Exact cluster size distribution in the one-dimensional Ising model

M. B. Yilmaz and Frank M. Zimmermann*
Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA
(Received 30 March 2004; revised manuscript received 3 September 2004; published 28 February 2005)

The exact solution for the cluster size distribution in the one-dimensional Ising model is obtained. In the thermodynamic limit the result is a simple analytical formula which gives the normalized number of clusters of different sizes. The analytical prediction is compared with Monte Carlo simulations and the energy dependence of the distribution is studied.

DOI: 10.1103/PhysRevE.71.026127 PACS number(s): 05.50.+q, 75.10.Hk

I. INTRODUCTION

Since its first introduction in 1920 by Lenz [1] and in 1925 by Ising [2], the Ising model has been the basis of many theoretical studies to describe cooperative phenomena. The Ising model can be described as a collection of interacting "up" or "down" magnetic spins at fixed lattice sites, or equivalently, as a lattice gas consisting of lattice sites which are either occupied or unoccupied. Interactions between lattice sites are restricted to nearest neighbors. The exact solution in one dimension was provided by Ising and in two dimensions, for the special case of zero applied magnetic field, by Onsager [2,3]. However, the one-dimensional problem does not show a phase transition which many scientists hoped to study in more detail. As a result, more effort has been made in studying the two- and three-dimensional versions of the Ising model than the one-dimensional model. Yet, the one-dimensional model has proved useful in studying many natural phenomena in surface science, nanophysics, and biophysics [4–9]. The fact that certain aspects of the one-dimensional problem have a known exact solution render the one-dimensional Ising model a particularly attractive model description of one-dimensional or quasi-one-dimensional phenomena. Many crystal surfaces exhibit one-dimensional features, such as the terrace step edges of vicinal surfaces, or surface reconstructions featuring one-dimensional chains or rows which can act as one-dimensional lattice gas systems for adsorption of atoms and molecules. For example, the Si(001)2×1 surface consists of long, parallel rows of Si dimers with reactive dangling bonds, which serve as adsorption sites. The one-dimensional lattice gas model successfully describes the statistical mechanics of H atoms adsorbed on the Si(001) surface [4–7]. Single walled carbon nanotubes readily incorporate atoms and molecules into their interior, serving as a conduit for one-dimensional liquids and interacting gases. Maniwa *et al.* have used the one-dimensional Ising model to study orientational ordering of C_{70} inside single-wall carbon nanotubes [8]. A one-dimensional Ising-type model has even led to a semiquantitative understanding of a process as complex as DNA denaturization [9].

*Corresponding author. Electronic address: fmz@physics.rutgers.edu

The problem of cluster size distribution is of fundamental interest and has been studied intensively in Ising or similar models [10–15]. In order to arrive at quantitative predictions, many of the previous studies have relied on numerical simulations. However, the benefit of knowing an exact solution becomes apparent when studying a system's behavior in the vicinity of a critical point, where quantities such as correlation lengths and mean cluster sizes diverge to infinity. In such cases, the required system size for well-converged numerical simulations is expected to also diverge to infinity, together with the computational cost. Here, we provide the general solution for the cluster size distribution in one dimension.

II. MODEL

For definiteness, we cast the problem in terms of the one-dimensional lattice gas, which maps onto the one-dimensional Ising spin model. In this model, we consider a row of n lattice sites with periodic boundary conditions, where each site is either occupied or unoccupied. Adjacent occupied sites interact through the clustering energy ϵ. Here, an occupied cluster consists of a chain of adjacent occupied sites terminated by unoccupied sites on both ends. Unoccupied clusters are defined accordingly. We solve the problem in the canonical ensemble, i.e., keeping the number of occupied sites n_1 constant.

First, we review the solution for the total number of occupied clusters q using the maximum term method [16], adapted to periodic boundary conditions. The partition function can be written as

$$Q = \sum_q \frac{n}{q} g_1 g_2 e^{-E/(k_B T)}, \tag{1}$$

where q is the number of occupied clusters, and g_1 and g_2 are the degeneracy factors associated with energy

$$E = -\epsilon(n_1 - q). \tag{2}$$

The sum is over all possible values of q. The negative sign in front of ϵ in Eq. (2) indicates that the interaction between occupied sites is attractive for positive ϵ. The degeneracy factor g_1 is the number of ways of arranging n_1 occupied lattice sites into q clusters, given by

$$g_1 = \frac{(n_1 - 1)!}{(n_1 - q)!(q-1)!}. \tag{3}$$

Similarly, g_2 is the number of ways of arranging the unoccupied lattice sites into q unoccupied clusters,

$$g_2 = \frac{(n_0 - 1)!}{(n_0 - q)!(q-1)!}, \tag{4}$$

where $n_0 = n - n_1$ is the number of unoccupied lattice sites. The factor of n in Eq. (1) represents the number of ways of arranging these configurations with respect to the lattice. Last, we correct for overcounting related to the periodic boundary conditions. Depending on which of the q clusters in a given configuration appears "first" with respect to an arbitrary lattice site marker, we can divide the set of ng_0g_1 configurations into q subsets, all of which are clearly identical. Each configuration was therefore counted q times, which is corrected by the factor $1/q$ in Eq. (1).

To solve for the thermal equilibrium value q, we use the maximum term method, i.e., the partition function is replaced by its largest term, which is exact in the thermodynamic limit. Maximizing the summand of Eq. (1) with respect to q and solving the resulting quadratic equation in the thermodynamic limit ($n \to \infty$), the term

$$q = \frac{\sqrt{n^2 + 4n_0 n_1 (e^{\epsilon/k_B T} - 1)} - n}{2(e^{\epsilon/k_B T} - 1)} \tag{5}$$

gives the largest contribution to the sum in Eq. (1). Normalizing this with respect to the number of sites n, this becomes

$$\alpha = \frac{\sqrt{1 + 4\theta_0 \theta_1 (e^{\epsilon/k_B T} - 1)} - 1}{2(e^{\epsilon/k_B T} - 1)}, \tag{6}$$

where $\alpha = q/n$ is the normalized number of clusters, $\theta_0 = n_0/n$ is the fraction of unoccupied sites, and $\theta_1 = n_1/n$ is the fraction of occupied sites.

To determine the distribution of cluster sizes, we rewrite the partition function in terms of the cluster size distribution. We define q_l as the number of occupied clusters with length l, where l indicates the number of occupied sites within the cluster. Then each set of occupied clusters $\{q_k\}$ must satisfy the constraints

$$n_1 = \sum_{k=1}^{\infty} k q_k \tag{7}$$

and

$$q = \sum_{k=1}^{\infty} q_k. \tag{8}$$

Inserting these two equations into Eq. (2), the total energy for a particular set of $\{q_k\}$ becomes

$$E = -\epsilon \sum_{k=1}^{\infty} (k-1) q_k. \tag{9}$$

We now calculate the degeneracy factor for a given set of $\{q_k\}$. Let us first consider the case where all clusters are distinguishable, i.e., all clusters have different lengths. In the periodic system there are n possibilities for choosing the position of the first cluster. Since, by definition, each occupied cluster is placed in between two unoccupied sites, there are now $n_0 - 1 = n - n_1 - 1$ positions for the remaining $q-1$ clusters, giving $(n_0 - 1)!/(n_0 - q)!$ possibilities for arranging these $q-1$ clusters. However, since all clusters of the same size are indistinguishable, we have to divide by the number of permutations within each cluster group of size k. For a given set of occupied clusters $\{q_k\}$, the number of different arrangements is then

$$g_3 = \frac{n\left(n - \sum_{k=1}^{\infty} k q_k - 1\right)!}{\left(n - \sum_{k=1}^{\infty} (k+1) q_k\right)! \prod_{k=1}^{\infty} q_k!}. \tag{10}$$

The partition function can now be simply written as

$$Q = \sum_{\{q_k\}} g_3 e^{-E/(k_B T)}, \tag{11}$$

where the sum is over all sets of clusters $\{q_k\}$ satisfying Eqs. (7) and (8).

To determine the number of clusters with length l we again use the maximum term method. Because the values q_k are constrained by Eqs. (7) and (8), we introduce Lagrange multipliers and take partial derivatives with respect to q_l. First, taking the logarithm of the summand of Eq. (11), we write

$$\frac{\partial}{\partial q_l}\left[\ln g_3(\{q_k\}) - \frac{E(\{q_k\})}{k_B T} + \lambda_1\left(n_1 - \sum_{k=1}^{\infty} k q_k\right) + \lambda_2\left(q - \sum_{k=1}^{\infty} k q_k\right)\right] = 0, \tag{12}$$

where λ_1 and λ_2 are the Lagrange multipliers corresponding to Eqs. (7) and (8), respectively. Using the forms given in Eqs. (9) and (10), the above set of equations simplifies to

$$q_l = S_1^l S_2, \tag{13}$$

where

$$S_1 = \frac{(n - n_1 - q)}{(n - n_1 - 1)} e^{\epsilon/(k_B T)} e^{-\lambda_1}, \tag{14}$$

and

$$S_2 = (n - n_1 - q) e^{-\epsilon/(k_B T)} e^{-\lambda_2} \tag{15}$$

are constants independent of l. Inserting Eq. (13) into Eqs. (7) and (8) and solving for S_1 and S_2 we obtain

$$S_1 = 1 - \frac{q}{n_1}, \tag{16}$$

and

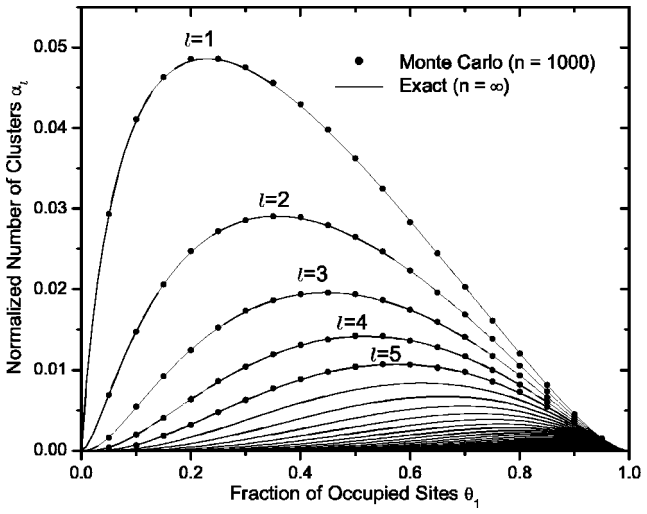

FIG. 1. Comparison of the number of different size clusters obtained from a Monte Carlo simulation on 1000 sites and by using our exact result, Eq. (19). The obtained plots are for $\epsilon/(k_BT)=2$.

$$S_2 = \frac{q^2}{n_1 - q}. \tag{17}$$

Then, Eq. (13) can be written as

$$q_l = \frac{q^2}{n_1}\left(1 - \frac{q}{n_1}\right)^{l-1}. \tag{18}$$

In the thermodynamic limit, the relevant quantity for the occupied cluster size distribution is the normalized cluster size distribution. Dividing Eq. (18) by the number of sites n, we reach the final result,

$$\alpha_l = \frac{\alpha^2}{\theta_1}\left(1 - \frac{\alpha}{\theta_1}\right)^{l-1}, \tag{19}$$

where l is the cluster size, $\theta_1 = n_1/n$ is the fraction of occupied sites, and $\alpha = q/n$ is the normalized number of clusters given by Eq. (6). The mean cluster size is $\langle l \rangle = \theta_1/\alpha$, and the most probable cluster size (l maximizing α_l) is $l=1$ for all occupation fractions and temperatures.

Due to the occupied-unoccupied symmetry of the lattice gas model, the normalized unoccupied cluster size distribution is given by the equivalent expression

$$\beta_l = \frac{\alpha^2}{\theta_0}\left(1 - \frac{\alpha}{\theta_0}\right)^{l-1}, \tag{20}$$

where $\theta_0 = n_0/n$ is the fraction of unoccupied sites. (The number of unoccupied clusters is clearly the same as the number of occupied clusters α.)

In Fig. 1 we show a comparison of the results obtained by using Eq. (19) and Monte Carlo simulations of the cluster size distribution for a clustering energy of $\epsilon/(k_BT)=2$. The Monte Carlo simulations are performed on a one-dimensional chain of 1000 lattice sites with periodic boundary conditions. Using the Metropolis scheme [17], we sampled 5×10^5 accepted trial moves, following equilibration of a random initial configuration by 10^4 accepted trial moves.

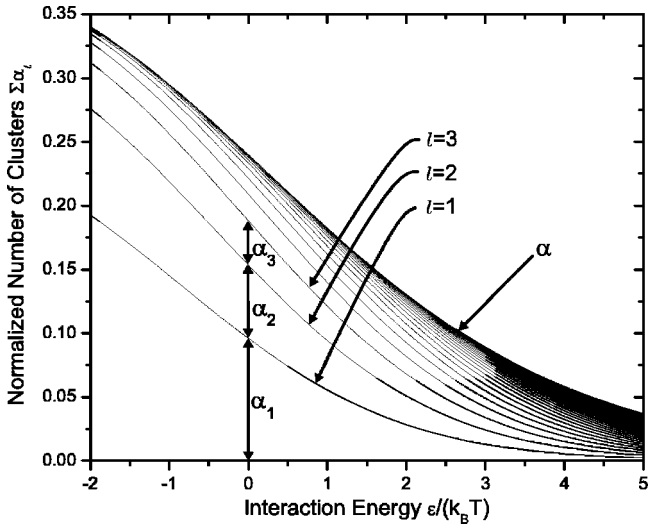

FIG. 2. Number of clusters of size $\leqslant l$ as a function of the clustering interaction energy. Negative energies correspond to repulsion (anticlustering). The plots are obtained for a fraction of occupied sites $\theta_1=0.6$. The vertical distance between lines $l-1$ and l is the number of clusters α_l with length l.

In Fig. 2, we show how the distribution is affected by the interaction energy. The number of single occupied sites (clusters with length $l=1$) shows a monotonic decrease with increasing clustering attraction because more and more longer clusters are formed at the expense of the singles. Clusters with length $l \leqslant (3\theta_1 - 1)/(1-\theta_1)$ display a similar behavior. On the other hand, if $l > \max[1, (3\theta_1-1)/(1-\theta_1)]$, the number of clusters first increases with increasing energy but peaks at

$$\frac{\epsilon}{k_BT} = \ln\left\{\frac{(l-1)[l+1-(l+3)\theta_1]}{4\theta_1}\right\}, \tag{21}$$

decreasing again thereafter. This result could, e.g., be used for designing a system that maximizes the number of clusters of a given length. In the figure, the topmost line shows the normalized number of clusters α given by Eq. (6). In the limit $T\to0$, all curves α_l approach zero as $e^{-\epsilon/(k_BT)}$ and the mean cluster size $\langle l \rangle$ diverges as $e^{\epsilon/(2k_BT)}$, indicating critical behavior at zero temperature.

As an example of a physical system that can be described by such a one-dimensional lattice gas model, we consider the hydrogen covered Si(001)2×1 surface. The surface is covered with parallel rows of silicon dimers, each of which can be unoccupied or (singly or doubly) occupied by hydrogen atoms. Most of the adsorbed hydrogen atoms are paired up on Si dimers due to an attractive pairing interaction of about 0.3 eV [19]. These doubly occupied dimers also have a propensity to form one-dimensional clusters along the dimer row [18,20]. Both pairing and clustering interactions are important for the understanding of the hydrogen adsorption/desorption processes [5,6,19]. Hu *et al.* [18] have carried out

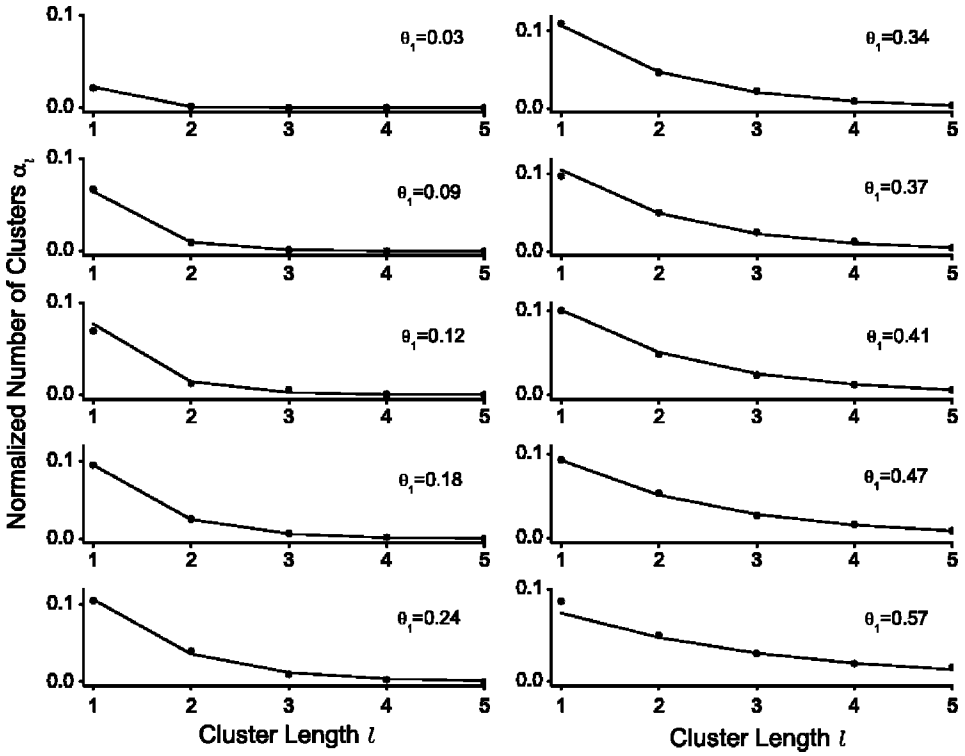

FIG. 3. Cluster size distribution of paired hydrogen atoms adsorbed on one-dimensional dimer rows of the Si(001) surface. Data points are experimental values measured by Hu, Biedermann, Knoesel, and Heinz [18]. Lines are the predictions of the lattice gas model [Eq. (19)] simultaneously fit to all data sets using $\epsilon/(k_BT)$ as the only adjustable parameter. The best fit is obtained for $\epsilon/(k_BT)=0.64$.

a careful scanning tunneling microscopy study of the cluster size distribution for the H/Si(001) system for ten different hydrogen surface coverages ranging from 0.03 to 0.59 monolayer. At each coverage a surface area corresponding to 3×10^4 dimers was sampled. Because of the absence of an analytical solution of the cluster size distribution derived above, the authors had to resort to comparing the observed distributions with the results of Monte Carlo simulations. Using the exact solution we are now in a position to directly fit Eq. (19) to the observed cluster size distribution of Ref. [18] to determine the clustering energy, and to perform a chi-squared analysis to quantitatively characterize the agreement between the lattice gas model and the physical system.

First, we performed individual least-squares fits at each hydrogen coverage, using only the cluster interaction $\epsilon/(k_BT)$ as adjustable parameter. For θ_1 in Eq. (19), we used the coverage of doubly occupied sites as measured by Hu *et al.* [18], thus ignoring any possible clustering interactions involving singly occupied sites [21]. This set of fits yields an average interaction energy $\epsilon=(0.71\pm0.08)k_BT$. The uncertainty reflects the standard deviation of the mean of ten measurements. A simultaneous least squares fit to all ten measured cluster size distributions, corresponding to hydrogen coverages ranging from 0.03 to 0.59 monolayer, yields $\epsilon/(k_BT)=0.64$. Both of these values are in agreement with the Monte Carlo study of Hu *et al.*, who suggested a value of $(0.8\pm0.2)k_BT$. A comparison of the fit to the experimental data measured by Hu *et al.* is given in Fig. 3. The visual quality of the fit indicates that the model is indeed an excellent description of the physical system. However, if the discrepancies between the experimental distributions and the model prediction were purely due to statistical sampling, a chi-squared value close to unity would be expected [22]. Our somewhat larger value of $\chi^2=8.3$ indicates that factors beyond statistics are likely to be responsible for the small differences between the model fit and the experimental data. Possible factors include the neglect of certain interactions in the model (such as next-nearest-neighbor interactions, interactions involving singly occupied dimers, or inter-row interactions), as well as uncertainties in the experimental data due to the inevitable presence of a small number of surface defects. Fitting Eq. (19) to the experimental data thus allows us to not only accurately determine $\epsilon/(k_BT)$ and to conclude that the one-dimensional lattice gas model is an excellent description of the physical system, but also to quantify the quality of agreement.

In conclusion, we have given the exact solution for the cluster size distribution in the one-dimensional Ising model. The result [Eq. (19) in conjunction with Eq. (6)] is an analytical formula which, compared to Monte Carlo simulations, affords a much simpler and more accurate means for analyzing experimental data and making predictions.

ACKNOWLEDGMENTS

We are indebted to S. Lou for valuable discussions, and gratefully acknowledge funding from the National Science Foundation through CAREER Award No. 9733701.

[1] W. Lenz, Physik **21**, 613 (1920).
[2] E. Ising, Z. Phys. **31**, 253 (1925).
[3] L. Onsager, Phys. Rev. **65**, 117 (1944).
[4] Y. L. Yang and M. P. D'Evelyn, J. Vac. Sci. Technol. A **11**, 2200 (1993).
[5] M. C. Flowers, N. B. H. Jonathan, A. Morris, and S. Wright, J. Chem. Phys. **108**, 3342 (1998).
[6] F. M. Zimmermann and X. Pan, Phys. Rev. Lett. **85**, 618 (2000).
[7] M. B. Yilmaz, A. Rajagopal, and F. M. Zimmermann, Phys. Rev. B **69**, 125413 (2004).
[8] Y. Maniwa, H. Kataura, K. Matsuda, and Y. Okabe, New J. Phys. **5**, 127 (2003).
[9] R. M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985).
[10] C.-K. Hu, Phys. Rev. B **34**, 6280 (1986).
[11] J. Marro and R. Toral, Physica A **122**, 563 (1983).
[12] R. Toral and C. Wall, J. Phys. A **20**, 4949 (1987).
[13] X. Campi, H. Krivine, and J. Krivine, Physica A **320**, 41 (2002).
[14] A. V. Bakaev and V. I. Kabanovich, Phys. Lett. A **188**, 97 (1994).
[15] J. Vavro, Phys. Rev. E **63**, 057104 (2001).
[16] T. L. Hill, *Statistical Thermodynamics* (Addison-Wesley, Reading, MA, 1960).
[17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[18] Z. Hu, A. Biedermann, E. Knoesel, and T. F. Heinz, Phys. Rev. B **68**, 155418 (2003).
[19] M. P. D'Evelyn, Y. L. Yang, and L. F. Sutcu, J. Chem. Phys. **96**, 852 (1992).
[20] J. J. Boland, Phys. Rev. Lett. **67**, 1539 (1991).
[21] Due to strong pairing of adsorbed hydrogen atoms on a single dimer, there are few singly occupied dimers. Therefore ignoring interactions involving singly occupied dimers introduces little error to the final results.
[22] J. R. Taylor, *An Introduction to Error Analysis* (University Science Books, Sausalito, CA, 1997).