Cluster distribution in mean-field percolation: Scaling and universality

Joseph Rudnick and Paisan Nakmahachalasint
Department of Physics, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095-1547

George Gaspari
Department of Physics, University of California at Santa Cruz, Santa Cruz, California 95064

(Received 31 December 1997)

The partition function of the finite q-state Potts model in the limit $q\rightarrow 1$ is shown to yield a closed form for the distribution of clusters in the immediate vicinity of the percolation transition. Various important properties of the transition are manifest, including scaling behavior and the emergence of the spanning cluster. [S1063-651X(98)15311-4]

PACS number(s): 64.60.Ak, 02.50.−r, 02.70.Lq, 05.70.Jk

I. INTRODUCTION

One of the aspects of bond percolation that has captured the imagination of researchers is the collection of scaling properties that emerge in the vicinity of the transition at which the spanning cluster emerges [1]. These scaling properties manifest themselves in various correlation functions. They also control the distribution of clusters in the vicinity of the percolation transition. An adjunct of these scaling properties are quantities that are *universal* at the percolation transition. Universality reflects the insensitivity of behavior at and near the transition to details, and follows from the dominating influence of long range correlations. In this regard, the percolation transition is an example of a critical point. The generating function for cluster distributions (the Laplace transform of the cluster distribution function) plays the role of the thermodynamic potential.

The extension of scaling to finite systems is, in principle, as straightforward for the percolation transition as for conventional critical phenomena. Finite size scaling [2] is based on the assumption that the extent of a system, L, has ''naive'' scaling dimensions, in that it enters into thermodynamic functions only in combination with the correlation length ξ through the ratio L/ξ.

There are, however, difficulties in the application of standard field-theoretical methods to the study of the percolation transition in a finite system. Bond percolation reduces to a ϕ^3 theory [3]. Because of this, there is no immediately obvious way to construct the version of mean-field theory that commonly applies to finite systems [4]. The integral over the exponential of the free energy is not nominally convergent. However, previous research [5,6] has led to the development of a method for the construction of the generating function for percolation on a finite lattice in the case of both long- [7,8] and short-range percolation. This approach avoids the singularities that are inherent in the standard implementation of the integral corresponding to the mean-field limit of the finite size partition function. However, the derivation is not rigorous, and, for this reason, there is room to argue that the problem of the field theory of percolation on a finite lattice has not been entirely solved.

As noted above, the field-theoretic approach yields the generating function for cluster statistics. However, the ''holy grail'' of cluster statistics is the distribution function n_m^c, the average number of clusters containing m sites. Given a sufficiently well-characterized generating function, one can, in principle, reconstitute the distribution function, by a form of the inverse Laplace transformation. The fact that the generating function is consistent with scaling leads to general inferences with regard to the form of the cluster distribution function. However, detailed and specific behavior of the cluster size distribution can only be derived from a specific form for the generating function.

In this paper, we extract, from previously derived results for the generating function of a finite percolating system, a closed-form expression for the cluster distribution function. This closed form displays key attributes of the transition, including consistency with scaling and the separation from the distribution of a single large cluster that, in an infinite system, becomes the spanning cluster.

The closed form described above is tested against the results of simulations, and agreement is documented. In the case of systems for which the percolation transition has occurred, the convergence between theory and simulations is slow. However, the evidence for convergence is sufficiently strong that we are confident in the accuracy of the theoretical model.

The present calculation is relevant to long-range percolation, in that bonds can form between sites separated by an arbitrarily large distance. This model is a poor representation of the kind of bond percolation that occurs in most physical settings. On the other hand, there is every reason to believe that it generates the proper mean-field limit for short-ranged percolation [9]. According to well-accepted results based on scaling considerations, the mean-field theory of the percolation transition is asymptotically accurate for short-ranged percolation in more than six spatial dimensions [3]. Finally, the strict mean-field model of percolation is also a model of interest in its own right in the context of the theory of random graphs [10].

In this infinite-range model of percolation, the probability that a bond connecting two sites is ''active'' scales inversely with the number of sites in the system. On the other hand, any such bond can form between any two sites in the system, in contrast to short-range versions of the model, in which active bonds connect only those sites that are reasonably close to each other. If the probability P that a given bond is active is of the form $P=p/N$, where N is the number of sites, the percolation transition occurs when $p=1$. Critical exponents at the transition take on their mean-field limits. To be specific, the correlation function exponent ν is equal to $\frac{1}{2}$, while the anomalous dimension exponent η is equal to zero. The thermodynamic exponents are all expressible in terms of ν and η through scaling and hyperscaling relations. When the system's dimensionality enters a scaling or hyperscaling relationship, it is commonly set equal to the upper critical value of 6. It should be noted that the ''proper'' dimensionality for this system is infinity, in that, in the thermodynamic limit, the effective coordination number is infinity.

The analysis of this model exploits the connection between the percolation generating function and the statistical mechanics of the one-state limit of the q-state Potts model [11] established by Fortuin and Kasteleyn [12]. A number of field-theoretical treatments of percolation are based on the above relation [3].

II. BRIEF REVIEW OF SCALING

The generating function for cluster sizes is given by

$$F(p,h)=\sum_m n_m^c(p)e^{-mh}, \tag{2.1}$$

where $n_m^c(p)$ is the ensemble average of the number of clusters containing m sites [12]. In the vicinity of the percolation transition ($p\approx p_c$, where p_c is the critical probability) the generating function takes on the scaling form

$$F(p_c(1+\Delta p),h)\rightarrow|\Delta p|^{2-\alpha}f\left(|\Delta p|^{\beta+\alpha-2}h,\frac{\Delta p}{|\Delta p|}\right). \tag{2.2}$$

The exponents α and β, which correspond to the thermodynamic exponents α and β [13], control the asymptotic behavior of various aspects of the cluster size distribution. For example, the lth moment of the cluster distribution function, given by

$$\langle m^l\rangle=\frac{\Sigma_m m^l n_m^c}{\Sigma_n n_m^c}=\frac{(-1)^l\frac{d^l}{dh^l}F(p,h)\Big|_{h=0}}{F(p,0)}, \tag{2.3}$$

acquires a scaling form that is readily obtained from Eq. (2.2).

In the mean-field limit, the two exponents α and β take on the following values:

$$\alpha=2-d\nu=2-6\times\frac{1}{2}=-1, \tag{2.4}$$

$$\beta=\nu\frac{d-2+\eta}{2}=\frac{1}{2}\times\frac{6-2}{2}=1. \tag{2.5}$$

The scaling form in Eq. (2.2), along with the relationship between the generating function and the cluster size distribution [Eq. (2.1)] implies the following cluster size distribution:

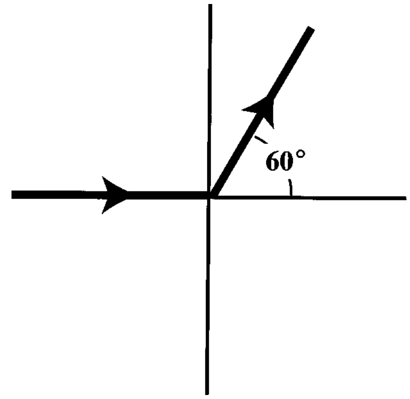

FIG. 1. The contour utilized in the evaluation of the integral in Eq. (3.3).

$$n_m^c=|\Delta p|^{4-2\alpha-\beta}\mathcal{X}\left(m|\Delta p|^{2-\alpha-\beta},\frac{\Delta p}{|\Delta p|}\right). \tag{2.6}$$

When there is a finite number of sites in the system, the cluster size distribution incorporates the number of sites, N, by taking on the more general scaling form

$$n_m^c=|\Delta p|^{4-2\alpha-\beta}\mathcal{X}\left(m|\Delta p|^{2-\alpha-\beta},\frac{|\Delta p|^{d\nu}}{N},\frac{\Delta p}{|\Delta p|}\right). \tag{2.7}$$

In the mean-field limit, the exponent ν is equal to $\frac{1}{2}$, and, as noted above, the dimensionality is set equal to 6. The way in which the number of sites, N, enters reflects the fact that it scales as the dth power of the linear extent L of the system.

III. CLUSTER SIZE DISTRIBUTION

The distribution of cluster sizes in mean-field percolation follows directly from the application of the connection to the Potts model. This result was based on an analysis of the mean-field version of the q-state Potts model in the limit $q\rightarrow1$ [5]. In the calculation leading to a closed-form expression for the generating function of the mean-field Potts model, limits were taken in the proper order, although the final result was obtained with the use of nonrigorous arguments. Making the replacements

$$p=1+N^{-1/3}t, \tag{3.1}$$

$$h=HN^{-2/3}, \tag{3.2}$$

then the generating function takes the following form:

$$F(p,h)\rightarrow\int_{-\infty}^{\infty}d\Delta\left(\left\{\int_0^{\infty}\exp\left[\frac{-(L-t)^3}{6}-\frac{t^3}{6}-\Delta L\right]dL\right\}\mathrm{Im}\times\ln\left\{\int_c\exp\left[(\Delta+H)x+\frac{x^3}{6}\right]dx\right\}\right)+K_c. \tag{3.3}$$

The contour integration in Eq. (3.3) is over a contour in the complex x plane that extends from $-\infty$ on the real axis to ∞ along a curve, making an angle of 60° with respect to the positive real axis. See Fig. 1. For details, see Ref. [5]. The Appendix contains a heuristic derivation of the expression in

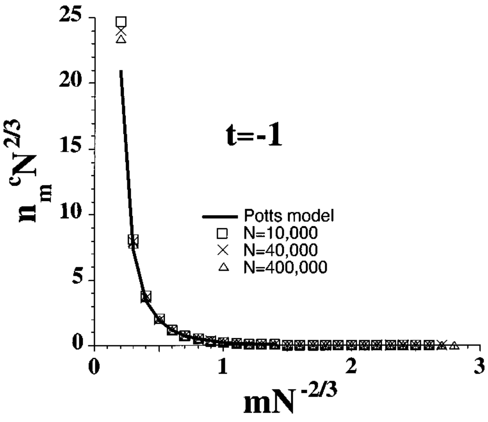

FIG. 2. The cluster size distribution n_m^c, multiplied by $N^{2/3}$, plotted against $mN^{-2/3}$, where m is the size of the cluster and N is the number of sites in the system. The graph in this figure is for $t=-1$, where the quantity t is defined in Eq. (3.1). The system is close to the percolation transition, but the transition has not yet been reached. Note the excellent agreement between the solid curve, representing the predictions of Eq. (3.4), and the results of simulations for $N=10\,000$, 40 000, and 400 000. The quantities plotted in this and in all subsequent figures are dimensionless.

Eq. (3.3). This derivation is based on the evaluation of the Ginzburg-Landau form of the partition function of the q-state Potts model.

The inversion of this function is straightforward to carry out. Shifting the integration variable by Δ, rotating by 90° in the complex plane, multiplying by e^{-imh} and integrating, one immediately obtains

$$n_m^c = N^{-2/3}\int_{\infty}^{\infty} d\Delta\left(\exp\left[\frac{-(mN^{-2/3}-t)^3}{6}-\frac{t^3}{6} -\Delta mN^{-2/3}\right]\mathrm{Im}\,\ln\left\{\int_c \exp\left[\Delta x+\frac{x^3}{6}\right]\right\}\right). \quad (3.4)$$

Expression (3.4) represents the central analytical result reported here. It embodies the expected scaling form of the cluster size distribution, and represents the mean-field limit of the distribution of cluster sizes in the case of short-range bond percolation. As such, it ought to yield the distribution of cluster sizes on a lattice in more than six dimensions, six being the upper critical dimension for short-range bond percolation [3]. In addition, it constitutes the ''zeroth order'' distribution, about which one expands to obtain the cluster size distribution in bond percolation in lower dimensionality. Two characteristics of the distribution function in Eq. (3.4) are worthy of note. First, there is the fact that it *is* in a closed form. Second, there is the fact that the expression contains irreducible complexities, in that the integrals involved in its evaluation cannot be evaluated in terms of elementary or special functions. Nevertheless, the evaluation of the cluster size distribution in large but finite realizations of the mean-field version of percolation has been reduced to quadratures, and this is an interesting result.

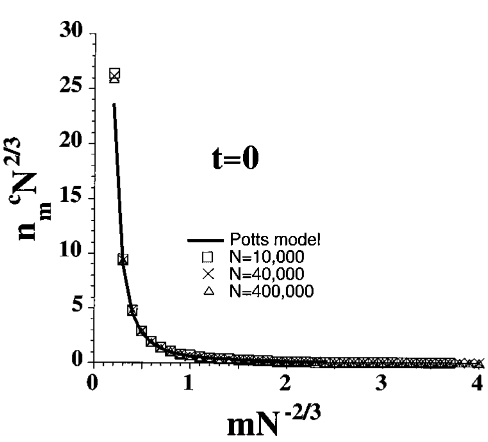

FIG. 3. The cluster size distribution when $t=0$. In the bulk limit, this is the exact location of the percolation transition.

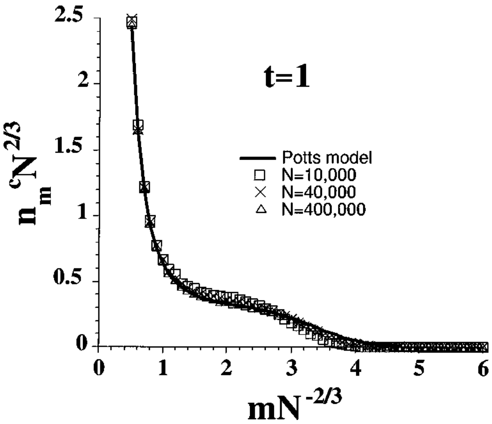

FIG. 4. The cluster size distribution when $t=1$. The system is just above the percolation transition, and the incipient spanning cluster has begun to emerge. The signature in the distribution function is a barely visible feature.

The next step is to test the validity of Eq. (3.4). We have measured the distribution of cluster sizes for mean-field bond percolation on systems with various numbers of sites, N. The results are displayed in Figs. 2–7. The fit between the simulations and Eq. (3.4) is excellent below the percolation transition. When $t>0$, so that the threshold for percolation in the ''thermodynamic limit'' has been exceeded, a feature appears in the distribution in the form of a peak in the upper

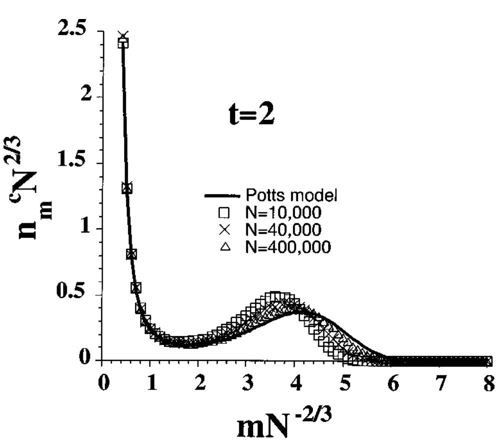

FIG. 5. The cluster distribution when $t=2$. Now the peak for the spanning cluster is becoming distinct. The agreement between the analytical prediction and the results of simulations is not nearly as good in the vicinity of this peak as elsewhere in the figure. However, the agreement improves with increased system size.

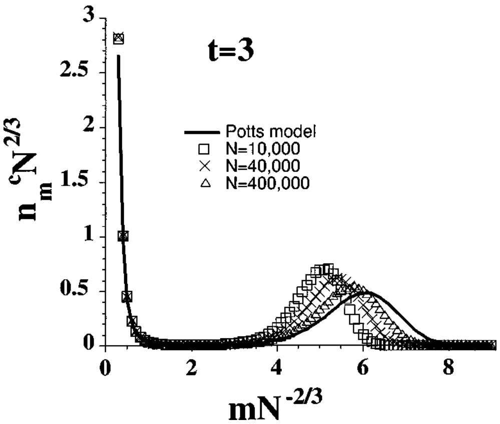

FIG. 6. The cluster distribution when $t=3$. The spanning cluster peak is well separated from the rest of the distribution. The agreement between analysis and simulations is not good in the vicinity of the peak, but, as previously, it improves with increasing system size. The tendency strongly indicates convergence.

reaches of the distribution. This peak—which can be demonstrated to have an integrated weight of unity when t is large and positive—corresponds to the contribution of what becomes the spanning cluster in the limit of an infinite system. As can be seen in Fig. 6, perfect agreement with simulations is not achieved for any of the systems explored. On the other hand, there is clear evidence for convergence between expression (3.4) and the results of numerical calculations as the number of sites increases to fairly large values. We are confident that a system with the sufficient number of sites will have a cluster distribution that is governed by Eq. (3.4).

A possible explanation of the slow convergence may lie in the behavior of the ''next-to-leading'' interaction vertex, the generator of leading order corrections to scaling. If we take into account the fourth order coupling in a ϕ^3 model, the mean-field equation of state in the infinite system has the form

$$t\phi - w\phi^2 + u_4\phi^3 = 0. \tag{3.5}$$

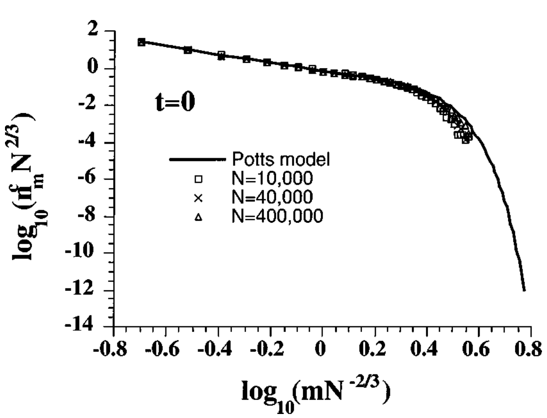

FIG. 7. A log-log plot of the distribution at the percolation transition ($t=0$). In the infinite system, this plot would have the form of a straight line. In the finite system, a power law is obeyed until $n \propto N^{2/3}$. This behavior is evident in the figure, and is displayed by both the analytical form and the results of simulations. As in the previous figures, the agreement between analysis and simulations is best for the largest systems. The logarithms are base 10.

When the interaction strength u_4 is small, the nonzero solution of this equation is

$$\phi = \frac{t}{w} + \frac{u_4}{w}\phi^2 \approx \frac{t}{w} + \frac{u_4 t^2}{w^3} = \frac{t}{w}\left(1 + \frac{u_4 t}{w^2}\right). \tag{3.6}$$

The correction term in the brackets on the last line of Eq. (3.6) vanishes as $t \to 0$. In terms of the mean-field correlation length, it dies off as ξ^{-2}. This last result follows from the fact that the correlation length exponent ν is equal to $\frac{1}{2}$ in the mean-field limit. In the case of a finite system, we can replace this rate of attenuation with $L^{-2} \propto N^{-2/d}$. We replace the dimensionality of the system by 6, in line with the considerations annunciated above. Thus, it is reasonable to expect that finite system results will converge to those expected to hold asymptotically as $N^{-1/3}$. Thus data for a system with 400 000 sites will be closer to the asymptotic predictions than the data generated by a system with 40 000 sites, by a factor of $10^{1/3} \approx 2.15$. Unfortunately, attempts to fit the convergence to a power law in the size of the simulated system did not lead to definitive results, so there is no unambiguous evidence for behavior associated with corrections to scaling. Nevertheless, we are of the opinion that this represents the most likely explanation for the difference between our analytical predictions and the results of our numerical investigations.

IV. CONCLUSIONS

Here we report a closed-form expression for the cluster distribution function of a system in the immediate vicinity of the percolation transition. This expression constitutes the lowest order, mean-field approximation to the cluster distribution function for short-ranged bond percolation. Simulations on a system for which the closed form ought to represent an exact result produce results in agreement with predictions based on that form. The next step is to correct the cluster distribution function by taking into account the effects of local fluctuations. That project is now underway.

APPENDIX: CLUSTER DISTRIBUTION FROM THE MEAN-FIELD PARTITION FUNCTION OF THE $1+\epsilon$ STATE POTTS MODEL

The derivation of the generating function (3.3) in Ref. [5] is long and fairly technical. In addition, the discussion in that reference relies on particular aspects of the ''strict'' mean-field version of the percolation problem, such as the fact that ''active'' bonds can connect sites separated by arbitrarily great distances. This appendix contains a heuristic derivation of that result that starts from the standard field-theoretical formulation of the q-state Potts model. The steps leading to an expression that is equivalent to Eq. (3.3) do not have the same solid mathematical foundation as the arguments in Ref. [5]. However, they are somewhat easier to follow. Furthermore, they lend themselves to the sort of generalization that allows for the treatment of short-ranged bond percolation. The authors are confident in the eventuality of a more complete and rigorous version of the discussion below.

The starting point in the mean-field calculation is the Ginzburg-Landau expression for the partition function of the q-state Potts model [3]

$$\int e^{[-r(s_1^2+s_2^2+\cdots+s_q^2)-w(s_1^3+s_2^3+\cdots+s_q^3)+hs_1]N} \times ds_1\ldots ds_q\,\delta(s_1+\cdots+s_q). \quad \text{(A1)}$$

The quantity N is the number of sites. In the case of a continuum system, this factor is replaced by the total volume. The δ function in Eq. (A1) ensures that the net projection of the system on all states is equal to zero. The phase transition that this model undergoes in the limit $h=0$ is from a ''paramagnetic'' phase, in which all states are equally occupied, and in which all s_i's are equal to zero, to a ''ferromagnetic'' phase, in which one of the states is singled out. In this phase, $s_1\neq 0$, and $s_2=s_3=\cdots=s_q=-s_1/(q-1)$.

To carry out the analysis we replace the δ function in Eq. (A1) by its Fourier representation. Then the partition function is given by

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} d\Omega \int \exp\{[-r(s_1^2+s_2^2+\cdots+s_q^2)-w(s_1^3+s_2^3+\cdots+s_q^3)+hs_1+i\Omega(s_1+\cdots+s_q)]N\}\prod_{i=1}^{q} ds_i$$
$$=\int \frac{d\Omega}{2\pi}\int e^{-rNs_1^2-wNs_1^3+hNs_1+i\Omega s_1}ds_1\left[\int e^{-rNs^2-wNs^3+i\Omega s}ds\right]^{q-1}. \quad \text{(A2)}$$

The generating function for percolation cluster statistics is the result of taking the derivative with respect to q of the partition function of the q-state Potts model at $q=1$ [11]. The quantity of interest is, then,

$$\int_{-\infty}^{\infty}\frac{d\Omega}{2\pi}\left\{\int e^{-rNs_1^2-wNs_1^3+hNs_1+i\Omega s_1}ds_1 \times \ln\left[\int e^{-rNs^2-wNs^3+i\Omega s}ds\right]\right\}. \quad \text{(A3)}$$

For future reference, we note that, if the logarithmic term in Eq. (A3) is replaced by a constant, then the integral over s_1 and Ω yields a fundamentally uninteresting expression. This is because the integration over Ω leads to the generation of the delta function $\delta(s_1)$. What this means is that any operation on the argument of the log that introduces a constant multiplier produces a contribution to the percolation cluster generating function that can be discarded.

We now perform some changes of variable in the integration in the argument of the logarithm. Replacing the integration variable s by $s(Nw)^{-1/3}$ and then shifting the new s to $s-rN^{-1/3}w^{-2/3}/3$, we transform the argument of the logarithm to

$$\int \exp\{-s^3+s[i\Omega N^{-1/3}w^{-1/3}+\tfrac{1}{3}r^2w^{-4/3}N^{2/3}]\}ds, \quad \text{(A4)}$$

where constant multiplicative factors have been discarded.

Rescaling the integration variable s_1 by the factor $(Nw)^{-1/3}$, we find, for the integral over that variable,

$$(Nw)^{-1/3}\int \exp\left[-\left(s_1+\frac{1}{3}rN^{1/3}w^{-2/3}\right)^3 + \frac{1}{3}r^2N^{2/3}w^{-4/3}s_1+\frac{r^3}{27}Nw^{-2}+hN^{2/3}w^{-1/3}s_1 + i\Omega(Nw)^{-1/3}s_1\right]ds_1. \quad \text{(A5)}$$

We now shift the Ω so as to eliminate the term $\frac{1}{3}r^2sw^{-4/3}N^{2/3}$ in the exponential in Eq. (A4), and a similar term in Eq. (A5). Then we replace Ω in the integrations by $\Omega(Nw)^{1/3}$. The argument of the logarithm is, then,

$$\int e^{-s^3+i\Omega s}ds. \quad \text{(A6)}$$

The resulting expression for the percolation generating function is, then, equal to

$$\int_{-\infty}^{\infty}\frac{d\Omega}{2\pi}\left\{\int ds_1\exp\left[-\left(s_1+\frac{1}{3}rN^{1/3}w^{-2/3}\right)^3 + \frac{r^3}{27}Nw^{-2}+hN^{2/3}w^{-1/3}s_1+i\Omega s_1\right]\ln\left[\int e^{-s^3+i\Omega s}ds\right]\right\}. \quad \text{(A7)}$$

In order to ensure that the generating function is a real quantity, we now specify that quantity to be the real part of the above expression. It is also important to note that integration contours in the integral over s must be chosen with care. The default choice will be a stationary phase contour that approaches the positive real axis as on the right hand side of the complex plane. The integration over s_1 is simplified as the result of the argument below.

At this point we argue that the integrals over s and Ω yield zero when the integration variable s_1 is negative. The essence of the argument is that in that case the integration over Ω can be closed in the lower half plane, and that the logarithm contains no singularities there. The regularity of the integrand in the Ω integration is readily established along the negative imaginary axis. A general demonstration has not been accomplished. However, we are bolstered in our belief that the assertion above is true for two reasons. The first is that the result that this leads to is identical in all important aspects to a result previously obtained with the use of an entirely different approach [5]. Additionally, the cluster distribution function that would have resulted if one were to allow negative values of s_1 to contribute to the expression above predicts clusters consisting of a negative number of sites. The absence of a firm justification for this portion of the derivation is the weak point in the present discussion. However, we are confident in both the correctness of the result and that a convincing demonstration of this portion of the derivation can be constructed.

As the final stage in the development of our expression for the generating function for percolation cluster statistics, we rotate the contour in the integration over Ω. Instead of integrating from $-\infty$ to ∞, we integrate over Ω from $i\infty$ to $-i\infty$. Recalling that the real part of the expression is the contribution of interest, we end up with the following result for the percolation generating function:

$$\int_{-\infty}^{\infty} \frac{d\Omega}{2\pi} \int_0^{\infty} ds_1 \exp\left[-\left(s_1 + \frac{1}{3} r N^{1/3} w^{-2/3}\right)^3 + \frac{r^3}{27} N w^{-2} + h N^{2/3} w^{-1/3} s_1 + \Omega s_1\right] \times \mathrm{Im}\ \ln\left[\int e^{-s^3+\Omega s} ds\right]. \quad \text{(A8)}$$

The integral over the variable s follows a stationary phase contour. As the two last steps in the creation of the expression for the cluster statistic generating function, we replace the integration variable s by $-s$ and the integration variable Ω by $-\Omega$. With an appropriate choice of the third order coupling strength w, a shift in the integration variable Ω and some changes of notation, we are led to Eq. (3.3). It is worth noting that the scaling implicit in Eq. (3.3) can be seen explicitly in result (A8).

[1] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1992), and references therein.

[2] See V. Privman, in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, edited by V. Privman (World Scientific, Singapore, 1990), p. 1, and references cited therein.

[3] A. B. Harris, T. C. Lubensky, W. K. Holcomb, and C. Dasgupta, Phys. Rev. Lett. **35**, 327 (1975).

[4] J. Rudnick, H. Guo, and D. Jasnow, J. Stat. Phys. **41**, 353 (1985).

[5] J. Rudnick and G. Gaspari, J. Stat. Phys. **42**, 833 (1986).

[6] J. Rudnick, J. Morris, and G. Gaspari, Phys. Rev. B **49**, 1016 (1994).

[7] P. Erdös and A. Rényi, Publ. Math. Inst. Hungarian Acad. Sci. **5**, 17 (1960) (in Hungarian).

[8] F. Y. Wu, J. Phys. A **15**, L365 (1982).

[9] See M. Kac, in *Statistical Physics, Phase Transitions and Superfluidity*, edited by M. Chrétien, E. P. Gross, and S. Deser (Gordon and Breach, New York, 1968), Vol. 1, p. 241.

[10] B. Bollobás, *Random Graphs* (Academic, London, 1985).

[11] R. B. Potts, Proc. Cambridge Philos. Soc. **48**, 106 (1952).

[12] C. M. Fortuin and P. W. Kasteleyn Physica (Amsterdam) **57**, 536 (1971).

[13] H. E. Stanley, *Phase Transitions and Critical Phenomena* (Oxford University Press, New York, 1987).