16

The Dimensions of Cosmic Fractals

Reuben Thieberger, Edward A. Spiegel, Leonard A. Smith

Abstract

Chaotic behavior is often described by *flows*, which are vector fields in appropriate spaces. Vector fields, such as the magnetic field in the sun, are important in astrophysics, and they are typically chaotic. Thus, when these fields are solenoidal, they display a mixture of regular and chaotic portions that qualitatively rationalize the spotty appearance of the solar magnetic field. A less tangible example is the flow that can be imagined to sweep galaxies forward in cosmic time. The evidence suggests that this flow forms a fractal object whose lacunae or voids are analogues of the stellar spots.

The textures of these objects will help the cosmologist in understanding some large-scale dynamics. We describe and assess the cosmologists' schemes to quantify the galactic distribution by determining its dimensions. Theoretical sets and a set of galaxy positions measured at the Observatory of Nice provide the illustrations. We find that the data suggest a value close to 2, which is rather larger than the dimension of the distribution of galaxies than has been generally accepted. The methods described for finding such results may be helpful in other astrophysical studies involving complex sets.

Cosmic Chaos

Reliable identification of chaotic behavior in astrophysical systems is usually much more difficult than for controlled laboratory systems. In astrophysics as in any *observational* science, it is hard to develop the extensive data sets needed to detect chaos. It is also difficult to isolate the phenomena of central interest without contamination from other processes in observational sciences.

Consider pulsars, for example. The flashes seen from pulsars are believed to emanate from polar caps of rotating neutron stars. The rotation rates of these stars decrease slowly with time, but the rates of decrease normally fluctuate slightly. To decide whether these fluctuations are deterministic, we need a good series of measures of the rotation rate. Those might be obtained from the blinking light from the stars' polar caps, which are thought to arise like flashes from light houses. But the neutron stars are intrinsically inconstant, like most stars. They may well emit their light chaotically; hence it is hard to say from the observations whether the rotational braking, or the emission process, or both are chaotic.

Even though it is difficult to establish the existence of deterministic temporal chaos in astrophysics, the basic ideas of chaos theory are helpful in astronomical thinking. They can rationalize apparently aperiodic behavior that we encounter in the universe as deterministic processes. Astromathematicians no longer need to invoke the *deus ex machina* called noise to understand aperiodicity. In chapters 17 and 18 by Buchler and Regev in this volume, one sees examples of how astrophysicists are incorporating the ideas of temporal chaos into their work.

Beyond this, the ideas of chaos theory can influence our vision of the mathematical objects with which we describe the contents and structure of the universe. Chaos theory helps us to cope with irregular temporal behavior; it provides us with ways to think about complexity in spatial structure; it provides tools for quantifying complicated behavior. This chapter is about some of the issues of such quantification. We shall describe methods for determining the dimensions of fractal sets and illustrate their use with theoretical sets and an astrophysical example.

Chaotic Vector Fields

Erratic temporal behavior is but a superficial manifestation of chaos. Astrophysicists can avail themselves of other, far-reaching aspects of chaos theory. For example, the property of being chaotic can apply to mathematical structures such as vector fields, not only to temporal behavior itself. This is a useful notion, and it is easy to grasp.

The trajectory of a particle is described by giving its position, x, as a function of time, t. Such information often emerges in the solution of a differential equation of the form

$$\dot{x} = \boldsymbol{v}\,(x, t). \tag{1}$$

Here the vector field $\boldsymbol{v}$ is a specified *flow* that carries the particle about. Equation 1 is called a dynamical system by mathematicians. Fluid dynamicists might think of the vector field $\boldsymbol{v}$ as the velocity field for the flow of a real fluid. For them though, the content of Eq. 1 is kinematic, and it describes how a fluid particle moves. A trajectory of Eq. 1 would be called a *streak line* in fluid dynamics.

Suppose that we are working in a three-dimensional (3D) space, as for ordinary physical flows. A way to visualize the nature of a trajectory $\boldsymbol{x}(t)$ is to cut the space with a plane into portions called left and right. Whenever the trajectory crosses this plane, going from left to right, for example, we mark the place. If a particle moves for a long time, its path will typically pierce the surface of section quite a few times, given a certain amount of judgment in choosing this *Poincaré surface*. If you put the plane where the orbits are plentiful, you will see something like what you see on cutting open a coaxial cable.

In learning to use the Poincaré section, it is well to practice on simple examples. For a periodic orbit, the surface of section may consist of just one point because the particle keeps coming back to the same place. Indeed, so long as the set of marked points is finite, it must always be the trace of a periodic orbit. Other simple situations may arise. For example, suppose that the orbit is wrapped on the surface of a torus. The points in the surface of section will then lie on a closed curve. Orbits with such simple spoors are often called regular.

When the series of points in the surface of section does not repeat, go off to infinity, nor remain on a simple closed curve, we must have a case of highly complicated motion. The points often do not fill the surface of section nor any piece of it, so they may be said to form an object of fractional dimension, as we shall explain presently. The motion in such cases may be called chaotic. Since the motion is dictated by $\boldsymbol{v}$, we may consider that $\boldsymbol{v}$ itself is the chaotic object. The notion that chaos is basically the property of a vector field is of importance in many disciplines. For astrophysics, a telling example is the solar magnetic field.

In a certain sense, what we see of the solar magnetic field is a surface of section. Where the field is strong, it inhibits the convective motions. Near the solar surface, such motions are responsible for the outward flux of heat. Hence, magnetic inhibition of this flow makes for relatively dark regions called sunspots. So the places where strong ordered fields protrude from the solar surface are distinguished by being darker than their surroundings.

To understand this as a surface of section, think of any snapshot of the solar magnetic field $\mathbf{B}(\boldsymbol{x}, t)$ as a flow as in Eq. 1, with a fictitious particle moving along some trajectory. The parameter that tells us how far along in a trajectory of the field the particle has gone is, of course, not the real time, for that is fixed. Rather, we introduce a fictitious time, s, increasing along the trajectory of a particle in the flow $\mathbf{B}(\boldsymbol{x}, t)$. Then the position of the particle, for fixed t, traces out a path governed by the equation

$$\frac{dx}{ds} = \mathbf{B}\,(x(s), t). \tag{2}$$

Such trajectories of $\mathbf{B}$, for fixed t, are called the *streamlines* of $\mathbf{B}$.

We can make a surface of section for the streamlines of $\mathbf{B}$ and get some idea of its topology in this way. And there is no reason we should not use a spherical surface to make our surface of section. Then we can hope to get something that looks like the solar surface for a reasonable choice of **B**. The regular parts of the field correspond to flux tubes that would appear as sunspots, while the chaotic portions resemble the general field over the solar surface.

Another way to think theoretically in terms of a surface of section is to use a return map. That is, instead of following a trajectory round and round waiting for it to cross the Poincaré surface, we can derive or invent a rule that tells us where the system will next cross the surface, given the location of its previous crossing. Such a rule saves us the effort of having to solve a differential equation like Eq. 2 to understand the structure of the field.

Although we do not know how to make a solar return map (as yet), we can anticipate the real thing by simply looking at typical return maps that have been constructed in the study of Hamiltonian chaos. (We specify Hamiltonian chaos because magnetic fields are divergence-free, so we need to use the so-called area-preserving maps of Hamiltonian dynamics.) Such maps generally produce distributions of magnetic field that look qualitatively quite right. From such maps, we typically get intense concentrations into spots of very regular field surrounded by a background chaotic field. The spots in this theoretical image correspond to flux tubes of the basic vector field, at least in this imagery *(29)*. In Fig. 1 we show a section of a simple area-preserving return map invented by M. Hénon

In studying such maps, without going into complicated physical processes, we can see the possible qualitative structure of solar fields. That is, chaos theory, with no appeal to any physical processes, predicts that spots ought to exist for any generic magnetic field other than the highly special cases that are constructed for classrooms and certain carefully designed devices. Without this realization, we might have been tempted to give special credence to models that predict a field whose surface of section is spotty. But since this aspect of vector fields is generic, we realize that any sensible theory ought to predict sunspots. The theory of chaotic vector fields suggests, moreover, that such surfaces of section ought to have spots of smaller and smaller sizes. What we need to look for are the quantitative aspects of such fields as a test of our theories, as indeed some people have been doing *(30)*.

Similar general conclusions may change our picture of other cosmic arrangements such as the spatial distribution of galaxies. It is now believed that the luminous portions of galaxies contain only a very small fraction of the total mass in the universe. Therefore, a galaxy is like the test particle whose motion is described in Eq. 1. The flow, $\mathbf{v}$, in that equation is presumably determined by the invisible mass of the universe, but the way in which this happens may well surprise us if we can unravel it. Naturally, to be a bit more precise, we should write the four-dimensional version (at least) of Eq. 1 for galaxies. Then what we would be observing is a surface of section made with a space-like hypersurface, which for present purposes is a surface of (nearly) constant time.

Observational limitations lead us to think differently about the construction of such surfaces of section in cosmology than in more tractable examples. Instead of a galaxy going round and round on a single trajectory, we adopt an approach like that used to visualize flows by fluid dynamicists. We start out with a uniform distribution of particles and let them all run according to Eq. 1. Then we mark the points where they cross a particular space-like hypersurface. That is, we take a photograph at some moment. If the cosmic flow is generic, we naturally would not be surprised to see a fractal distribution of the test particles or galaxies. This is, indeed, what the observations indicate. Again, we appreciate that a theory has to do more than just predict this outcome if it is to command our respect.

Cosmic Cascades

When a swarm of points is swept along by a flow, we shall see the points rearrange themselves as we look at successive time-like cuts of

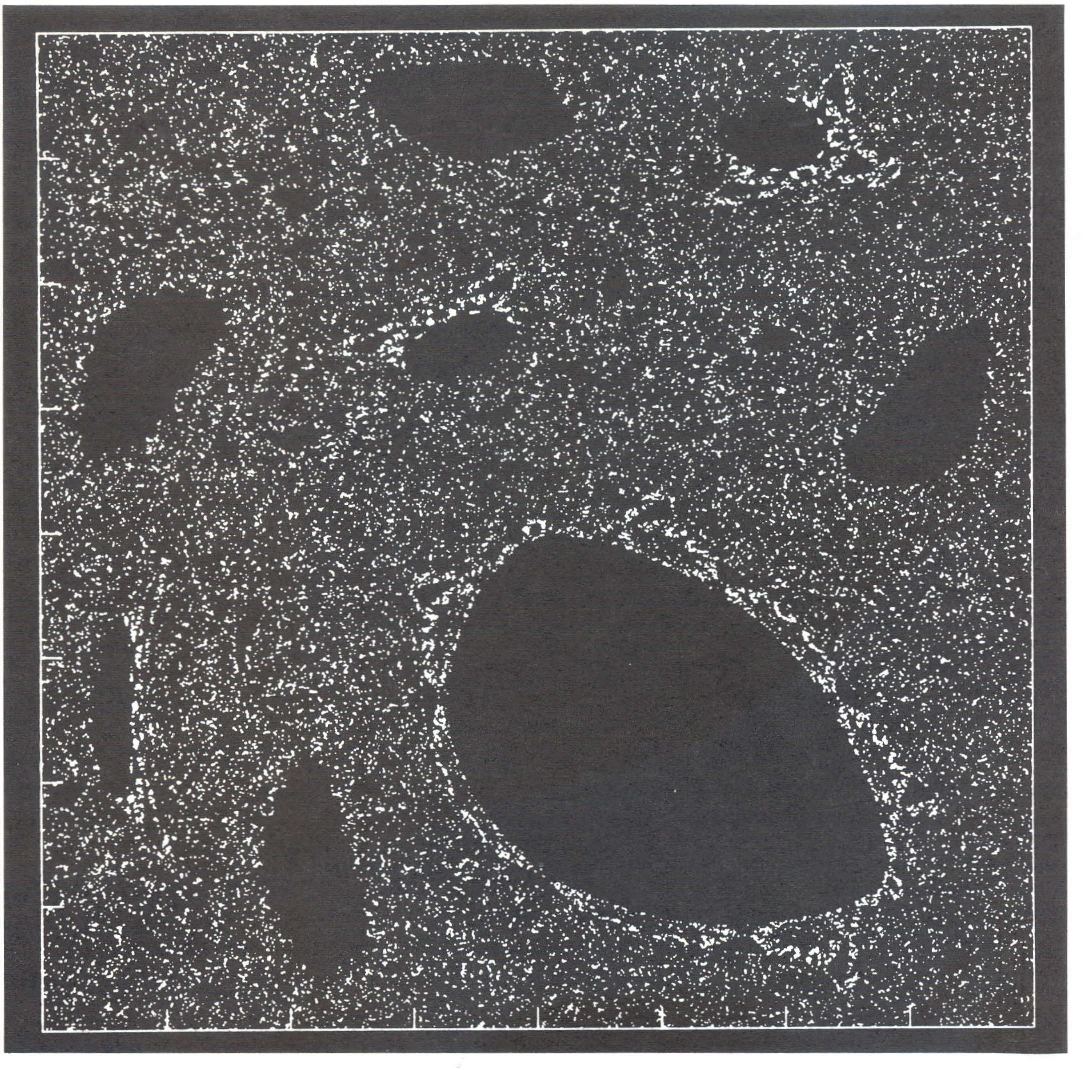


Fig. 1. A chaotic trajectory from the area preserving *Hénon map* of chaos theory [e.g., *(29)*]. The dark areas are the integrable or regular regions of the field, which we liken to flux tubes.

their world lines. If the flow has chaotic parts, we can see the emergence of a hierarchical structure from an initially rather uniform configuration. This begins with a slight redistribution of points leading to the first traces of clustering on a particular scale. In the simplest cases, these first clusters occur either on the largest or the smallest scale allowed by physical or geometrical constraints of the real world.

In a flow where the first inhomogeneities appear on the largest scales, we first see large clusters appear. Within these, the distribution of points is initially rather uniform. These first clusters of points will behave like the initial set, and clusters will form within them. In the simplest possible cases, a self-similar hierarchy of clusters will emerge from this geometrical cascade. On the other hand, if the initial clusters are formed at the smallest scale allowed by the physics, we may see them coming together to form larger structures that cluster together in their turn. The continuation of this latter process is called an inverse cascade. If the initial scale is determined by some forcing extrinsic to the process of interest, we may get cascading in both senses at once.

This way of thinking about how a flow moves points about is often much more convenient than trying to deal with the flow directly. Mandelbrot *(17)* considers the way that Cantor first made his hierarchical sets as a form of cascade. Students of turbulence have used such thinking for years, beginning with the work of Novikov and Stewart *(20)*. Unno *(31)* sees cascades throughout natural history, referring to Hoyle's *(11)* discussions of the hierarchical formation of cosmic bodies in the astronomical case. Although the mathematical rationalization for this way of thought is still in its infancy, Kida *(12)* has used a model equation for the cascade of fluid turbulence related to the so-called Kolmogorov-Ford equation of statistics. The same kind of modeling may be helpful in the cosmological cascade as well.

Consider a great cosmic ball of matter subject to the Hubble expansion of the universe. As cosmologists often do, let us think of the motion of the galaxies in comoving coordinates, in which the uniform Hubble expansion has been subtracted out of the flow. With the appearance of the first large-scale inhomogeneity in the cosmic fluid, we get our first generation of objects (clumps, aggregates, groupings, or what you will). These cosmic lumps are objects with characteristic scale l_1. In the ball there are N_1 such entities.

In the following stage, the N_1 objects create the next generation. Suppose that the first generation spawns a total of N_2 second generation objects of size l_2. If we let this process run its course through a great many generations, we shall have a basis for writing an equation for N_n, the number of objects in the n^{th} generation. Let the creation of objects of the n^{th} generation take, on average, a time τ_n. Then, we can write a so-called master equation of this form:

$$\frac{dN_n}{dt} = -\frac{N_n}{\tau_n} + \zeta_n \frac{N_{n-1}}{\tau_{n-1}}, \tag{3}$$

where ζ_n is the mean number of offspring produced by the objects in the $(n-1)^{th}$ generation.

The meaning of the first term on the right in this equation is that the members of the n^{th} generation are destroyed with a characteristic lifetime τ_n. If that were the only term, we would find that $N_n \propto \exp(-t/\tau_n)$. But the dying members of each generation may be used to make the next generation. That is the meaning of the second term on the right of Eq. 3. To use this equation, we have to provide appropriate formulae for ζ_n and for τ_n. In the simplest cases,

$$\zeta_n = \left(\frac{l_{n-1}}{l_n}\right)^3. \tag{4}$$

The exponent 3 in this formula indicates that the objects are embedded in a three-dimensional space. Kida sought greater generality by replacing the 3 in Eq. 4 with a parameter, s. It appears that he wanted the freedom to choose s, not an integer, in modeling some features of turbulence. As we see next, the outcome of Sreenivasan's measurements on turbulent flows suggest that 3 is a good choice (see the appendix to this volume for references).

Suppose that we have statistically steady turbulence. Then, if we apply the steady state version of Eq. 3, we have

$$\frac{N_n}{\tau_n} = \zeta_n \frac{N_{n-1}}{\tau_{n-1}}. \tag{5}$$

For the example of homogenous turbulence, Kolmogorov *(13)* has provided a similarity theory that gives us $\tau_n \propto l_n^{2/3}$. With this result, we can find a *particular* solution of Eq. 5 with $N_n \propto l_n^{-7/3}$.

To extract the meaning of this result, we observe that the total mass M_n in the n^{th} generation is (to within a constant factor) $\rho_n l_n^3 N_n$, where ρ_n is the mass density of the typical object in its generation. In laboratory turbulence, the motion is effectively incompressible, so ρ_n is a constant. If the mass involved in the cascade were conserved, M_n would be constant and independent of n. This would require that $N_n \propto l_n^{-3}$. In that case, the objects in any generation would fill a finite volume. Instead, we have $l_n^3 N_n \propto l_n^{2/3}$, so the volume occupied by each generation is tending to zero with decreasing l_n. This means that less and less mass is participating in the cascade as we go down through the generations. There must be *lacunae* where the cascade has been arrested.

We could even improve on the exponent of 7/3 in the solution. Kolmogorov's formula for the average lifetime of an eddy is implicitly based on the assumption that the eddies are space-filling. Now that we have the first estimate of the departure from this condition, we can correct the expression for τ_n and get an improved estimate to replace the 7/3. However, that estimate is close enough for our present purposes, so we shall not pause to do this here.

The fact that the total volume occupied by the smallest of the objects goes to zero has an interesting interpretation. If, in ordinary space, a set of points fills a piece of that space and so has a finite volume, we would say in ordinary language that this object is three-dimensional. When the volume of the object vanishes, we are tempted to quantify this by assigning a dimension to the object that is lower than 3. We shall be going further into this in "Dimension," pages 206–208. For now, let us call the exponent in the power law for N_n the dimension of the set of objects. As we shall see, there are many ways to define the dimension of a set of points, so this is really just one particular dimension; perhaps we could call it the cluster dimension. It resembles a dimension discussed by Unno *(31)*.

In the case of turbulence, the dimension of the set of points on which the dissipation is appreciable in the experiments is quite close to 7/3 as Sreenivasan has found. This encourages us to think that Eq. 5 may provide some useful insights into other cascades. We are therefore tempted to try it out on the cosmic cascade. But before doing that, let us remark that there is a more general solution to Eq. 5 in a geometrical cascade, when the ratio of successive sizes is a constant, that is, ζ_n does not depend on n. In that case, we may reasonably suppose that

$$\tau_n = const.\, l_n^{\kappa}, \qquad (6)$$

where the parameter κ is characteristic for the particular cascade. Then we may readily verify the existence of the following solution to Eq. 5:

$$N_n = (l_n/\mathrm{L})^{-d}\Psi\,[\log(l_n/\mathrm{L})], \qquad (7)$$

where

$$d = 3-\kappa, \qquad (8)$$

and

$$\Psi(x) = \Psi(x + \log\beta), \qquad (9)$$

where $\beta = \zeta_n^{1/3}$ and L is a unit of length.

The function Ψ measures the *lacunarity* of fractals. It generalizes Mandelbrot's definition of the term as a constant prefactor in the statistical description of fractals. Instead, we may call Ψ a prefunction. So far, it remains an arbitrary function, found only by observing real fractals. The ripple described by Ψ can also be found in inhomogenous fractals or multi-fractals as some call them. This ripple has long been noticed both in mathematical *(18)* and physical *(23)* theories involving hierarchies and cascades. In connection with his early model of the turbulent cascade, Novikov *(19)* pointed out that the ripple would occur in turbulence in principle, but would not be detectable. However, it may have been seen in turbulent flows *(28)*. It will take many more data than we have now, or a new way to analyze them, to see the ripple in statistical analysis of the distribution of galaxies.

To formulate a cascade model for a cosmic process like the creation of the hierarchical structure of the galaxy distribution *(32)*, we need to look harder at Eq. 3. First, we have to depart from the usual imagery of fluid turbulence, in which eddies of a given size, l_n, are normally considered to be effaced after a typical time, τ_n, when they give rise to smaller eddies. In the usual situation of fluid turbulence, we need an energy source to obtain a statistically steady state. In the cosmic example, we need to allow for the possibility that the eddies persist for much longer than this decay time. That is, they may have produced the next generation without being destroyed themselves. If the memory of the early generations can linger on in the cosmic structures, we can expect to observe the coexistence of objects on all scales. We want to suggest a way of doing this that seems suited to some versions of cosmic clustering theory.

If only the first term on the right of Eq. 3 existed, with no feeding from previous generations, we would find that N_n decays exponentially. And for turbulence, this makes good sense. The linear parts of the turbulent processes with constant coefficients do give an exponential behavior like this. For any process where such exponential behavior is expected, Eq. 3 is likely to be helpful. A measurement of d will give a value of κ, which will give clues to the basic mechanism causing the cascade. But the universe is expanding, hence it is time dependent. Therefore, describing the cosmic cascade calls for a modification of this feature of Eq. 3.

Some theoretical pictures of the cosmic cascade suggest that the decay of individual structures in the Hubble flow may be algebraic rather than exponential. For example, a density anomaly in a self-gravitating medium grows as a power of t, as in Lifshitz's well-known work *(21)*. Then we ought to try a master equation, which in the absence of cascading, gives $N_n \propto t^{-\alpha}$. Such behavior comes from an equation like

$$\tau_n \frac{dN_n}{dt} = -N_n^{\mu},$$

where $\mu = (\alpha + 1)/\alpha$ and τ_n is the time constant for the n^{th} generation. This version mollifies the sudden-death, exponential aspect of Eq. 3. When we also include a cascade term, we get this equation:

$$\tau_n \frac{dN_n}{dt} = -N_n^{\mu} + \zeta_n N_{n-1}^{\mu}, \qquad (10)$$

if we assume that the exponent in the decay rate is the same for each generation.

Suppose that galaxies were formed in a burst of activity some billions of years ago. They are then approximately frozen into the Hubble flow. Though clustering is still going on, we shall be cavalier and look at the equivalent of Eq. 5 for this modified situation. That is, we look at the steady-state version of Eq. 10, $N_n^{\mu} = \zeta_n N_{n-1}^{\mu}$. In the case of a geometrical cascade, we get the solution Eq. 7 again, with the condition Eq. 9. But this time, for the dimension derived from the size distribution, we get $d = 3\alpha/(\alpha + 1)$ instead of Eq. 8.

According as β is greater or less than unity, we have a cascade or an inverse cascade. That is, either big objects spawn smaller ones, or small ones gather together to make bigger ones. Both possibilities have been considered for galaxy formation, but the appearance of β in the size distribution $N_n(l_n)$ may help settle this matter by observational means. Unfortunately, the ripple in the sizes implied by the periodic nature of Ψ is hard to detect without abundant and accurate data, and that is why it is so little discussed. But de Vaucouleurs *(32)* has suggested that a size ripple indeed exists in the clustering of galaxies. His interpretation of the observed clustering of galaxies is interesting, and it suggests a value of β in excess of unity. If that is correct, it implies that a direct cascade is responsible for the cosmic clustering. This naturally needs confirmation, but already we see how observations together with the new developments in mathematical thinking may lead us to some significant astrophysical conclusions, if we can solve the calibration problems. In perceiving evidence of a size ripple in the galaxy distribution, de Vaucouleurs has, in effect, descried an indication of the existence of the cosmic voids *(14)*.

Even more can be learned from comparing Eq. 7 to observations. If the hierarchical arrangement of galaxies is gravitational in origin, then we can say something about the possible values of α. Those values are model dependent and may be different according as the universe is closed or open. In particular,

$$\alpha = \frac{d}{3 - d}. \qquad (11)$$

Hence the clustering dimension, d, will give us rather direct information on global cosmic properties.

Although this is the simplest possible kind of theory, it serves our purpose well, for it brings out nicely how parameters of the distribution of galaxies in space actually can emerge explicitly from theoretical calculations. Next, we want to explain how one goes about getting such parameters, especially the dimension. We begin that task by reviewing some background on statistical physics.

Statistical Tools and Methods

We have spoken of sets of points that can be observed in astronomy such as the points in which the solar magnetic field pierces the solar surface or the way galaxies lie in space at fixed time. We need ways to characterize such sets of points, both to describe the flow fields that lie behind their arrangements and to test theories that we hope will bear on such questions. If we know too little about such details, we can fall back on a less detailed kind of discussion, as in "Cosmic Cascades," pages 200–204, and use some statistical property of the sets such as the distribution of cluster sizes. In simple cases, a statistical distribution of points may be characterized by one or two parameters, including the dimension. In fact, there are many statistical quantities that have been used to extract dimensions from cosmic fractals. In this chapter, we introduce some of the main statistical descriptors of fractal sets that may be distilled down into the dimensions of the sets, along with one or two other parameters such as the period of Ψ.

To describe some of the statistical methods that may have use in cosmic applications, we shall suppose that the basic notion of a space of dimension D is already intuitively plain when D is an integer. We shall consider a set of points such as idealized galaxies, embedded in such a space of integral dimension D. When the set is generated numerically or observed, the total number of points in the set, J, is finite; ideally, J is very large. For convience, we can label the points by an index, $j = 1, 2, \ldots, J$. The data of the problem are the coordinates, $\mathbf{r}_j$, in the space to which the J points belong. We shall assume that the embedding space is Euclidean and the coordinates are Cartesian.

When dealing with a set of points such as one pictures for the atoms in a gas or liquid, one of the first quantities that one tries to determine is the density. This may at first seem a simple thing to work with, but it becomes elusive in lacunar sets such as we are now discussing. In these cases we have a problem underlined in cosmology by de Vaucouleurs *(32)* almost twenty years ago: "What precisely do we mean by average density? What is the evidence to support the notion that a mean density can be defined?" In short, we shall have to watch out. We shall simply use physical intuition to avoid the pitfalls of this topic, for we cannot here enter into the mathematics called *measure theory* that is appropriate to the general discussion. [For an intermediate approach, see *(16)*.]

Thus forewarned, we draw a spherical surface of radius λ around any point in the set. This spherical surface is called S^{D-1}, in our space of integral dimension D. (This space is called R^D.) Following current usage, we shall refer to the interior of the sphere as a ball. It is useful to make this distinction between the spherical surface and its interior, though it is not always done in ordinary scientific discussion when the meaning is clear from the context. We let V (λ) be the volume of the ball of radius λ whose surface is S^{D-1}.

Now, to define the number density function $n^{(1)}(\mathbf{r})$ we count the number of points in the ball centered on the point at $\mathbf{r}$. Designate that number N (λ). For a continuous medium, we would define the density $n^{(1)}(\mathbf{r})$ as the small-λ limit of N (λ)/V(λ). The awkwardness in this is like the one we face in defining the density of a real fluid composed of atoms. In that case, as here, we dare not take the limit of λ going all the way to zero, when N(λ) is not large. For the fractal sets that we are dealing with in the cosmological case, the inhomogeneity on all scales is so great that $n^{(1)}(\mathbf{r})$ is an even more awkward quantity to deal with. Because of the voids in the galaxy distribution *(14)*, this definition is not only sensitive to λ, the radius of the ball; $n^{(1)}$ also depends on $\mathbf{r}$, even for a statistically homogenous system. This is what led Pietronero *(22)* to suggest the introduction of other statistical quantities into the description of the galaxy distribution.

Just as $n^{(1)}(\mathbf{r})$ gives the density of points at $\mathbf{r}$, there is a function $n^{(2)}(\mathbf{r}'', \mathbf{r}')$, which gives the density of *pairs* of points separated by Δ =

$\mathbf{r}' - \mathbf{r}''$. To get $n^{(2)}(\mathbf{r}, \mathbf{r} + \Delta)$, suppose that there is already a point in the set at $\mathbf{r}$. Then, $n^{(2)}(\mathbf{r}, \mathbf{r} + \Delta)$ is the number of points whose displacement vector from $\mathbf{r}$ lies inside a sphere of (small) radius δ, centered on $\mathbf{r} + \Delta$, divided by the volume within the sphere of radius δ, and multiplied by the density at $\mathbf{r}$.

As in statistical physics, we define the *pair distribution function* $\tilde{g}$ through the relation

$$n^{(2)}(\mathbf{r}', \mathbf{r}'') = n^{(1)}(\mathbf{r}')\, n^{(1)}(\mathbf{r}'')\, \tilde{g}(\mathbf{r}', \mathbf{r}'' - \mathbf{r}'). \quad (12)$$

If no point in the set is special, we may assume that the pair distribution function has this form:

$$\tilde{g}(\mathbf{r}', \mathbf{r}'' - \mathbf{r}') = g(\mathbf{r}'' - \mathbf{r}'). \quad (13)$$

The density at $\mathbf{r}''$, given that a point is already at $\mathbf{r}'$, is therefore $n^{(1)}(\mathbf{r}'')g(\mathbf{r}'' - \mathbf{r}')$.

We have already said that the idea that no points are special is risky for complicated distributions of points. A point at the edge of a large void may be different than other points in some ways. But now, we have pushed our use of the ideas of homogeneity back one level to the relations between pairs of points, and we shall see where this gets us. It is reasonable to adopt the convenience of choosing a particular origin, at $\mathbf{r}'$, for example. Then the conditional density becomes $n^{(1)}(\mathbf{r}'')\, g(\mathbf{r}'')$ and Eq. 12 becomes

$$n^{(1)}(\mathbf{r}'')\, g(\mathbf{r}'') = \frac{n^{(2)}(\mathbf{0}, \mathbf{r}'')}{n^{(1)}(\mathbf{0})}. \quad (14)$$

In the same spirit, let us consider the sets in which there are, on average, no preferred directions. We may then assume that the dependencies in this expression do not involve angles. Hence there is a function $\Gamma(r)$ such that

$$n^{(1)}(\mathbf{r})\, g(\mathbf{r}) = \Gamma(r), \quad (15)$$

where r is the magnitude of $\mathbf{r}$. The function Γ is equal to that introduced by Pietronero *(22)*, albeit defined slightly differently. Γ may be used to define a dimension of a set of points, as we shall see in "Dimension," page 206. But other, more popular, statistical moments have been widely used.

The correlation is a very frequently studied object in all sorts of statistical investigations. This is a standard quantity defined, for example, by Peebles *(21)* for cosmological purposes. In the study of fractals, the related correlation integral $C(r)$ is used by Grassberger and Procaccia *(8)*, defined as the

$$\lim_{J \to \infty} J^{-2} \cdot [\text{number of pairs of points } (i,j)],$$

such that the separations between the members of each pair, $|\mathbf{r}_i - \mathbf{r}_j|$, is less than r. They call C an integral because it is the integral over the ball of radius r of the correlation itself. We mention this as an intuitive way to think about the correlation, whose usual definition may seem less vivid. In statistical physics, this definition is

$$\xi(r) = \frac{n^{(2)}(\mathbf{r}', \mathbf{r}'')}{n^{(1)}(\mathbf{r}')\, n^{(1)}(\mathbf{r}'')} - 1, \quad (16)$$

where $r = |\mathbf{r}' - \mathbf{r}''|$. But this brings out the relation to Γ, whose use in the study of fractals has been urged by Pietronero.

Finally, we conclude this small selection of statistical tools for analyzing sets of points by mentioning the ways that cosmologists have confronted the problem of limited data samples in calculating statistical quantities like ξ. A good example is provided by a scheme described in Peebles' book *(21)* on cosmology; this scheme is used in many discussions of galaxy distributions. A brief discussion may help in assessing the results under discussion.

To illustrate the use of this amelioration scheme, we describe it for the pair correlation determination. The first step, of course, is to derive the distances between the J^2 pairs of points in the set. Let $N_r^{(2)}$ be the number of such pairs with separation in the interval $[r - dr, r + dr]$. Points too close to the edge of the set may be excluded *(6)*. Next generate a random set with the same boundaries and the same number of points as in the data set. If we combine this set with the data, we can calculate the distance between a randomly selected set of J^2 pairs in the combined set. In doing this, we always choose the first point from the data and the second from the random set. Let $M_r^{(2)}$ be the number of such pairs, with separation in the interval $[r - dr, r + dr]$. We then calculate

$$\Xi(r) = \frac{N_r^{(2)}}{M_r^{(2)}} - 1 . \tag{17}$$

This quantity is the ameliorated pair correlation function.

We may use similar amelioration methods to determine a pair distribution function. A fiducial random background set may be used as described. We are then able to calculate

$$\Gamma(r) = n_0 \, [1 + \Xi(r)], \tag{18}$$

where n_0 is the integral of Γ over the ball of radius r.

Another way to reduce the problem caused by finite data sets has been to use coarse-graining. For example, some workers use a coarse-grained Γ defined by *(6)*

$$\Gamma^*(r) = V^{-1} \int_V \Gamma(r')\, dr', \tag{19}$$

where V is the volume within a sphere of radius r. But problems remain, whether we use $\Gamma(r)$ or $\Gamma^*(r)$; the finite sets that arise from observations present practical difficulties.

Perhaps, methods like these have been used because the occurrence of significant fluctuations in plots of various statistical moments vs. r were thought to be mainly a result of insufficient data samples. The amelioration methods tend to wipe out such fluctuations. Although small samples may contain spurious fluctuations, *real fluctuations* are also inherent to the statistical moments of fractal sets. As we saw in "Cosmic Cascades," the ripples in such moments contain valuable information. Hence, we should stress that in describing these amelioration methods we are not advocating them. Indeed, we hope that soon there may be sufficient cosmological data to permit the dropping of amelioration. If we sometimes will use it here, it is to facilitate comparison with current work in cosmology.

Dimension

In the early stages of development of a subject, one prefers to paint with a broad brush. So it is useful to try to characterize the statistical distributions of points by a few parameters such as their dimensions. There is more than one kind of dimension that can be associated with a set of points, but for the simplest sets, the various dimensions that are conventionally defined are normally not very different in value. The evidence is that the cosmic fractals are not too far from this simplest state, so we need not go into the finer points of dimension determination. Nevertheless, it is best to begin with the most direct method, though it may not be the easiest one to carry out.

To describe some of the dimensions that may have use in cosmic applications, we continue to suppose that our set is embedded in a space of integral dimension D. We shall reserve the symbol D to denote the dimension of the set of points under study such as the points in a surface of section.

An operational definition of the dimension of the set of J points, assuming the J is as large as is required, is provided by the following simple procedure. We recall that $N(\lambda)$ is the number of points contained in S^{D-1}, the surface of the sphere in D dimensions, and introduce the quantity $\langle N \rangle(\lambda)$, which is the average of N over the points in the set. If we measure this quantity over a good range of values of λ, we can make a reasonable determination of the dependence of $\langle N \rangle$ on λ. For these results, we make a plot of log $\langle N \rangle$ vs. log λ. We shall call the slope of the line that best fits this plot (allowing for the intrinsic ripples) the fractal dimension of the set.

This prescription makes the notion of fractal dimension operational, but it may not provide a very convenient way to determine the dimension. First of all, there are pitfalls when we deal with the limitations of real data, as detailed in the various books on the subject. In particular, the plot is limited in the available scales of λ, above by the size of the system and below by the limit of resolution. Furthermore, the best fit is normally not a line, but a line with superposed wiggles, as we can anticipate from the discussion of "Cosmic Cascades," pages 200–204. Inevitably, other definitions of dimension have been proposed, and these are based on asymptotic properties of various statistical moments of sets of points.

After all, the quantity $\langle N \rangle$ is a statistical moment on the set, and there is no reason why dimensions based on any other moments, such as those described in "Statistical Tools and Methods," pages 204–206, could not be defined. Renyi actually did just that and Halsey *et al.* *(10)* have described the usefulness of some of these dimensions and their interrelations. In recent years, much attention has been devoted to the distinction among these generalized dimensions for inhomogeneous fractal sets, or multi-fractals.

We may also base the definition of dimension on the conditional density function, Γ. This quantity has the units of number density. So the quantity $\int_{V(r)} \Gamma(r) d^{D}r$ is analogous to the number of points in the ball. For small r, we may expect it to grow like r to a power; that power could be used to define a dimension. However, an advantage in using a moment like Γ to define a dimension is that we avoid doing the integral. For power laws, the integration simply raises the exponent by D. So it is natural to define a dimension based upon Γ in this way:

$$D_{\Gamma} = \mathrm{D} - \lim_{r \to 0} \frac{\log \Gamma(r)}{\log r}. \tag{20}$$

In evaluating D_{Γ}, we have to make allowances for the lack of resolution, owing to the finiteness of the data sample. In practical terms, we should think of Γ as having an expansion like

$$\log \Gamma = -(\mathrm{D} - D) \log r + \cdots \tag{21}$$

for small r. The dots stand for higher terms in the expansion. We expect the first of such higher terms to describe the ripple associated with hierarchical processes *(27, 28)* or fluctuations associated with the finite data sample. For self-similar structures we can readily see why, in the following manner.

Self-similarity of a structure implies that statistical moments are scale invariant. If $C(r)$ is any such moment, we take self-similarity to mean that there are constants A and B such that

$$C(r) = AC(Br). \tag{22}$$

This functional relation has the solution

$$C(r) = \Psi(\log r)\, r^{-D_C}, \tag{23}$$

where $D_C = \log A / \log B$ and Ψ, like its counterpart in Eq. 7, is an arbitrary function satisfying $\Psi(\log r) = \Psi(\log r + \log B)$. In this case, there is only one higher term in the expansion Eq. 21, and that is the ripple we have already discussed *(28)*. However, for real data sets, there will be, in effect, corrections because of sampling errors, system boundaries, or systematic inhomogenities, and these may obscure the lacunarity of the set. In smoothing the data, we tend to obliterate such sources of error, but we eliminate the lacunarity ripple as well.

We may also define dimensions based on the various correlations, and the pair correlation is the one most commonly employed. It may appear from Eq. 18 that a dimension determined from the slope of a $\log \Xi$-$\log r$ plot would not be different from D_{Γ}, in the limit of small r. However, we find that, in practice, the power law dependence of Ξ may differ significantly from that of Γ. To make the comparison, we expand the correlation function for small r in terms of $\log r$. For the ameliorated version, this gives and expression like

$$\log \Xi = -\gamma \log r + \cdots, \tag{24}$$

where γ is a constant that is characteristic of the set. We then define the correlation dimension as

$$D_{\Xi} = \mathrm{D} - \gamma. \tag{25}$$

This is the quantity ν defined by Grassberger and Procaccia *(8)*.

We shall be comparing the values of these two dimensions for the cosmic clustering. Of course, it is always uncomfortable to compare just two methods when there is a chance that they may disagree. To help in deciding between them, we first try them on sets for which we know the fractal dimensions on theoretical grounds. And, in the belt and suspenders mode, we shall also use a third functional on fractal sets described by Badii and Politi *(1)*. Following them, we select a subset from the data. From each point in this subset, we find the distances P to K of the remaining points. Let the minimum among these distances be called P_{min}. Badii and Politi show that the quantity D_P in

$$P_{min}(K) = const. \times K^{-\frac{1}{D_P}} \tag{26}$$

gives a good approximation to D. By varying K, one can therefore get an estimate of D.

We now turn to the determination of these three dimensions in a few specific cases. We find that these are more straightforward to obtain than d of "Cosmic Cascades," Eq. 8. But, in the cosmic context, we shall mention some possible values for d also.

Theoretical Examples

To illustrate and compare the dimensions, we find them for two theoretical fractal sets in this section. In each case, the fractal dimension is known on theoretical grounds, and this is useful in assessing the methods. We shall, however, not rederive the theoretical results for the fractal dimensions since they are explained in books *(2)*.

(i) We start with an interval of length unity and mark the end points and the midpoint. Then we make two copies of this interval and scale them to a new length, $\sigma < 1$. We attach these two copies onto the original interval by attaching the midpoint of each to one of the endpoints of the original unit interval. Next we pivot each secondary with respect to the primary interval to some selected value θ. Then we make four scaled copies of the second generation, each with length σ^2. We attach these by their midpoints to the four free endpoints and set the angles at the contact points all equal to θ. We continue this to the n^{th} generation, where n is very large. The 2^{n-1} sticks of length σ^{n-1} will have 2^n free endpoints. These points are the model fractal. [A realization in a three dimensional space is described in *(27)*.]

The only differences between the present construction *à* la Barnsley *(2)* and that of Groth *et al.* *(9)* are that the latter chose the angles at the attachment at random and worked in three dimensions. In Fig. 2, we show such a set for (constant) $\theta = 22.5°$ and $\sigma = 0.7$. The density of points in the set is qualitatively indicated as if the points were individually bright. A bright region in the picture is very dense and a dark one is void. Two qualitative features of this set emerge on inspection: the presence of voids (or lacunae) and the filamentary distribution of the highest density of points. In these respects, the set resembles the observed distribution of galaxies in producing both voids and dense filaments. Such features may be caused mechanistically as in some theories, but they are a natural feature of a hierarchical universe *(32)*.

Coming back to this particular fractal set, we recall that, on theoretical grounds, its fractal dimension is

$$D = \ln 2 / \ln \frac{1}{\sigma}, \tag{27}$$

provided that the set is totally disconnected *(2)*. Qualitatively, we may say that the set is disconnected when the different parts produced by the generating algorithm do not overlap. It is not hard to see that Eq. 27 follows on applying the *ansatz* Eq. 22 to $N(r)$, on which the definition of fractal dimension is based. When the parts of the fractal set begin to overlap, we expect Eq. 22 to fail, and so too will Eq. 27. The formal definition of a connected fractal, based on such an idea of overlap, can be found in Barnsley's book *(2)*. A mathematical discussion of these points is given in *(4)*.

Notice that although D does not explicitly depend on θ, the largest σ for which Eq. 27 holds does depend on θ. It is not hard to work out that for $\theta = 0, \sigma_{max} = 1/2$ and for $\theta = \frac{\pi}{2}, \sigma_{max} = \frac{1}{\sqrt{2}}$. We illustrate the transition from a disconnected to a connected region in Fig. 3, where we fix $\theta = 0.40$ and vary σ. Clearly, as σ increases, we go from unconnected to totally connected cases.

When we are within the totally disconnected region of a fractal, and for some connected cases as well, then different values of θ will correspond to different lacunarities, as indicated by the appearance of voids in these sets. We can see this in Fig. 4, where we draw just a portion of the relevant (x, y) plane and use a specific $\sigma = 0.693$ with two different values of θ. We see that sets with the same dimension give quite distinct maps for the two

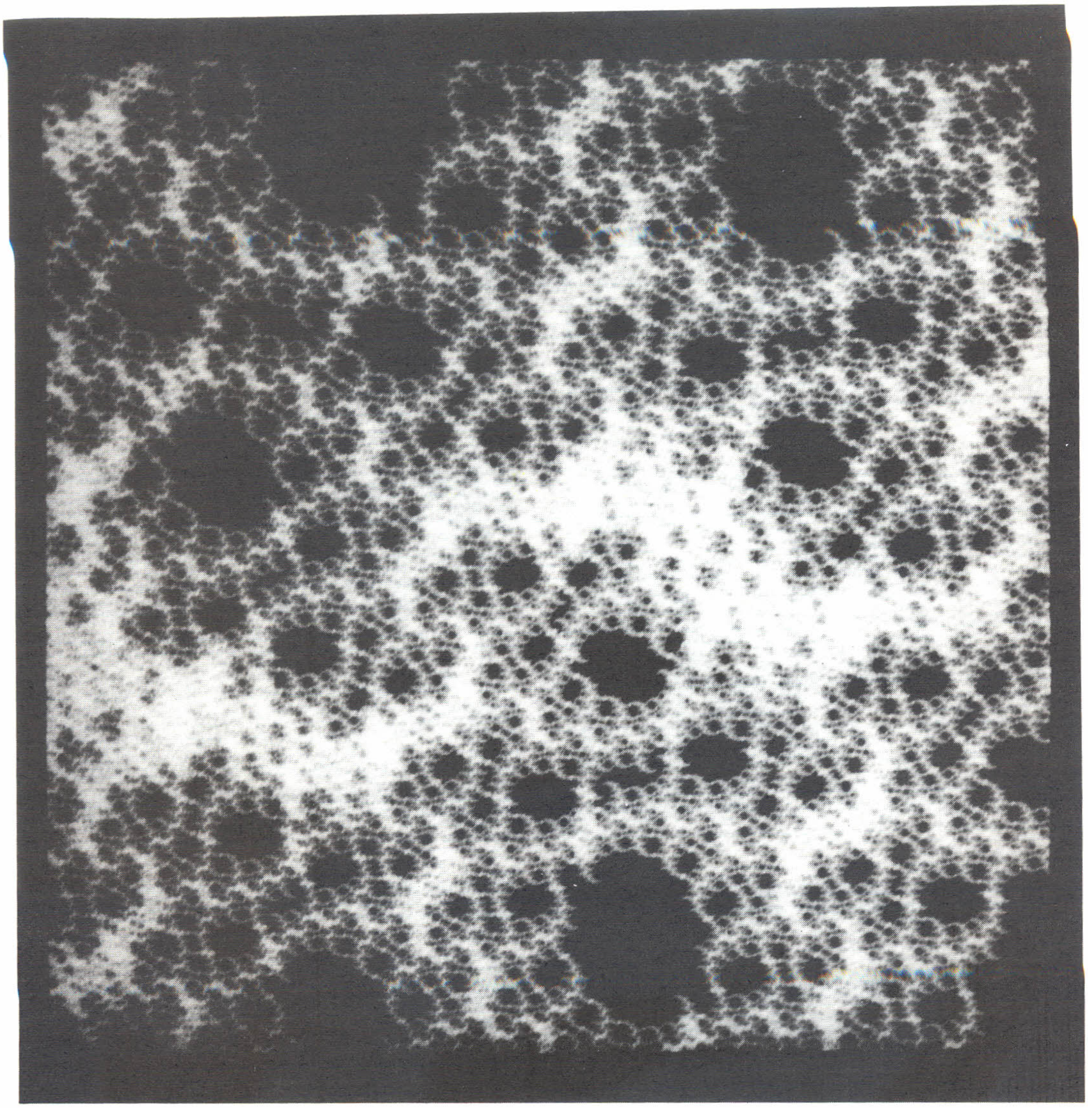


Fig. 2. The set of points generated by the map described in the text for $\theta = 22.5°$. The black spots are voids where the map generates no points unless one is introduced initially, in which case, a regular orbit arises.

different lacunarities.

We give the results of dimension calculations in this case for $\sigma = 0.636$, and $\theta = 0.40$ rad, which have been used in Fig. 3(1). As can be seen in the figure, this case is totally disconnected and Eq. 27 gives $D = 1.532$. We took 7990 points and obtained these values:

$$D_{\Xi} = 1.51 \pm 0.01$$

$$D_{\Gamma} = 1.56 \pm 0.01$$

$$D_{P} = 1.59 \pm 0.03$$

(ii) Another well-studied set for which we know the fractal dimension is an equilateral triangular Sierpinski gasket. To construct it, we pick a point inside the triangle. Then we select one vertex at random and mark the point midway between the chosen point and the selected vertex. This midpoint becomes the new chosen point. Again, we select a vertex at random and find the point halfway between it and the second interior point. This process is continued, and the points generated

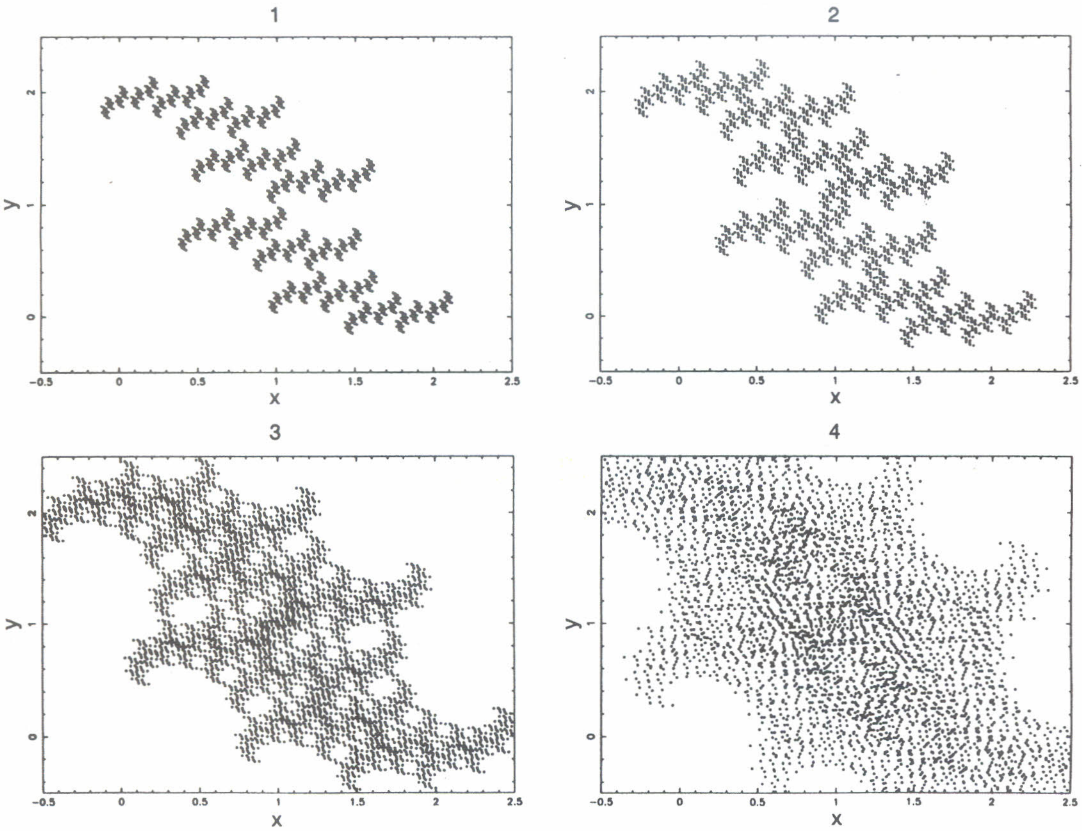


Fig. 3. Transition from a disconnected to a connected fractal set for the set described in the text with $\theta = 0.4$ rad and for (1) $\sigma = 0.636$, (2) $\sigma = 0.665$, (3) $\sigma = 0.700$, (4) $\sigma = 0.750$.

in this way form the set shown in Fig. 5A, where we have discarded the first dozen points. The theoretical dimension of this object is $D = \ln 3/\ln 2$, close to 1.58 *(17)*.

The three dimensions for 7500 points are:

$$D_\Xi = 1.37 \pm 0.02$$

$$D_\Gamma = 1.62 \pm 0.02$$

$$D_P = 1.60 \pm 0.03$$

We see that, in this case, D_Ξ is noticeably different from D_Γ, which is close to the known fractal dimension. This is consistent with our experience in such tests: D_Γ is typically close to D for self-similar sets, with D_P almost as close, but D_Ξ is sometimes well off D. Next, we turn to the case of the galaxy distribution, where these effects are more pronounced. But in passing, let us mention that the Sierpinski construction need not be restricted to triangles.

The distribution of Fig. 5B is a generalization of the Sierpinski gasket to the case where it is generated by the vertices of a regular pentagon. As in the triangular case, the image of a given point is the point halfway between a randomly chosen one of the five vertices. The mapping is continued, and a fractal set is generated (actually, it is a so-called fat fractal). This object shows ripples in its statistics.

Galaxies on the Sky

For many years, the positions of the galaxies on the celestial sphere have been measured by astronomers. This has never been an easy task. Because of our own position within a spiral galaxy, the distribution of galaxies on the celestial sphere is given an apparent large-scale nonuniformity. But if we restrict our at-

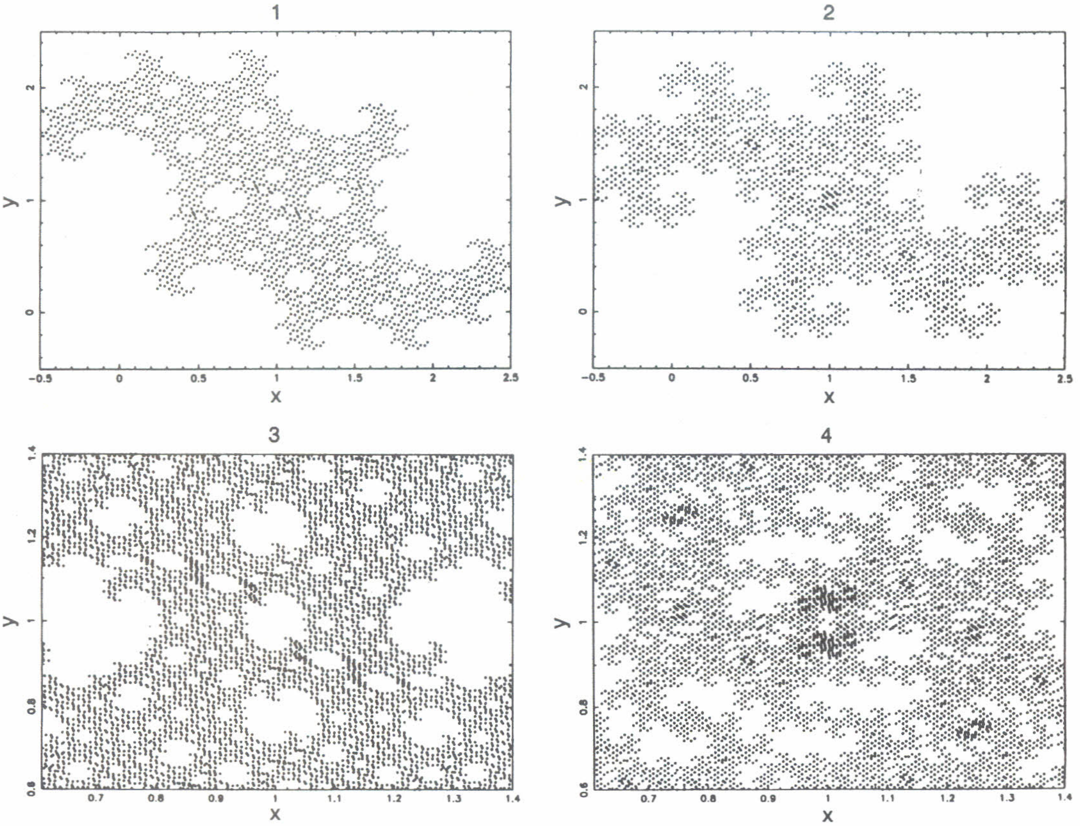
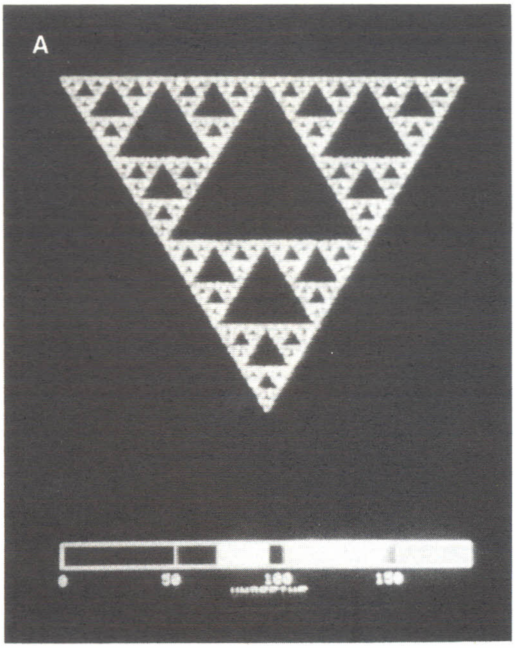


Fig. 4. The map for $\sigma = 0.693$ with two different lacunarities: (1) $\theta = 0.523$, with the whole range shown; (2) $\theta = 0.785$ with the whole range shown; (3) $\theta = 0.523$, with restricted range; (4) $\theta = 0.785$ with restricted range.

tention to selected regions of the sky, information bearing on the true current distribution of these markers in the cosmic flow may be derived. The story is well described in the books by Peebles who, with various collaborators, has been studying the statistics of the measured galaxy coordinates on the celestial sphere. These coordinates are simply two angles on S^2, in the language we have been using. So when Peebles and co-workers derive correlations of the second and higher order, the separation variable is no longer linear distance, but an angle on the sky called ϑ. To keep this in mind, we use the cosmologist's designation of ω for the pair correlation. (However, we do not use a special notation for Γ when using angular coordinates.)

Determinations of the plot of log ω vs. log ϑ have revealed a nonintegral slope. This has led Mandelbrot *(16)* to suggest that the galaxies, regarded as points, form a fractal set. This observation nicely focuses our thinking about hierarchical universes and the cascade that lies behind the galaxy distribution. The implied task is to derive the basic properties of the cosmic fractal. Fortunately, Peebles has long since begun this task, so we can follow his lead. Here we undertake the limited goal of describing how to get dimensions for cosmic fractals. In doing this, we treat the coordinates of the galaxies on the sky as given columns of numbers and, with regret, omit the story of how difficult they are to obtain and how much they have improved in recent decades.

There are several sets of measurements that are relevant. To illustrate the determination of the dimensions of the cosmic fractal, we use a set of coordinates compiled by a group at the Nice Observatory led by A. Bijaoui *(25)*, who kindly provided us with their



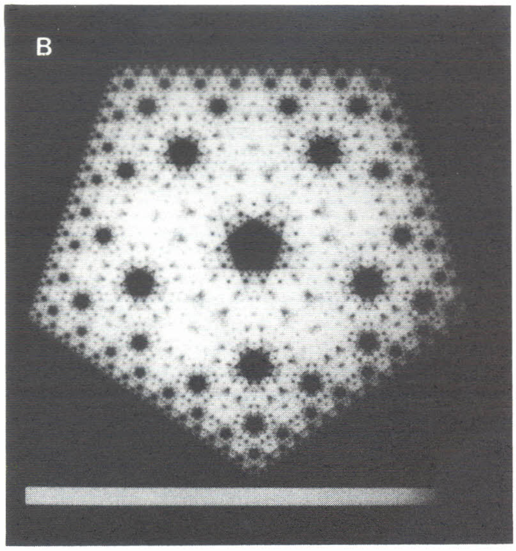


Fig. 5. (A) The Sierepinski triangle with 6000 points. (B) A pentagonal analogue of (A).

data in a convenient form. Although this set has only about 7500 galaxies from a single Palomar plate as yet, the high quality of the measurements and the prospect of further data of this kind make it worthwhile to expend some effort on them.

The most extensive statistical studies use the pair correlation $\omega(\vartheta)$. It was early known that the log-log plots of this function showed a straight line portion for small ϑ and a break to a less simple behavior. The break is known as the "knee" in this subject. Figure 6 is a plot of log ω vs. log ϑ for the Nice data showing the knee. The slope of the straight line portion of the plot is called γ and from these data we find $\gamma = 0.88 \pm 0.03$. For the fit, the values of $\omega(\vartheta)$, between $\vartheta_1 = 0.0001$ rad and $\vartheta_f = 0.006$ rad were used. For larger ϑ we obtain the usual picture of a breakdown of the fit due to a more rapid decrease in $\omega(\vartheta)$ *(9)*.

The problem with the knee is that its location changes with the size of the sample. This has been recently studied with data from a new position catalogue *(7)*. We show some values of ϑ_f, the position of the knee, for various data samples in Table 1; the first two values are from the quoted work and the third comes from Fig. 6. Not only do the break points or knees change with the sample size, the slopes in the linear portion also differ noticeably. Thus the measured dimension is $D_\Xi = 2 - \gamma$, which clearly depends on the sample size.

It was this uncomfortable dependence of ϑ_f on sample size that led Pietronero to suggest that Γ, the conditional density function, might provide a more robust statistical object to characterize galaxy positions. While the work performed by him and his collaborators seems to bear this out, the close connection between Γ and the pair correlation in Eq. 17 seemed to suggest that the dimensions D_Γ and D_Ξ would not differ by much. Indeed, Pietronero called the slope of the correlation function and "intrinsic quantity," and assumed that it is not much affected by sample size. On the other hand, it has been suspected *(32)* that galaxy "clusters occasionally overlap," so, given the lesson of "Theoretical Examples," pages 208–210, we may expect that this is not so.

The dimension that is normally accepted for galaxies on the sky is obtained from the correlations measured by Peebles and collaborators *(9)* from data in the Shane-Wirtanen catalogue. Following the suggestion of Mandelbrot, from their value of γ, one gets a value

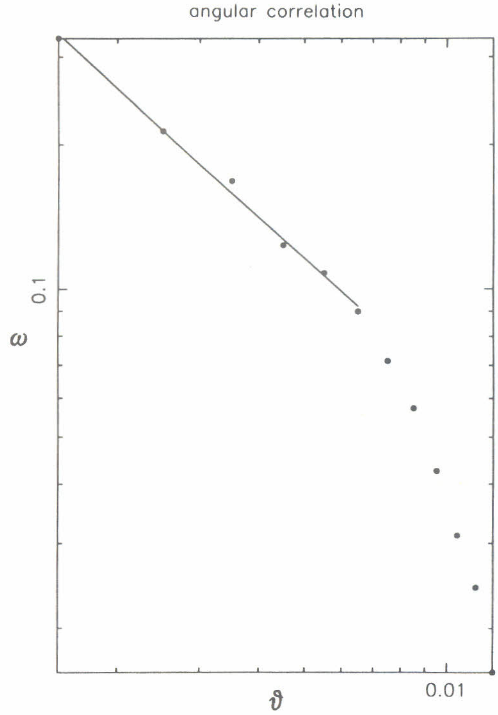


Fig. 6. Angular correlation ω, using the catalogue of *(26)*.

of about 1.2. The value we get from the Nice data is

$$D_{\Xi} = 1.12 \pm 0.03. \quad (28)$$

(We might have called this D_{Ω}, but the notation would have become too cumbersome.)

In Fig. 7 we show a plot of $\Gamma(\vartheta)$ vs. ϑ. From these results, we get $D_{\Gamma} = 1.88 \pm 0.01$. The values of $\Gamma(\vartheta)$ used for this fit were from the domain from $\vartheta_i = 0.0001$ rad to $\vartheta_f = 0.015$ rad. The wider range of usable ϑ supports Pietronero's *(22)* proposal to use Γ rather than Ξ. However, we should mention that changing ϑ_f, or using $\Gamma^*(\vartheta)$ instead of $\Gamma(\vartheta)$, results in small differences that are larger than the statistical error. Therefore, we prefer a more conservative error estimate and write

$$D_{\Gamma} + 1.88 \pm 0.03. \quad (29)$$

We may mention that, for this data set, the calculations done without amelioration give essentially the same dimensions. Finally, we add that, for the Nice data, we also get

$$D_{P} = 1.89. \quad (30)$$

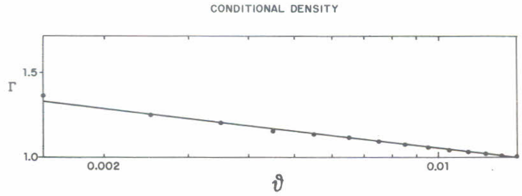


Fig. 7. Conditional density, Γ, using the catalogue of *(16)*.

The dimension of the set in Fig. 2 also has this value.

We reach the evident conclusion that D_{Γ} and D_{Ξ} are disparate. Table 1 indicates that D_{Ξ} increases with increasing sample size. It is not guaranteed that D_{Γ} will not change when the sample size is increased, but the studies done so far suggest that such a change will not be substantial unless the region of the sky used is unusual in some way. Completion of the Nice study will provide the data needed to be firm on this point. But the relation between ω (or ξ) and Γ make it unlikely that the discrepancy could be due to physical causes like the inhomogeneity of the cosmic fractal.

Wiedenmann and Atmanspacher *(32)*, whose recent study points to a discrepancy such as we have found, attempt to reconcile the two dimension values by associating them to two different domains of ϑ. This view is reminiscent of the interpretation that has been placed on observations of the solar granulation *(24)*. In any case, in the appropriate domain of ϑ, they get $D_{\Gamma} = 1.89$, consistent with our results, with data from the Zwicky catalog.

Another recent cosmic dimension determination, for the Virgo cluster of galaxies *(5)*, leads to the high value of 1.98 ± 0.23. Of course, this is a special region of the sky and only 200 galaxies were used. However, the clustering method employed gives a dimen-

Table 1. Values of ϑ_f, the position of the knee, for selected data samples.

γ	ϑ_f	No. of plates	Calculation
0.63	0.030	24	*(7)*
0.71	0.015	6	*(7)*
0.88	0.006	1	Present work

sion in the spirit of d in "Cosmic Cascades," pages 200–204.

A question that must be probed more deeply before we can make too much of all this, is that of the actual dimensions of the galaxy distribution in three-dimensional space. We need to remember that, in most of the statistical studies of the cosmic fractal, the embedding dimension D is 2. To determine what the effect of the projection onto the sky is on the true dimension of the galactic set is a difficult matter, since galaxies have a wide range of intrinsic luminosities. The correction will be positive, but we do not know its magnitude. We can only surmise that the true dimension in three-dimensional space is greater than the value measured on the sky. (We ignore cosmological corrections that are likely to be quite small for existing data.)

Some hints about the three-dimensional effects may be contained in the study of de Vaucouleurs *(32)* who made reasonable assumptions about the three-dimensional aspects of the objects under study in assessing the data. His plot of the log of the mean density of a ball of radius r vs. log r, has the slope -1.7. This implies a dimension $3-1.7 = 1.3$. Since de Vaucouleurs used a single particle density, we infer a dimension compatible with D_Ξ, conceivably marginally in excess of it because of the differences in D. On the same page, he also gave a plot of what corresponds to a plot of log N_n vs. log l_n, in the notation of "Cosmic Cascades." The negative slope of that curve is slightly in excess of 2—around 2.1. So there is also a result in reasonable agreement with D_Γ, slightly in excess of our 1.9. (There is even a third graph on the page suggesting a lacunarity ripple.)

Obviously, without considerable further measurement and discussion, no definite conclusions will be possible. But we do have enough information to propose that the difference between D_Ξ and D_Γ is significant and is a result of the cosmic fractal being "connected" *(2)* (de Vaucouleurs speaks of "interlocking" clusters). This effect is no doubt enhanced by projection effects. It would be valuable if we could ultimately discern how much of the observed cluster overlap is due to projection and how much is of the nature of the connectedness we encountered in the study of theoretical fractals. Observations are helping to unravel such effects *(14)*.

The structure of inhomogeneous fractal sets can be probed by studying differences among the generalized (or Renyi) dimensions *(10)*. It may be possible in some cases to remove significant differences among dimensions by retreating to higher embedding dimensions, D. Thus, one might conjecture that inhomogeneous sets may be understood as homogeneous sets projected from some higher dimensional space. The extent to which this is possible seems at present unknown, but we may suspect that similar possibilities for interpreting the difference between D_Γ and D_Ξ exist for the cosmic fractal. That is, the difference is in a sense physical and could result in large measure from projection effects. In that case, comparing D_Ξ and D_Γ may be of some use in ongoing attempts to deduce statistical information on three-dimensional distribution of galaxies. For now, however, we do suggest that the fractal dimension associated to the clustering of galaxies on the sky is rather closer to 1.8 than to 1.2 and that it is the former value that ought to be used in evaluating theoretical predictions such as those of "Cosmic Cascades."

Conclusion

In its modern incarnation, chaos theory began with the study of erratic, which is to say, aperiodic and unpredictable behavior. We now know a lot about how this sort of thing can come out of equations of various kinds. In many fields, including astrophysics, practitioners are trying to decide how to use the new understanding of these equations to unravel some of their own problems. The aim is to take observations of complex temporal variability and to try to understand them in terms of equations that are known to produce aperiodicity. That part of the subject is covered by other contributors to this volume and in many standard works on chaos. In brief, the idea is to describe the variability of a system by letting it move along a trajectory in a space whose coordinates characterize the state of the system. In certain cases, where the data cover a number of cycles of the system adequately, it has been possible to extract such trajectories by ingenious methods. Then, one can hope to get some idea of the kind of mathematical model that might generate a trajectory that resembles the one reconstructed from the observations. At the very least, one may hope to discover how many state variables are required to permit the prediction of the future of the system in principle.

When the number of variables, call it F, needed for complete predictability turns out to be finite, we have a deterministic system. In fact, for a given F, the prediction is possible for only a limited time in practice, and that time decreases with increasing F, tending to be effectively zero at some upper, critical value of F. The critical value, F_c, will depend on existing computational means. When it is quite large, predictability becomes a hopeless prospect, as people have realized for centuries.

Once F comes near to that critical value, and sometimes even sooner, we give up trying to predict in detail, use statistical methods instead, and call the system stochastic. In this uncomfortable situation we study systems with conveniently small F, replacing the missing variables by something called noise. This process has worked well in a number of problems, such as continuum physics. The transition between the two situations is discussed in a number of standard works and by others in this symposium.

When the effective F is quite small, we can make detailed studies and develop a deep understanding of the trajectories of the system, even to the point of grasping the topology of the orbits. In these circumstances, the issue of predictability in practice can now be probed. In the ideal situation, we can come to something as clear as the flow of Eq. 1. But sometimes, without making us pass through this rite of unraveling an erratic time signal, nature presents us with flows or trajectories that are revealing and fundamental. In the two cases we have discussed most here, we are led directly to some kinds of Poincaré maps or surfaces of section. Of course there may be other examples that are so simple that we hardly think in these terms. But when there is a wealth of structure, our interest perks up. Even then, we might not think in terms of Eq. 1 in a case like the galaxy distribution were it not for the sensitive awareness engendered by the excitement over chaos theory.

We have written much more about the galaxy distribution than about the other more evident example, the sunspot topology, because we know much less about the galaxy distribution. Naturally, in that case, it is easier to be expansive. de Vaucouleurs *(32)* has cautioned us with this quote from Otto Struve: "...the observer knows too many facts to be satisfied with any theory." Cosmic data are, however, beginning to accumulate, and our outlook will surely be changing rapidly.

As we have seen, it is not at all obvious what to do with the time slice of the cosmic flow that we do have at our disposal. We have attempted to describe some of the available tools and the issues involved in their use. We have concluded, with some surprise, that even the currently accepted value of the fractal dimension of the distribution of galaxies has to be reexamined; indeed, we think it has to be seriously revised. But the truth is, that this was not the reason for our delving into the statistics of galaxy distributions.

We set out in this study to look for the lacunarity ripple in the statistics that we have repeatedly mentioned. It is not surprising that we have not detected it with the low-order statistics of quantities like Γ or ξ. It is quite clear that to find the scaling factor β we need to go to higher moments of the galaxy distribution *(27)*. To do that we need more data. There can be no doubt that they are forthcoming, and we can then hope to proceed along the lines hinted at in "Cosmic Cascades."

In these circumstances, it is not appropriate for us to start discussing the theoretical consequences of lower order statistics, until some more is known of the higher terms in the developments. In "Cosmic Cascades," we merely gave a sample of the most primitive of possible theories, with no mechanistic aspects, just to give an idea of how the lore of the fractal may begin to help us understand the flow that propels our part of the universe. We have not, and could not, give any indication of the diversity of opinion on the theoretical side. We did, however, emphasize that many of the results that theories may claim to produce are generic properties of any sensible models and not to be fussed over unduly. We hope that no one will overlook this message or take it amiss.

But what of the sun, which is reaching a maximum of activity as we finish this report? We have introduced the notion *(29)* that when you cut open any vector field and see what looks like the innards of a coaxial cable, you are looking at the surface of section of the equivalent flow. This vision leads to the understanding that spots and a fractal structure are normal for most magnetic fields. In this case, the observations are abundant, and the sort of issue that hampers the galaxian discussion ought not to pose a real problem. The real issue is how to think about such problems theoretically. Normally, one tries to attack this by leading the magnetic field through the full dance of the equations of magnetofluid-dynamics. But there may be simpler steps to take as hinted in "Cosmic Cascades."

In oceanography and planetary sciences, the subject of vortex merger *(15)* has intrigued all who encountered it. It is not understood in detail (and it has been detailed only in two dimensions), but it does seem to be an inverse cascade process *par excellence*. Can we make a Kolmogorov-Ford-Kida type of model for this process? Can we do something similar for the spots on the sun that are mathematical analogues of vortices? These are the issues that we think are of interest in understanding the sunspot morphology. They loom large now that our minds have been opened to these prospects by thinking about the implications of chaos for astrophysics. Indeed, the main impact of that subject is the tremendous increase in our mental arsenal for formulating questions that previously seemed too complex to express. That perhaps is the main message of this volume.

Acknowledgments

In preparing this chapter we have had help of diverse kinds from many people to whom we express our gratitude. These include Jim Anderson, Albert Bijaoui, Jacqueline van Gorkom, Jim Peebles, Itamar Procaccia, Oded Regev, Ed Shaya, Charles Tresser and Bill Young. We are especially grateful to Wang Quingda for his kind collaboration in the use of AIPS for making figures.

Support for this work came from the National Science Foundation under grant PHY87-04250 the Air Force under grant AFOSR89-0012 and the NASA Cooperative Agreement NCC5-29. We thank the Geophysical Fluid Dynamics Program at Woods Hole Oceanographic Institution for its hospitality in the summer of 1989.

References

1. Badii, R., and A. Politii, *Phys. Rev. Lett.* **52,** 1661 (1984).
2. Barnsley, M., *Fractals Everywhere* (Academic Press, San Diego, CA, 1988).
3. Barrow, J. D., S. P. Bhavsar, D. H. Sonoda, *M.N.R.A.S.* **210,** 19p. (1984).
4. Bedford, T. J., *J. Lon. Math. Soc.* **2,** 89 (1986).
5. Chikawa, M., W. Unno and M. Yuasa, *Sci. Technol.* **1,** 35 (1989).
6. Coleman, P. H., L. Pietronero, R. H. Sanders, *Astron. Astrophys.* **200,** L32 (1988).
7. Collins, C. A., N. H. Heydon-Dumbleton, H. T. MacGillivray, *M.N.R.A.S.* **236,** 7p. (1989).
8. Grassberger, P., and I. Procaccia, *Phys. Rev. Lett.* **50,** 346 (1983).
9. Groth, E. J., P. J. E. Peebles, M. Seldner, R. M. Soneir, *Sci. Am.* **237,** 76 (1977).
10. Halsey, T. C., M. H. Jensen, L. P. Kadanoff, I. Procaccia, B.I. Shraiman, *Phys. Rev.* **A33,** 1141 (1986).
11. Hoyle, F., *Ap. J.* **118,** 513 (1953).
12. Kida, S., *Prog. Theor. Phys.* **67,** 1630 (1982).
13. Kolmogorov, A. N., *C. R.(Doklady) Acad. Sci. U.S.S.R.* **30,** 301 (1941).

14. de Lapparent, V., M. J. Geller, J. P. Huchra, *Ap. J.* **302,** L1 (1986).
15. McWilliams, J. M., *J. Fluid Mech.* **146,** 21 (1984).
16. Mandelbrot, B. B., *Fractals: Form, Chance and Dimension,* (W. H. Freeman and Co., New York, 1977).
17. Mandelbrot, B. B., *The Fractal Geometry of Nature* (W.H. Freeman and Co., New York, 1982).
18. Mulhaupt, A. P., *Phys. Lett.* **122A,** 403 (1987); Mulhaupt, A. P., *Phys. Lett.* **204A,** 151 (1987).
19. Novikov, E. A., *Sov. Phys. (Doklady)* **1,** 497 (1966).
20. Novikov, E. A., and R.W. Stewart, *Isv. Akad. Nauk. S.S.S.R. Seria Geofiz.* **3,** 408 (1964).
21. Peebles, P. J. E., *The Large Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980).
22. Pietronero, L., *Physica* **144A,** 257 (1987).
23. Pietronero, L., and E. Tosatti, Eds., *Fractals in Physics* (North-Holland, Amsterdam, 1986).
24. Roudier, T., and R. Muller, *Solar Phys.* **107,** 11 (1986).
25. Slezsak, E., G. Mars, A. Bijaoui, C. Balkowski, P. Fontanelli, *Astron. Astrophys. Suppl. Ser.* **74,** 83 (1988).
26. Smith, L. A., *Phys. Lett.* **133A,** 283 (1988).
27. Smith, L. A., Thesis, Columbia University (1987).
28. Smith, L. A., J.-D. Fournier, E.A. Spiegel, *Phys. Lett.* **114A,** 465 (1986).
29. Smith, L. A., and E.A. Spiegel, *Chaotic Phenomena in Astrophysics*, J. R. Buchler and H. Eichhorn, Eds., Ann. N.Y. Acad. Sci. **97,** 61 (1987).
30. Tarbell, T., S. Ferguson, A. Title, G. Scharmer, P. Brandt, private communication.
31. Unno, W., *Sci. Technol.* **1,** 41 (1989).
32. De Vaucouleurs, G., *Science* **167,** 1203 (1970).
33. Wiedenmann, G., and H. Atmanspacher, Preprint (Max Planck Institute, Garching, FRG, 1989).