Finite Size Scaling in Ecology

Jayanth R. Banavar,[1] Jessica L. Green,[2] John Harte,[3] and Amos Maritan[4]

[1]*Department of Physics and Center for Materials Physics, The Pennsylvania State University, University Park, Pennsylvania 16802*
[2]*Department of Nuclear Engineering, University of California, Berkeley, California 94720*
[3]*Energy and Resources Group, University of California, Berkeley, California 94720*
[4]*International School for Advanced Studies (S.I.S.S.A), Via Beirut 2-4, 34014 Trieste, Italy*
and INFM and the Abdus Salam International Center for Theoretical Physics, Trieste, Italy

Finite size scaling techniques are applied to ecology and are shown to predict relationships between power-law exponents as well as a connection between scaling phenomena in physics and ecology. The framework presented here ought to be useful for testing the hypothesis of scale invariance in the spatial distribution of species across a landscape.

PACS numbers: 87.23.Cc

The postulate of self-similarity, which is well understood in the context of critical phenomena in statistical physics [1], naturally leads to power-law singularities of generic quantities. Recently, empirical power-law phenomena in ecology have been shown to follow from the simple assumption of self-similarity or scale invariance in the spatial distribution of species across a landscape. In particular, Harte *et al.* [2] derived the species-area relationship [3,4]

$$S \sim A^z, \tag{1}$$

where S is the number of species sampled in a patch of area A, and the exponent z is typically around 0.25. This relationship is both widely observed in nature and frequently utilized in conservation biology [3,4]. Using the same self-similarity assumption, Harte *et al.* [2] also derived the species-abundance distribution $P_i(n)$, which defines the probability that any given species on a biome A_i has n individuals. Numerical analysis of $P_i(n)$ derived from the Harte model showed an algebraic dependence,

$$P_i(n) \sim n^c, \tag{2}$$

for n values sufficiently below the modal abundance with a dropoff for larger values of n [2].

The powerful machinery developed for critical phenomena may be applied to ecological data for purposes of testing the hypothesis of scale invariance, predicting the relationship between exponents, and deducing universal behavior. The fluctuations of physical quantities diverge at the critical point only in the thermodynamic limit, and finite size scaling [5] is a useful analysis technique for systems that are not infinite in spatial extent. Recently, this technique has been used to suggest a similarity between magnetic systems (finite size systems) and turbulent flow (finite Reynolds number) [6]. In the context of ecology, the total number of individuals of all species in an area A_i, N_i is finite and $P_i(n)$ ought to exhibit finite size scaling,

$$P_i(n) = \frac{1}{n} f\left(\frac{n}{N_i^\phi}\right), \tag{3}$$

where ϕ is a crossover scaling exponent. The scaling function f is universal (the same within a universality class)

and crosses over from a power law of the argument with an exponent equal to $c + 1$ to a rapidly decreasing function as its argument increases through a value around 1. ϕ is a nontrivial exponent, whose value is not necessarily equal to 1. Indeed, universality classes in ecology could be identified by first assessing whether finite size scaling holds by collapsing distinct abundance distributions measured in different areas and with different total numbers of individuals on to one universal scaling curve (an example is shown in Fig. 1) and then compiling the list of associated exponents. In addition, there ought to be well-defined relationships between the scaling exponents.

We now turn to an illustration of these ideas in the context of the Harte model [2]. The model postulates that the fraction of species in an area A that is also found in
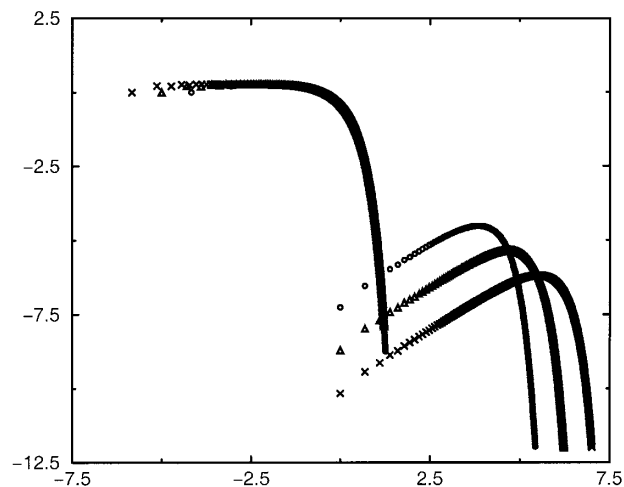

FIG. 1. Scaling collapse of species-abundance distributions. The three curves on the right are log-log plots of $P_0(n)$ vs n for the self-similar model of Harte *et al.* [2] for a z value around 0.4, and they correspond to the total number of individuals being equal to 1024, 4096, and 16384. The curve on the left shows a scaling collapse log-log plot of all the data in the three curves. The x coordinates are $\theta = (n/N^\phi)$ and the y coordinates are $[nP_0(n)/\theta^{c+1}]$ with the ϕ and c exponents predicted from theory.

    0031-9007/99/83(20)/4212(3)$15.00

one-half of that area is independent of A. Specifically, if a species is known to be in area A, the probability that, under bisection, it will be found in at least a specific one of the two resulting subareas is denoted by a. It was shown by Harte *et al.* [2] that this self-similarity postulate leads to an algebraic species-area relationship with the exponent

$$z = -\frac{\log(a)}{\log(2)}. \tag{4}$$

To obtain the distribution of the abundance of individuals within species, Harte *et al.* derived a recursion relation for $P_i(n)$, the probability that if a species is found in a patch of area A_i, then it contains n individuals. The notion of a smallest patch size, $A_m = A_0/2^m$, where A_0 is the total area, was introduced [2] so that $P_m(1) = 1$; there is, on average, just one individual of whatever species is present in a given unit patch of this smallest size. (Of course, not all species are found with equal probability in this smallest size patch.) Furthermore, the mean total number of individuals $N_0 = 2^m$, $P_i(n) = 0$ for $n \geq 2^{m-i} + 1$ and $\sum_n P_i(n) = 1$. One then obtains [2] the following recursion relation:

$$P_i(n) = xP_{i+1}(n) + (1 - x)\sum_{k=1}^{n-1} P_{i+1}(n - k)P_{i+1}(k), \tag{5}$$

where

$$x = 2(1 - a). \tag{6}$$

We begin by noting that the total number of individuals must be equal to the product of the number of species and the average number of individuals per species. Thus $N_0 \sim A_0$ must equal the product of A_0^z (the number of species according to the species-area relationship) and $\langle n \rangle$, whose scale is set by N_0^ϕ, leading to the equation

$$\phi + z = 1, \tag{7}$$

a relationship that is valid even when the Harte model is generalized to one in which a given census patch area is partitioned into b subareas (b is equal to 2 in the original model).

We now turn to an analysis of the recursion relation [Eq. (5)]. We define a new variable $Q_i(n) \equiv P_{m-i}(n)$ and the associated generating function $R_i(\lambda) = \sum_{n=1}^{\infty} Q_i(n)e^{-n\lambda}$, with $\lambda \geq 0$, to recast Eq. (5) in the form

$$R_{i+1}(\lambda) = xR_i(\lambda) + (1 - x)[R_i(\lambda)]^2, \tag{8}$$

where $R_i(0) = 1$ (in order to satisfy the normalization condition), whereas the condition in the smallest patch of $Q_0(n) = \delta_{n,1}$ becomes $R_0(\lambda) = e^{-\lambda}$. The fixed points of the above renormalization group equation for $R_i(\lambda)$ in the physical interval between zero and 1 are zero, which is an attractive fixed point and 1, which is repulsive. For a given value of λ, one has a specific initial condition, and one obtains a renormalization group trajectory flowing towards the attractive fixed point.

The scaling form for $P_0(n)$ postulated in Eq. (3) is equivalent to

$$Q_i(n) = \frac{1}{n} f\left(\frac{n}{2^{i\phi}}\right), \tag{9}$$

which leads to a scaling form for

$$R_i(\lambda) = h(\lambda 2^{i\phi}). \tag{10}$$

The function h is related to the function f by the relation

$$h(y) = \int_0^\infty \frac{dx}{x} e^{-xy} f(x) \tag{11}$$

with the scaling form for $R_i(\lambda)$ holding in the limit of λ tending to zero, with i much greater than 1, so that $\lambda 2^{i\phi}$ is maintained at a fixed value. Using the scaling form for $R_i(\lambda)$, its recursion relation [Eq. (8)] may be recast in the form

$$h(2^\phi \epsilon) = xh(\epsilon) + (1 - x)[h(\epsilon)]^2, \tag{12}$$

where $\epsilon = \lambda 2^{i\phi}$.

We will proceed now to deduce the consequences of Eq. (12) on the behavior of h for small and large arguments. When $\epsilon = 0$, one obtains

$$h(0) = xh(0) + (1 - x)[h(0)]^2, \tag{13}$$

so that $h(0)$ equals zero or 1. The value consistent with $R(\lambda = 0) = 1$ is the latter. Assuming that $h(\epsilon)$ is analytic in the vicinity of $\epsilon = 0$ and using Eq. (12), one immediately finds that

$$2^\phi = 2 - x. \tag{14}$$

Using Eqs. (4) and (6), Eq. (14) gives back the scaling relationship of Eq. (7). When ϵ tends to infinity in Eq. (12), h reaches a fixed point, this time equal to zero. Thus for very large ϵ, $h(\epsilon)$ tends to zero and Eq. (12) simplifies to its linearized form

$$h(2^\phi \epsilon) \sim xh(\epsilon). \tag{15}$$

This implies

$$h(\epsilon) \sim \epsilon^{-(1+c)}, \tag{16}$$

with

$$x = 2^{-\phi(1+c)}. \tag{17}$$

This asymptotic behavior of the function h leads to the function $f(x)$, for small argument, behaving as

$$f(x) \sim x^{(1+c)}. \tag{18}$$

For the Harte model [2] (with only two partitions), there is only one independent exponent and $c + 1 = -[\log(2 - g)/\log(g)]$, where $g = 2^\phi = 2^{1-z}$. The analytic predictions are nicely born out by numerical calculations (Fig. 1). For a generalized Harte model with more than two partitions, the exponent c depends not only on z but also on the probability that a species is found in just one of the partitions of the census patch area.

Finally, we relate our scaling results to those presented in Bramwell *et al.* [6]. These authors demonstrated the scaling collapse of magnetic systems and turbulent flow by plotting $\sigma_i P_i(x)$ vs (x/σ_i), where $P_i(x)$ is a probability density function characterizing either system at scale i and σ_i as its standard deviation. To see the connection of their finding to our Eq. (3), we first note that the following recursion relation describes the variance of the solution to Eq. (5):

$$\sigma_i^2 = (2 - x)\sigma_{i+1}^2 + (2 - x)2^{(m-1-i)}x(1 - x). \quad (19)$$

This yields an exact expression for the variance of $P_i(n)$,

$$\sigma_i^2 = \sum_{j=0}^{m-i-1} (2 - x)^j (2 - x)^{(m-i-1)}x(1 - x). \quad (20)$$

From this solution, it follows that, for $m - i \rightarrow \infty$,

$$\frac{\sigma_i}{\sigma_{i+1}} \rightarrow 2 - x. \quad (21)$$

In particular, for $(2 - x)^{(m-i)}$ much larger than 1, $(\sigma_i/\sigma_{i+1}) \approx 2 - x$. Combining Bramwell *et al.*'s [6] scaling collapse law with Eq. (21) results in another form of the recursion relation, again valid for $(2 - x)^{m-i}$ much larger than 1,

$$P_i(n) \approx (2 - x)P_{i-1}[(2 - x)n]. \quad (22)$$

Casting the recursion formula (22) in terms of the universal scaling function (3) yields the relationship expressed in Eq. (14), $2 - x = 2^\phi$; hence Eq. (3) is equivalent to Bramwell *et al.*'s scaling collapse.

It is noteworthy that the shape of Bramwell *et al.*'s universal probability distribution describing both turbulent power fluctuations and magnetic ordering is remarkably similar to the solution of the recursion relation, Eq. (5), derived by Harte *et al.* [2]. However, the physical relationship between the variable log abundance and either the rate of turbulent power fluctuation or magnetic ordering in Bramwell *et al.* [6] remains unexplained.

In summary, we have shown that finite size scaling techniques applied to ecology provide a powerful theoretical tool for predicting relationships among power-law exponents, as well as a connection between scaling phenomena in ecology and physics.

[1] K. G. Wilson, Rev. Mod. Phys. **55**, 583–600 (1983); H. E. Stanley, Rev. Mod. Phys. **71**, S358–366 (1999).

[2] J. Harte, A. Kinzig, and J. Green, Science **284**, 334–336 (1999).

[3] R. M. May, in *Ecology and Evolution of Communities,* edited by M. L. Cody and J. M. Diamond (Belknap Press, Cambridge, MA, 1975), pp. 81–120.

[4] M. L. Rosenzweig, *Species Diversity in Space and Time* (Cambridge University Press, Cambridge, England, 1995).

[5] M. E. Fisher, in *Critical Phenomena,* edited by M. S. Green (Academic, New York, 1971).

[6] S. T. Bramwell, P. C. W. Holdsworth, and J.-F. Pinton, Nature (London) **396**, 552–554 (1998).