISTRIBUTION OF DIAMETER WEIGHTED BY MASS. Observe that $\mathrm{Nr}(\Lambda>\lambda)$ counts the number of lines above line λ in a list that starts with the largest contact cluster, continues with the next largest, etc. But we shall momentarily have to attribute to each contact cluster a number of lines equal to its mass. The resulting relation is easily seen to be

$$\mathrm{Wnr}(\Lambda>\lambda) \propto \lambda^{-D+D_c}$$

2.13. THE MASS EXPONENT $Q=2D_c-D$

Denote by $\mathcal{F}$ a fractal of dimensionality D, constructed recursively with $[0,\Lambda]$ as initiator, and take its total mass to be Λ^D. When $\mathcal{F}$ is a Cantor dust, Chapter 8 of Ref. 2 shows that the mass in a disc of radius $R<\Lambda$ centered at 0 is $M(R) \propto R^D$. More accurately, the quantity $\log[M(R)R^{-D}]$ is a periodic function of $\log_b(\Lambda/R)$, but we shall not dwell on these complications because they vanish when the fractal is modified so that all $r>0$ are admissible self-similarity ratios.

We know that $M(R) \propto R^D$ also applies to the Koch curve of (Chapter 6 of Ref. 2). Furthermore, this formula extends to the recursive islands and clusters of this chapter, with D replaced by D_c. In all cases, the mass in a disc of radius R centered at 0 takes the form

$$M(R,\Lambda) = R^{D_c}\phi(R/\Lambda),$$

with ϕ a function deducible from the shape of $\mathcal{F}$. In particular,

$$M(R,\Lambda) \propto R^{D_c} \text{ when } R \ll \Lambda,$$
$$\text{and } M(R,\Lambda) \propto \Lambda^{D_c} \text{ when } R \gg \Lambda.$$

Now consider the weighted average of $M(R)$, to be denoted by $\langle M(R) \rangle$, corresponding to the case when Λ is variable with the widely spread-out hyperbolic distribution $\mathrm{Wnr}(\Lambda>\lambda) \propto \lambda^{-D+D_c}$. We know that $1 \leq D_c < D \leq 2$. Excluding the combination of $D=2$ and $D_c=1$, $0<D-D_c<D_c$. It follows that

$$\langle M(R) \rangle \propto R^Q \text{ with } Q=2D_c-D>0.$$

When the disc's center is a point of $\mathscr{F}$ other than 0, the prefactor changes, but the exponent is unchanged. It also remains unchanged by averaging over all positions of the center in $\mathscr{F}$, and by the replacement of [0,1] by a different initiator. Usually, an arc of random size Λ is also of random shape. But the above formulas for $M(R,\Lambda)$ apply to $\langle M(R,\Lambda)\rangle$ averaged over all shapes. The final result is unchanged.

REMARK. The preceding derivation does not refer to the clusters' topology: they can be loops, intervals, trees, or anything else.

CONCLUSION. The formula $\langle M(R)\rangle \propto R^Q$ shows that, when Λ is hyperbolically distributed, hence of very wide statistical scatter, one of the essential roles of dimensionality is taken up by an exponent *other than* D. The most natural exponent is $2D_c - D$, but different weighting function give different Q's.

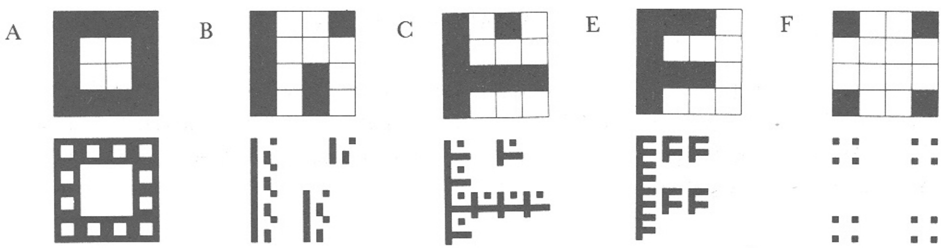
WARNING: NOT EVERY MASS EXPONENT IS A DIMENSION. The combined quantity Q is important. And, since it is a mass exponent, it is tempting to call it a dimensionality, but this temptation has no merit. Mixing many clusters with identical D_c but varying Λ leaves D_c unchanged, because fractal dimensionality is *not* a property of a mixed population of sets. Dimension is but a property of an individual set. Both D and D_c are fractal dimensionalities, but Q is not.

More generally, many areas of physics involve relations of the form $\langle M(R)\rangle \propto R^Q$, but such a formula does *not by itself* guarantee that Q is a fractal dimensionality. And calling Q an *effective dimensionality*, as some authors have proposed, is an empty gesture because Q does not possess any of the other properties that characterize D (for example, Ref. 2 shows that sums or products of D's have a meaning with no counterpart in the case of Q). Moreover, this empty gesture has proven a source of potential confusion.

2.14. NONLUMPED CURDLING CLUSTERS

We now proceed to describe two additional methods for generating contact clusters. One is based on curdling and applies for $D<2$, while the other is based on Peano curves and applies for $D=2$. The reader interested in percolation may skip this section and the next.

First, let us replace the Koch construction by the natural generalization of Cantor curdling to the plane. As illustration, consider the following five generators, with the next construction stage drawn underneath



In all these cases, the limit fractal is of zero area and contains no interior point. Its topology can take diverse forms, determined by the generator.

With generator A, the precurd of every stage k is connected, and the limit fractal is a curve, an example of the very important Sierpiński carpet examined in Chapter 14.

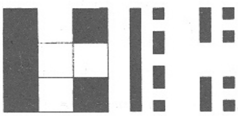
With generator F, the precurd splits into disconnected portions, whose maximum linear scale steadily decreases as $k \to \infty$. And the limit fractal is a dust, akin to the Fournier model of Chapter 9 of Ref. 2.

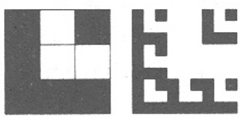
The generators B, C and E are more interesting: in their case, the precurd splits into pieces to be called *preclusters*. Each stage can be said to transform every "old" precluster by making it thinner and wigglier, and to give birth to "new" preclusters. Nevertheless, by deliberate choice of generators, each newborn precluster is entirely contained in a single smallest cell in the lattice prevailing before its birth. By contrast with the "cross lumped clusters" of the next section, the present ones are to be called "nonlumped." It follows that the limit contact clusters have a dimensionality of the form $\log N_c / \log b$, where N_c is an integer *at most* equal to the number of cells in the generator's largest component. This maximum is attained for generators B and C, for which the contact clusters are, respectively, intervals with $D_c = 1$ and fractal trees with $D_c = \log 7 / \log 4$. But the fractal based on the generator E does not attain this maximum: in its case, the F-shaped preclusters keep splitting into parts, and the limit, again, is made of straight intervals with $D_c = 1$.

Drawing the collection of cells of side b^{-k} intersected by a contact cluster, the diameter-number relation and the other results of the preceding sections extend unchanged.

2.15. CROSS LUMPED CURDLING CLUSTERS

Next, let the generator of plane curdling takes either of the following shapes, with the next construction stages drawn to the side





Both cases exhibit massive "cross lumping," meaning that each newborn precluster combines contributions coming from several smallest lattice cells prevailing before its birth.

In the Koch context, an analogous situation prevails when the teragons are allowed to self-contact, resulting in the merger of small cluster teragons. In either case, the analysis is cumbersome, and we cannot dwell on it. But $Nr(\Lambda > \lambda) \propto \lambda^{-D}$ remains a valid relation for small λ.

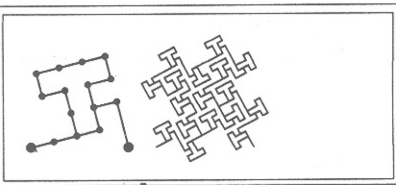
However, if one attempts to estimate D from this relation, without excluding the large λ's, the estimate is systematically biased and smaller than the true value.

Novel features arise concerning the quantity b^{D_c}: it need not be an integer deducible from the generator by simple inspection, but it may be a fraction. The reason is that every contact cluster combines: (a) an integer number of versions of itself, downsized in the ratio $1/b$, and (b) many downsized versions due to lumping, which involve smaller ratios of the form $r_m = b^{-k(m)}$. The dimensionality-determining equation $\Sigma r_m{}^D = 1$ given on page 56 of Ref. 2, when rewritten in terms of $x = b^{-D}$, takes the form $\Sigma a_m x^m = 1$. Cases where $1/x$ is an integer can only occur as exceptions.

2.16. KNOTTED PEANO MONSTERS, TAMED

A plane-filling collection of clusters ($D=2$) cannot be created by curdling, but I found an alternative approach, using Peano curves beyond those that are tamed in Chapter 7 of Ref. 2. As the reader must recall, Peano curves with self-avoiding teragons create river and watershed trees. But some other Peano curve teragons (for example the teragons in Plate 63, assuming that the corners are not rounded off) are simply chunks of lattice. As the construction proceeds, the open lattice cells separated by such curves "converge" to an everywhere dense dust, e.g., to the points for which neither x nor y is a multiple of b^{-k}.

Between these extremes stands a new interesting class of Peano curves. Their generators are exemplified by the following, shown together with the next step



This class of Peano curves is now ready to be tamed. We observe that each point of self-contact "knots off" an open precluster, which may acquire branches and self-contacts, sees chunks of itself "knotted away," and eventually thins down to a highly ramified curve that defines a contact cluster. A cluster's diameter Λ, defined as in previous sections of this chapter, is fixed from the moment of birth: roughly equal to the side of the square that "seeded" this cluster. Its distribution is ruled by the familiar relation $Nr(\Lambda > \lambda) \propto \lambda^{-2}$.

Observe in passing that, while Koch contact clusters are limits of recursively constructed curves, the present clusters are limits (in a peculiar sense) of the open components of the *complement* of a curve.

3. *BERNOULLI PERCOLATION CLUSTERS*

Whichever method is used to generate fractal contact clusters that satisfy $D=E$ and

$D_c < D$, they provide a geometric model that had been lacking in the study: Bernoulli percolation through lattices. J. M. Hammersley, who posed and first investigated this problem, did not inject Bernoulli's name in this context, but fractal percolation Chapter 23 of Ref. 2 makes the full term unavoidable here.

DEFINITIONS. The simplest percolation problem for $E=2$ is bond percolation on a square lattice. We construct a large square lattice with sticks made either of insulating vinyl or of conducting copper. A *Bernoulli lattice* obtains if each stick is selected at random, independently of the other sticks, the probability of choosing a conducting stick being p. Maximal collections of connected copper or vinyl sticks are called copper or vinyl *clusters*. When the lattice includes at least one uninterrupted string of copper sticks, the current can *flow through* from one side of the lattice to the other, and the lattice is said to *percolate*. (In Latin, *per* = through, and *colare* = to flow.) The sticks in uninterrupted electric contact with the top and bottom sides of the lattice form a "percolating cluster," and the sticks actually active in conducting form the percolating cluster's "backbone."

The generalization to other lattices, and to $E>2$, is immediate.

CRITICAL PROBABILITY. Hammersley's most remarkable finding concerns the special role played by the *critical probability* p_{crit}. This quantity enters in when the Bernoulli lattice's size (measured in numbers of sticks) tends to infinity. One finds that, when $p>p_{crit}$, the probability that there exists a percolating cluster increases with lattice size, and tends to 1. When $p<p_{crit}$, to the contrary, the probability of percolation tends to 0.

Bond percolation on square lattices being such that either copper or vinyl must percolate, $p_{crit} = 1/2$.

ANALYTICAL SCALING PROPERTY. The study of percolation long devoted itself to the search for analytic expressions to relate the standard quantities of physics. All these quantities were found to be *scaling*, in the sense that the analytic relations between them are given by power laws. For $p \neq p_{crit}$, scaling extends up to an outer cutoff dependent on $p-p_{crit}$ and denoted by ξ. As $p \rightarrow p_{crit}$, the cutoff satisfies $\xi \rightarrow \infty$. Physicists postulate (see Stauffer 1979, p. 21) that $\langle M(R,\Lambda) \rangle$ follows the rule obtained in section 2.13.

3.2. THE CLUSTERS' FRACTAL GEOMETRY

THE CLUSTERS' SHAPE. Let $p=p_{crit}$, and let individual sticks decrease in size while the total lattice size remains constant. The clusters become increasingly thin ("all skin and no flesh"), increasingly convoluted, and increasingly rich in branches and detours ("ramified and stringy"). Specifically the number of sticks situated outside of the cluster, but next to a stick within the cluster, is roughly proportional to the number of sticks within the cluster[6].

HYPOTHESIS THAT CLUSTERS ARE FRACTALS. It is natural to conjecture that the property of scaling extends from analytic properties to the clusters' geometry. But this idea could not be implemented in standard geometry, because the clusters are *not* straight lines. Fractal geometry is of course specifically designed to eliminate such difficulties: thus, I conjectured that clusters are representable by fractal σ-curves satisfying $D=2$ and $1<D_c<D$. This claim has been accepted, and found to be fruitful. It is elaborated upon in Chapter 36 of Ref. 2.

To be precise, scaling fractals are taken to represent the clusters that are *not* truncated by the boundary of the original lattice. This excludes the percolating cluster itself. (The term *cluster* has a gift for generating confusion!) To explain the difficulty, start with an extremely large lattice, pick a cluster on it, and a smaller square that is spanned by this cluster. By definition, the intersection of this cluster and the smaller square includes a smaller percolating cluster, but in addition it includes a "residue" that connects with the smaller percolating cluster through links *outside* the square. Note that neglect of this residue creates a downward bias in the estimation of D_c.

VERY ROUGH BUT SPECIFIC NONRANDOM FRACTAL MODELS. To be valid, the claim that any given natural phenomenon is fractal must be accompanied by the description of a specific fractal set, to serve as first approximation model, or at least as mental picture. My Koch curve model of coastlines (Chapter 6 of Ref. 2), and the Fournier model of galaxy clusters (Chapter 9 of Ref. 2), demonstrate that a rough nonrandom picture may be very useful. Similarly, I expect recursively constructed contact clusters (like those introduced in this chapter) to provide useful fractal models of the ill-known natural phenomena that are customarily modeled by Bernoulli clusters.

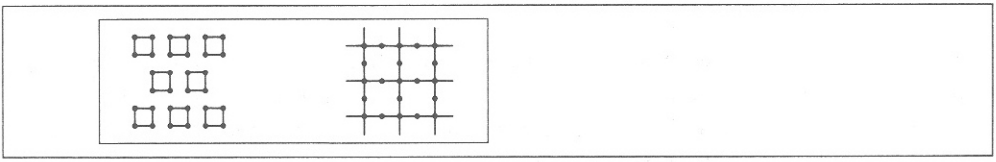
However, the Bernoulli clusters themselves are fully known (at least in principle), hence modeling them via explicit recursive fractals is a different task. The Koch contact clusters I studied are not suitable, due to dissymetry between vinyl and copper, even when there are equal numbers of sticks of both kinds. Next examine the knotted Peano curve clusters. Take an advanced teragon, and cover the cells to the left of the curve with copper, and the other cells with vinyl. The result involves a form of percolation applied to lattice cells (or to their centers, called sites). The problem is symmetric. But it differs from the Bernoulli problem, because the configuration of copper or vinyl cells are *not* the same as in the case of independence: for example, 9 cells forming a supersquare can all be of copper or vinyl in the Bernoulli case, but not in the knotted Peano curve case. (On the other hand, both models allow groups of 4 cells forming a supersquare to take any of the possible configurations.) This difference has far-reaching consequences: for example, neither copper nor vinyl percolate in the Bernoulli site problem with $p=1/2$, while both percolate in knotted Peano clusters, implying that $1/2$ is a critical probability.

The list of variants of Bernoulli bond percolation is already long, and can easily be lengthened further. And I have already examined many variants of recursively constructed fractal contact clusters. The detailed comparison of these lists is unfortunately complicated, and I shall not dwell on it here.

Let me therefore be satisfied with stating the loose conclusion that significant fractal essentials of the Bernoulli percolation problem seem to be illustrated by nonrandom space-filling σ-clusters defined earlier in this chapter. This model's principal weakness is that it is completely indeterminate beyond what has been said. It can accommodate any observed degree of irregularity and fragmentation. On the matter of topology, see Chapter 14 of Ref. 2.

MODEL OF CRITICAL CLUSTERS. Specifically, consider the critical clusters, defined as the clusters for $p=p_{crit}$. To represent them, a recursive σ-cluster is extrapolated as indicated in earlier sections of this chapter. Then it is truncated by stopping the interpolation so that the positive inner cutoff is the cell size in the original lattice.

MODELS OF NONCRITICAL CLUSTERS. To extend this geometric picture to noncritical clusters, that is, to clusters for $p \neq p_{crit}$, we seek fractals with a positive inner cutoff and a finite outer cutoff. Analysis calls for the largest copper cluster's extent to be of the order of ξ when $p<p_{crit}$, and to be infinite when $p>p_{crit}$. Either outcome is readily implemented. For example, one can start with the same generator as in the preceding subsection. But, instead of extrapolating it naturally, one initiates it with either of the following shapes



SUBCRITICAL CLUSTERS. The initiator to the left, which is geared towards $p<p_{crit}$, is made of squares of side $\xi/2$. Now let the chosen generator be positioned *in* through each initiator's left side, and *out* through the other sides. The initiator square will transform into an atypical cluster of length ξ, surrounded by many typical clusters of length $<\xi$.

SUPERCRITICAL CLUSTERS. The initiator to the right, which is geared towards $p>p_{crit}$, is made of those lines of the initial square lattice, whose x or y coordinates are even integers. Four links radiate from each node whose coordinates are even integers; the chosen generator is always positioned to the left. In the special case when the coastline generator involves no loops nor dangling links, the resulting picture is a de-randomized and systematized variant of a crude model of clusters based solely on "nodes and links."

Observe that the fractal geometric picture deduces the noncritical clusters from the critical ones, while physicists prefer to consider the critical clusters as limits of the noncritical clusters for $\xi \to \infty$.

3.3. CRITICAL BERNOULLI CLUSTERS' D_c

The value of D_c is immediately inferred from either the exponent $D/D_c=E/D_c$ in the formula for $Nr(M>\mu)$, or the exponent $Q=2D_c-D=2D_c-E$ in the formula for $\langle M(R)\rangle$. Using the Greek letters τ, δ and η with the meanings customary in this context, we find $E/D_c=\tau-1$ and $2D_c-E = 2-\eta$. That is,

$$D_c = E/(\tau-1) = E/(1+\delta^{-1}),$$
$$\text{and } D_c=1+(E-\eta)/2.$$

Due to known relations between τ, δ and η, the above formulas for D_c are equivalent. Conversely, their equivalence does not reside in physics alone, because it follows from geometry.

Independently of each other, Refs. 18, 19 and 20 obtain the same D_c. They start from the properties of clusters for $p>p_{crit}$, hence express their result in terms of different critical exponents (β, γ, ν and σ). These derivations do not involve a specific underlying fractal picture. The dangers inherent in this approach, against which we warned earlier in this chapter, are exemplified by the fact that it misled Stanley[7b] into advancing Q and D_c are equally legitimate dimensionalities.

For $E=2$, the numerical value is $D_c=1.89$. It is compatible with the empirical evidence, as obtained by a procedure I often use for fractals. It is based on an asymptotic form of the similarity dimension. Pick r, which need *not* be of the form $1/b$ (b an integer). Then take a big eddy, which is simply a square or cubic lattice of side set to 1. Pave it as best you can with subeddies of side r, count the number N of the squares or cubes that intersect the cluster, and evaluate $\log N/\log(1/r)$. Then repeat the process with each nonempty subeddy of side r by forming subsubeddies of side r^2. Continue as far as feasible. The most meaningful results obtain when r is close to 1. Some early simulations[21,22] gave the biased estimate $D^+\sim1.77$, but large simulations[23] confirm D.

The biased experimental D^+ is very close to Q, hence briefly seemed to confirm the theoretical arguments Refs. 7a and 21, which were both in error in claiming that the dimensionality is Q. The error was brought to my attention by S. Kirkpatrick. A different and even earlier incorrect estimate of D is found in Ref. 6.

4. *THE CYPRESS TREES OF OKEFENOKEE*

When a forest that is not "managed" systematically is observed from an airplane, its boundary is reminiscent of an island's coastline. Individual tree patches' outlines are extremely ragged or scalloped, and each large patch is trailed by satellite patches of varying area. My hunch that these shapes may follow the Richardson and/or generalized Korčak laws, is indeed confirmed by an unpublished study of the Okefenokee swamp by H. M. Hastings, R. Monticciolo & D. VunKannon. The patchiness of cypress is great, with $D\sim1.6$; the patchiness of broadleaf and mixed broadleaf trees is much less pronounced, with D near 1. My informants comment on the presence of an impressive variety of scales both on personal inspection and on examination of vegetation maps. There is an inner cutoff of about 40 acres, probably a consequence of aerial photography.

1 Much of this text was available in private memoranda since 1978, the bulk being identical to Chapter 13 of Ref. 2 written in August 1981. However, the presentation has been indirectly affected by the long range program of research I pursue with A. Aharony and Y. Gefen. In order to obtain an approximately self-contained work, an introduction and notes were added. The mathematicians' term *dimension* was replaced throughout by the physicists' *dimensionality*. Typographical slips were corrected, including an infortunate one: in the first formula on p. 129 of Ref. 2, and in the line above the formula, τ was (of course) a misprint for $\tau-1$.

2 B. B. Mandelbrot, *The Fractal Geometry of Nature*, (Freeman, San Francisco, Cal. 1982).

3 Before scaling became widely used in statistical physics I used it in many other fields, including linguistics and economics (see Chapters 37 and 38 of Ref. 2).

4 B. B. Mandelbrot, *Les objets fractals: form, chance & dimension*, (Flammarion, Paris, 1975).

5 The need for fractals in the description of percolation clusters first became apparent in the fall of 1975, when I visited Rutgers University to give a physics colloquium on fractals, and was shown Leath's Monte Carlo simulations. Unfortunately, Leath went on to estimate D incorrectly[6] as equal to the total dimensionality of all the clusters (D=2), rather than the dimensionality of an individual cluster. In the spring of 1976, I discussed with H. E. Stanley the dimensionality-like exponent $2D-D_C$ that he had identified[7a], and I agreed much too hastily that this exponent was likely to be a fractal dimensionality, which it is not. These false starts led to some initial confusion about the fractal aspects of percolation. I regret having further contributed to it by being slow in publishing an explicit fractal model.

6 P. L. Leath, Phys. Rev. B **14**, 5046 (1976).

7 a) H. E. Stanley, R. J. Birgenau, P. J. Reynolds, and J. F. Nicoll, J. Phys. C9, L553 (1976). b) H. E. Stanley, J. Phys. A, **10** L211 (1977).

8 K. G. Wilson, Revs. Mod. Phys. **47**, 773 (1975).

9 B. B. Mandelbrot, *Fractals: Form, Chance and Dimension*, (Freeman, San Francisco, Cal. 1977).

10 A. Kapitulnik, A. Aharony, G. Deutscher and D. Stauffer, to appear.

11 R. F. Voss, R. B. Laibowitz and E. I. Alessandrini, Phys. Rev. Lett. **49**, 1441 (1982).

12 A. Kapitulnik and G. Deutscher, Phys. Rev. Lett. **49** 1444 (1982).

13 Y. Gefen, A. Aharony, B. B. Mandelbrot and S. Kirkpatrick, Phys. Rev. Lett. **47**, 1771 (1981).

14 Y. Gefen, B. B. Mandelbrot and A. Aharony, Phys. Rev. Lett. **45**, 855 (1980).

15 Y. Gefen, A. Aharony and B. B. Mandelbrot, J. Phys. A (in the press) (1983).

16 Y. Gefen, Y. Meir, B. B. Mandelbrot and A. Aharony, Phys. Rev. Lett. **50** , 145 (1983).

17 J. Korčak, Bull. Inst. Int. Stat. **111**, 295 (1938).

18 R. J. Harrison, G. J. Bishop and G. P. Quinn, J. Stat. Phys., **19**, 53 (1978).

19 S. Kirkpatrick, *Les Houches Summer School on Ill Condensed Matter 1978*, edited by R. Balian *et al.* (North-Holland, Amsterdam, 1979), p. 323.

20 D. Stauffer, Phys. Rep. **54**, 1 (1979).

21 B. B. Mandelbrot, Ann. Israel Phys. Soc. **2**, 226 (1978).

22 J. W. Halley and T. Mai Phys. Rev. Lett. **43**, 740 (1979).

23 D. Stauffer, Z. Physik B, **37**, 597 (1980).