Cluster size and boundary distribution near percolation threshold*

P. L. Leath
Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830
and Physics Department, Rutgers University, New Brunswick, New Jersey 08903 †
(Received 9 February 1976)

It is shown that the shape of the large, random clusters, near the critical percolation concentration c_0, is such that their mean boundary $\langle b \rangle$ is proportional to their mean bulk $\langle n \rangle$ and this is illustrated by an argument which shows that the dimension of the boundary is the same as that of the bulk. The resulting ratio $\langle b \rangle / \langle n \rangle$ is simply related to the critical concentration c_0. The detailed results of a Monte Carlo calculation, previously reported, are given for $c < c_0$ on a simple square lattice; they yield an empirical formula for the probability distribution $\mathcal{P}(n, b)$, for finding a cluster of size n and boundary b, that is proportional to a Gaussian in b/n, which is independent of concentration and which narrows to a δ function at $b/n = \alpha_0$, $n \to \infty$. The asymptotic behavior of the Gaussian form gives the critical exponents $\beta = 0.19 \pm 0.16$, and $\gamma = 2.34 \pm 0.3$, and α_0, gives the critical concentration $c_0 = 0.587 \pm 0.14$, in agreement with previous determinations.

I. INTRODUCTION

In a recent Letter,[1] we reported a new limiting behavior in the shape of the large fluctuations near percolation threshold, which gives rise to the critical behavior.

We report here the details of the Monte Carlo calculations, the primary results of which were given in Ref. 1. Essentially, we find a Gaussian form for the probability distribution $\mathcal{P}(n,b)$, for large n, of finding a single non-null fluctuation extending over n sites, and bounded by b sites. This Gaussian form sharpens to a δ function about a characteristic ratio $\alpha_0 = (b/n)_0$ as n approaches infinity, in a way which allows one to calculate the critical concentration $c_0 = (1+\alpha_0)^{-1}$, and the critical exponents. The probability $\mathcal{P}(n,b)$, of course, contains not only information on the critical point and the critical exponents in its moments, but also much information on the other, nonuniversal, features of the system near the critical point, which are not usually calculated by other techniques.

For simplicity, we study here the site-percolation problem on a simple, square lattice where each site is randomly occupied (or not) with a probability c [or $(1-c)$], which is independent of the occupation of other sites. The common characteristic of these so-called "percolation problems" in nature are the direct relations between the physical quantities and various moments of the cluster size distribution, a relation which gives rise to singular behavior of these quantities about the percolation threshold. A very nice discussion of a percolation problem and its critical behavior has been given by Dunn *et al.*,[2] in terms of a dilute Ising model at low temperatures $k_B T \ll J$. The dilute Ising Hamiltonian $\mathcal{H}$, in a magnetic field H, is

$$\mathcal{H} = -J \sum_{\{i,j\}} \eta_i \eta_j (\sigma_i \sigma_j - 1) - mH \sum_i (\sigma_i - 1) , \quad (1)$$

where η_i is 1 (or 0) depending upon whether the site is occupied (or vacant) by a magnetic atom, where J is the nearest-neighbor ferromagnetic exchange, and $\{i,j\}$ is the set of all pairs of nearest-neighbor sites on the lattice, where $\sigma_i (= \pm 1)$ is the spin variable on site i, and where m is the magnetic moment. In the limit of low temperatures $k_B T \ll J$ every spin in a given magnetic cluster will almost certainly be parallel although the spins in separated clusters (with no nearest-neighbor connections) may not be aligned with each other. We shall only be considering concentrations $c < c_0$ below percolation threshold where there is no infinite cluster. In this limit it may be shown[2] that the free energy per site is given by

$$f(c,T,H) = -k_B T \left\langle \frac{1}{n} \ln \left[1 + \exp\left(-\frac{2mHn}{k_B T} \right) \right] \right\rangle , \quad (2)$$

where $\langle \cdots \rangle$ is the average over all configurations and sizes n of non-null clusters. From this expression, it follows[2] that

$$\left. \frac{d^l f}{dH^l} \right|_{H=0} = -k_B T a_l l! \left(\frac{m}{k_B T} \right)^l \langle n^{l-1} \rangle , \quad (3)$$

where a_l is the coefficient of x^l in the expansion of $\ln[1 + \frac{1}{2}(e^{-2x} - 1)]$, so that, in particular, the magnetic susceptibility exponent γ will be determined from the first moment

$$\chi \propto \langle n \rangle \underset{c \to c_0}{\sim} |c - c_0|^{-\gamma}, \quad (4)$$

and other critical exponents will be determined by studying the exponents of the singular part of other moments $\langle n^l \rangle$.

The quantities of direct physical importance are thus the moments $\langle n^l \rangle$ as a function of concentration c. These moments are simply integrals over the size probability distribution $\mathcal{P}(n)$,[3]

$$\langle n^l \rangle = \sum_{n=1}^{\infty} n^l \mathcal{P}(n) . \tag{5}$$

We shall, in this paper, be concerned primarily with the analytic form of the probability distribution functions. Following the general arguments of de Gennes *et al.*,[4] we can write an exact, general expression for $\mathcal{P}(n)$ on any network:

$$\mathcal{P}(n) \equiv \sum_b \mathcal{P}(n,b) = \sum_b M(n,b)\, c^{n-1}(1-c)^b \tag{6}$$

for $c < c_0$, where $\mathcal{P}(n,b)$ is the probability that a given non-null cluster will have n occupied, connected sites, isolated from the rest of the network by b vacant, bounding sites. For example, in Fig. 1 there is shown a cluster of $n = 193$ connected sites, bounded by 169 vacant sites. Note, in particular, that there are internal, as well as external, boundaries. Clearly $\mathcal{P}(n,b)$ is proportional to c^{n-1}, since $(n-1)$ sites (beyond the one assumed) must be occupied, and proportional to $(1-c)^b$, since the b bounding sites must be vacant. The coefficient $M(n,b)$ of this proportionality is just the number of distinct clusters (beginning with one specified site) containing n sites and bounded by b sites which can be drawn on the network. The coefficient $M(n,b)$ contains all the information on the lattice structure; an exact calculation of M would represent an exact solution of the percolation problem. In particular, the form of $M(n,b)$ determines the critical concentrations, and the critical exponents.

Nevertheless, some limited information can be obtained from Eq. (6) which must apply to all lattices.[1] For example, since the probability distribution must be normalized, we can write the identity

$$1 = \sum_{n=1}^{\infty} \mathcal{P}(n) = \sum_{n=1}^{\infty} \sum_b M(n,b)\, c^{n-1}(1-c)^b \tag{7}$$

for $c < c_0$, which we can multiply by c and differentiate with respect to c to obtain

$$1 = \sum_{n=1}^{\infty} \sum_b M(n,b)\left[nc^{n-1}(1-c)^b - bc^n(1-c)^{b-1}\right] \tag{8}$$

or

$$1 = \langle n \rangle - [c/(1-c)]\langle b \rangle . \tag{9}$$

Thus we find the general relation

$$c = \langle n-1 \rangle / \langle n+b-1 \rangle . \tag{10}$$

This relation simply says that the ratio of the average number of occupied sites (minus the one site assumed occupied) in a cluster to the average number of associated (occupied plus bounding) sites of a cluster is just the concentration of occupied sites. Near the critical concentration, $\langle n \rangle$ diverges with critical exponent γ; thus, from Eq. (10) we can see that $\langle b \rangle$ also must diverge with this same exponent γ, or that $\langle b \rangle$ becomes proportional to $\langle n \rangle$. It should be noted that Stauffer and Domb[5] have previously suggested that b should be proportional to n near c_0. Near the critical concentration c_0, we find (as was previously shown in Ref. 1) specifically that

$$\langle b \rangle_{c_0} = \alpha_0 \langle n \rangle_{c_0} , \tag{11a}$$

where

$$c_0 = (1+\alpha_0)^{-1} . \tag{11b}$$

This result that $\langle b \rangle \to \alpha_0 \langle n \rangle$ is indicative of the fact that the boundaries of the large clusters are of the *same dimension* as that of the bulk or volume of the clusters. For example, large two-dimensional (three-dimensional) clusters have edges (surfaces) which are proportional to their area (volume), which indicates that the edges (surfaces) are also two dimensional (three dimensional) in some sense. To understand more fully this result, consider the two-dimensional fluctuations on a square lattice. First, we associate a randomly chosen number between 0 and l with each site in the square lattice. Then, we plot these numbers on a perpendicular line below each lattice point, as shown in Fig. 2. These random points are then connected by lines between neighbors as shown. The result is the irregular surface of Fig. 2. Now, we imagine that all of the valleys of this surface are filled with, say, water to a level cl (where c is the concentration), as measured downward from

FIG. 1. Randomly generated cluster of 193 occupied, connected sites (●), surrounded by 169 vacant, boundary sites (○).

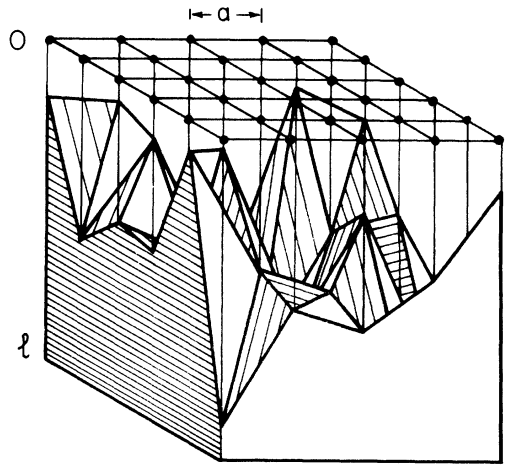

FIG. 2. Random numbers between 0 and l are plotted vertically below the square lattice (lattice constant a). The resulting, neighboring points are connected to form an irregular surface.

the origin. The boundary of the clusters is then measured by the intersection of this irregular surface with the water level. The dimension of the boundary is one less than that of the irregular surface[6]; thus, we must see how this irregular surface can be considered three dimensional.

With only a trivial generalization of Mandelbrot's analysis[6] of the dimension of the irregular west coast of Britain, we can define the Hausdorff-Besicovitch[7] or "fractal"[6] dimension of the surface as follows. Imagine that the approximate area A of a large section of this irregular surface were measured by marching a distortable, three-pronged caliper of constant area g^2 across this surface. The fractal dimension d of the surface is then defined in general by

$$A = Kg^{2-d},$$

where K is a constant. This definition of d is useful if d is independent of g, which turns out to be the case here, when the lattice spacing a is small compared to g and when the typical peak height is large compared to g, that is, when $l \gg g \gg a$. In this limit, the normal to the plane defined by the three points of the calipers makes an angle θ to the vertical l axis, where θ is of order g/l. Thus, we find that, to order l/g, A is given by

$$A \simeq A_N l/g,$$

where A_N, a constant, is the projection of the area A onto the plane of the lattice. And, since l is independent of g, we find, from Eq. (12), that the fractal dimension of the surface is three, and finally then that the corresponding dimension of the boundary of the clusters must be two. A similar proof exists in any dimension that the fractal dimension of the boundary of these random clusters is equal to the dimension of the lattice.

Finally, in order to motivate the data analysis of Sec. II, let us consider briefly the case of the Cayley tree, which was solved exactly by Fisher and Essam.[8] In this case all clusters of n sites are topologically equivalent with a boundary of $b = \alpha_0 n + 2$ lattice sites, where $\alpha_0 = z - 2$. Thus $M(n,b)$ is of the form

$$M(n,b) = \delta_{b,\alpha_0 n+2} m(n), \tag{12}$$

so that the probability function $\mathcal{P}(n)$ becomes, from Eqs. (6) and (12),

$$c\mathcal{P}(n) = m(n) A(c)^n (1-c)^2, \tag{13}$$

where $A(c) = c(1-c)^{\alpha_0}$. Equation (7) then can be written

$$c = \sum_{n=1}^{\infty} c(n) = \sum_{n=1}^{\infty} m(n) A(c)^n (1-c)^2, \tag{14}$$

so that we obtain, by dividing by $(1-c)^2$ and taking l derivatives of each side

$$m(l) = \frac{1}{l!} \frac{d^l f}{dA^l}\bigg|_{A=0 \text{ or } c=0}, \tag{15}$$

where $f = c(1-c)^2$. But, since $A(c) = c(1-c)^{\alpha_0}$ we obtain the closed form

$$m(n) = \frac{(n+b-2)!\,(b+2n-2)}{n!\,b!}, \tag{16}$$

where $b = \alpha_0 n + 2$. For large n, we can use Stirling's approximation to find the asymptotic form of $m(n)$:

$$m(n) \underset{n\to\infty}{\sim} \frac{1+c_0}{[2\pi(1-c_0)]^{1/2} n^{3/2}} \left(\frac{1}{c_0}\right)^{n-1} \left(\frac{1}{1-c_0}\right)^{b}, \tag{17}$$

where $c_0 = (1+\alpha_0)^{-1}$. Thus, for large n,

$$\mathcal{P}(n) \sim \frac{K}{n^{3/2}} \left(\frac{A(c)}{A(c_0)}\right)^n \frac{(1-c)^2}{c}, \tag{18}$$

where K is a constant depending upon c_0. The positive l moments $\langle n^l \rangle$ in Eq. (5) thereby all diverge at $c = c_0$, where $A(c)$ reaches its maximum value of 1, with critical exponents which are easily shown to be $2L - 1$,

$$\langle n^l \rangle \sim (c_0 - c)^{-(2L-1)}. \tag{19a}$$

This result is in agreement with the previous result of Harris.[9] Thus, for example, $\gamma = 1$, and since[1,10]

$$\langle n^l \rangle \sim (c_0 - c)^{\beta - l(\beta+\gamma)}, \tag{19b}$$

we find $\beta = 1$, also. From the general scaling re-

lations for ordinary phase transitions[11] and the established connection with the percolation transition,[12] we can assume

$$\alpha + 2\beta + \gamma = 2 , \tag{20a}$$

so that $\alpha = -1$. The gap exponent

$$\Delta = \beta + \gamma \tag{20b}$$

turns out to be $\Delta = 2$.

In real lattices, of course, $M(n,b)$ is no longer diagonal; that is, for each n there is a distribution of allowed values of b. This distribution is discussed in Sec. II.

II. MONTE CARLO CALCULATIONS

Previous Monte Carlo calculations[13-22] have been performed to measure critical behavior but Dean and Bird,[13] in particular, measured directly the mean cluster size $\langle n \rangle$ and the mean-square cluster size $\langle n^2 \rangle$ versus concentration for atoms randomly placed in finite sections (typically 62 500 lattice sites) of several two- and three-dimensional lattices. From Dean and Bird's calculations it is possible to extract values for the critical indices of $\langle n \rangle$ and $\langle n^2 \rangle$.[17,22] The approach of the work reported here is different in that we measured the probability disbributions $\mathcal{P}(n)$ and $\mathcal{P}(n,b)$ directly and that we avoided systematic errors associated with finite lattice sections by generating and measuring one isolated cluster at a time; in this sense our Monte Carlo calculation is like that of Frisch *et al.*[14]

The computer program used a pseudo-random number generating subroutine (called FLTRN in the IBM system at Oak Ridge National Laboratory) to randomly decide whether each site was occupied or vacant. The site at the origin (see Fig. 3) was always considered occupied. The program then went outward, one shell of sites at a time as indicated in Fig. 3, deciding randomly whether the sites were vacant or occupied until the cluster of sites connected via nearest-neighbor occupations became isolated (i.e., until an entire shell of vacant or disconnected sites could be drawn around the cluster). Then the program counted the number of sites in the cluster, the number of boundary sites (including internal boundaries), and the rms size (radius of gyration) of the cluster. These data were obtained for 101 086 clusters at $c = 0.50$ and for 24 310 clusters at $c = 0.55$. The execution time was between 1 and $1\frac{1}{2}$ h in each of the two cases. The results are as follows.

The normalized size probability distribution function $\mathcal{P}(n)$ obtained is shown versus n in Fig. 4, for the two concentrations $c = 0.50$ and $c = 0.55$. As the critical concentration (at $c_0 \simeq 0.59$) is approached, the large-n tail of the size probability distribution grows so that the positive moments of the distribution diverge at $c = c_0$. The vertical bars on a few selected points represent the statistical error in the histogram bins of width 10.

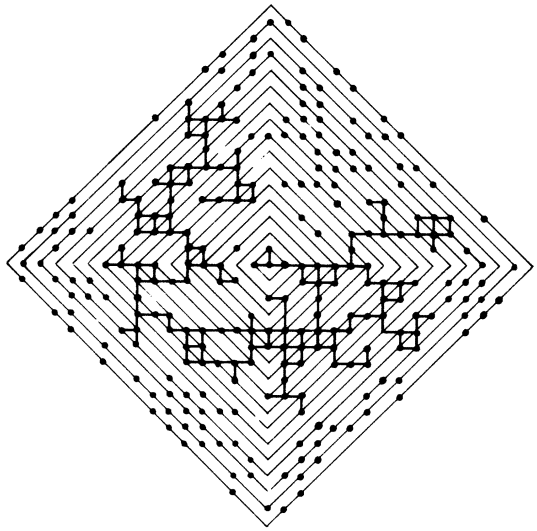

FIG. 3. Randomly generated cluster of occupied sites (●), connected by neighbor bonds (heavy lines). The cluster is generated shell by shell (light diagonal diamonds) until the cluster randomly terminates. The unconnected, occupied sites (●) belong to other clusters and are ignored in this calculation.

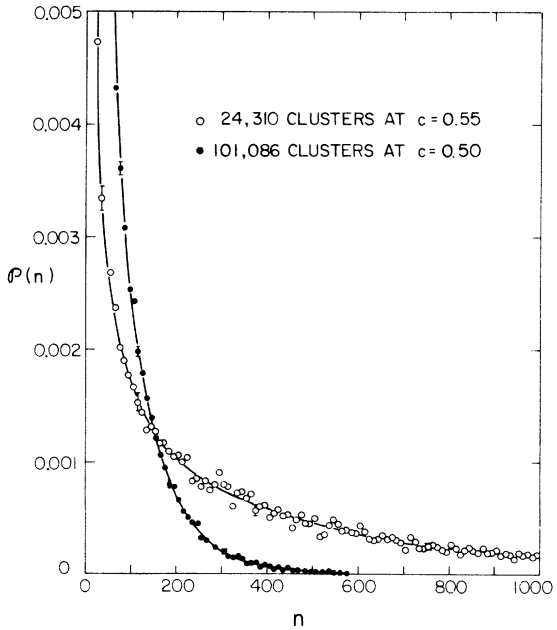

FIG. 4. Probability $\mathcal{P}(n)$ of randomly generating a non-null cluster of size n vs n as determined from the Monte Carlo calculation (data points). The statistical errors are shown as vertical bars.

The mean number of boundary sites $\langle b(n)\rangle$ was indeed found to be proportional to n for large n as shown in Fig. 5. That is,

$$\langle b(n)\rangle = a(c)\, n\,, \tag{21}$$

where the proportionality constant $a(c)$ decreases with increasing concentration, and tends toward the critical value $\alpha_0 = (1-c_0)/c_0$ as the concentration c approaches the critical concentration c_0.

The radius of gyration of the clusters $\langle R^2(n)\rangle^{1/2}$, which is a measure of their linear extent, does not seem to be simply proportional to the square root of the area or $n^{1/2}$ for the two-dimensional clusters as might have been expected. Rather we find the power-law behavior

$$\langle R^2(n)\rangle^{1/2} \propto n^{\rho(c)}\,, \tag{22}$$

where the power $\rho(c)$ is a decreasing function of concentration, as can be seen in the data of Fig. 6; also $\rho(c)$ seems to be heading toward 0.5 at $c = c_0$, which would be true if the large clusters were two-dimensional in their extent.

In order to study in detail the behavior of large clusters, it would be most useful to find the form of the coefficient $M(n,b)$ in Eq. (6). Our first thought was to simply find $\mathcal{P}(n,b)$, then divide by $c^{n-1}(1-c)^b$ to directly obtain $M(n,b)$. This approach turns out to be not so practical, since $c^{n-1}(1-c)^b$ is such a tiny number for typical clusters of size 500 and for $c \simeq 0.5$. Thus, we try to guess that part of the form of $M(n,b)$ which grows so large and counteracts the smallness of $c^{n-1}(1-c)^b$. The best guess we found was motivated by the Cayley-tree solution above. This guess leads to the following definition:

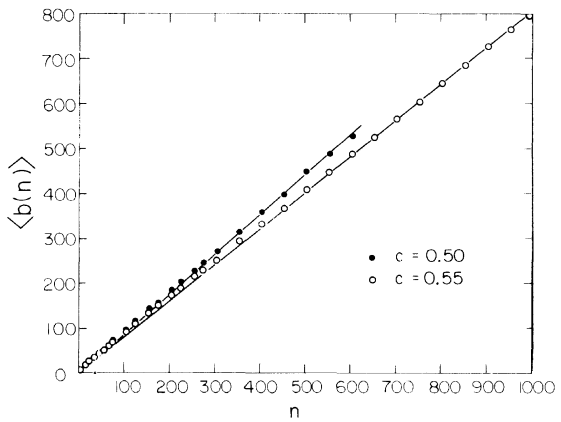

FIG. 5. Mean boundary $\langle b(n)\rangle$ of the clusters vs n from the Monte Carlo calculation (data points) and the straight-line fit to the data.

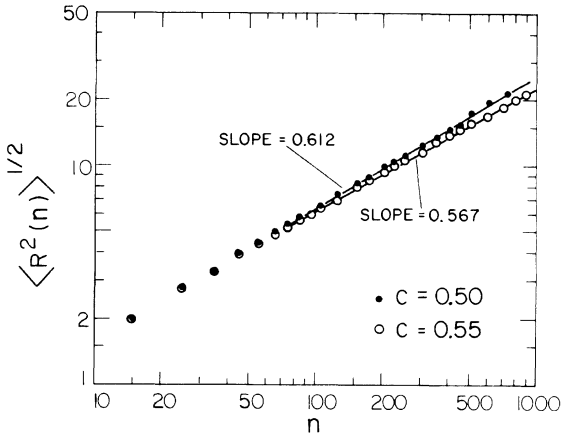

FIG. 6. Root-mean-square size $\langle R^2(n)\rangle^{1/2}$ of the clusters vs n from the Monte Carlo calculation (data points) and the log-log straight line fit to the data.

$$M(n,b) \equiv m(n,\alpha)\, K n^{-\chi} \left(\frac{1}{c_\alpha}\right)^{n-1} \left(\frac{1}{1-c_\alpha}\right)^{\alpha n}\,, \tag{23}$$

where $\alpha = b/n$, $c_\alpha = (1+\alpha)^{-1}$, K, and χ are constant parameters to be determined from the data, and where $m(n,b)$ is, for each n, a normalized function of b, to be determined. Whether this definition is useful can only be determined at this stage from the Monte Carlo data. With this definition, the probability distribution becomes

$$\mathcal{P}(n,b) = m(n,\alpha)\, K n^{-\chi}\, (c/c_\alpha)^{n-1}[(1-c)/(1-c_\alpha)]^{\alpha n}. \tag{24}$$

In order to verify form (6) for the concentration dependence of $\mathcal{P}(n,b)$, and to find K and χ in Eq. (24), we then proceeded to weigh each (n,b) cluster generated by the factor $\{(c/c_\alpha)^{n-1}[(1-c)/(1-c_\alpha)]^{\alpha n}\}^{-1}$ and to integrate the resulting histogram over b to obtain

$$I(n) = \sum_{b/n} K n^{-\chi}\, m(n, b/n) = K n^{-\chi}\,, \tag{25}$$

since $m(n,b/n)$ is, by definition, normalized. The resulting integral obtained is shown in Fig. 7, where $I(n)$ is plotted versus n for the data at concentrations 0.50 and 0.55. The two curves coincide out to about $n = 300$ where the number of points per bin in the $c = 0.50$ calculation fall to a mere handful and statistical significance is poor. This coincidence verifies the general form (6) of the concentration dependence. In addition, a least-squares fit of the curve to the form $Kn^{-\chi}$, for n in the range $75 < n < 905$, gives $\chi = 0.97 \pm 0.007$ and $K = 0.257 \pm 0.01$ as previously reported. This fit is best shown in

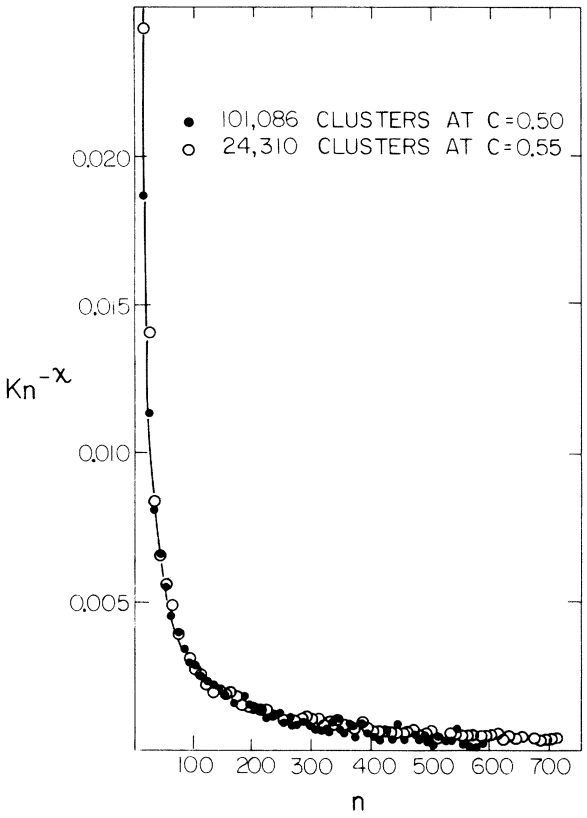

FIG. 7. Weighted probability $I(n)=\sum_b \mathcal{P}(n,b) \times\{(c_\alpha/c)[(1-c_\alpha)/(1-c)]^\alpha\}^n = Kn^{-\chi}$ vs n, from the Monte Carlo calculation (data points).

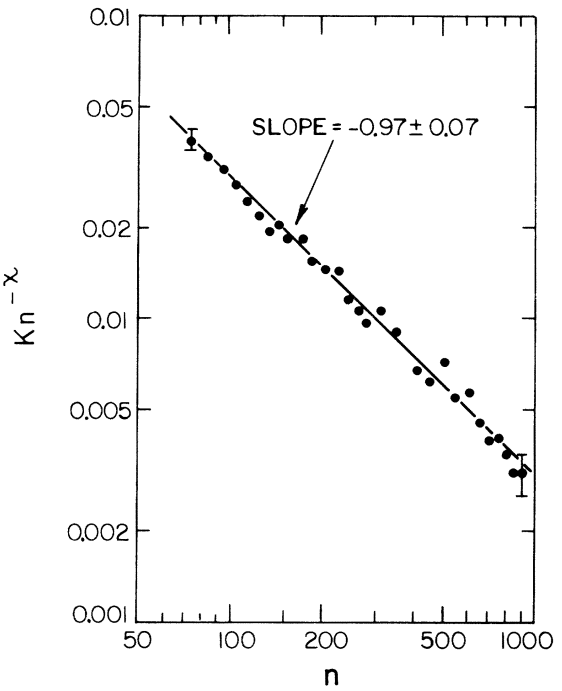

FIG. 8. Log-log plot of the data in Fig. 7 including the straight-line fit to the data.

the straight line, log-log plot in Fig. 8 of the $c=0.55$ data. The quoted errors are statistical.

Next we explicitly found $Kn^{-(\chi-1)}m(n,b/n)$, by weighing each cluster in $\mathcal{P}(n,b)$ with the factor $n\{(c/c_\alpha)^{n-1}[(1-c)/(1-c_\alpha)]^{\alpha n}\}^{-1}$ to obtain directly results typified by those shown in the histograms of Fig. 9, as previously reported,[1] for $n=105$ and $n=205$. The histogram binwidths are $\Delta n=10$ and $\Delta(b/n)=0.01$, and the statistical error bars are as shown at a few selected points.

It is noted that the error bars of Fig. 9 get very large for small b/n (at $c=0.50$) since the *a priori* probability of generating these clusters is very small (i.e., the factor $(c/c_\alpha)^{n-1}[(1-c)/(1-c_\alpha)]^{\alpha n}$ is very small), this effect, which could lead to systematic errors if the data were not properly treated, disappears as the critical concentration is approached. For example, there were no clusters (of the 101 086 generated) found at $c=0.50$ with $200\leq n<210$ and with $b/n<0.76$; although the curve was filled out with the $c=0.55$ data.

The shape of the resulting histogram is very suggestive of a Gaussian form for $m(n,b/n)$ vs b/n. The least-squares fit of the data to Gaussian forms for $n=105$ and $n=205$ is illustrated in Fig. 9 where the smooth curves are the Gaussian fits. But are the curves Gaussian? The most convincing way to plot the data we have found is given in

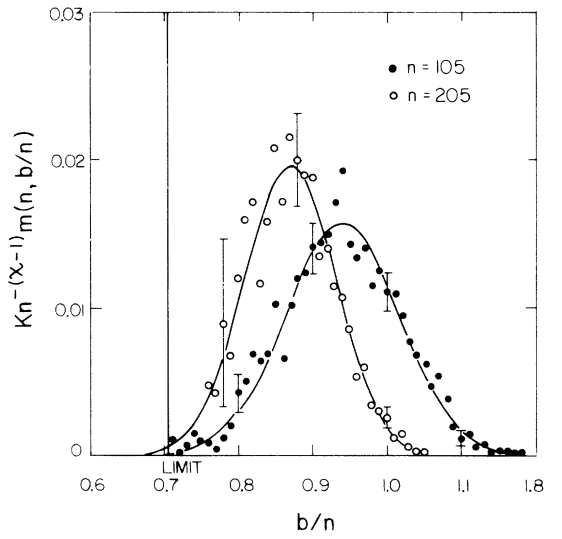

FIG. 9. Form of $m(n,b/n)$ vs b/n from the Monte Carlo calculation (data points) at $c=0.50$ for histogram bins centered at $n=105$, and $n=205$, the vertical line at $b/n=0.704$ represents the limiting δ function, extrapolated from the data as $n\to\infty$, and the Gaussian fit to the data (smooth curves).

Fig. 10 for $n = 105$, where an arithmetic probability plot of the integrated data $\int_{-\infty}^{b/n} m(n,\alpha)\,d\alpha$ is given. This plot is one such that the error function appears as a straight line, with the mean position of the Gaussian occurring where the straight line crosses 0.50. These data are the same as those plotted in Fig. 9; the fluctuations near the ends of the straight line in Fig. 10 are the statistical fluctuations in the tails of the Gaussian seen in Fig. 9. The straightness of this plot indicates its degree of Gaussian character.

Two important systematic features of the Gaussian curves of Fig. 9 are evident. The Gaussian mean shifts toward lower values of b/n, and the Gaussians narrow as n increases. The n dependence of the mean $\mu(n)$ and the standard deviation $\sigma(n)$ are easily seen in the log-log plots of Fig. 11. The straight line for the standard deviation $\sigma(n)$ in Fig. 11(a) indicates that the Gaussian narrows according to the formula

$$\sigma(n) = B/N^{\phi}, \tag{26}$$

where a least-squares fit yields[1] that $B = 0.250 \pm 0.096$ and $\phi = 0.40 \pm 0.04$. The downward shift of the Gaussian with increasing n is such that a limit is approached as $n \to \infty$. This limit can be obtained by finding the value of the limit $\alpha_0 = (b/n)_0$ such that $\mu(n) - \alpha_0$ has the least-squares deviations from a straight-line logarithmic plot. This fit, shown in Fig. 11(b), indicates that $\mu(n)$ is of the form

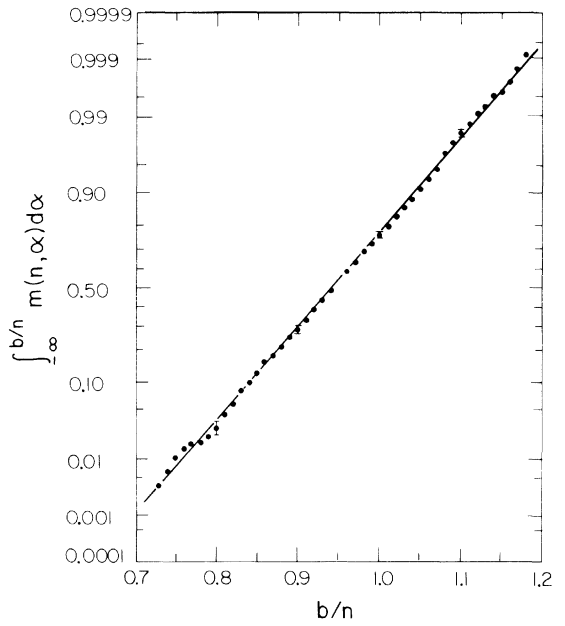

FIG. 10. Arithmetic probability plot of $\int_{-\infty}^{b/n} m(n,\alpha)\,d\alpha$ vs α, for the $n = 105$ data of Fig. 2. And the fit to an error function (straight line).

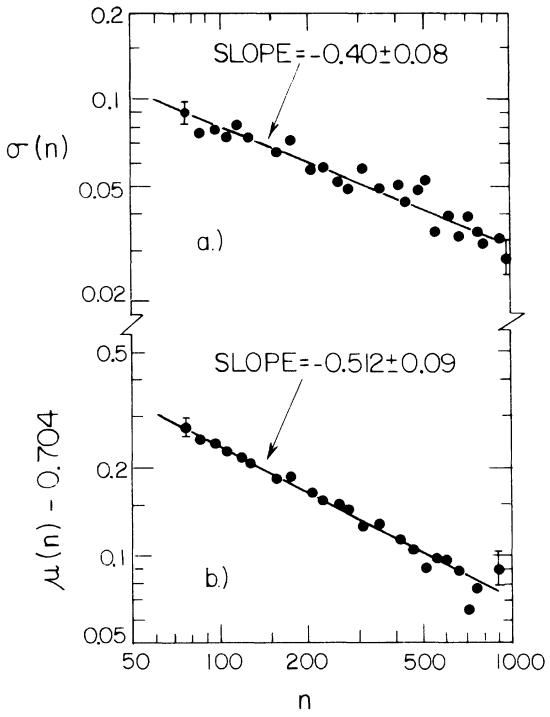

FIG. 11. (a) Log-log plot of $\sigma(n)$ vs n from the Monte Carlo calculation (data points), and the straight-line fit to the data. (b) A log-log plot of $\mu(n) - 0.704$ vs n from the Monte Carlo calculation (data points), and the straight-line fit to the data.

$$\mu(n) = \alpha_0 + A/N^{\psi}, \tag{27}$$

where the least-squares fit gives[1] $\alpha_0 = 0.704 \pm 0.04$, $A = 2.53 \pm 0.5$, and $\psi = 0.512 \pm 0.09$. From the value of α_0 and Eq. (11b) we obtain immediately the critical concentration $c_0 = (1 + \alpha_0)^{-1} = 0.587 \pm 0.014$, which is in agreement with previous determinations[14-16,19,23,24] of the critical point as shown in Table I.[13] The position of the δ-function limit gives directly only the b/n ratio for very large clusters whereas Eq. (11b) requires the ratio of the averages $\langle b \rangle / \langle n \rangle$ over clusters of all sizes at $c = c_0$; but as c approaches c_0 the large clusters dominate so that the identity of the two α_0 factors can be made. One of the most important results of this calculation is that *all* very large clusters that occur for $c < c_0$ have the same b/n ratio, *independent of concentration*.

In summary, the coefficient $m(n,\alpha)$ has been found to be of the Gaussian form

$$m(n,\alpha) = (2\pi)^{-1/2}\sigma(n)^{-1} \exp\left(-\frac{[\alpha - \mu(n)]^2}{[2\sigma(n)^2]}\right), \tag{28}$$

where $\mu(n)$ and $\sigma(n)$ are as given by Eqs. (26) and (27).

TABLE I. Comparison with previous estimates of c_0 for the site problem on a simple, square lattice.

Reference	Method	c_0
14	Monte Carlo	0.581 ± 0.015
15	Monte Carlo	0.580 ± 0.018
16	Monte Carlo	0.591 ± 0.001
23	Series (ratio)	0.580 ± 0.020
23	Series (matching lattice)	0.590 ± 0.010
19	Monte Carlo	0.593 ± 0.005
24	Experiment	0.587 ± 0.005
This work	Monte Carlo	0.587 ± 0.014

III. CRITICAL EXPONENTS

In order to test whether this empirical formula is sensible and to understand more fully its implications, we proceed to find the most singular part of the Lth moment $\langle n^L \rangle$ of the cluster size distribution, namely,

$$\langle n^L \rangle \sim \int_0^\infty dn\, n^L \int_{-\infty}^\infty d\alpha\, \mathcal{P}(n, \alpha). \tag{29}$$

From Eqs. (22) and (26), this integral has the most singular part

$$\langle n^L \rangle \sim \int_0^\infty dn\, n^{L-\chi+\phi} g(n), \tag{30}$$

where

$$g(n) = \int_{-\infty}^\infty e^{-F(\alpha)}\, d\alpha \tag{31}$$

and

$$\begin{aligned} F(\alpha) = {} & (n^{2\phi}/2b^2)(\alpha - \alpha_0 - A/n^\psi)^2 \\ & + n[\alpha \ln\alpha - (\alpha+1)\ln(\alpha+1) \\ & + (\alpha+1)\ln(\alpha_c+1) - \alpha \ln\alpha_c], \end{aligned} \tag{32}$$

with $c \equiv (1+\alpha_c)^{-1}$. In order to obtain the asymptotic form of $g(n)$, we use the method of steepest descents. First, we assume $\phi < \frac{1}{2}$ (which *is* the result in our Monte Carlo calculation) and thereby, for the purpose of locating the value of α for which $F(\alpha)$ is maximum [by setting $F(\alpha) = 0$], we ignore the $n^{2\phi}$ term relative to the linear n term. Then we ignore the lower-order A/n^ψ term. We find trivially that the maximum value of $F(\alpha)$ occurs at $\alpha = \alpha_c$, where F has the value

$$F(\alpha_c) = n^{2\phi} x^2 (2B^2 c_0^4)^{-1}, \tag{33}$$

where $x = c_0 - c$ measures the deviation from the critical concentration. At $\alpha = \alpha_c$, the second derivative F'' is given by

$$F''(\alpha_c) = n[\alpha_c(1+\alpha_c)]^{-1}. \tag{34}$$

Thus, the integral $g(n)$ becomes

$$g(n) \sim \int_{-\infty}^\infty e^{-F(\alpha_c)}\, e^{-\Delta^2 F''(\alpha_c)/2}\, d\Delta \tag{35a}$$

or

$$g(n) \sim \exp[-n^{2\phi}(2B^2 c_0^4)^{-1} x^2]\, \alpha_c^{1/2}(1+\alpha_c)^{1/2} n^{-1/2}. \tag{35b}$$

Then, we find, from Eq. (30),

$$\langle n^L \rangle \sim \int_0^\infty dn\, n^{L-\chi+\phi-1/2}\, e^{-n^{2\phi} x^2 D}, \tag{36}$$

where D is a constant. Now, we change variables, letting $u = n^{2\phi}$, with the result

$$\langle n^L \rangle \sim \int_0^\infty u^{(L-\chi-\phi+1/2)/2\phi}\, e^{-Dx^2 u}\, du, \tag{37}$$

which can be evaluated analytically with the resulting singular part

$$\langle n^L \rangle \sim x^{-(L+\phi+1/2-\chi)/\phi}. \tag{38}$$

Since the $L = 0$ moment gives β and the $L = 1$ moment gives γ, as in Eq. (19b), we set

$$(\chi - \phi - \tfrac{1}{2})/\phi = \beta \tag{39}$$

and

$$\phi^{-1} = \beta + \gamma = \Delta. \tag{40}$$

The values obtained above for χ and ϕ then give the values $\beta = 0.19 \pm 0.16$, $\gamma = 2.34 \pm 0.3$, and $\Delta = 2.53 \pm 0.3$ which are in general agreement with previous results[17,18,20-22,25,26] as shown in Table II (with the rather larger statistical error bars reported here). The free energy per site[2] is proportional to $\langle n^{-1} \rangle$; thus the exponent $(2 - \alpha)$ is obtained by setting $L = -1$ in (38) or the basic scaling law

$$\alpha + 2\beta + \gamma = 2 \tag{41}$$

follows directly from the empirical formula here. The value of α thus obtained is also given in Table II. Perhaps, it should also be noted that if one accepts the standard relations[2] $2 - \alpha = d\nu$, $\eta = 2 - \alpha/\nu$, and $\delta = (2 + d - \eta)/(-2 + d + \eta)$, one indirectly obtains the remaining static exponents $\nu = 1.36 \pm 0.2$, $\eta = 0.28 \pm 0.2$, and $\delta = 13.3 \pm 7$ from this calculation.

It is important to note that, in the above formulas (39)–(40), the mean-field exponents $\beta = \gamma = 1$ [as in Eq. (19)] will be obtained only if $\phi = \frac{1}{2}$ and $\chi = \frac{3}{2}$. The value $\phi = \frac{1}{2}$ would correspond to the usual central limit behavior where the standard deviation in the Gaussian form shrinks as $n^{-1/2}$. The true critical exponents arise because the form of the limiting behavior differs somewhat from that given by the usual limit theorems. It should also be noted that in evaluating the most singular part of the integral (31) we assumed that ϕ was less than $\frac{1}{2}$ in order to neglect certain

TABLE II. Comparison with previous estimates of the critical exponents for the two-dimensional lattices.

Method	α	β	γ
Series			
Site (square) (Ref. 25)			2.1 ±0.2
Site (triangular) (Ref. 26)		0.14 ±0.03	
Bond (triangular) (Ref. 17)	−0.7 ±0.2	0.148±0.004	1.85±0.2
Monte Carlo			
Site (square) (Ref. 18)		0.14 ±0.02	1.9 ±0.2
Site (square) (Ref. 17)		0.16 ±0.02	2.1 ±0.2
Site (square) (Ref. 20)		0.14 ±0.03	2.38±0.05
Site (square) (Ref. 22)		$\gtrsim$0.1	2.23±0.2
This work	−0.72±0.4	0.19 ±0.16	2.34±0.3

terms. At the critical situation (or dimension) where ϕ is exactly one-half, these extra terms will give corrections to the behavior calculated here.

Finally, to clarify the relation of this to previous work it should be noted that the number of boundary sites b (or perimeter, as it is sometimes called) includes internal as well as external surface sites and should not be confused with the external surface only which plays a role in various cluster theories.[27] Also, for clarity, we should note that empirically we found that $m(n, b/n)$ is proportional to a Gaussian in b/n, given by Eqs. (26)—(28) independent of concentration. Nevertheless, the full probability distribution (n/b), given by Eq. (24) is proportional not only to $m(n, b)$, but also to the factor $(c/c_\alpha)^{n-1}[(1-c)/(1-c_\alpha)]^{\alpha n}$ which (for large n) becomes a Gaussian in b/n, centered at $\alpha=(1-c)/c$ and with a standard deviation proportional to $n^{-1/2}$. Clearly the competition of the product of these two limiting behaviors gives rise to much of the interesting behavior seen here.

IV. CONCLUSIONS

We have found a new limiting behavior in the shape of the large random clusters that occur near percolation threshold. This very irregular shape is characterized by a boundary to the clusters which is of the same fractal dimension as the dimension of the bulk of the clusters. The ratio of the boundary to bulk seems empirically to be distributed in a way proportional to a Gaussian (independent of concentration) which narrows to a δ function for very large clusters. The position of the δ function gives the critical concentration and the narrowing gives the static critical exponents (although the poor statistics here prevented a better determination of these exponents than those previously reported). The empirical formula presented here contains much more information, however, than just the critical point and the critical exponents; it is hoped that others will test and use other features of this result.

ACKNOWLEDGMENTS

The author gratefully acknowledges the hospitality of the Metals and Ceramics Division, Oak Ridge National Laboratory during a summer visit and especially the many useful discussions with Dr. J. S. Faulkner, Dr. W. H. Butler, and Dr. J. Olson. He also acknowledges helpful discussions with Dr. J. Sak, Dr. B. Mandelbrot, and Dr. T. Lubensky, and would like to thank G. Reich for help with the final stages of the data analysis.

*Supported in part by National Science Foundation Grant No. DMR 72-03230-AO1.

†Permanent address.

[1]P. L. Leath, Phys. Rev. Lett. 36, 921 (1976).

[2]A. G. Dunn, J. W. Essam, and J. M. Loveluck, J. Phys. C 8, 743 (1975).

[3]It should be noted that $\mathcal{P}(n)$ is the probability that a particular site in a non-null cluster is in a cluster of size n. Thus, if one had a large sample of randomly occupied sites, we could write $\mathcal{P}(n)=nN_n/\sum_n nN_n$, where N_n is the number of clusters of size n in the sample. Therefore, for example, $\langle n\rangle=\sum_n n^2N_n/\sum_n nN_n=S(c)$, where S is the mean cluster size as defined in Ref. 8.

[4]P. G. deGennes, P. Lafore, and J. P. Millot, J. Phys. Chem. Solids 11, 105 (1959).

[5]D. Stauffer, J. Phys. C 8, L172 (1975); and C. Domb, J. Phys. C 7, 2677 (1974).

[6]B. Mandelbrot, *Les Objects Fractals* (Flammarion, Paris, 1975); B. Mandelbrot, Science 156, 636 (1967).
[7]F. Hausdorff, Math. Ann. 79, 157 (1919); A. S. Besicovitch, *ibid.* 110, 321 (1935).
[8]M. E. Fisher and J. W. Essam, J. Math. Phys. 2, 609 (1961).
[9]A. B. Harris, Phys. Rev. B 12, 203 (1975).
[10]K. G. Wilson and J. Kogut, Phys. Rev. 12, 75 (1974).
[11]M. E. Fisher, Rev. Mod. Phys. 46, 597 (1974).
[12]P. W. Kasteleyn and C. M. Fortuin, J. Phys. Soc. Jpn. Suppl. 26, 11 (1969); C. M. Fortuin and P. W. Kasteleyn, Physica (Utr.) 57, 536 (1972).
[13]For a review for these methods, see J. W. Essam, in *Phase Transition and Critical Phenomena,* edited by C. Domb and M. Green (Academic, New York, 1972), Vol. 2, Chap. 6.
[14]H. L. Frisch, E. Sonnenblick, V. A. Vyssotosky, and J. M. Hammersley, Phys. Rev. 124, 1021 (1961).
[15]P. Dean, Proc. Camb. Philos. Soc. 59, 397 (1963).
[16]P. Dean and N. F. Bird, National Physical Laboratory Report No. Ma61, 1966 (unpublished); Proc. Camb. Philos. Soc. 63, 477 (1967).
[17]A. B. Harris, T. C. Lubensky, W. K. Holcomb, and C. Dasgupta, Phys. Rev. Lett. 35, 327 (1975).
[18]This result was obtained from the data of Ref. 13 by Harris *et al.* as reported in Ref. 17.
[19]D. G. Neal, Proc. Camb. Philos. Soc. 71, 97 (1972).
[20]M. E. Levinshtein, B. I. Shklovsky, M. S. Shur, A. L. Efros, Zh. Eksp. Teor. Fiz. 69, 386 (1975) [Sov. Phys.-JETP 42, 197 (1976)].
[21]S. Kirkpatrick, Phys. Rev. Lett. 36, 69 (1976).
[22]D. Stauffer, Phys. Rev. Lett. 35, 394 (1975).
[23]M. F. Sykes and J. W. Essam, Phys. Rev. 133, A310 (1964).
[24]B. P. Watson and P. L. Leath, Phys. Rev. B 9, 4893 (1974).
[25]M. F. Sykes, J. L. Martin, and J. W. Essam, J. Phys. A 6, 1306 (1973).
[26]M. F. Sykes, M. Glen, and D. S. Gaunt, J. Phys. A 7, L105 (1974).
[27]K. Binder and D. Stauffer, J. Stat. Phys. 6, 49 (1972).