5-14-1998

The two-dimensional Ising model: mean-field and Monte Carlo results for equilibrium and nonequilibrium transitions

Kevin Scott Brown

The two–dimensional Ising model: mean–field and Monte Carlo results for equilibrium and nonequilibrium transitions

K. S. Brown
Department of Physics and Astronomy
Louisiana State University, Baton Rouge, LA 70803
(May 14, 1998)

Abstract

After a brief introduction to the general features of thermodynamic phase transitions, the equilibrium Ising model is introduced. Mean–field techniques are applied to the model and compared to exact solution; the results emphasize the failure of mean–field theory near the critical point. A detailed Monte Carlo study of the model for two types of dynamics and five algorithms is carried out and the results partially uphold the notion that the static critical properties of the Ising model are independent of the dynamics, provided the end result of that dynamics is an ensemble of states distributed according to the Boltzmann distribution in equilibrium with a heat bath of temperature T. A two–temperature Ising model with nonequilibrium (kinetic) phase behavior is then introduced, and a Monte Carlo study of its static critical properties is conducted.

I. INTRODUCTION

The construction of a simple mathematical model of a complex system is the primary path by which we gain understanding of the world. If we want to study a chemical reaction, for instance, we throw a lot of atoms together in a box, decide on some simple rules by which they interact, and then see what happens. We have a phenomenological theory to describe the behavior we see; it is called thermodynamics. We also have a microscopic theory to predict the behavior; it is called statistical mechanics. It is true that these atoms — Sodium, Hydrogen, whatever we choose — obey Newtonian mechanics at one level and quantum mechanics at another. We *could* write Schrödinger's equation for 10^{23} atoms, but we hardly think anyone would care to try. The point is we can use thermodynamics and statistical mechanics to understand the bulk behavior of our atoms, even though we know they obey the fundamental principles of quantum and Newtonian mechanics. Thermodynamics and statistical mechanics are particularly useful in describing the organization of our atoms into phases, like solid, liquid, and gaseous.

Thermodynamics tells us that if we leave our box of atoms to themselves, they will come to thermal equilibrium. What if we want to study the "atom box" when, say, there is an energy flux through it? This is clearly no longer an equilibrium situation. Can we still use the same techniques? What is the effect of the nonequilibrium perturbation on the system? It is these kinds of questions that gave rise to the field of nonequilibrium thermodynamics and statistical mechanics. However, there is a deeper question we could ask: what if the objects *can't* be described by quantum or classical mechanics? We would like to know how the tendencies of our system to organize itself into phases are dependent on the set of laws they obey. Throwing out the need for the kinds of problems classically studied in statistical mechanics opens up a new universe of problems; change the atoms to people and the rule of interaction to the communication of a virus, and we can study the spread of an epidemic. Now put cars in the box, make traffic laws the interaction rules, and we have a traffic flow model. Change the cars to pieces of paper that show ownership in a company and change the interactions to the means by which those pieces of paper are bought and sold, and we have a financial market.

It should be evident that there is *no way* you could effectively write Schrödinger's equation for a group of cars on a road or stocks on the NYSE; how do we include the driver and the seller? Despite their resistance to obey our sacred fundamental theories, these kinds of systems still show self–organized phase behavior, often much richer than the case for our equilibrium atoms in a box. One can see evidence of this kind of behavior in Refs. [4–11], as well as hundreds of other papers that have sprung from this field. It is this ability to draw an analogy between systems classically studied by physics, like electron gases, molecular solids, and liquids, to systems that at first seem completely outside the realm of physics, like traffic, disease, and finance, that brings nonequilibrium statistical mechanics firmly into the realm of the physicist. In addition to the ability to describe and predict more varieties of natural phenomena, a study of nonequilibrium systems of the kind described above hopes to give us even more fundamental information. We seek to understand how the behavior we observe is dependent on the fundamental theories underlying that behavior. In other words, how much do Hamiltonian and quantum mechanics matter?

The Ising model is one of the most famous systems in all of statistical physics; it was introduced to study the order–disorder transition in ferromagnets. Lenz introduced the model in 1925 [1] and gave it to his student Ising, from which the model eventually derived its name. Ising solved the model in one dimension and zero magnetic field, and found it had a phase transition only at zero temperature. [2] In 1944, Lars Onsager was able to explicitly sum the partition function for the $d = 2$ model in zero field [3], and thereby derive all the critical exponents as well as any other desired thermodynamic quantity. There is still no exact solution for the 2D model in nonzero magnetic field, as is the case for the 3D Ising model in zero field or otherwise. The Ising model is normally studied with two types of dynamics; one (Glauber [17]) in which spins anchored to a periodic lattice spontaneously flip from up to down, and another (Kawasaki [18]) in which adjacent spins exchange their values. We study both sets of dynamics in later sections.

Part of the continued popularity and theoretical mileage of the Ising model is due to the growing field of nonequilibrium physics. Nonequilibrium statistical physics concerns itself with model systems whose dynamics violate detailed balance, thereby enjoining the static critical properties of the system (steady–state averages, the analog of thermal averaging in

the equilibrium case) to its dynamical properties. Nonequilibrium models are being used to describe chemical adsorption/deposition [4,5], forest fires [6–8], capital exchange [9], traffic flow [10], and the growth of warring civilizations [11], to cite a few of the more exotic examples.

There is a tremendous variety of nonequilibrium Ising models. [12] Some mix Glauber and Kawasaki dynamics [48–50]; often this mix is in the form of spin flips at one temperature and spin exchanges at a different temperature. Others models feature spatially varying heat baths [13,14], and some couple each spin to two different heat baths. [15] We consider a specific case of a general class of nonequilibrium models, called driven diffusive systems [39], that have become recently popular. Studies of driven diffusive models can be found in Refs. [39,41], and information and results specific to the model we consider can be found in Refs. [37,39,40,45].

The reason for such an intensive study of so many Ising models, many of which seem quite similar, is because a general theory for nonequilibrium systems in not yet in place. It is hoped that the study of these different models will bring to light some of the general features of nonequilibrium phase transitions. Much of the work in studying these systems consists in transplanting ideas from the study of equilibrium physics and carrying the analogy as far as possible. It is here where field theory, the Renormalization Group, and Monte Carlo computer simulation become of vital importance; they are so useful in the equilibrium case that we cannot help but turn to them when we attempt to study nonequilibrium transitions.

II. THERMODYNAMIC PHASE TRANSITIONS

Before embarking on a detailed study of several flavors of the Ising model, some comments on general features of phase transitions, specifically continuous phase transitions, are in order. In the discussion that follows, we make use of two concrete examples. One is the transition of water from the liquid to gaseous phases, and the other is the heating of a ferromagnet through the Curie temperature.

The basic indicator of a phase transition is a singularity in some derivative of the free energy, as some parameter of which the free energy is a function is varied. In our two example systems, we would probably choose to express the free energies as $F_{\rm water}(T,V,p)$ and $F_{\rm mag}(T,M,H)$, respectively. Singularities could show up in derivatives of $F_{\rm water}$ and $F_{\rm mag}$ as any one of the parameters shown are varied, but we will concentrate on changes in T. These transitions are classified as first–order or continuous, depending upon whether or not a latent heat is involved upon crossing the phase plane boundary. If we heat a flask of water at some pressure less than the critical pressure, it will require the addition of 2.257 MJ/kg to cross into the gaseous phase. Conversely, if we cool water vapor at $p < p_c$, that same amount of heat will be released upon condensation. However, if we heat iron in zero magnetic field, the spontaneous magnetization of our sample decreases continuously, finally reaching zero at $T_c = 1043$ K, the Curie point. There is no latent heat involved in this transition from the ferromagnetic to paramagnetic state, so the transition is deemed continuous. Similarly, if we heat a flask of liquid water at $p = p_c = .323$ g cm^{-3}, we can see a continuous transition in water at $T_c = 647$ K, the liquid–gas critical point.

If we conduct experiments such as the ones above, heating iron in zero field or water

at p_c, we would also notice huge fluctuations in some of the system characteristics as we came close to the critical point. For example, we would notice a divergence in the magnetic susceptibility and specific heat in the iron and in the compressibility and specific heat in the water. These quantities, the susceptibility, specific heat, and compressibility, are related to fluctuations in the magnetization, internal energy, and density, respectively. Near the critical temperature, in the thermodynamic limit (an infinite–sized system), we could characterize these divergences as power laws as follows:

$$\chi, \kappa \sim |T - T_c|^{-\gamma}, \tag{1}$$

$$C \sim |T - T_c|^{-\alpha}, \tag{2}$$

$$\langle m \rangle, \langle \rho \rangle \sim (T_c - T)^{\beta}, \tag{3}$$

where ρ in the liquid–gas system is defined as

$$\rho \equiv \rho_{\text{liq}} - \rho_{\text{gas}}.$$

The fact that m in the iron and ρ in the water behave similarly is due to the fact that they are both order parameters. An order parameter is a quantity used to characterize a thermodynamic phase transition. About the only restrictions on the order parameter are that is be nonzero in one phase and zero in the other and be a smooth (continuous) function. The spontaneous magnetization is obviously a good order parameter, and ρ as defined above is similarly suitable as an order parameter for the liquid–gas system.

In addition to the power law divergences above, spatial correlations in our system get tremendously large as we approach T_c. We see this physically in terms of length scales over which local information is propagated near the critical point. For example, if we try to make a small local change in the magnetization in our iron sample at the critical point, this perturbation can be felt across the entire sample. In fact if we define, with ϕ representing the order parameter, the two–point connected correlation function[1]

$$G_c^{(2)} = \langle \phi(0)\phi(r) \rangle - \langle \phi \rangle^2,$$

then it behaves as follows near T_c

$$G_c^{(2)} \sim \frac{e^{-r/\xi}}{r^{d-2+\eta}},$$

where η is another critical exponent, d is the spatial dimensionality of the system, and ξ is the correlation length. The value of ξ gives us information about these spatial correlations, and it turns out that it is a fundamental length scale in the problem. Near the critical temperature, ξ also diverges as

$$\xi \sim |T - T_c|^{-\nu}. \tag{4}$$

---

[1]The word "connected" arises for graph theoretical reasons.

Having defined all of our critical exponents, we could set out to measure them, and such experimental data certainly exist for a variety of models and real systems alike. If we were to do this, we would find another striking property; the phenomenon of universality. Universality refers to the fact that sometimes strikingly dissimilar systems have the same critical exponents. For example, the 3D Ising model of a ferromagnet and a 3D binary fluid have the same critical exponents. We are careful to use the phrase "3D" in both cases, because the values of critical exponents are strongly dependent on the dimensionality of the system in question. We can present a heuristic argument for the notion of universality, but upon first reading one is likely to think it a difficult pill to swallow. The argument proceeds as follows. The power law divergences we encounter in quantities like the magnetic susceptibility or specific heat result from large scale correlations; a small change can be felt throughout the entire experimental sample, which can lead to wild fluctuations in response functions like χ and C. These huge correlations imply that microscopic details are beginning to become less important; the range of forces is no longer as crucial, since there is a much larger *macro*scopic length scale (namely ξ) beginning to dominate, and interparticle forces normally operate on *micro*scopic scales. As we get closer to T_c, ξ is diverging to infinity. Thus, right at the critical point, the macroscopic length scale ξ is completely dominant and not even finite, and the microscopic length scales play no role. This being the case, what can impact the critical properties of our sample system? Only things that have no length scale. This rules out forces of any kind, but it certainly allows for symmetries and conservation laws to be important, and this is indeed the case.

III. THE MODEL

The Ising model is intended to represent a spin-$\frac{1}{2}$ ferromagnet. It is easy enough to write down the Hamiltonian:

$$\mathcal{H} = -J\sum_{\langle ij\rangle}\sigma_i\sigma_j - h\sum_i \sigma_i, \quad \sigma = \pm 1. \tag{5}$$

Here, the first sum is over nearest–neighbor pairs and the second is over all spins. The σ variable represents spin up or spin down, with the $\frac{1}{2}$ being absorbed into the constants J and h. J is a coupling constant and h is a magnetic field variable, although they both have units of energy as written above. If $J > 0$, the lower energy state between pairs of spins is the state of parallel alignment, indicating a ferromagnetic interaction. $J < 0$ implies an antiferromagnetic interaction.

The Ising model is a special case of several more general classes of models, and we mention one here. The Hamiltonian for the q–state Potts model may be written as follows [16]:

$$\mathcal{H}_P = -J\sum_{\langle ij\rangle}\left[q\delta_{\sigma_i\sigma_j} - 1\right], \quad \sigma = 1, 2, \ldots, q.$$

The sum is again over nearest–neighbor pairs. Basically, the model allows objects to be in any one of q distinct states and assigns one value of energy to pairs of similar objects and a different value to pairs of dissimilar ones. The $q = 2$ case corresponds to the zero–field Ising

model (Eqn. (5) with $h = 0$), the $q \to 1$ case is bond percolation, and the $q \to \infty$ limit is often of interest in exact calculations.

The Ising model is a purely classical statistical model. Other than the fact that we call the objects "spins," there is absolutely nothing quantum mechanical about the Ising model as described above. In fact, since all operators in $\mathcal{H}$ are classical and therefore commute, the Heisenberg equations of motion predict that the model never does anything at all! It is this lack of intrinsic dynamics that lead us to imagine that our spin system (lattice) is coupled to a heat bath of temperature T. This heat bath then induces changes in the spin configurations, yielding a dynamics. Several types of dynamics have been introduced for the Ising model, the most famous being Glauber (spin–flip) dynamics [17] and Kawasaki (spin–exchange) dynamics [18]. However, if we study an equilibrium model, we must remember the mechanism for causing transitions has no effect on the static critical properties, assuming that mechanism satisfies detailed balance. If this is the case, any dynamics we choose generates states distributed according to the Boltzmann distribution in thermal equilibrium with the aforementioned heat bath.[2] One goal of this work is to verify the aforementioned statement for the equilibrium Ising model.

IV. SOLVING THE MODEL

Our goal in studying the Ising model is, of course, to be able to have some sort of compact representation of its equilibrium behavior. We could then obtain an exact, closed–form equation for any thermodynamic function, in addition to values for the critical exponents. The most obvious object, in the context of statistical mechanics, from which we could extract all this information is the partition function. Ising computed the partition function for the zero-field $d = 1$ model and found a transition only at $T = 0$ [2]. After this, the task of performing the sum in Eqn. (5) becomes prohibitively difficult. Onsager did so for the zero-field $d = 2$ model [3], but no one has yet followed suit and solved the $d = 2$ model in nonzero field or the $d = 3$ model, nonzero field or otherwise.

It is because the partition function is so difficult to evaluate for $d > 1$ that other techniques become of prime importance in understanding the behavior of model systems such as the Ising model. Two of the most widely used are mean–field theory (MFT) and Monte Carlo (MC) simulation. The Renormalization Group (RG) [19,20] is also quite useful and important in performing calculations, but RG calculations are beyond the scope of this paper. We will save MC for another later section and focus on MFT for now.

There are many names associated with variants of MFT, including Weiss [21], Bragg–Williams [22], Bethe [23], and Kikuchi [24] . Some versions of MFT work to simplify the partition function [21,23], while others attempt to use combinatorial arguments to estimate the free energy directly [22,24], and still others attempt to write a master equation for

---

[2]This may seem like a restriction, but it is in some sense a blessing; we do not have to simulate a real dynamics, with phonon collisions, photon–induced transitions, etc. We may choose a very simple, though unrealistic dynamics, provided it still generates states distributed according to the appropriate equilibrium (Boltzmann) distribution.

localized configurations and then search for steady–state solutions [25,26][3]. At the heart of all the various mean–field treatments lies a "totalitarian" view of the physics. At one stage or another, MFT determines single–object (whether it be a spin or a cluster) behavior according to a local average. All flavors of MFT ignore high–order correlations, often leading to their failure near critical points. However, MFT is by no means a useless technique; it has several virtues. For one, it tends to be much easier to formulate than RG or an exact solution. In addition, there are some systems that are "mean–field–like," in the sense that they are accurately described by some type of MFT; examples include the normal metal to superconductor transition and the smectic-A to smectic-C transition in liquid crystals [16]. Finally, while MFT may be quantitatively incorrect, it can often predict some qualitative features of a model's phase diagram.

V. MFT FOR THE ISING MODEL I

Our starting point for MFT as it relates to the Ising model will be the Weiss molecular field (WMF) approximation [21]. We assume that we have a lattice of coordination number z, filled with N spins, one spin per lattice site. We begin with the Ising Hamiltonian, Eqn. (5):

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i.$$

Though all results in this paper concern the zero–field model, for the mean–field treatments an infinitesimal field h is carried until the end of the calculations. The reason for this is that the magnetization, defined thermodynamically as the field derivative of the free energy, cannot be properly defined otherwise. We can always subsequently take the limit as $h \to 0$ to investigate any spontaneous magnetization.

In the WMF approximation, we replace the interaction at site i by an effective field, composed of an interaction with the real field h and a field created by the z nearest neighbors of site i. Thus we may write for the action of the effective field h_e on one spin

$$\begin{aligned} h_e \sigma_0 &= J(\sigma_1 + \cdots + \sigma_z)\sigma_0 + h\sigma_0 \\ &= (Jz\bar{\sigma} + h)\sigma_0 \end{aligned}$$

The variable $\bar{\sigma}$ is a sort of local magnetization; it assigns to each nearest neighbor spin a value equal to the average over the near–neighbor cluster.

This approximation allows us to rewrite the Hamiltonian as

$$\mathcal{H}_W = -h_e \sum_{i=1}^{N} \sigma_i$$

---

[3]The field theoretic approaches described in [25,26] reduce to the Kikuchi method in the equilibrium case, but they can be used in the nonequilibrium (dynamical) case as well, making them useful for current problems in kinetic phase transitions.

$$= -Jz\bar{\sigma}\sum_{i=1}^{N}\sigma_i - h\sum_{i=1}^{N}\sigma_i.$$

Now we use $\mathcal{H}_W$ to try to calculate Z, the partition function:

$$\begin{aligned} Z &= \sum_{\text{(states)}} e^{-\beta\mathcal{H}_W} \\ &= \sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_N=\pm1} e^{\beta Jz\bar{\sigma}\sum_{i=1}^{N}\sigma_i} e^{\beta h\sum_{i=1}^{N}\sigma_i} \\ &= \sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_N=\pm1}\prod_{i=1}^{N} e^{\beta(Jz\bar{\sigma}+h)\sigma_i}. \end{aligned}$$

Notice how the WMF approximation has helped us. We have effectively turned one N–body problem into N identical one–body problems. Writing out a few of the sums above will convince one that

$$Z_N = \left[2\cosh\beta\left(Jz\bar{\sigma}+h\right)\right]^N. \tag{6}$$

Now we can easily calculate the free energy per spin:

$$f = -\frac{1}{\beta N}\ln Z_N = -\frac{1}{\beta}\ln\left[2\cosh\beta\left(Jz\bar{\sigma}+h\right)\right].$$

To go any further, we apply a self–consistency equation. We relate the local magnetization $\bar{\sigma}$ to the global magnetization as follows:

$$\bar{\sigma} = \frac{\langle M\rangle}{N} = m.$$

This is certainly a logical step, given our interpretation of $\bar{\sigma}$. Now we can find the equation of state by using a thermodynamic relationship between m and f. We compute

$$\begin{aligned} m = \bar{\sigma} &= -\frac{\partial f}{\partial h} \\ &= \tanh\beta\left(Jz\bar{\sigma}+h\right), \end{aligned}$$

and see that our equation of state is

$$\bar{\sigma} = \tanh\beta\left(Jz\bar{\sigma}+h\right). \tag{7}$$

Now we may set $h = 0$ to study the onset of a spontaneous magnetization, giving

$$\bar{\sigma} = \tanh\left(\beta Jz\bar{\sigma}\right).$$

Given a recursive equation such as the one above, we would normally attempt to use some sort of root–finding procedure like Newton–Raphson or the shooting method. However, this equation is simple enough that it permits a graphical solution, illustrated in Fig. 1. One can see that the intersection of the two functions gives us either one root with multiplicity three

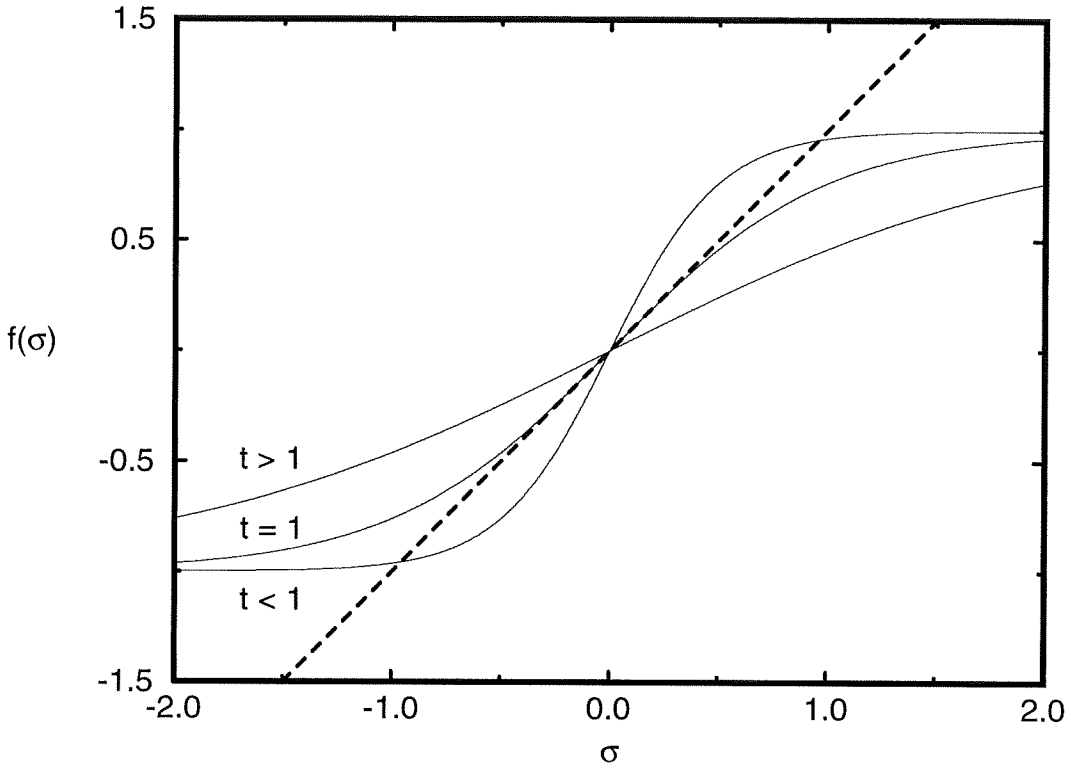

FIG. 1. Graphical method for solving Eqn. (7) with $h = 0$. Three different hyperbolic tangent curves for different values of $t = T/T_c$ are shown. The dotted line is the graph of $f(\bar{\sigma}) = \bar{\sigma}$. Since the slope of $\tanh(ax)$ near zero is ax, if $a > 1$ the two functions will intersect three times at 0 and $\pm\bar{\sigma}(T)$, and if $a < 1$ they intersect once (with multiplicity three) at 0. Thus $a = 1$ defines the critical temperature, where the three roots first coalesce.

or three distinct roots. The point at which the three roots first coalesce and $\bar{\sigma}$ is identically zero is the critical point. We determine that value as follows. The slope of $\tanh(x)$ near $x = 0$ is one, and that value decreases monotonically away from zero. Therefore, if the value of βJz is greater than one, the tanh curve crosses the $y = \bar{\sigma}$ curve both at $\bar{\sigma} = 0$ and again at some value $\pm\bar{\sigma}(T)$. However, if βJz is less than one, the only crossing point is at $\bar{\sigma} = 0$. Thus we see

$$\beta_c Jz = 1,$$

from which we can infer

$$T_c = z\frac{J}{k_B} = z$$

in units where J/k_B is one, which is a unit we will use throughout this paper. Applied to the $d = 2$ Ising model, we have determined that no spontaneous magnetization exists for $T > 4.0$, corresponding to a paramagnetic phase, and that for $T < 4.0$ we have ferromagnetic

ordering giving rise to a temperature–dependent magnetization that increases toward $\bar{\sigma} = 1$ as $\beta \to \infty$.

Now that we are armed with an expression for T_c and an equation of state, we want to set out to determine the critical exponents under the WMF approximation. For this purpose, it is convenient to define different critical exponents above and below T_c. Using the magnetic susceptibility as an example, we define [27]

$$\chi \sim (T_c - T)^{\gamma'}, \quad T < T_c$$
$$\chi \sim (T - T_c)^{\gamma}, \quad T > T_c,$$

with similar definitions for other critical exponents, except an exponent such as β which has no meaning above T_c. While these exponents are believed to be identical [28], we have no *a priori* reason to suspect such a result, and the procedure for using the mean–field equation of state to calculate exponents above and below T_c is somewhat different. We concentrate first on the unprimed exponents.

There are a couple of different ways to calculate α, and we will use the definition of C_V in terms of Z. Recall

$$C_V = k_B \beta^2 \left(\frac{\partial^2 \ln Z}{\partial \beta^2} \right)$$

Equation (6) gives us Z, and if we perform this calculation we discover

$$C_V = N k_B (J z \bar{\sigma} \beta)^2 \mathrm{sech}^2 (\beta J z \bar{\sigma}) .$$

Since $\bar{\sigma} = 0$ for $T > T_c$, we discover $C_V = 0$ for $T > T_c$, so of course $\alpha = 0$.

To calculate γ, we use the definition of the zero–field magnetic susceptibility

$$\chi = \lim_{h \to 0} \frac{\partial \bar{\sigma}}{\partial h}.$$

In addition, we use the Taylor series for $\tanh(x)$:

$$\tanh(x) = x - \frac{1}{3} x^3 + \frac{2}{15} x^5 - \cdots, \tag{8}$$

which for small x we may approximate as

$$\tanh(x) \approx x.$$

We now return to the equation of state, Eqn. (7), which we may rewrite as

$$\tanh \left(\frac{T_c}{T} \bar{\sigma} + \frac{h}{T} \right).$$

For $T \gtrsim T_c$, $\bar{\sigma}$ is small and we may use the approximation in Eqn. (8). This allows us to solve for $\bar{\sigma}$, resulting in

$$\bar{\sigma} = \frac{h}{T - T_c},$$

and we easily see

$$\chi \sim \frac{1}{T - T_c},$$

leading us to conclude $\gamma = 1$. Do not, however take the coefficient on $(T - T_c)^{-\gamma}$ seriously; we have played fast and loose with constants and units, which is supported by the fact that χ as written above has the wrong units. We try to indicate this by using the symbol "$\sim$" instead of something stronger.

Now let us calculate these same exponents for $T < T_c$, along with β. Below T_c, $\bar{\sigma} \neq 0$, so we cannot use quite the same methodology as before. We can still make use of the Taylor polynomial expansion for $\tanh(x)$, however (Eqn. (8)). For $T \lesssim T_c$, $\bar{\sigma}$ is small and we may write for the equation of state

$$\bar{\sigma} \approx \frac{T_c}{T}\bar{\sigma} - \frac{1}{3}\left(\frac{T_c}{T}\right)^3 \bar{\sigma}^3. \tag{9}$$

Solving for $\bar{\sigma}$ involves dividing by $\bar{\sigma}$, but since it is nonzero this does not pose a problem. We discover

$$\bar{\sigma} = \pm\sqrt{3}\,(T_c - T)^{\frac{1}{2}} \left(\frac{T^2}{T_c^3}\right)^{\frac{1}{2}}, \tag{10}$$

which allows us to deduce $\beta = \frac{1}{2}$. The double–valuedness of $\bar{\sigma}$ results from the choice of two equivalent branches for the magnetization: net spin–up or net spin–down. Now put this expression to work. To calculate γ', it is convenient (following Ref. [27]) to use the relation

$$\tanh(x + y) = \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x)\tanh(y)},$$

which allows us to rewrite the equation of state as

$$\bar{\sigma} = \frac{\tanh\left(\frac{\bar{\sigma}}{\widetilde{T}}\right) + \tanh\left(\frac{h}{T}\right)}{1 + \tanh\left(\frac{\bar{\sigma}}{\widetilde{T}}\right)\tanh\left(\frac{h}{T}\right)},$$

where we have made the definition $\widetilde{T} = T/T_c$. We can rearrange this equation to obtain

$$\tanh\left(\frac{h}{T}\right) = \frac{\bar{\sigma} - \tanh\left(\frac{\bar{\sigma}}{\widetilde{T}}\right)}{1 - \bar{\sigma}\tanh\left(\frac{\bar{\sigma}}{\widetilde{T}}\right)}.$$

If we make the definition

$$h' = \tanh\left(\frac{h}{T}\right), \tag{11}$$

the left hand side of the equation is cast into a simpler form. Additionally, using Eqn. (8), we can write for $T \lesssim T_c$

$$h' = \frac{\bar{\sigma}\left(1 - \frac{1}{\tilde{T}}\right) + \bar{\sigma}^3\left(\frac{1}{3\tilde{T}^3}\right)}{1 - \bar{\sigma}\left(\frac{\bar{\sigma}}{\tilde{T}} - \left(\frac{\bar{\sigma}}{3\tilde{T}}\right)^3\right)}.$$

For small x,

$$\frac{1}{1-x} \approx 1 + x,$$

so we may write

$$h' = \bar{\sigma}\left(1 - \frac{1}{\tilde{T}}\right) + \bar{\sigma}^3\left(\frac{1}{3\tilde{T}} + \frac{1 - 1/\tilde{T}}{\tilde{T}}\right) + \mathcal{O}(\bar{\sigma}^5). \tag{12}$$

Now the utility of Eqn. (11) becomes clear. Notice that for small h/T,

$$\frac{\partial\bar{\sigma}}{\partial h'} = \frac{\partial\bar{\sigma}}{\partial h}\frac{\partial h}{\partial h'} \approx \frac{1}{T}\frac{\partial\bar{\sigma}}{\partial h},$$

This allows us to use Eqn. (12) , which is expressed in terms of h', to calculate χ. Take the partial derivative with respect to h of both sides to obtain

$$\frac{1}{T} = \frac{\partial\bar{\sigma}}{\partial h}\left(1 - 1/\tilde{T}\right) + 3\bar{\sigma}^2\left(\frac{1}{3\tilde{T}^3} + \frac{1 - 1/\tilde{T}}{\tilde{T}}\right)\frac{\partial\bar{\sigma}}{\partial h} + \mathcal{O}(\bar{\sigma}^4)\frac{\partial\bar{\sigma}}{\partial h}.$$

We can easily solve for the field derivative of $\bar{\sigma}$, which, as $h \to 0$, is the zero field susceptibility. h doesn't appear anywhere, so there is no need to take the limit. If we also use Eqn. (10), which implies

$$\bar{\sigma}^2 \sim \left(1 - 1/\tilde{T}\right),$$

we finally arrive at the result

$$\chi \sim \frac{1}{T}\left[(1 - 1/\tilde{T}) + 3(1 - 1/\tilde{T})\left(\frac{1}{3\tilde{T}^3} + \frac{1 - 1/\tilde{T}}{\tilde{T}}\right) + \mathcal{O}\left((1 - 1/\tilde{T})^2\right)\right]^{-1}$$

from which we can see $\gamma' = \gamma = 1$.

It remains for us to calculate α'. The easiest way to compute C is normally to use the thermodynamic definition

$$C = \frac{\partial U}{\partial T},$$

but we have the free energy and not the internal energy. For computational ease, we will postpone the calculation of α' to the next section, where we introduce an alternate, though equivalent, approach to the WMF equation of state.

VI. MEAN FIELD THEORY FOR THE ISING MODEL II

In this section we discuss another approach to the Ising model, that of Bragg–Williams (BW) theory [22]. BW is appealing because it allows us to extract all of the thermodynamic information from the model, like the free energy and an equation of state, but we do so directly, rather than deriving these quantities from the partition function.[4] Since BW gives us identical information to the WMF approximation treated in the last section, We will be light on the algebra and hit only the high points.

Start with the thermodynamic definition of the free energy:

$$F = \langle E \rangle - TS = \langle \mathcal{H} \rangle - TS.$$

We thus need expressions for the internal energy and the entropy. The entropy is easy, since the number of states is just a binomial distribution. If we denote the number of up spins and down spins as $N_\uparrow$ and $N_\downarrow$ respectively, then

$$S = \ln\left(\frac{N!}{N_\uparrow! N_\downarrow!}\right),$$

where we have set $k_B = 1$ and N is the total number of spins. Applying Stirling's approximation and using the relations

$$\begin{aligned} N_\uparrow &= \tfrac{N}{2}(1+m) \\ N_\downarrow &= \tfrac{N}{2}(1-m), \end{aligned}$$

we would eventually arrive at the expression

$$\frac{S}{N} = \frac{1}{2}(1+m)\ln\left(\frac{2}{1+m}\right) + \frac{1}{2}(1-m)\ln\left(\frac{2}{1-m}\right).$$

Now we go after the internal energy. We want to evaluate

$$\langle \mathcal{H} \rangle = \langle -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \rangle \tag{13}$$

for a fixed value of the per–spin magnetization m. For a fixed magnetization, we have the relation

$$\sum_{i=1}^{N} \sigma_i = mN,$$

---

[4]Though we found an expression for f in the WMF approximation, we will see it does not agree with the BW f. This is because we suffer from multiple counting in the WMF treatment, which leave things like the equation of state unaffected but does affect things sensitive to how we count, like the entropy.

which we can apply to Eqn. (13). If we in addition rewrite the first average in terms of the possible nearest neighbor pairs that can occur, the equation for $\langle \mathcal{H} \rangle$ becomes

$$\langle \mathcal{H} \rangle = -hmN - J\left(\langle N_{\uparrow\uparrow} \rangle + \langle N_{\downarrow\downarrow} \rangle - \langle N_{\uparrow\downarrow} \rangle\right),$$

where the signs on the pair averages arise because like spins give rise to negative energy and unlike spins contribute positive energy. Notice that so far, no approximations have been made. Now is where the BW approximation comes in. We write

$$\begin{aligned}
\langle N_{\uparrow\uparrow} \rangle &= \tfrac{z}{2} N_{\uparrow} P_{\uparrow\uparrow} \approx \frac{z}{2} N_{\uparrow} \frac{N_{\uparrow}}{N} \\
\langle N_{\downarrow\downarrow} \rangle &= \tfrac{z}{2} N_{\downarrow} P_{\downarrow\downarrow} \approx \frac{z}{2} N_{\downarrow} \frac{N_{\downarrow}}{N} \\
\langle N_{\uparrow\downarrow} \rangle &= z N_{\uparrow} P_{\uparrow\downarrow} \approx z N_{\uparrow} \frac{N_{\uparrow}}{N}
\end{aligned}$$

which shows that we have replaced the pair probabilities $P_{\uparrow\uparrow}, P_{\downarrow\downarrow}$, and $P_{\uparrow\downarrow}$ by expressions that assume all spins are uncorrelated. This is the key approximation. With it, we can go on to discover

$$f = -hm - \frac{1}{2} J z m^2 - \frac{T}{2}(1+m) \ln\left(\frac{4}{1-m^2}\right) - \frac{Tm}{2} \ln\left(\frac{1-m}{1+m}\right), \tag{14}$$

leading to an equation of state of the form

$$m = \tanh\left(\frac{Jzm + h}{T}\right),$$

which is identical to Eqn. (7) derived in the WMF approximation.

We promised a value of α', so here it is. The portion of Eqn. (14) that is the internal energy is, assuming $h = 0$, (notice we have switched back to $\bar{\sigma}$, just for consistency's sake):

$$E = -\frac{1}{2} J N z \bar{\sigma}^2 = -\frac{1}{2} T_c N \bar{\sigma}^2,$$

from which we easily compute

$$C = \frac{\partial E}{\partial T} = -T_c \bar{\sigma} \frac{\partial \bar{\sigma}}{\partial T}.$$

We can return to Eqn. (10) to compute the temperature derivative of $\bar{\sigma}$, and use it to substitute in for $\bar{\sigma}$ itself as well. Doing some more algebra finally gives us the result

$$C = \frac{3}{2} N \left[\frac{1}{\tilde{T}^4} - (1 - 1/\tilde{T})\left(\frac{2}{\tilde{T}^3}\right)\right],$$

which shows that C remains finite as we approach T_c. We also see

$$\lim_{T \to T_c} C = \frac{3}{2} N,$$

which would have included a k_B if we had not normalized it to one earlier. Thus, $\alpha' = 0$, and the heat capacity does not diverge at all.

Source	T_c	α	β	γ	ν
Weiss, BW	4.0	0 (jump discont.)	$\frac{1}{2}$	1	$\frac{1}{2}$
Bethe	2.885...	0 (jump discont.)	$\frac{1}{2}$	1	$\frac{1}{2}$
Exact (Onsager)	2.269...	0 (log div.)	$\frac{1}{8}$	$\frac{7}{4}$	1

TABLE I. Table of values for critical exponents calculated from mean field theory. The "source" column describes the form of approximation, explained in detail in the text. The row labeled "exact" is from Onsager's exact solution of the 2D model.

VII. MFT VS. ONSAGER

Table I shows the results of our mean field calculations. Notice we have included a value for ν here, though we did not calculate one. We could have extracted a value for ν from a variational field theory (Landau theory), but that would require too large a digression to simply calculate ν. Instead we used the Josephson law, $\nu d = 2 - \alpha$, with the additional knowledge that this hyperscaling relation is valid for MFT only if $d \geq 4$. We therefore used $\alpha = 0$, which we calculated, and the value $d = 4$. In the table, we have also included the critical temperature in another approximation, the Bethe approximation [23]. In a sense, the Bethe approximation is one step up from the BW and WMF treatments. It tries to include one more level of correlations than does BW and WMF. We give a short treatment of the Bethe approximation in appendix A. We can see MFT does not do such a great job of describing the critical behavior, if we compare our results to the exact values of Onsager [3]. The Bethe approximation does slightly better with T_c, but the exponents have not changed from the BW and WMF cases at all. All of the exponents differ from their exact values, and MFT predicts no singularity in the specific heat at all! This is a common feature of MFT near a critical point. All mean–field theories ignore some higher–order correlations, which become increasingly important as we move towards a critical point. We can, of course, carry approximations to even higher order, but the improvement in values for the exponents is not usually commensurate with the exponential increase in algebra.

VIII. THE MONTE CARLO METHOD

We hope we have demonstrated that man cannot live on MFT alone. One of the most popular and useful techniques for understanding the behavior of model systems is the Monte Carlo (MC) method of computer simulation. Monte Carlo has several advantages. Computers are as fast and as cheap today as they have ever been; this allows the collection of a tremendous amount of data in a relatively short amount of time. This is quite important, since the quality of our MC results depend on the statistical errors in our data sample. MC is also useful in many systems where MFT is exceedingly difficult or impossible; the MC method can give accurate results for models that either have no Hamiltonian (like site percolation) or have a Hamiltonian that is difficult to write down. The Monte Carlo method is not without drawbacks, however. To run Monte Carlo on a model is not the same as solving the model exactly. For example, notice in table I that all of the $d = 2$ Ising model exponents

are simple rational fractions. [5] Monte Carlo cannot ever tell us for sure if an exponent is a rational fraction; we might be able to provide a physical argument for why an exponent assumes a particular value, but Monte Carlo can't unequivocally verify this. Regardless of this detraction, Monte Carlo is without a doubt one of the most important computer methods in all of physics, and its use is certainly not confined to the field of statistical physics.

In essence, the MC method is just a clever (or sometimes not too clever) way of sampling phase space to calculate a steady–state average. If our lattice were small, we could simply enumerate all the N possible states, calculate for each state the energy E and the quantity X in which we are interested, and then compute directly

$$\langle X \rangle = \frac{\sum_{i=1}^{N} X_i e^{-\beta E_i}}{\sum_{i=1}^{N} e^{-\beta E_i}}, \tag{15}$$

giving us the thermal average of X at any temperature which we choose. However, even for a modest sized lattice of Ising spins (a square lattice of linear dimension 16 spins, for example), the computation of the energy of all the states would require of order 10^{80} floating point operations! The MC strategy is therefore to try to produce a set of phase space points $\{\alpha_i\}$, generated according to a probability distribution $P(\alpha_i)$ and to only evaluate the necessary quantities at this select group of configurations, thus replacing Eqn. (15) by

$$\overline{X} = \frac{\sum_{i=1}^{M} (X_{\alpha_i} e^{-\beta E_{\alpha_i}})/P(\alpha_i)}{\sum_{i=1}^{M} \left(e^{-\beta E_{\alpha_i}}/P(\alpha_i) \right)},$$

which we want to equal the thermal average of X at the appropriate temperature [28,30]. One simple choice for P would be the normalized Boltzmann distribution itself; then a convenient cancellation occurs and we find

$$\overline{X} = \frac{1}{M} \sum_{i=1}^{M} X_{\alpha_i},$$

a simple arithmetic average. This choice of P produces an algorithm known as Metropolis [29], a name nearly synonymous with MC itself.

When we try to develop the evolution algorithm for our model, there are many choices for the transition probability that carries us from state to state; the choice of Metropolis ensures a nice cancellation in our estimator for $\langle X \rangle$, but it is by no means the only choice. However, not all of these choices are valid, since we place two constraints on the behavior of the transition probability W. First of all, W must be chosen so that we can reach any allowed state of the system from any other allowed state; this is called the accessibility criterion. The accessibility requirement ensures that we do not restrict ourselves to some select subset of the system under study but can, in principle, visit any allowed state of the system during a simulation. The second restriction we impose is the condition of detailed

---

[5] Thus leading to a famous folklore theorem that *all* critical exponents are rational fractions, though there is no compelling reason to believe this is so.

balance, sometimes called microreversibility. We can always write a master equation for the change in the probability P_a of observing state a

$$\frac{dP_a}{dt} = \sum_{a'} W(a' \to a)P(a') - \sum_{a} W(a \to a')P(a),$$

and we recognize that the left hand side of this equation equals zero in a steady state. However, detailed balance dictates a much stronger constraint; that of *term–by–term* equality in the two sums, which we may express as follows:

$$W(a \to a')P_{\text{eq}}(a) = W(a' \to a)P_{\text{eq}}(a'),$$

where a and a' are any two allowed states of the system. This condition is enough to guarantee that our choice of W leads to the Boltzmann distribution after a sufficient amount of "time" has elapsed. This is a necessary condition if we want to study any system which eventually comes to thermal equilibrium.

The question of the independence of configurations now arises. Obviously if we make a transition to a new state by flipping a single spin, our new state is not statistically independent from our previous state, and this new state should not be used in determining our error. In effect, some time has to elapse (i.e., we have to flip a large number of spins) before the system "forgets" its previous state. One way to determine this correlation time is to calculate the time–dependent correlation function

$$C_M(0,t) = \frac{\langle M(0)M(t)\rangle - \langle M(0)\rangle\langle M(0)\rangle}{\langle M^2(0)\rangle - \langle M(0)\rangle^2}, \tag{16}$$

which is clearly equal to one at $t = 0$, and then decays to zero when $M(t)$ and $M(0)$ become statistically independent. The expected decay of this function is exponential:

$$C_M(0,t) \sim e^{-t/\tau_M},$$

where τ_M is defined to be the correlation time.[6] Unfortunately, like almost everything else, τ_M diverges near a critical point, a phenomenon known as critical slowing down. In fact, the expected divergence is

$$\tau_M \sim \xi^z, \tag{17}$$

where z is a dynamic critical exponent; it depends on our choice of evolution algorithm. Developing so–called acceleration algorithms, which are MC methods where z is made small, is still a lively research area. The exploration of various algorithms and the effect of z on our ability to collect good data in a MC simulation is a large focus of this paper.

---

[6]One might expect that correlations among other system features, like the internal energy, would also have an associated τ. This is in fact the case, but it is the order parameter that is normally the slowest relaxing quantity at a phase transition, so when we speak of "the" correlation time, we refer to the value of τ measured with respect to the order parameter.

IX. EXTRACTING INFORMATION FROM SIMULATION

Before we charge headfirst into MC simulation of the Ising model, some words need be said about how we can get any information from a simulation at all. After all, everything we have said about the divergence of system characteristics, like Eqns. (2), (3), (1), (4), and (17), is phrased in terms of infinite systems, i.e. systems in the thermodynamic limit. We cannot hope to simulate an infinite system on a computer. However, thanks primarily to Michael Fisher, we can extract information about an infinite system at criticality from finite systems. Suppose we have a quantity A that follows a power law near T_c with an exponent ϕ, which may be positive or negative. Then we may write

$$A \sim (T - T_c)^{\phi}, \quad T \gtrsim T_c$$

assuming an infinite size system.[7] Recall that ξ, the correlation length, diverges near T_c as

$$\xi \sim (T - T_c)^{-\nu}, \quad T \gtrsim T_c.$$

Using this equation, we could just as easily characterize the divergence of A as

$$A \sim \xi^{-\phi/\nu}, \quad T \gtrsim T_c.$$

Here is the finite–size scaling ansatz [32,31]. Suppose that for a finite system of linear dimension L

$$A = \xi^{-\phi/\nu} f\left(\frac{\xi}{L}\right), \quad T \gtrsim T_c \tag{18}$$

which is reasonable (or was to Fisher!) if we recognize that near the transition, L and ξ are the only relevant length scales in the problem, by virtue of their macroscopic size. From now on, we just do algebra; the functional form is the intellectual leap. Notice we can rewrite Eqn. (18) as

$$A = L^{-\phi/\nu}\left[\left(\frac{\xi}{L}\right)^{-\phi/\nu} f\left(\frac{\xi}{L}\right)\right], \quad T \gtrsim T_c.$$

We can go still further and write

$$A = L^{-\phi/\nu} \tilde{A}_+\left(\frac{\xi}{L}\right), \tag{19}$$

---

[7]Notice we have stipulated that we are slightly above T_c; this is not central to the discussion, but because the *coefficient* on the power law is in general different above and below T_c, we avoid having to bother with writing two finite–size scaling functions by only considering behavior above T_c. Since ϕ will normally be the same above and below T_c, these arguments apply for temperatures close to but below T_c as well.

where

$$\lim_{L\to\infty} \tilde{A}_+(x) \propto x^{-\phi/\nu}, \tag{20}$$

which ensures that we recover the infinite–size scaling form when we take the thermodynamic limit. An alternate and probably more common statement of Eqns. (19) and (20) is [32]

$$A = L^{-\phi/\nu}\tilde{A}_+\left(L^{1/\nu}(T-T_c)\right), \quad T \gtrsim T_c$$

$$\lim_{L\to\infty} \tilde{A}_+(x) \propto x^{\phi}.$$

There is nothing new here, just a little fiddling with the argument of the scaling function.

Let us emphasize the power of these finite–size scaling forms. If we are able to get some idea of where the transition temperature is, perhaps from the location of a peak in C or χ or the disappearance of the order parameter, we can then collect data near this point. If we then plot

$$AL^{\phi/\nu} \quad \text{vs.} \quad L^{1/\nu}|T-T_c|,$$

then we can treat ϕ, ν, and T_c as fit parameters which we vary until all the data collapse onto a single curve. In addition, if we already know what T_c is to good accuracy, we see that at T_c,

$$A \sim L^{-\phi/\nu}, \tag{21}$$

and the ratio ϕ/ν can be determined from the slope of a simple log–log plot! [36] This all sounds very simple, but there are complications. There are correction to scaling terms that enter into the picture if L is too small; the assumption with these scaling forms is that we are probing a region when ξ is very large compared to the interatomic (lattice) spacing, but small compared to L. Obviously in a square lattice of size $L = 4$ that condition does not hold. There are other difficulties as well, but we will treat more of those as they arise in later sections.

Another comment about finite–size scaling deserves attention. It is certainly difficult to produce a scaling plot with three different fit parameters! It would be nice, under certain circumstances, to be able to determine T_c from some other method, and then simply collect data at T_c and apply Eqn. (21). At first, one might think that T_c could be determined from the location of the peak of χ or C in a finite size system; however, a glance at Fig. 2 shows the difficulty with this procedure. The figure highlights two finite–size effects; a *finite–size shift* in the infinite system T_c, and a *finite–size rounding* in T_c [31], evidenced by the shift of the peak of χ as we increase L and the broadening of the hump respectively. These two effects diminish as we increase L, but we would like simulate systems as small as possible to save computational time. Fortunately, there is another way. Consider the fourth–order Binder cumulant [33]

$$U_L(M) = 1 - \frac{\langle M^4\rangle}{3\langle M^2\rangle^2}, \tag{22}$$

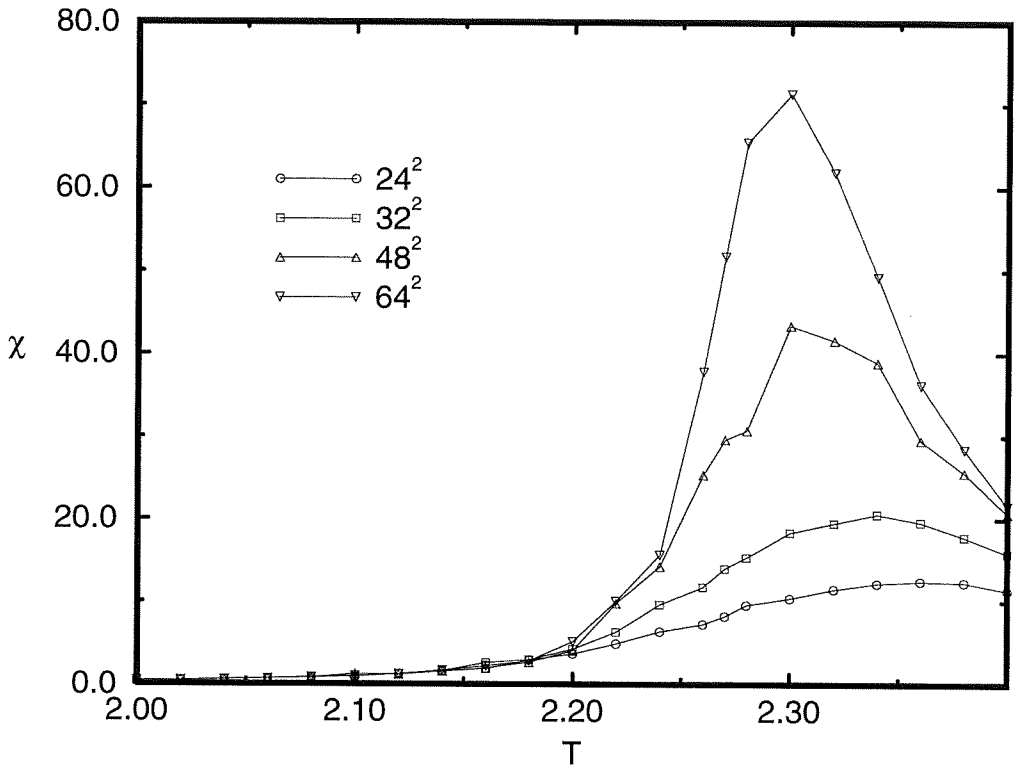

FIG. 2. The magnetic susceptibility of the Ising model as a function of temperature. A rough estimate of T_c can be gleaned from the position of the peak in χ, but the broadening and shifting of the peak as L is changed (discussed in the text) make locating T_c accurately using this method almost impossible. The legend in the figure gives the size of the lattice (number of spins) from which each curve was derived. The lines are drawn as a guide to the eye.

which analyzes the deviation of M from Gaussian behavior. These cumulants are bounded; for $T < T_c$ and $L \gg \xi$, $U_L \to U_\infty = 2/3$, while for $T > T_c$ and $L \gg \xi$, $U_L \to 0$ [30,33]. In addition, cumulants for different values of L cross at a fixed point value U^*. The value of T at which they cross is T_c.[8] In addition, the cumulants are surprisingly insensitive to the system size; systems of a size unsuitable for conventional finite size scaling can have cumulants that give quite accurate results for T_c.

---

[8]Though these cumulants do not work for all model systems, they do for a surprising number, and not just equilibrium Ising models. U_L will resurface again in the $\beta_x - \beta_y$ model considered later in this paper.

X. PRACTICAL CONSIDERATIONS FOR SIMULATION OF FINITE SYSTEMS

We already know that an infinite lattice of Ising spins below the critical temperature will exhibit a spontaneous ordering, leading to a nonzero magnetization even in zero field. However, will the net magnetization be spin up or spin down? To resolve this question in the infinite lattice, we impose an infinitesimal ordering field h on the system. We first take the thermodynamic limit and *then* the zero field limit; this assures the system will pick a direction for M and stay with it. The probability for the *infinite* system to switch over to the other branch of M is zero. Infinitesimals are fine in functional analysis, but they are less than useful for MC simulation. Since we do not impose an ordering field, we have to be careful how we compute averages of M.[9] To compute correctly, we map both branches into the positive M branch [30] by tabulating $|M|$. Though this leads to different coefficients for Eqns. (3) and (1), the critical exponents are unaffected [30]. In addition, all scaling plots we construct of $|M|$ show $|m|$, the absolute value of the per–spin (intensive) magnetization. Though we tend to slip and use either M for m or M for $|M|$ (depending on context), be aware of the practical difference and that we use $|m|$ from a practical standpoint.

Finally, we mention a relationship that we use to determine the dynamical exponent z. Recall the scaling form for the correlation time τ, Eqn. (17),

$$\tau \sim \xi^z.$$

At criticality, $\xi = \infty$. In a finite system, however, it makes no sense to speak of ξ greater than L, the linear dimension of our system. Therefore τ assumes the simple scaling form

$$\tau \sim L^z, \quad T = T_c, \tag{23}$$

a result similar in form to Eqn. (21). Of course this result is by no means accidental; if one puts τ in the role of A in the previous section Eqn. (23) presents itself in short order. We make use of this scaling result for all of our future estimates of z.

XI. DETERMINATION OF ERRORS AND FITTING PROCEDURES

Some statements with regards to the determination of errors for our simulation data are now pertinent. The time–dependent autocorrelation function $C_M(0,t)$ was introduced in Eqn. (16). One procedure for determining error in the data presents itself as follows. Compute the correlation time in a simulation by recording M from a few thousand (say n) trial spin flips and computing

$$C_M(t, t+k) = \frac{\sum_{t=0}^{n-1}[M(t) - \overline{M(t)}][M((t+k) \bmod n) - \overline{M(t)}]}{\overline{M^2(t)} - \overline{M(t)}^2}.$$

---

[9] E is quadratic in the spin variables, so it is not affected by the sign crossing of M.

Now measure the e–folding time of C_M.[10] One can then use this to assign a value for τ_M; we normally used five e–folding times of C_M, just to be safe. However, there are several disadvantages to this method. Primarily, the method is computationally intensive; computing a correlation function from a time series is an $\mathcal{O}(n^2)$ process. In addition, one would measure a different correlation time in every simulation, which yields a data file of variable length making data processing a more tedious task. The advantage of this method is that all samples can be used in determining the standard error of the mean, defined as

$$\sigma_x = \frac{\sigma}{\sqrt{N-1}},$$

where N is the number of *independent* samples.

There is another method one can use to estimate τ, and hence the standard error of the mean, and it is one we use for almost all errors computed in this paper, where appropriate. Suppose we have a time series for the magnetization M at equilibrium. It of course fluctuates around some $\overline{M}$. We want to use only the independent samples in the time series; they are all valid equilibrium measurements, but some samples are correlated. Assume also the time series is very long, i.e. we are in a regime where the law of large numbers holds. If we were to compute $\overline{M}$ and χ, and then do so again using only every tenth data point, for example, we would find very little difference in the two, assuming our now smaller sample is still relatively large. With regards to $\overline{M}$, the reason for this is that the fluctuations are basically random about the mean, and they all contribute equally to the average. The reason that χ changes little is that χ is a measurement of the *average fluctuation* of M from its mean value, which is also unaffected assuming the law of large numbers still holds. These predictions were tested with representative data sets from our Metropolis simulations, and the difference found to be one part in ten thousand or smaller. So we can use all our data in determining averages, but to calculate our errors we still need to know what τ is, and hence how many independent samples we have. We may do this as follows. [44] We compute the variance of our entire data set. Assuming the set has N members,

$$\sigma^2 = \frac{1}{N}\sum_{j=1}^{N}(M_j^{(0)} - \overline{M})^2, \tag{24}$$

where we employ the parenthesized superscript to indicate the iteration of a procedure now described. We bin the data into bins of increasing size, performing an average over each bin. Suppose at the i^{th} step our data set now has N_i members, corresponding to a bin size of b_i; then of course

$$\sigma_{i-1}^2 = \frac{1}{N_{i-1}}\sum_{j=1}^{N_{i-1}}(M_j^{(i-1)} - \overline{M})^2$$

$$\sigma_i^2 = \frac{1}{N_i}\sum_{j=1}^{N_i}(M_j^{(i)} - \overline{M})^2$$

---

[10] One could just as well measure the half–life, but we choose the e–folding time because of the expected exponential decay of the correlation function.

$$\sigma_{i+1}^2 = \frac{1}{N_{i+1}} \sum_{j=1}^{N_{i+1}} (M_j^{(i+1)} - \overline{M})^2$$

and so forth. The binning is continued until we feel the law of large numbers no longer applies, i.e. the bin sizes become of order the size of the time series. By the arguments in the previous paragraph, the sums in the previous four equations will remain approximately the same. Hence, we could reconstruct the variance in the previous iteration by a simple multiplication

$$\begin{aligned} \sigma_{i-1}^2 &= \frac{b_i}{b_{i-1}} \sigma_i^2 \\ &= \frac{b_i}{b_{i-1}} \frac{b_{i+1}}{b_i} \sigma_{i+1}^2 \\ &= \frac{b_{i+1}}{b_{i-1}} \sigma_{i+1}^2, \end{aligned}$$

from which we see the renormalizing factor is always the ratio of bin sizes. However, this is only true *if the data are uncorrelated.* Using this procedure, we can determine the correlation time by performing this iterative procedure until we *do* recover the variance we expect, at which point we may say the previous bin size was large enough to ensure uncorrelated samples. From this we may determine the number of independent samples by simply dividing the length of the time series by our computed bin size. Is this way we do not discard data points for averages, but still obtain an honest estimate of our errors. We computed $\overline{M}$ and $\sigma_{\overline{M}}$ this way and compared it to results obtained by not binning and merely using only those points deemed to be independent. The difference was again minimal, supporting the validity of this technique.

A quantity like $\overline{M}$ is an average over configurations, so we used the binning technique above to compute its standard error. However, to be as confident about our error as possible, we actually tabulated the mean of $\ln(\overline{M})$ over several data sets and used standard error analysis along with our known value of $\sigma_{\overline{M}}$ to determine the error bars on the logarithm of $\overline{M}$. The result of this computation for each L is the source of the error bars in Fig. 4. For C and χ, however, the computation of errors cannot proceed in the same fashion. C and χ were computed from the ensemble averages

$$\begin{aligned} C &= \tfrac{1}{T^2} \left(\langle E^2 \rangle - \langle E \rangle^2 \right) \\ \chi &= \tfrac{1}{T} \left(\langle M^2 \rangle - \langle M \rangle^2 \right) \end{aligned}$$

so multiple ensembles are needed to compute both mean values and errors in the mean. This is exactly what was done. The data were naturally collected in multiple sets (normally five or ten total), so a χ and a C were computed for each set. Then $\ln \chi$ and $\ln C$ were computed for each set, and the mean and standard error computed over the multiple sets. We are justified in treating each set as an independent ensemble, since a different random number seed was used for each time series. The results of these calculations appear as the error bars in Figs. 5 and 6. Unless otherwise noted, all errors were obtained using these two techniques.

We now discuss the fitting procedures used to determine values for critical exponents. For almost all results in this paper, we attempt to first determine T_c and then use the finite–size scaling form of Eqn. (21) to determine critical exponents. When one looks at Eqn. (21), a simple linear least squares fit to a plot of $\ln A$ vs. $\ln L$ seems a natural way to extract the exponents. However, a glance at Eqn. (23) should be enough to convince you of the flaw in this procedure. The growth of τ as we increase L means we will be dealing with fewer and fewer independent samples, which in turn dictates the growth of errors as L grows. The simple linear least squares method (LSQ) is a poor choice for this case, as LSQ is suitable for a data set where the errors in each point remain essentially constant. We therefore use a method that weights each point differently according to the error in that point. Assuming a linear model and minimizing χ^2 leads to formulas that resemble conventional LSQ, with the exception that the points are weighted by the inverse of their variance. [43] In effect, this allows much more "wiggle" in the fit line than conventional LSQ, leading to correspondingly larger, and we feel more honest, estimates of errors in the fit parameters. While there are some tacit assumptions made in this procedure (like the assumption that errors are normally distributed) that do not always make it suitable [43], we encounter none of these dangers here. When necessary, we will refer to this variance weighting method as weighted least squares (WLSQ). Some of the linear fits we will perform are suitable for conventional least squares, but we will always be explicit about whether we use the LSQ or WLSQ methods.[11]

XII. MC FOR GLAUBER DYNAMICS I: METROPOLIS

We concentrate first on MC simulation of the $d = 2$ Ising model, in zero applied field, with Glauber (spin–flip) dynamics [17]. For the case of Glauber dynamics, we imagine the coupling of our lattice to a heat bath of temperature T that induces spins to spontaneously flip from up to down or vice versa. In addition, the coupling of a spin to its nearest neighbors mitigates this flipping process; spins are more likely to flip if the result is an overall lower energy state for the system.

The first algorithm we apply to the model with Glauber dynamics is the famous algorithm of Metropolis [29]. The transition probability (i.e., the probability to flip a spin) is given by

$$P = \begin{cases} 1 & \text{if } \Delta E \leq 0 \\ e^{-\Delta E/T} & \text{otherwise} \end{cases} .$$

We may select the spins either sequentially or at random, but for all simulations we perform in this paper the spins are selected at random. On average, then, if we have a square lattice of L^2 spins, a flip attempt will occur at each site if we perform L^2 flips. For the Metropolis algorithm, this is what we define as one Monte Carlo update per site (L^2 flip attempts), which for brevity's sake we will always refer to as one Monte Carlo step (MCS).

---

[11] An alternative to this algebraic procedure (which a computer can easily perform) is a purely graphical technique (which requires a person), whereby one draws the steepest and shallowest fit lines consistent with the data. One then takes the mean of the two slopes as the best line, and the error is the absolute difference between the best line and either of the two extreme lines.

In addition, for all simulations we use periodic boundary conditions, which consists of thinking of the topmost and bottommost rows of the lattice as neighbors and similarly for the leftmost and rightmost columns. Thus a $d = 1$ linear chain becomes a ring and the $d = 2$ square lattice is wrapped onto the surface of torus. Though this is a convenient topological image, let us emphasize we are most certainly simulating a 2D system and not a 3D one; each spin on the lattice has four nearest neighbors as for a 2D square lattice. The reason for imposing these boundary conditions, rather than leaving the edges of the lattice free, is the following. It has been shown [32,42] that free edges give rise to additional surface terms in the scaling functions discussed above. For an L by L square lattice, the ratio of perimeter (free edge) sites to the total number of sites is

$$n = \frac{4L - 4}{L^2}.$$

Even for a system of $L = 32$, n is still about twelve percent! We would like to extract as much information out of small systems as possible, and these surface terms hurt our ability to do so. The addition of periodic boundary conditions to the problem eliminates this surface problem. However, there is a cost; with periodic boundary conditions, the two most distant spins in our system are $L/2$ blocks apart, rather than L as in the free system.

Another practical question with regards to the simulations now arises; what kind of configuration do we start with? We really can only be sure about the appearance of two configurations; the $T = 0$ Ising lattice is filled with spins of only one type, and the $T = \infty$[12] configuration corresponds to each spin being oriented up or down at random. We could therefore "warm up" from the $T = 0$ configuration or "cool down" from the $T = \infty$ configuration. We could certainly choose either method, but there are practical considerations that prevent us from using either. For the case of randomly oriented spins, performing simulations at finite temperature with this configuration corresponds to a drastic quench in temperature. The system does eventually equilibrate, but this relaxation becomes slower and slower as we approach T_c, leading to us having to wait longer and longer (in MCS) until we can begin to accumulate data. For the case of the perfectly ordered configuration, the problem is also one of slowness; once the system is "frozen" in this fashion, flipping enough spins to grow a sizeable domain is tremendously difficult, just from probabilistic considerations.

To avoid both of these problems, a hybrid of the two procedures is used. The system of a given size is placed in a random ($T = \infty$) configuration, and one long simulation in conducted in which no data were collected and a quench to a temperature $T_1 < T_c$ is performed. This configuration is then saved and used to begin a simulation at T_1; the ending configuration of the run at T_1 is then saved and used to begin a run at the next temperature $T_2 > T_1$. However, simulations at each additional temperature, beginning with T_2, begin with a "warmup" period of 5,000 - 10,000 MCS. This procedure is then continued for all subsequent simulations. The temperature increments are small to try to minimize effects attendant to establishing equilibrium at the new temperature. We could equivalently

---

[12]We mean by $T = \infty$ the precise statement $T > R$ for any $R > 0$.

develop a procedure to "cool" the system down, and we verified similar qualitative behavior for such a procedure.[13]

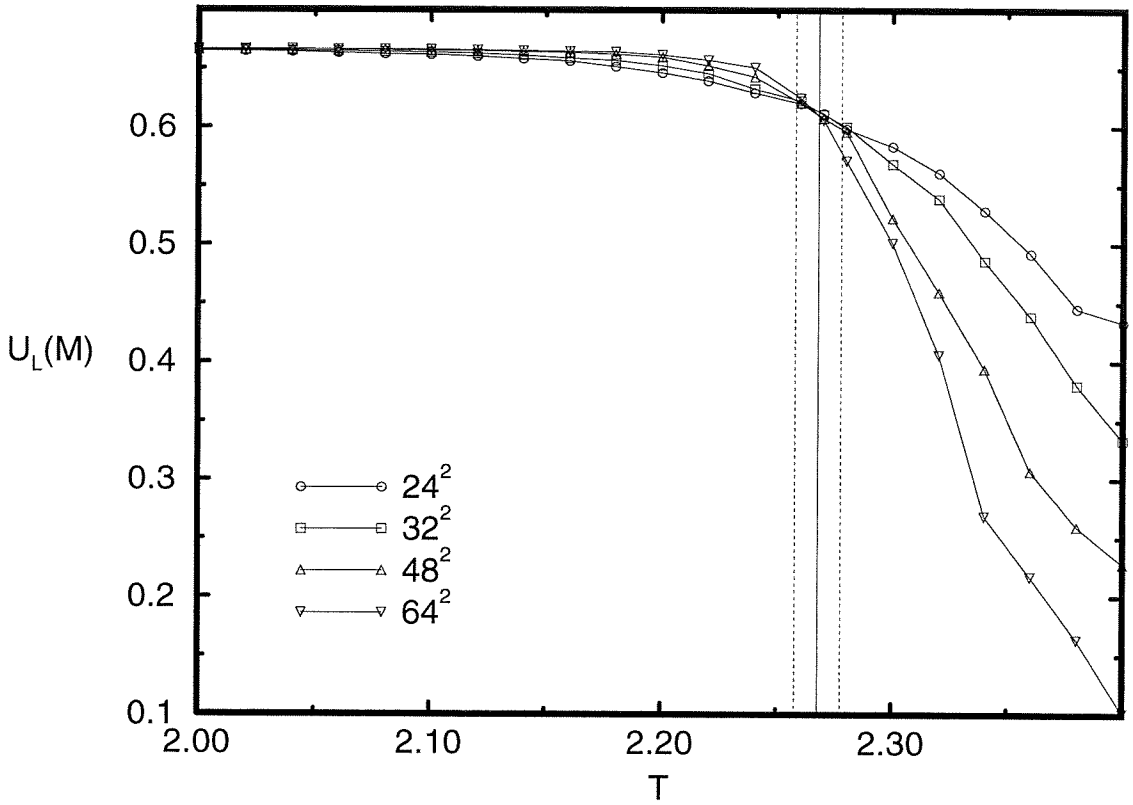

FIG. 3. Binder cumulants (Eqn. (22)) for the $d = 2$ equilibrium Ising model. The curves represent equilibrium averages over 4×10^4 configurations, with data obtained using the Metropolis algorithm. The solid line is the estimate of T_c, and the two dotted lines give the error in this estimate. From this figure, a value of $T_c = 2.268(10)$ was extracted. The lines drawn that connect the data are an aid to the eye.

We now present MC simulation data for the Metropolis algorithm. After placing rough bounds on the location of T_c, obtained from the position of the peak in the susceptibility, a finer temperature scan was conducted. A total of 4×10^5 MCS were run, over four separate data sets. However, to try to anticipate the effect of the correlation exponent z, data were only collected every 10 steps. This resulted in a final sample size of 4×10^4 configurations. From these data, we analyzed the crossing points of the cumulant $U_L(M)$ (Eqn. (22)) for L values of 24, 32, 48, and 64. Fig. 3 shows these cumulants. The best estimate for T_c was obtained by noting the crossing points of U_L of the two largest systems, size 48^2 and 64^2. Error estimates were obtained by noting other cumulant crossings. From this analysis, we determined a value for T_c of 2.268(10). Since these data were not judged to be of sufficient

[13]In general this statement does require the verification we conducted, since if the transition were first–order hysteresis effects would have played a role.

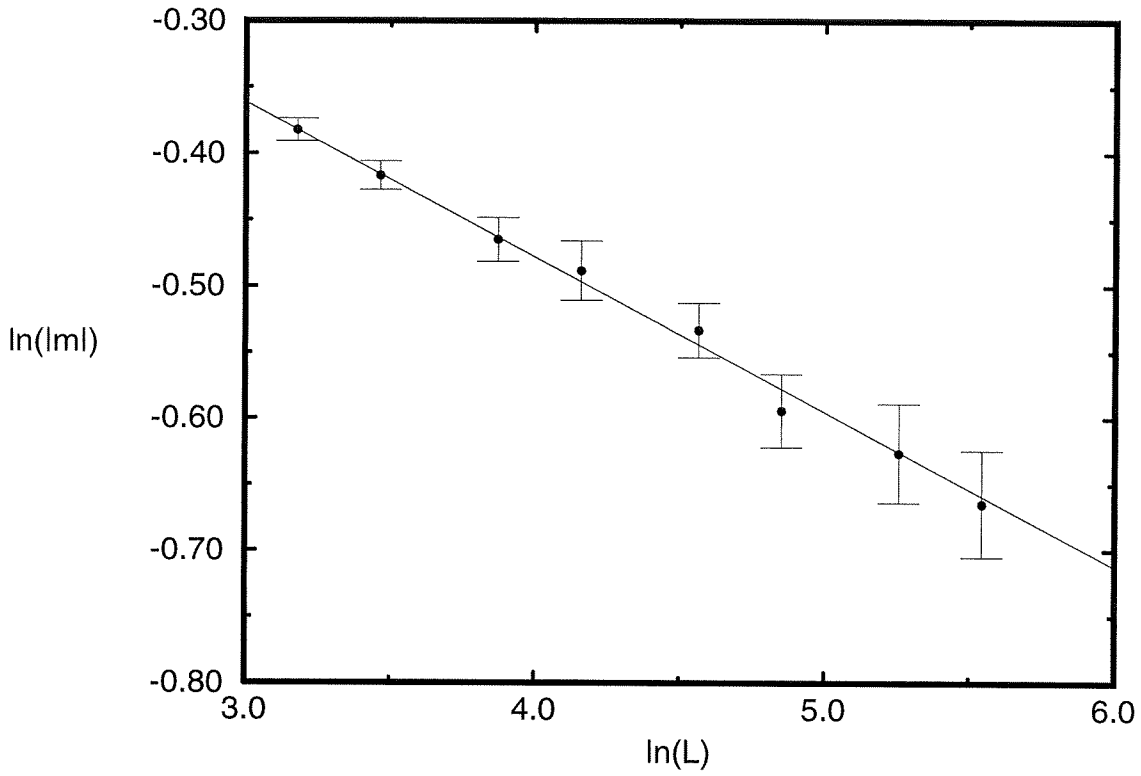

FIG. 4. Finite–size scaling plot of the absolute value of the magnetization per spin at the value $T_c = 2.268$, for square systems of linear dimension 24, 32, 48, 64, 96, 128, 192, and 256. The data were obtained for Glauber dynamics using the Metropolis algorithm; an ensemble size of 5×10^4 is shown. From a weighted least squares linear fit to the data (shown as the solid line) the value $\beta/\nu = 0.12(1)$ was extracted.

quality for conventional finite–size scaling, data were collected at the nominal critical point $T_c = 2.268$, for square systems of linear dimension 24, 32, 48, 64, 96, 128, 192, and 256. Log–log, and in the case of the specific heat log–linear, plots were then produced of $\langle M \rangle$, χ, and C, with the expectation that they would lie on a straight line with slope given as in Eqn. (21). The results of 5×10^4 configurations are shown in Figs. 4 - 6. Notice that for C, α is zero. The divergence is logarithmic, so the expected behavior is

$$C \sim C_0 \ln(L), \tag{25}$$

and the figure is a log–linear plot rather than a log–log one. The behavior of the errors in the points leads us to use the weighted least squares (WLSQ) procedure earlier discussed. From the linear fits, shown in the figures, we extracted the exponent values $\beta/\nu = 0.12(1)$ and $\gamma/\nu = 1.7(1)$, along with the value $C_0 = 0.44(3)$. ν was not measured, but rather the known value $\nu = 1$ for the $d = 2$ Ising model was used to conclude $\beta = 0.12(1)$ and $\gamma = 1.7(1)$. As one can tell by looking at Figs. 4 - 6, the quality of the fit varies. For M and χ it is relatively good, but quite poor for C. In fact, without prior knowledge of the value of α, one could not determine that a plot of $\ln C$ vs. $\ln L$ does not yield a straight line. This

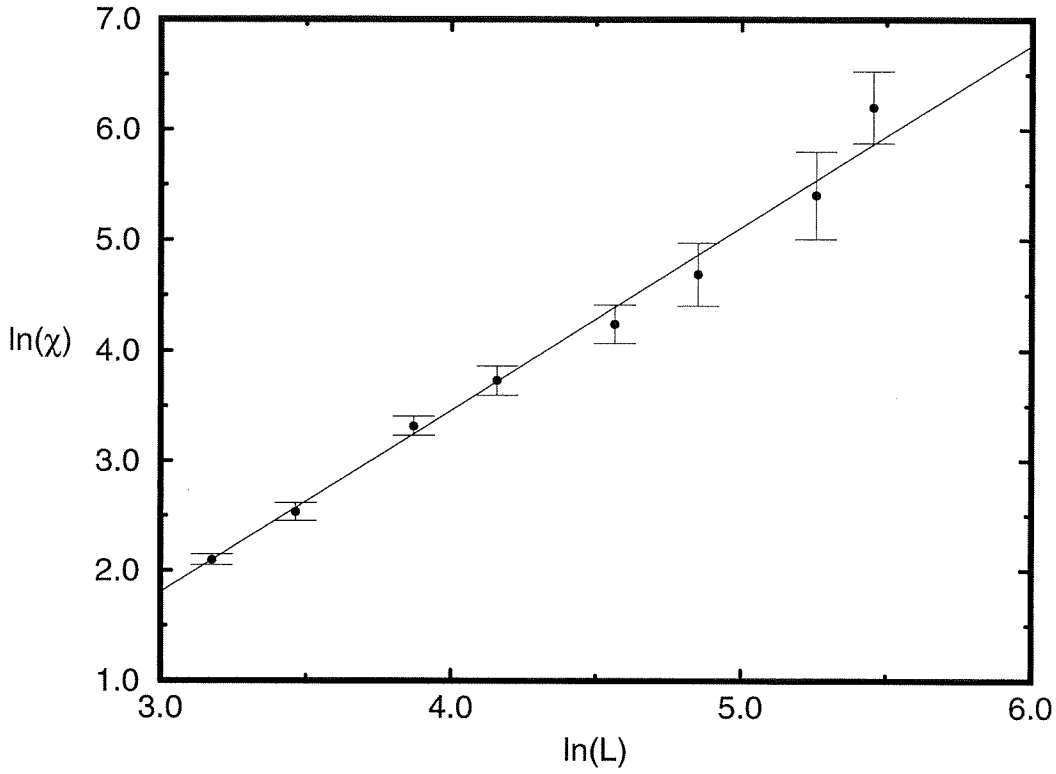

FIG. 5. Finite–size scaling plot of the magnetic susceptibility for the Metropolis algorithm. From a weighted linear fit (described in the text) to the data, shown as the solid line in the figure, the value $\gamma/\nu = 1.7(1)$ was extracted.

situation will change with the Wolff algorithm, however. It is also important to note that the errors become larger as L increases. This is not an artifact of the data analysis; it is a direct result of Eqn. (17) . As L increases, the number of independent samples contributing to the error decreases, and hence our error increases. To determine what kind of correlation time we expect to see as we increase L, we can measure z by computation of the correlation function, averaged over many trials. If we perform this calculation at the nominal critical temperature, we can then apply Eqn. (23) and determine the value of z from the slope of a log–log plot. We did so for the Metropolis algorithm, and we present the results in Fig. 7. The fit was performed with conventional least squares (LSQ) since the errors in the individual points are not known. As we noted before, computation of the correlation function for a large time series is a computationally intensive process. However, assuming a stationary distribution (as we have with the equilibrium time series for M), $C_M(t)$ is simply related to Fourier transforms of M. The derivation of this relationship is the focus of appendix B. This being the case, we can use an FFT algorithm to reduce the computation to an $\mathcal{O}(n \log_2 n)$ process. To produce Fig. 7, this was done for square systems of linear dimension 4, 8, 12, 16, 24, 32, 48, 64, 96, and 128. The correlation time for each system was averaged over 10^4 independent trials; the points denoted by an "x" in the figure were not included in the linear fit, as finite size effects seemed dominant there. The data in Fig. 7

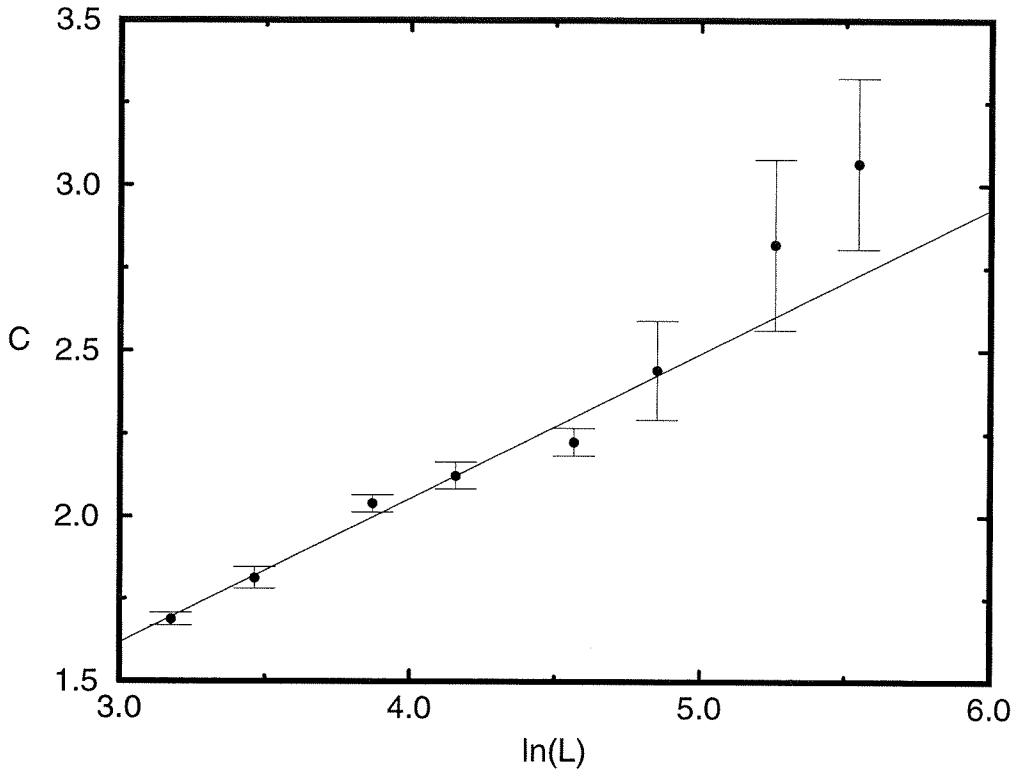

FIG. 6. Finite–size scaling plot of the specific heat for the Metropolis algorithm. From a weighted linear fit to the data (the solid line) the value $C_0 = 0.44(3)$ was extracted.

give a value $z = 2.19(7)$. The results from this section, as well as all other equilibrium MC exponents, are summarized in table II.

XIII. MC FOR GLAUBER DYNAMICS II: WOLFF

We now study the $d = 2$ model with Glauber dynamics using an algorithm introduced by Wolff [34]. The purpose of Wolff's algorithm, which is an improvement on an earlier algorithm of Swendsen and Wang [35], is to reduce the value of z. If we were able to do so, we would certainly reduce our errors and get data of higher quality than in the Metropolis case for a comparable amount of computer time. The idea behind the Wolff algorithm is this. Select a spin at random. Visit the four nearest neighbors of this spin and add them to a cluster with probability

$$P = \begin{cases} 1 - e^{-2/T} & \text{if } \sigma_i = \sigma_j \\ 0 & \text{otherwise.} \end{cases}$$

We then continue this process with each new site added to the cluster, until all perimeter sites of the cluster have been visited. This cluster of spins is then flipped *en masse*. The growth and flipping of a cluster is defined as one MCS in the Wolff algorithm. Data from

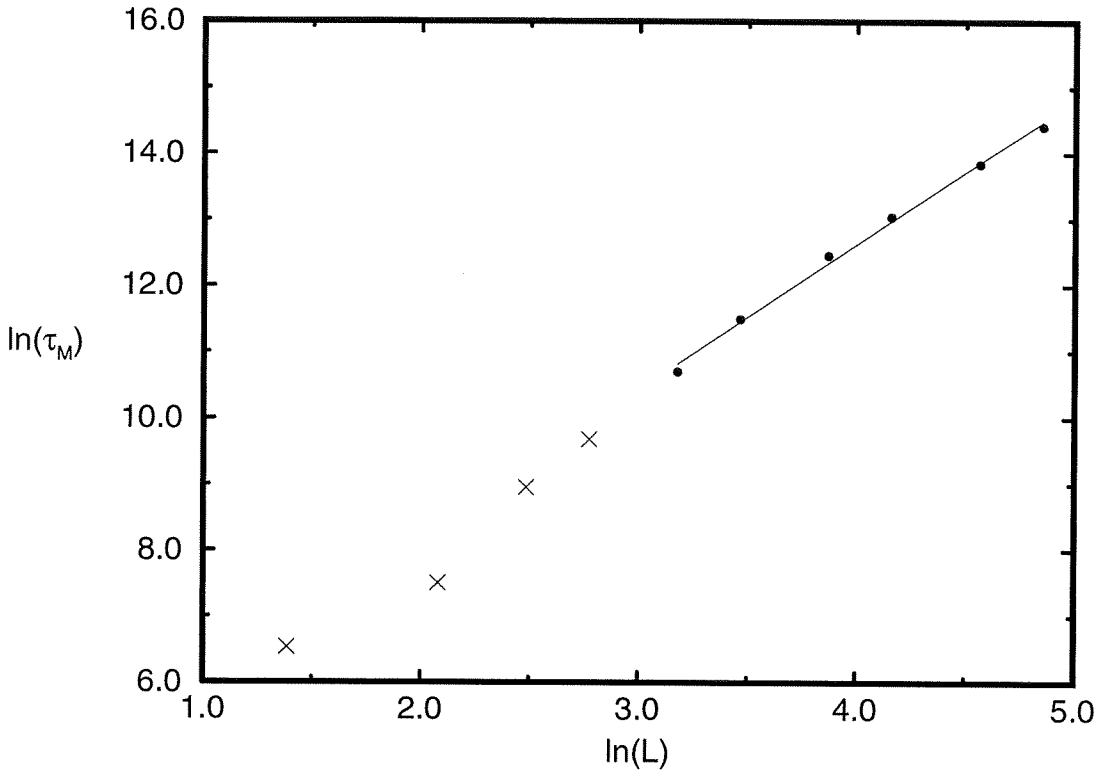

FIG. 7. Finite–size scaling plot of the magnetization correlation time, τ_M, at $T_c = 2.268$. τ_M was measured in units of spin flips for square systems of size 4^2, 8^2, 12^2, 16^2, 24^2, 32^2, 48^2, 64^2, 96^2, and 128^2. Data were averaged over 10^4 independent trials. The points marked with an "x" were not used in determining z; a simple LSQ linear fit (the solid line) to the remaining data yields a value $z = 2.19(7)$.

this algorithm representing 2×10^5 different configurations is presented in Figs. 8 - 12. We used square systems of linear size 24, 32, 48, 64, 96, 128, 192, and 256 to obtain these data. Instead of redetermining the critical point, the T_c determined in the Metropolis algorithm (2.268) was used again. From the finite size scaling data at this nominal value of T_c and the known result $\nu = 1$, we determined $\beta = 0.112(2)$, $\gamma = 1.70(1)$, $\alpha = 0$, and $C_0 = 0.465(3)$. An additional estimate of $\gamma = 1.655(4)$ is given by the finite size scaling behavior of the mean square magnetization of the flipped cluster, computed according to

$$\langle |c| \rangle = \frac{1}{L^2} \left\langle \left(\sum_{\sigma_i \in c} \sigma_i \right)^2 \right\rangle$$

where the sum is over all sites in the cluster.

Notice from Fig. 12 that the Wolff data are of sufficient quality to determine that C diverges logarithmically; the plot of $\ln C$ vs. $\ln L$ shows a more definite curvature than that of C vs. $\ln L$. While error bars are not included for any quantity except M (because only two independent sets of data were used), the increase in precision is clear. The quality of

	G_M	G_W	K_{nn}	$K_{nn+\infty}$	K_∞
T_c	2.268(10)	—	—	—	2.27(4)
C_0	0.44(3)	0.465(3)	—	—	—
α	0	0	—	—	—
β	0.12(1)	0.112(2)	—	—	0.19(3)
γ	1.7(1)	1.70(1)	—	—	1.7(1)
z	2.19(7)	≈ 0	3.6(8)	2.4(1)	1.5(3)

TABLE II. Summary table of values for critical exponents obtained from simulation of the equilibrium 2D Ising model. All exponents (the leftmost column) are as defined in the text. The algorithms are, from left to right, Glauber dynamics with Metropolis, Glauber with Wolff, Kawasaki dynamics with nearest neighbor exchanges, Kawasaki with mixed exchanges, and Kawasaki with arbitrary range exchanges. A value of "—" in the table indicates the value was not measured for that algorithm.

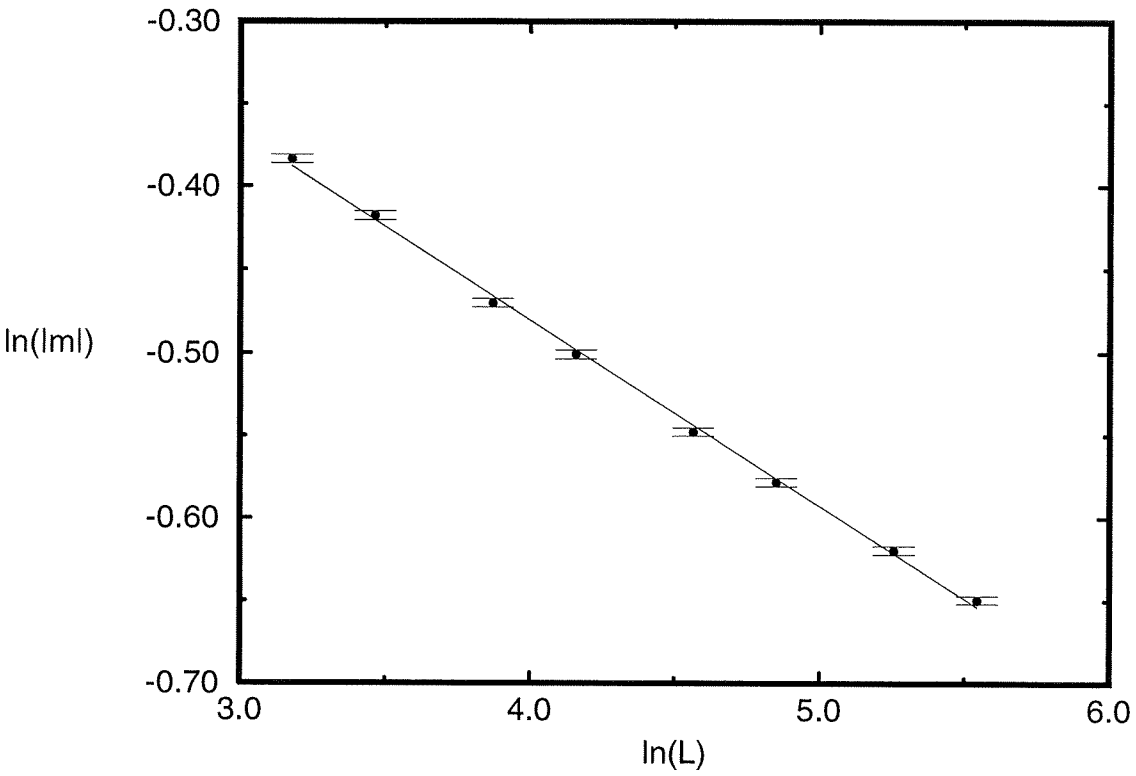

FIG. 8. Finite–size scaling plot of the absolute value of the magnetization, at $T_c = 2.268$, for square systems of linear dimension 24, 32, 48, 64, 96, 128, 192, and 256. The data were obtained for Glauber dynamics using the Wolff algorithm. An ensemble size of 2×10^5 is shown. A conventional least squares linear fit to the data (the solid line) yielded the value $\beta/\nu = 0.112(2)$.

fit is undeniably better than in the Metropolis case. While the Wolff results do reflect four times as many data as were taken in the Metropolis case, the reduction of error is much

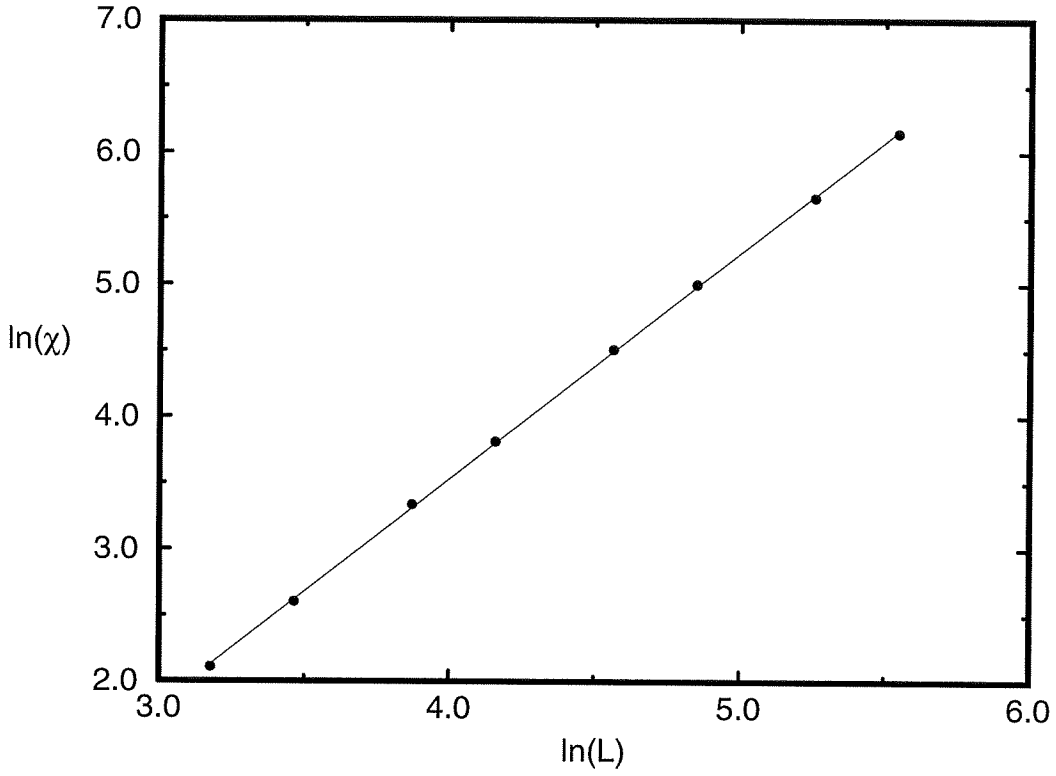

FIG. 9. Finite–size scaling plot of the magnetic susceptibility for the Wolff algorithm. From a simple linear (LSQ) fit to the data, shown as the solid line, the value $\gamma/\nu = 1.70(1)$ was extracted.

greater than twofold, which is what we would expect just from statistical considerations. The great virtue of this algorithm is the increase in *independent* samples via a reduction in the correlation time and the dynamical exponent z. Most striking is the trend in errors in Fig. 8. It appears the error is *nearly constant* over a wide range of L values; this would imply $z \approx 0$.[14] By measuring the bin size necessary to achieve uncorrelated samples, we arrived at a value for z consistent with 0, which is a tremendous reduction from the Metropolis case! Fig. 13 shows the results of our estimation of z. As in the Metropolis case some of the smaller L values were discarded, since finite size effects seemed significant there. We state in the captions for Figs. 8 - 11 that conventional least squares fits were used to calculate the values for the critical exponents; it is precisely our expectation that all errors in quantities measured using the Wolff algorithm are similar to those in Fig. 8 that allows us to fall back to such a method, rather than using the WLSQ method discussed earlier.

---

[14]To say $z = 0$ is not to imply the correlation time is only one update; z may not grow with L, but we have made no claim about the coefficient! On average, the correlation time was around 10 MCS for the Wolff data.

XIV. COMMENTS ON THE GLAUBER DYNAMICS DATA

It should be clear by glancing at table II that the use of the Wolff algorithm greatly improves our ability to measure accurately the values of critical exponents. Look back at table I for Onsager's exact results, and we see we have done well with Monte Carlo. Compare also the values of C_0 to the known value $C_0 \approx 0.4995$ [36]. When discrepancies between simulation and exact results are evident, we can offer three possible reasons for the differences. Firstly, we are not right at criticality; our estimate of T_c is slightly subcritical. Secondly, we have used an exact value of ν rather than a measurement; were we to have measured ν, it too would have an associated error and the errors in our other exponents would have reflected this. Thirdly, we have presented data representing only one decade in L; corrections to scaling become less important as L increases, so data for even larger values of L might yield slightly different — and certainly more accurate — values for the critical exponents.

These remarks deserve special application to the case of C_0; neither the Metropolis nor Wolff algorithms yielded a value of C_0 that agrees with the expected value within statistical error. The Metropolis result is hardly severe; the measured C_0 is only 2σ away from expectation, and the fit is so poor as to render our result almost meaningless. A WLSQ fit to the data set less the first two points yields $C_0 = 0.5(1)$, and this may be a more realistic estimate. The Wolff data are a bit more alarming; we are 10σ away from the expected value, and the points fit the line quite well. Here we must appeal to an both an insufficient amount of data and a slight underestimate of T_c. A conclusive statement about C_0 would require improvement in both these areas. However, despite these observations, both the reliability and utility of the Wolff algorithm are clear.

XV. MC FOR KAWASAKI DYNAMICS I: NEAREST NEIGHBOR EXCHANGES

Before moving on to a nonequilibrium variant of the Ising model, we must verify that when we do so we are really comparing apples and apples, so to speak. The reason for this is that the dynamics of the nonequilibrium model is basically Kawasaki (spin–exchange), introduced by Kawasaki to the equilibrium Ising model as a way to study spin diffusion in ferromagnets. [18] The essential thrust of Kawasaki dynamics is the introduction of a dynamics obeying a conservation law. The algorithm for conventional Kawasaki dynamics (which we will denote as K_{nn}) is as follows. Choose a spin and one of its nearest neighbors at random. Compute the energy change in the system if we exchanged the two spins; perform the exchange with probability

$$P = \begin{cases} 1 & \text{if } \Delta E \leq 0 \\ e^{-\Delta E/T} & \text{otherwise.} \end{cases} \tag{26}$$

Notice that in this algorithm, M is strictly conserved. The spin–exchange algorithm is essentially Metropolis, so we do expect a Boltzmann distribution in the long time limit. However, the low–temperature equilibrium states are very different; if we perform Kawasaki simulations with $M = 0$, the best the lattice can do at zero temperature is to minimize the interface between the spin up and spin down domains. This leads to a "striped" ordering;

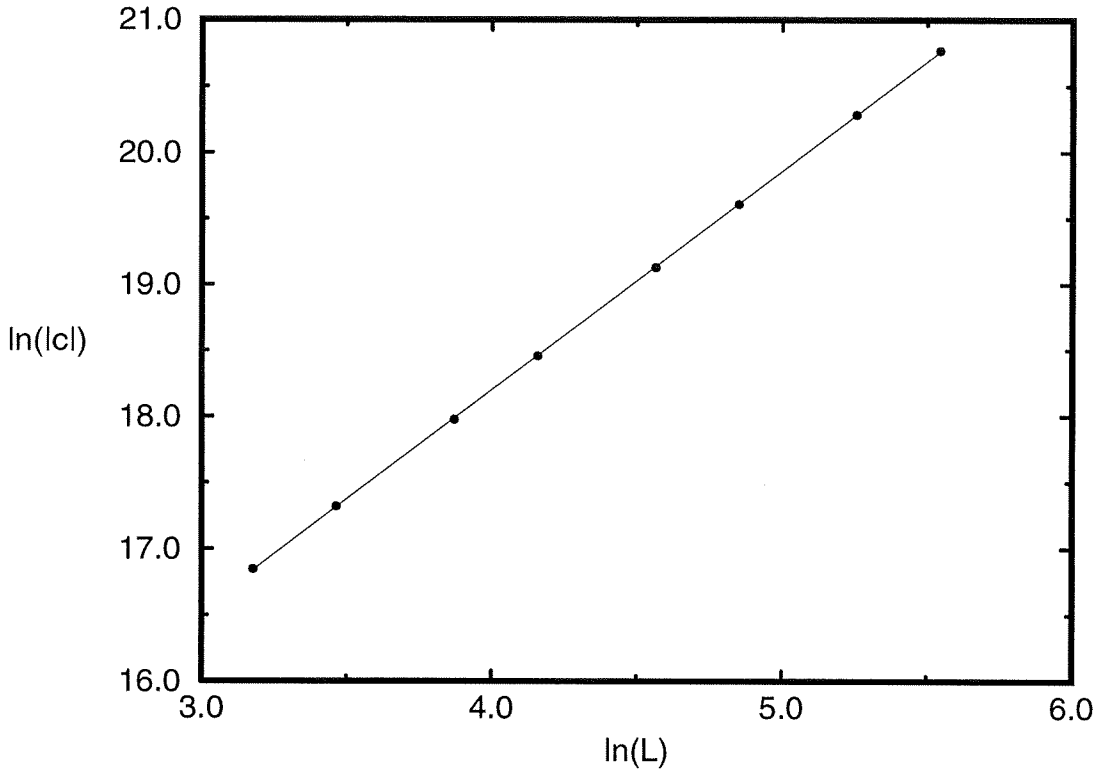

FIG. 10. Finite–size scaling plot of the mean square magnetization of the flipped cluster for the Wolff algorithm. This quantity is an estimator for χ. A simple linear (LSQ) fit to the data (the solid line) yielded the value $\gamma/\nu = 1.655(4)$.

at zero temperature there are only two domains with a perfectly sharp interface between them. If we believe the principle of detailed balance, then the static behavior of the Ising model should not change with this new dynamics; it remains to test this hypothesis.

The conservation of M adds another wrinkle to our analysis of the model; M is no longer a good order parameter. Instead, we use [37]

$$\Psi = S(2\pi/L_x, 0) + S(0, 2\pi/L_y), \tag{27}$$

which is the longest wavelength component of the structure function. The structure function is strongly related to scattering amplitude, which is what an experimentalist measures. If we define $\rho(q)$ as the Fourier transform of the charge density, then

$$S(q) = |\rho(q)|^2 = \rho(q)\rho^*(q).$$

For our purposes, some simplification can be applied to S to arrive at Eqn. (27). Instead of computing a two–dimensional Fourier transform, we compute two one–dimensional transforms. We first proceed from column to column in the lattice, summing up the magnetization in each column. The column magnetization, which we may think of as a kind of charge density, is multiplied by a phase factor; column k is multiplied by

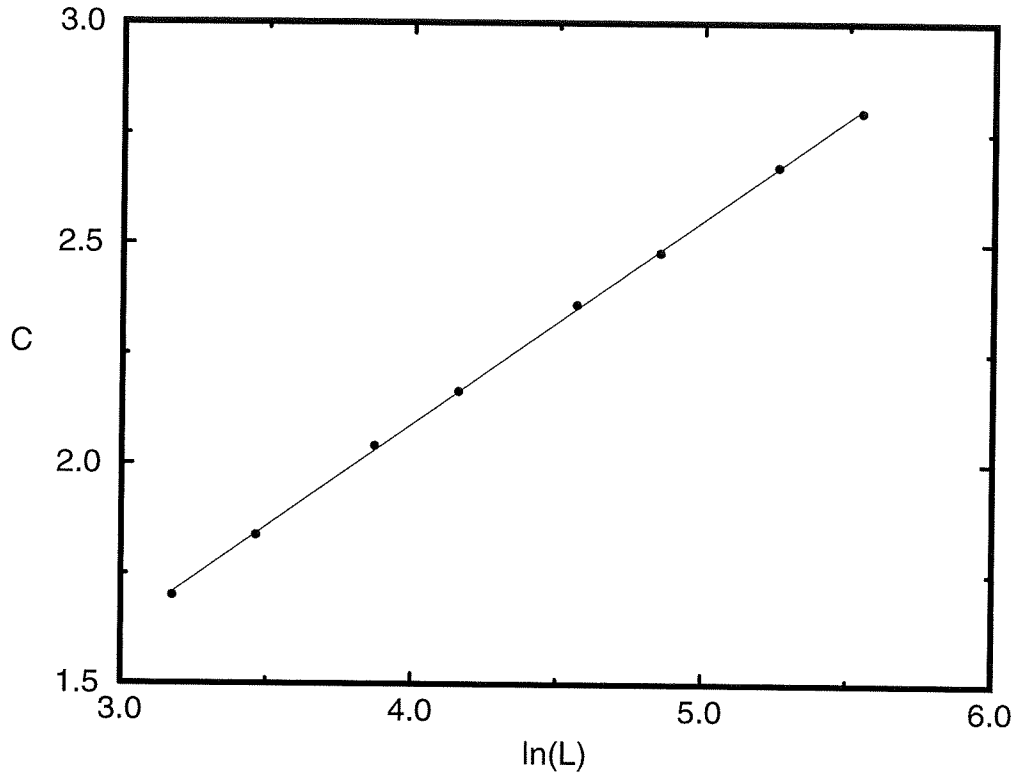

FIG. 11. Finite–size scaling plot of the specific heat for the Wolff algorithm. A linear fit to the data (the solid line) yielded $C_0 = 0.465(3)$.

$$e^{(2\pi/L_x)k}.$$

The sum over all columns of the column–wise "spin density" times the appropriate phase factor is the complex number $\rho(q)$; we compute its square modulus to determine $S(2\pi/L_x, 0)$. If we proceed in exactly the same fashion with the rows of the lattice and use the phase factor

$$e^{(2\pi/L_y)k}$$

for row k, we can compute $S(0, 2\pi/L_y)$. They are then combined to form the order parameter Ψ of Eqn. (27). The contribution along both axes is summed because the direction of ordering is arbitrary; the definition of Ψ above covers both cases. We also normalize Ψ so that it is zero at infinite temperature. Given this new definition of Ψ, we can also define a "susceptibility"

$$\chi = \langle \Psi^2 \rangle - \langle \Psi \rangle^2 \tag{28}$$

which is not as physically meaningful as the magnetic susceptibility, but does perform the same function; it characterizes fluctuations in the order parameter. Notice also that if we compute the *extensive* Ψ, that is we do not normalize the mean row and column

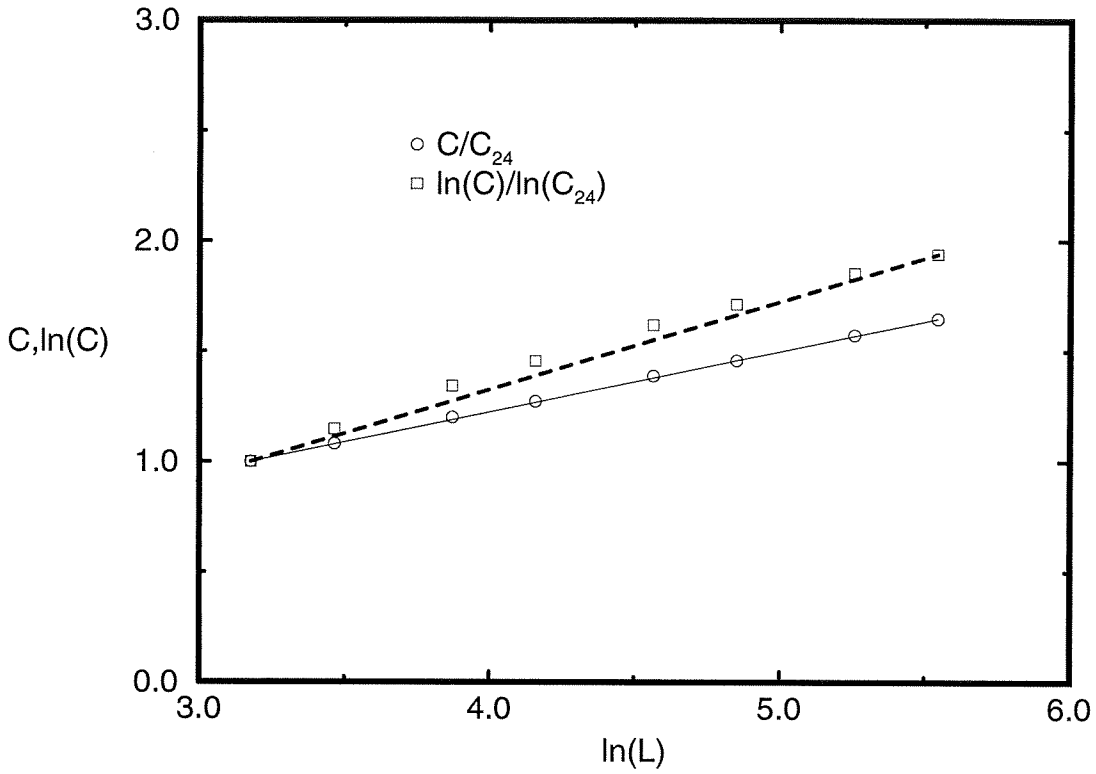

FIG. 12. Finite–size scaling plot of the specific heat for the Wolff algorithm. The more pronounced curvature in the log–log plot of C leads us to the conclusion $\alpha = 0$ and C has the scaling form given by Eqn. (25). The dotted and dashed lines are drawn to emphasize the straightness of the log–linear plot and the curvature of the log–log plot. In the interest of clarity, the two curves have been normalized by their smallest values; this is indicated in the legend.

magnetizations to be between ± 1, χ is properly extensive. We can again estimate T_c by considering cumulants of Ψ [37], naturally defined as

$$U_L(\Psi) = 1 - \frac{\langle \Psi^4 \rangle}{3\langle \Psi^2 \rangle^2}.$$

If one thinks about the progress of the Metropolis algorithm applied to Kawasaki dynamics, it is not difficult to see that many attempted exchanges will not be meaningful; exchanging two like spins accomplishes nothing. We therefore use a variable (also called continuous) time MC algorithm to increase computational efficiency. We keep lists of nearest neighbor pairs of antiparallel spins and make all exchanges from these lists; we thereby ensure every successful exchange results in a new (distinguishable) state of the system. The consequence of such a choice causes us to reset the time scale in such a way as to allow a variable–length time step. The idea is as follows. Let Γ_k be the rate for process k, and let N_k be the number of places where process k can occur. In the context of the Kawasaki simulations, the process (there is only one) is spin exchange and the place where it can occur

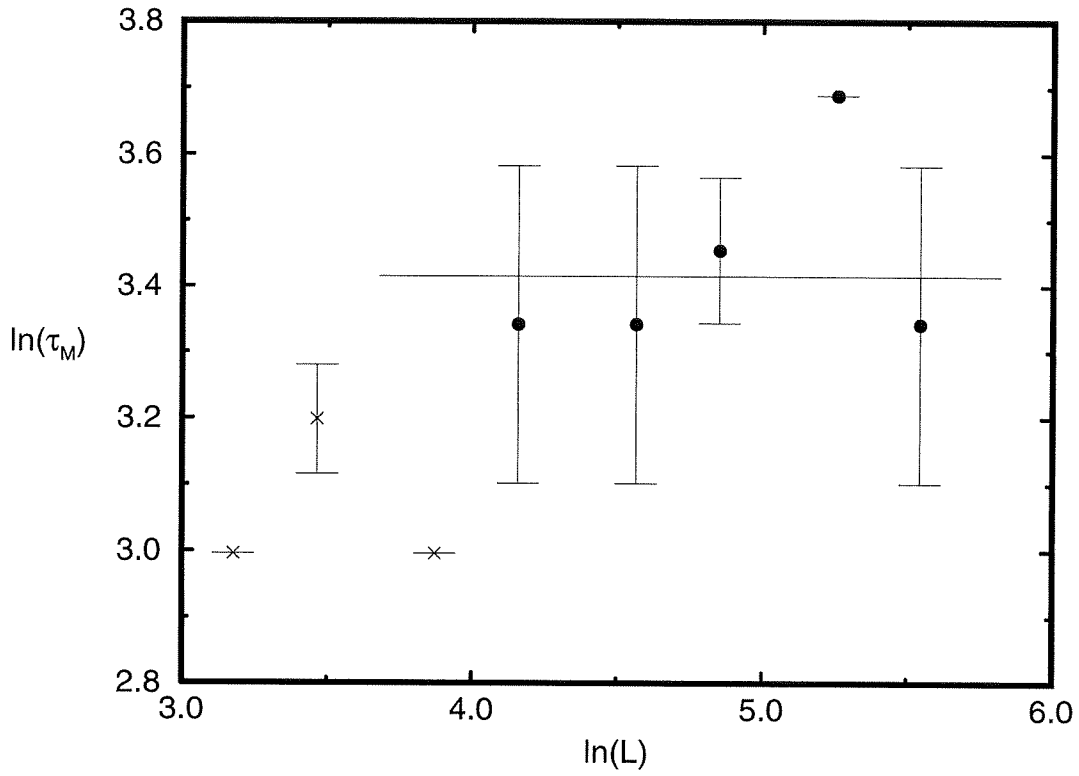

FIG. 13. Finite–size scaling plot of the magnetization correlation time τ_M for the Wolff algorithm. Values of τ_M for systems of size 24^2, 32^2, 48^2, 64^2, 96^2, 128^2, 192^2, and 256^2 were obtained by using the binning technique described in the text, at $T_c = 2.268$. The ensemble size is 2×10^5, but the large error bars result from the fact that this ensemble was compiled from only two simulations. The points labeled with an "x" seem to be strongly susceptible to finite size effects. The solid line indicates that z is consistent with 0.

is any two nearest–neighbor sites with dissimilar spins. The net rate for each process is

$$\Gamma_k N_k,$$

and if we have multiple processes the rate at which configurations transform, i.e. the rate at which *something* happens, is

$$\sum_k \Gamma_k N_k.$$

The probability that the system stays in its current state for a time t before transforming is

$$P(t) = e^{-rt},$$

and the probability that the system transforms using process k is

$$\frac{\Gamma_k N_k}{r}.$$

The waiting time distribution is easy to understand; this type of MC algorithm is used when the number of possible places for a process to occur is small. The total number of sites is quite large. Thus the probability of the event occurring is small, and we are led to Poisson statistics. In fact, the description of continuous time MC given above seems similar to radioactive decay; this is not accidental, and radioactive decay is a useful physical image to keep in mind while thinking about this kind of MC algorithm.

For the case of multiple processes, the implementation of this algorithm involves generating a time to transform, according to the above distribution, and then randomly choosing the process that causes the transformation. However, in the case of only one process, as we have in our Ising model simulations, the algorithm assumes a simple form. We only have one process, so set $\Gamma = 1$. Thus the transformation probability is

$$\frac{N}{r}.$$

For the case of Kawasaki dynamics, N is just the number of unlike pairs in the lattice. Now here's the trick; pick $r = N$. This reflects our knowledge that spin exchange is the the only process in the system that causes transformations. As a consequence, our waiting time distribution assumes the form

$$P(t) = e^{Nt}.$$

We simply increment time by $1/r = 1/N$, and then attempt to exchange two unlike spins.[15] The disadvantage here is that the Monte Carlo time changes as the system changes configurations; we can still define one MC step by accumulating the *fractional* time until our counter "rolls over" 1.0 units, so to speak, but the number of transformation attempts occurring in a Monte Carlo step is no longer fixed. Processes that push the MC time over one unit are aborted and the clock is reset, since the process actually would have occurred in the next step. When the number of possible sites for a transformation is small a variable time MC algorithm is much more computationally efficient than the more obvious (though conceptually simpler) method of choosing event locations completely at random — the analog of just sitting there and waiting for our Geiger counter to register a decay.

With order parameter in hand, we set out to investigate the static properties of the 2D model with K_{nn} dynamics. All simulations were carried out at zero total magnetization; this choice of M is common to all the Kawasaki algorithms discussed in this paper. Cumulants of Ψ are shown in Fig. 14. Square lattices of $L = 6$, 8, 12, 16, 24, and 32 are represented in this figure. We conducted a total of ten independent runs, each of 500,000 MCS; data were sampled every 10 MCS, leading to an ensemble size of 5×10^5 for use in computing averages. Even though we ran an order of magnitude more data than in the Glauber Metropolis case, extraction of a value of T_c from data in Fig. 14 would have been nigh impossible. Some cumulants do not cross at all, and many cross multiple times. The reason for the poor quality of the data became evident once we estimated the value of z, using the binning technique described previously. Our best estimate, shown in Fig. 15, is $z = 3.6(8)$. The

---

[15]We still have to satisfy Metropolis; otherwise we aren't doing the physics we want!

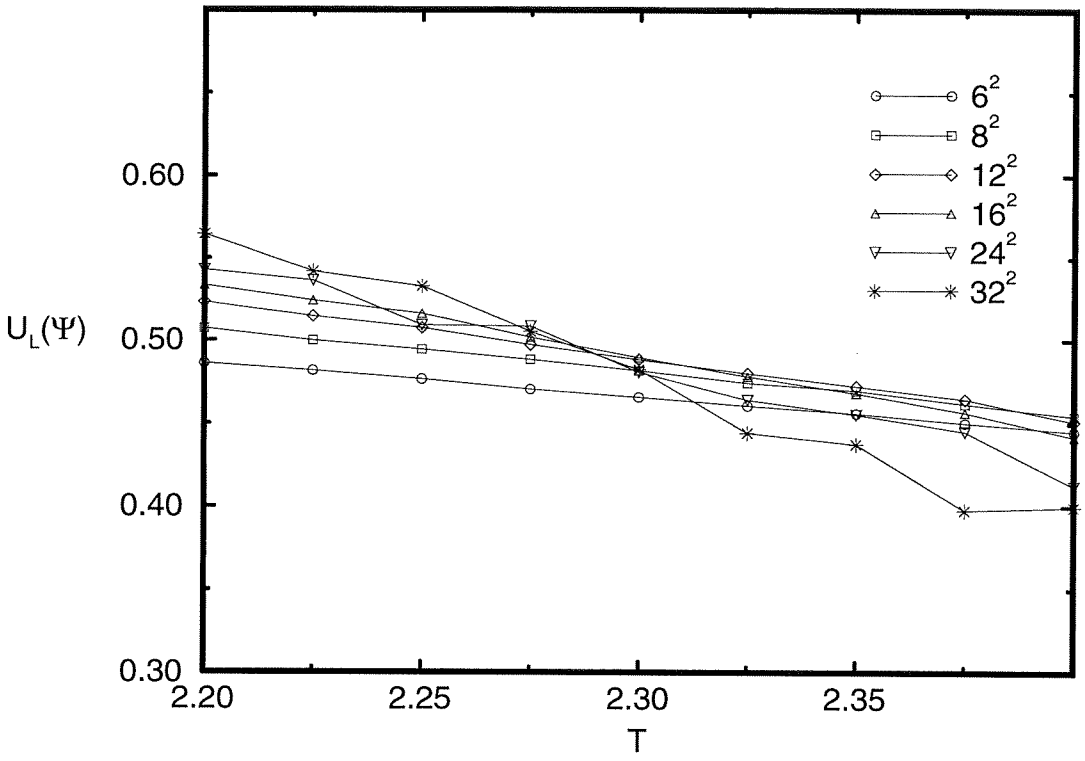

FIG. 14. Binder cumulants of Ψ; system sizes are denoted by the legend. The ensemble size is 5 $\times$ 10^5, and the data were compiled using Kawasaki dynamics with only nearest neighbor exchanges. The slowness of the purely diffusive dynamics leads to a rather large value for z, the dynamical critical exponent, which makes it difficult to collect a large amount of quality data. An estimate of T_c obtained from this figure would be poor at best.

slowness of the nearest neighbor exchange dynamics (which are purely diffusive) makes it virtually impossible to collect enough data to ensure good statistics, even for modest values of L like those shown in the figure.

XVI. MC FOR KAWASAKI DYNAMICS II: LONGER RANGE EXCHANGES

Since our goal in collecting data on the Ising model with Kawasaki dynamics is to verify our belief that the equilibrium properties of the model will be insensitive to the dynamics (assuming the dynamics obeys accessibility and detailed balance), we propose to take that one step further. In an effort to speed up the slow, diffusive dynamics of the K_{nn} algorithm, we decided to mix in some fraction of arbitrary range exchanges. We denote this mixed algorithm $K_{nn+\infty}$ and the case when all exchanges are infinite ranged as K_∞. The transition probability is still given by Eqn. (26), but now the spins selected for exchange need not be nearest neighbors. MC time is again redefined; the $K_{nn+\infty}$ uses a variable time MC algorithm similar to that used by K_{nn}, but it is slightly more complicated due to the fact that there are

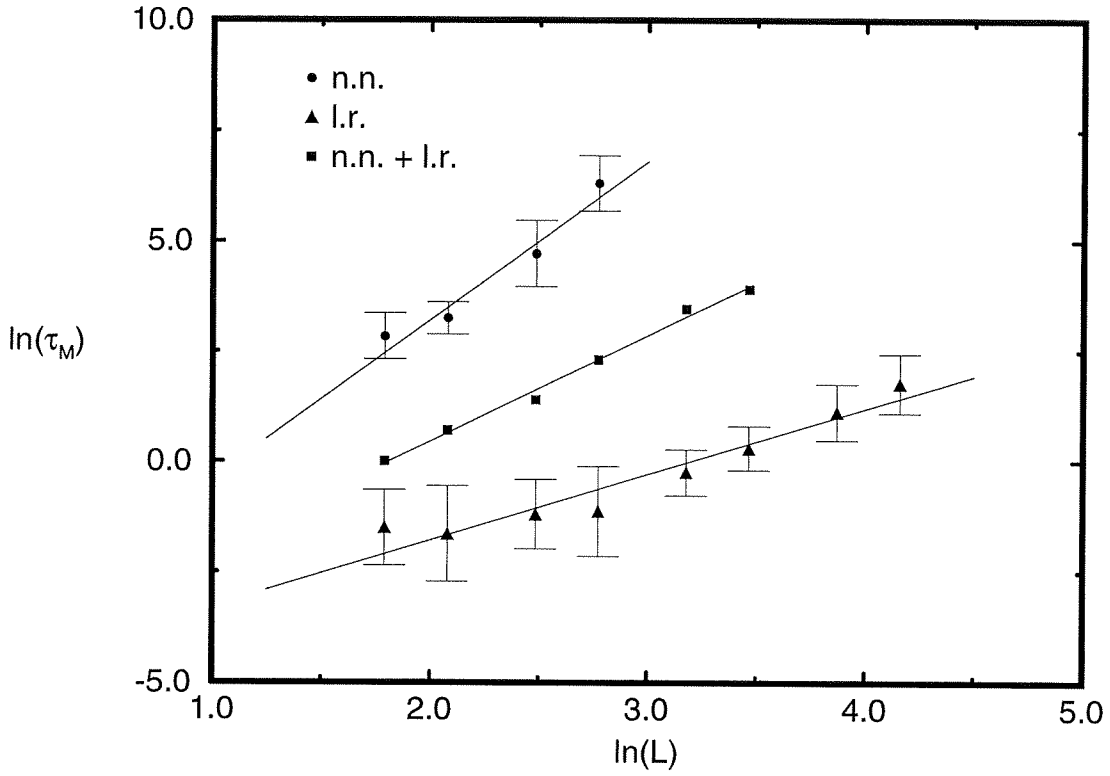

FIG. 15. Finite–size scaling plot of the magnetization correlation time for the $d = 2$ Ising model with Kawasaki dynamics for the K_{nn}, $K_{nn+\infty}$, and K_∞ algorithms, denoted in the legend by n.n., n.n.+l.r., and l.r. respectively. The K_{nn} and $K_{nn+\infty}$ data were compiled at $T = 2.275$, while the K_∞ data were taken at $T_c = 2.267$. τ_M was computed using the binning technique described in the text, with averages over 10 data sets for the K_{nn} and K_∞ algorithms and only 1 data set reported for $K_{nn+\infty}$. Weighted linear fits were used for the K_{nn} and K_∞ data sets, and a simple LSQ fit was used for the $K_{nn+\infty}$ data since the errors are unknown. These fits yield values of $z = 3.6(8)$ for K_{nn}, 2.4(1) for $K_{nn+\infty}$, and 1.5(3) for K_∞. The K_∞ data have been shifted down for clarity.

two processes occurring with different rates. We used the $K_{nn+\infty}$ algorithm only minimally, so a detailed explanation of its associated MC time is unnecessary. We do collect a good deal of data on the K_∞ algorithm, so an explanation of the definition of one MCS for this algorithm deserves mention. Two lists are kept; one of spin–up sites and one of spin–down. We perform all simulations at $M = 0$, so each list contains $L^2/2$ members. A site from each list is selected and an exchange is attempted. One MCS is defined as $L^2/2$ exchange attempts, since the probability P_a that we select two spins which are antiparallel if we choose the spins at random is

$$\begin{aligned} P_a &= P_\downarrow P_\uparrow + P_\uparrow P_\downarrow \\ &= \tfrac{1}{2}\tfrac{1}{2} + \tfrac{1}{2}\tfrac{1}{2} \\ &= \tfrac{1}{2}, \end{aligned}$$

since we are not concerned with which member of the pair is selected first. One MCS is similarly defined for simulations in which $M \neq 0$.

Figure 15 shows the results of our estimate of the correlation time for the $K_{nn+\infty}$ and K_∞ algorithms, extracted once again with the previously described binning technique. From linear fits to these data we discover $z = 2.4(1)$ for the mixed–range exchange algorithm and $z = 1.5(3)$ for the arbitrary–range exchange algorithm. Though we intuited the use of the K_∞ algorithm to reduce z independently, we later discovered that it is the algorithm of Tamayo and Klein. [38] In addition, they quote a value $z = 1.78(4)$ for their algorithm and present an argument that this z should be equal to $z_{K_{nn}} - 2$; our data support both these statements. We now proceed to quote results from our simulations using the K_∞ algorithm. The cumulants of Ψ are shown in Fig. 16. The parameters for these simulations are similar to those found in Fig. 14. Ten simulations of 500,000 MCS each were run for square systems

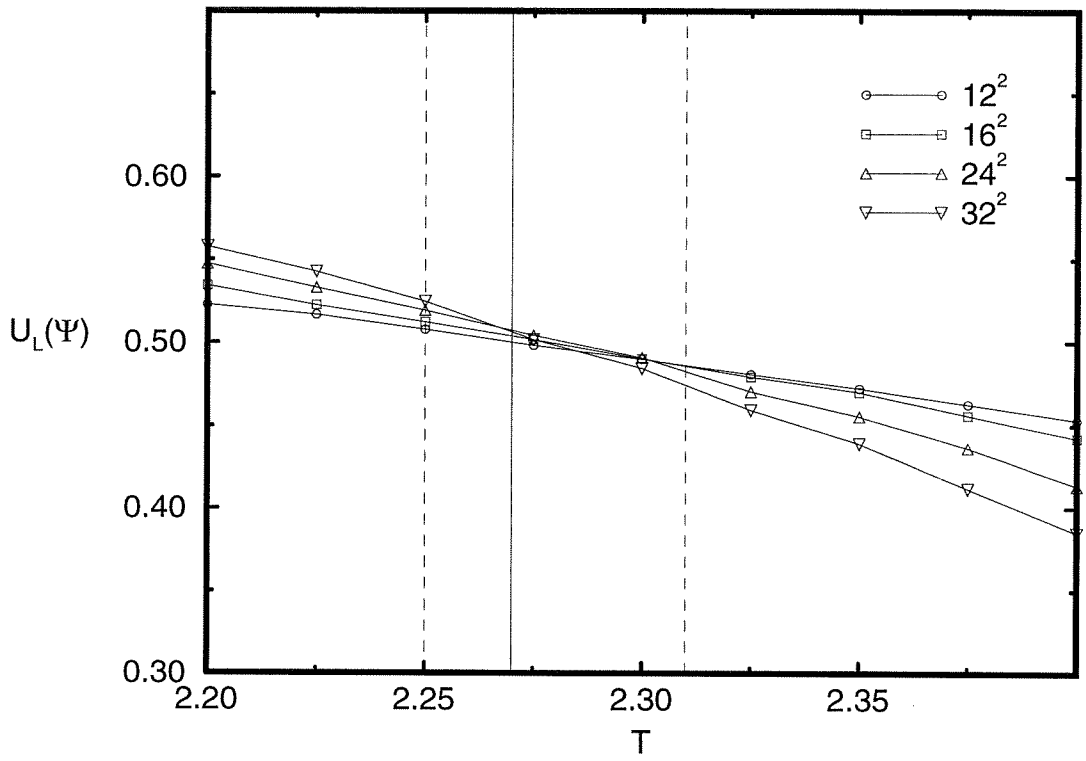

FIG. 16. Binder cumulants for the Ising model with Kawasaki dynamics. Systems simulated are denoted by the legend in the figure, and an ensemble of size 5×10^4 is represented. The solid line gives the best estimate of T_c, with the dotted lines indicating the uncertainty in T_c. From this figure a value of $T_c = 2.27(4)$ was extracted; the error in this number is more generous than that of the figure just so the error is symmetric about the estimate. The lines drawn connecting the data are an aid to the eye.

of size $L = 6$, 8, 12, 16, 24, and 32. Data were sampled once every 100 MCS, rather than once every 10 as was the case for the K_{nn} algorithm, leading to a total ensemble size of 5×10^4 configurations. Data from the size 6×6 and 8×8 simulations are not shown in this

plot, as finite size effects in those cases seemed prohibitively large. Clearly for a comparable amount of data the K_∞ algorithm yields much smoother results. Fig. 16 gives a value for T_c of 2.27(4), though for finite size scaling the value $T = 2.267$, the location of the crossing points of U_{24} and U_{32}, was used.

The results of simulations at the nominal critical value $T_c = 2.267$ are shown in Figs. 17 and 18. Simulation parameters are the same for these data as for the previous figure,

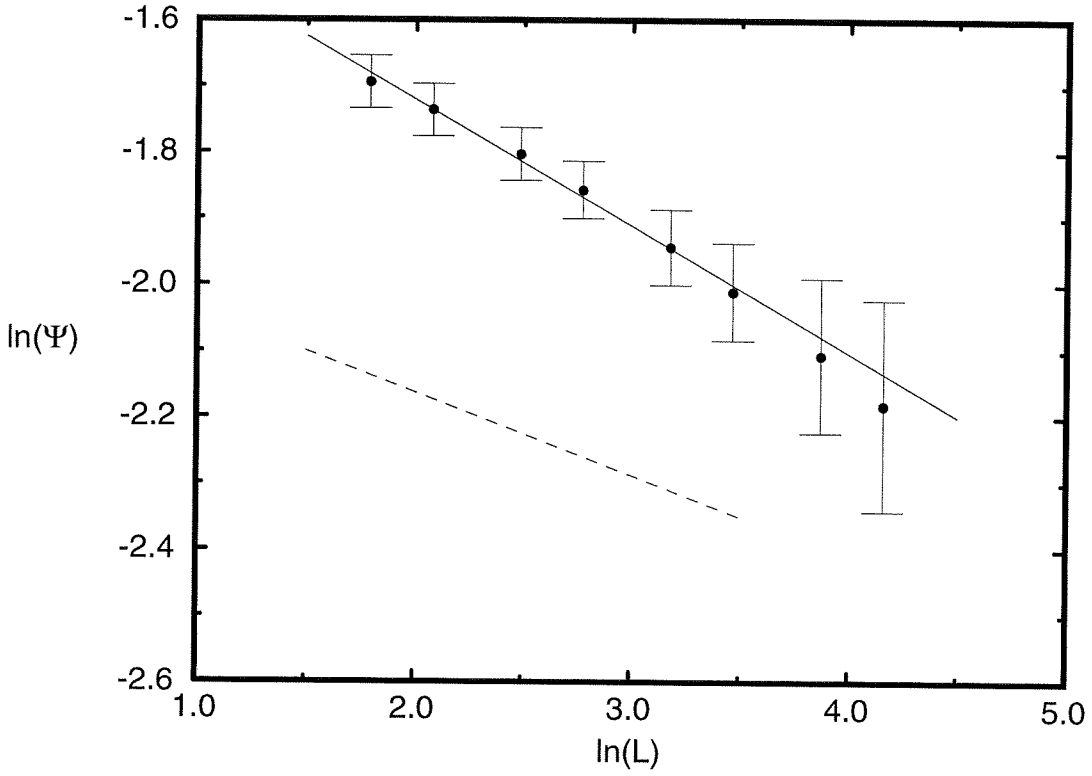

FIG. 17. Finite size scaling data for the order parameter Ψ as defined in the text. The K_∞ algorithm was used for systems of sizes 6^2, 8^2, 12^2, 16^2, 24^2, 32^2, 48^2, and 64^2 at $T_c = 2.267$. An ensemble of size 5×10^4 is represented. From a weighted linear fit to the data (the solid line) the value $\beta/\nu = 0.19(3)$ was extracted. The dotted line, shown for reference, has a slope of $-\frac{1}{8}$, the expected value of $-\beta/\nu$.

with the exception that square systems of size 48×48 and 64×64 were also simulated. From these figures and the suspected value $\nu = 1$ (assuming the critical point for Kawasaki dynamics is still an Ising critical point), we determined $\beta = 0.19(3)$ and $\gamma = 1.7(1)$. These results, as well as all other equilibrium MC simulation data, are summarized in table II.

XVII. COMMENTS ON THE KAWASAKI DYNAMICS DATA

A study of table II shows excellent agreement between measured values of T_c for the Glauber and Kawasaki cases. This confirms our prediction that the critical point would not

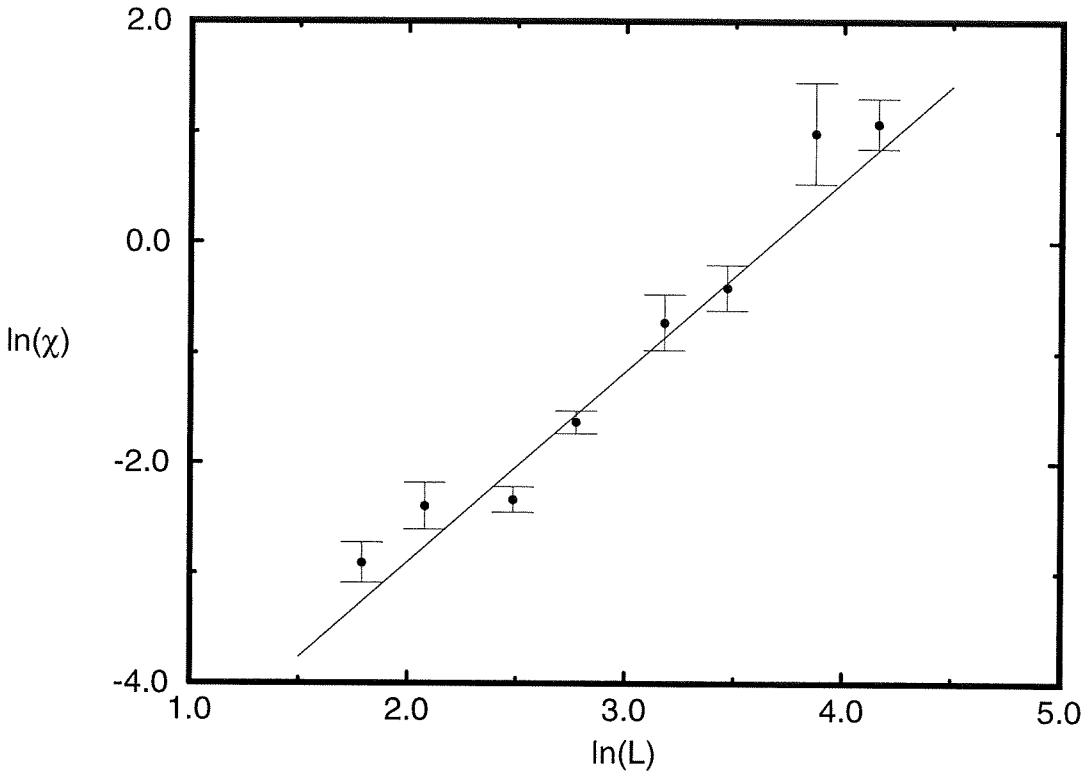

FIG. 18. Finite–size scaling data for the generalized susceptibility χ, defined in text equation (28). The data are for the K_∞ algorithm at $T_c = 2.267$. A weighted linear fit (WLSQ) to the data is the solid line in the figure; this fit yielded the value $\gamma/\nu = 1.7(1)$.

move if the dynamics were changed. The agreement with Onsager's exact value in table II is also good. There is also good agreement in values for the exponent γ for Glauber dynamics, Kawasaki dynamics, and exact solution. However, the discrepancy between measured values of β for the two sets of dynamics is larger than we would like. Our value of β for Kawasaki dynamics is approximately 2.5σ from the expected value, and a look at Fig. 17 shows that the trend in our data for Ψ is not consistent with the expected result (the dotted line in the figure). We conjecture that the behavior of Ψ as we pass through the transition is largely responsible. Ψ did not turn out to be a good order parameter. Though it scales near the transition somewhat within our expectation, its behavior near T_c is less than ideal. In the case of Glauber dynamics, our order parameter is insensitive to the spatial distribution of domains; an equal number of up and down spins has $M = 0$ regardless of how the spins are arranged. Ψ is much more sensitive to the distribution of domains. Even slightly above T_c large domains of up and down spins exist, and one can see from the definition of Ψ that these large domains make a sizeable contribution to Ψ's value. Thus Ψ is not exactly zero above the transition, as borne out by Fig. 19. In fact, even for a system of size 32×32 Ψ at

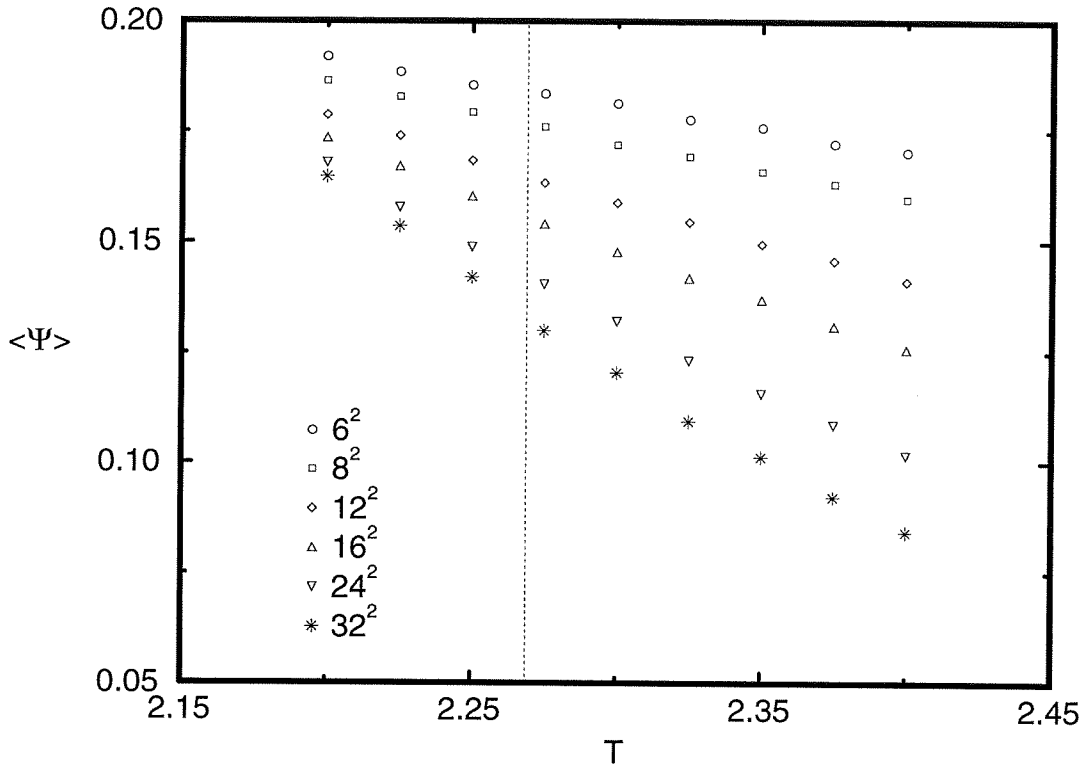

FIG. 19. The order parameter Ψ for the K_∞ algorithm, for systems of sizes denoted by the legend. The dotted line passes through T_c and is shown as an aid to the eye. Notice the less than ideal behavior of Ψ near T_c. Ψ does not go to zero at T_c; in fact, even for the largest system size shown in the figure Ψ at T_c is still greater than one fourth of its zero temperature value. Though Ψ at T_c is tending to zero as L is increased, the convergence is obviously quite slow with L, and it is much slower than the convergence of the order parameter for nonconserved dynamics. This behavior makes Ψ a less than ideal order parameter and contributes to the discrepancy between β measured in the conserved and nonconserved dynamics cases.

T_c is still more than a quarter of its maximum value.[16] The fact that Ψ does not go to zero fast enough near T_c also affects our results for γ, which is defined by Eqn. (28). Though the exponent extracted from Fig. 18 agrees well with both our previous simulation results and the expected exact value, the fit is poor, especially considering the fact that we have managed to reduce z to a value less than that of the Glauber case.

Due to the disappointing behavior of Ψ in characterizing the phase transition in the Ising model with spin exchange, we cannot say for certain that the critical point is an Ising critical point. The agreement in γ and T_c make us feel that this is indeed the case,

[16] Ψ was normalized to be zero at infinite temperature, which resulted in it having a value of four tenths at zero temperature.

but the measured value of β in the Kawasaki case prevents us from making a stronger statement. We can, however, make one very strong statement about using Monte Carlo on the Ising model. While there may not be a right or wrong choice for the simulation of the equilibrium model, there are certainly difficult and easy choices! Another look at table II should be enough to convince you that the Wolff algorithm is the clear choice for studying static properties of the model. The fact that one's errors do not grow with system size is tremendously helpful, especially since finite–size scaling is so effective in extracting critical exponents from simulation data.

XVIII. INTRODUCTION OF A NONEQUILIBRIUM ISING MODEL

In the previous sections we have shown that a simple change in the dynamics for the Ising model likely leaves the static properties unchanged; at the very least we may say we have proven any perturbations from Ising criticality are small. We now proceed to introduce a change in the dynamics of the model that will cause significant deviations from the critical behavior we have so far experienced. We have previously played with various dynamical algorithms, but they all still satisfied detailed balance. We now imagine a two dimensional lattice of Ising spins coupled to *two* different heat baths, of inverse temperatures β_x in one direction and β_y in the other direction. [40] We still conserve the magnetization, so we allow only spin exchanges. In addition, these exchanges are only allowed to be between nearest neighbors, for reasons which will soon be clear. Two antiparallel spins are selected and exchanged with probability

$$P = \begin{cases} 1 & \text{if } \Delta E \leq 0 \\ e^{-\beta_x \Delta E} & \text{if the pair is oriented in the } x \text{ direction} \\ e^{-\beta_y \Delta E} & \text{if the pair is oriented in the } y \text{ direction.} \end{cases}$$

It is clear from the definition of P above that exchanging spins that are not aligned along one of the two coordinate axes of the square lattice has no meaning. Thus only a spin and its four nearest neighbors are eligible for exchange. Defined this way, this β_x–β_y model is a specific case of a general class of driven diffusive models. A driven diffusive model is one in which the dynamics are diffusive (as for the equilibrium Kawasaki Ising model) and a driving (symmetry breaking) field acts in m dimensions of the total dimension d, while the other $d - m$ dimensions do not feel the effect of the driving field. [39] For the special case described above $d = 2$ and $m = 1$; the $m = 0$ case corresponds to equilibrium.

The addition of the second heat bath prevents the system from coming to thermal equilibrium. While it is true that for a steady state

$$\sum_a W(a \to a')P(a) = \sum_{a'} W(a' \to a)P(a'),$$

we no longer have term–by–term equality as before; consequently, we no longer know what the probability distribution P is. Though there are no equilibrium states in this model (except for the special case discussed below), there may still be a transition from an ordered to a disordered state. Our experience with the equilibrium Kawasaki model tells us we may use Ψ as defined in Eqn. (27) as our order parameter. Although we discovered Ψ is a poor

order parameter for the *equilibrium* Ising model with Kawasaki dynamics, it will prove more useful in this case. We can estimate the locations of transition points using cumulants of Ψ, defined exactly as before. However, the static properties of the model are now coupled to the dynamics. Changing the dynamics in effect gives us a new model, with different static and dynamical behavior. We also have a more complicated phase diagram, which we can depict in the β_x-β_y plane. We already know something about this phase diagram; along the line $\beta_x = \beta_y$, we have the equilibrium Ising model with Kawasaki dynamics. There is an Ising critical point at (0.4407,0.4407), with a disordered phase for $\beta_x = \beta_y < 0.4407$ and an ordered phase for $\beta_x = \beta_y > 0.4407$. The transition along this line is continuous.

The MC methodology is similar to the simulation of K_{nn}. Lists of pairs are kept and the variable time algorithm previously discussed is implemented. The system is started from a random configuration, and several thousand MCS (typically 10,000) are allowed for the system to come to a steady state characteristic of the given combination of β_x and β_y. Once this relaxation period is over, data are collected with the expectation of producing steady–state ensemble averages, since we are interested (for this work) in the static critical exponents. We concentrate on the transition when $\beta_y = 0$.

Before collecting data for various values of β_x, the subject of anisotropic finite–size scaling needs to be discussed. When we wanted to extract critical exponents for the equilibrium model from finite–size scaling plots, it was assumed all square systems were equivalent. Indeed they were; while we chose system sizes that would give nice spacing on a log–log plot, we were not restricted in any way and could have chosen square systems of whatever size we wanted. This is not the case with driven diffusive systems. The driving field (in our case the different temperatures) breaks a symmetry in the system. We introduced the spatial correlation function

$$G_c^{(2)} = \langle \phi(0)\phi(r) \rangle - \langle \phi \rangle^2$$

earlier in this work. Recall also the scaling form for $G_c^{(2)}$,

$$G_c^{(2)} \sim \frac{1}{r^{d-2+\eta}}, \quad T = T_c, \quad r \gg 1,$$

where d is the dimensionality of the system and η is another critical exponent. For the driven diffusive system, the symmetry–breaking effect causes us to have to define *two* exponents η, one in the direction of the driving field and one in directions perpendicular to the field. [39] In fact, if we want to be able to use finite size scaling to compare equivalent systems, we have to look at systems whose lengths scale as

$$L_x \times L_y^{1+\theta},$$

where θ is an anisotropy exponent. [41] Calculations [39] and simulations [45] suggest $\theta =$ 0.9 - 1.0, so we study systems of size 8×4, 12×9, 16×16, 24×36, and 32×64, in which the scaling is $L_x \times L_y^2$. [37] To emphasize again the degree to which this symmetry–breaking field has affected the scaling behavior, let us emphasize that this is *not* the same thing as studying the equilibrium Ising model on a rectangular lattice. There, the rescaling is trivial;

there is still only one *relevant* length scale in the problem.[17] This is not the case with the driven diffusive model.

We now present simulation data for the two–temperature nonequilibrium model. Systems

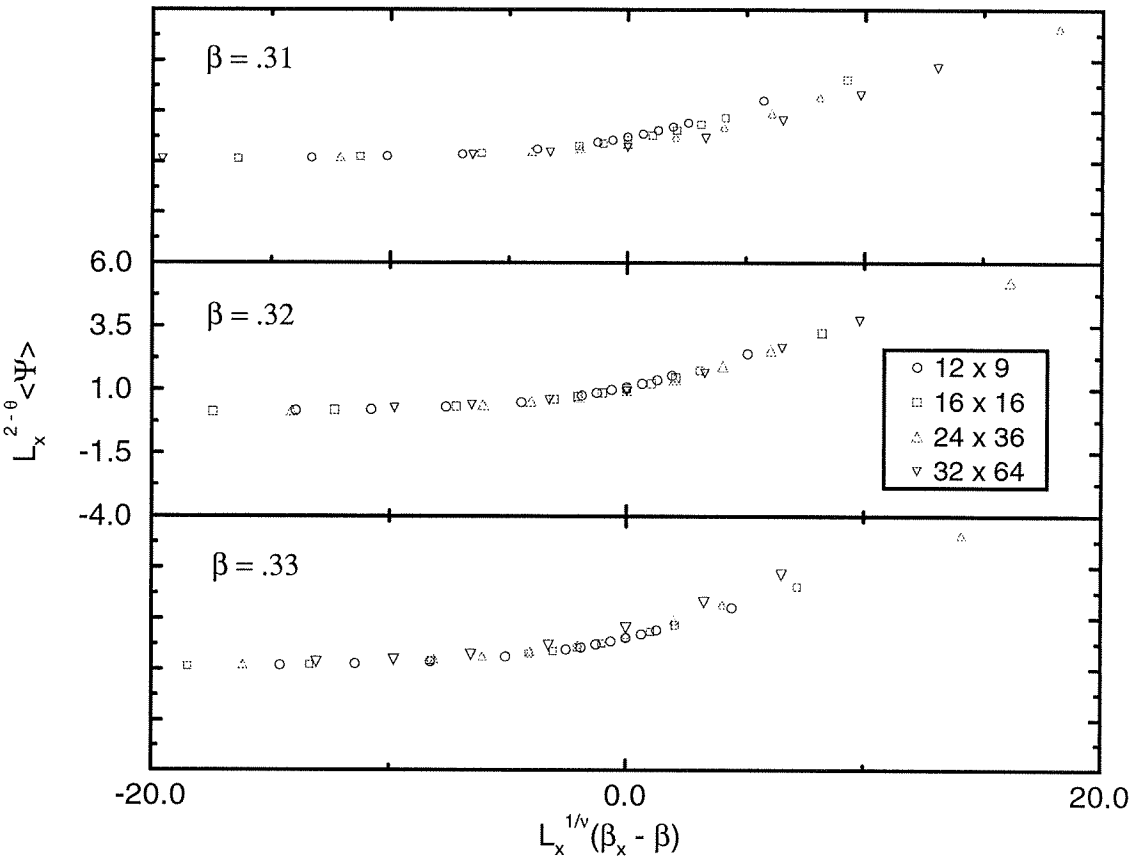

FIG. 20. Scaling plot of $\langle\Psi\rangle$, for the two–temperature nonequilibrium Ising model. The values of $\nu = 0.6$ and $\theta = 1.0$ are as described in the text. The parameter β is β_{xc}; the fixed value of β for each of the three plots is shown in the upper left corner of each graph. System sizes are denoted by the legend, and these data were compiled over 10^6 steady–state configurations. Notice the change in quality of the scaling as β_{xc} is varied, leading us to deduce $\beta_{xc} = 0.32(1)$.

of the sizes above were simulated for 10^7 MCS, though anticipating the production of correlated data led us to collect data once every 10 MCS, leading to a total ensemble size of 10^6. β_y was fixed at 0, and data at values of β_x ranging from 0.1 to 0.4 were collected. These data are presented in Figs. 20 - 23. Notice that the data for the 8×4 system are not present in any of the figures; this is because finite size effects made the data unsuitable for a scaling analysis. Cumulants of ψ were used to find a rough estimate for the value of $(\beta_{xc}, 0)$. A

[17]The word "relevant" may seem somewhat vague here; a more precise statement is that in the two–temperature model, there is an additional length scale relevant in the renormalization group sense.

more accurate estimate of the critical value of β_{xc} comes from a scaling plot of

$$L_x^{2-\theta}\langle\Psi\rangle \quad \text{vs.} \quad L_x^{1/\nu}(\beta_x - \beta_{xc}), \tag{29}$$

shown in Fig. 20. The value of θ is taken to be one, and $\nu = 0.6$ is taken from references [37,45]. The parameter β_{xc} is varied until the data collapse onto one curve; conservative error bars are determined by further altering β_{xc} until a noticeable decline in the quality of the scaling occurs. The data in Fig. 20 yielded a value of $\beta_{xc} = 0.32(1)$. A further scaling plot for Ψ, of the form in Eqn. (29), yields a value for the critical exponent β. This scaling plot is presented in Fig. 21; notice we have now separated the $\beta_x < \beta_{xc}$ and $\beta_x > \beta_{xc}$ data, with the expectation that the scaling function in the two cases will be completely different. The figure gives a value of 1.2(1) for β.

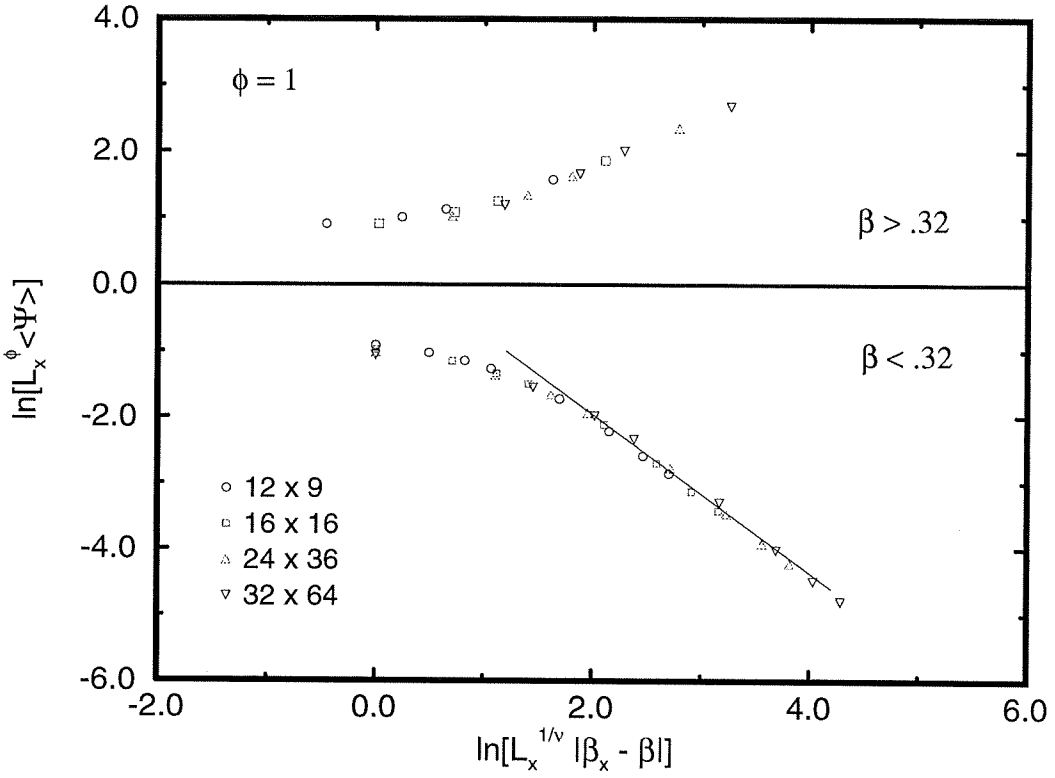

FIG. 21. A similar scaling plot to Fig. 20, this time with the data for $\beta < \beta_{xc}$ and $\beta > \beta_{xc}$ separated, in expectation of the different scaling functions in the two regimes. In this case, $\phi = 2-\theta$. The slope of the solid line, shown in the $\beta_x < .32$ window, yields the value of the critical exponent β. From that slope we determined $\beta = 1.2(1)$.

Once we have a value for β_{xc} we can proceed to perform conventional finite size scaling on other quantities. The scaling variables are

$$L_x^{\phi}A \quad \text{vs.} \quad L_x^{1/\nu}(\beta_x - \beta_{xc}),$$

where A is some quantity we are interested in (C or χ) and ϕ is an exponent that yields the best data collapse. Don't confuse this ϕ with the ϕ of Eqns. (18),(19), (20), and (21).

β_{xc}	α	β	γ
0.32(1)	0.24(6)	1.2(1)	0.48(6)

TABLE III. Static critical exponents for the nonequilibrium two–temperature Ising model. All values come from scaling plots, shown in Figs. 20 - 23. The value β_{xc} refers to the location of the transition when $\beta_y = 0$.

It is an unfortunate coincidence; there are so many Greek letters already reserved for other quantities that it is difficult to use anything else. Scaling plots of this form for C and χ appear in Figs. 22 and 23, respectively. Using again the value $\nu = 0.6$, values of $\alpha = 0.24(6)$ and $\gamma = 0.48(6)$ were determined. Remember that our ability to perform scaling plots, which appear very much like those described in earlier sections, hinges on the fact that we have studied systems whose sizes scale in the appropriate way, namely $L_x \times L_y^{1+\theta}$. Results from the nonequilibrium model are summarized in table III.

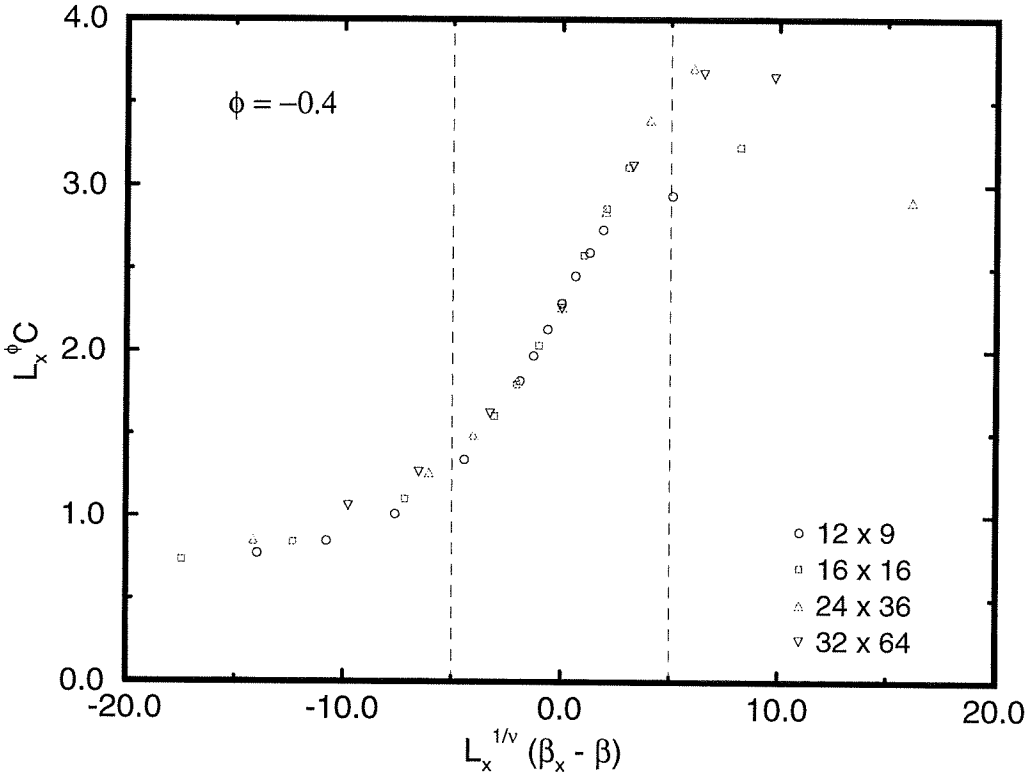

FIG. 22. Finite–size scaling plot for the specific heat C for the two–temperature nonequilibrium Ising model. Focus on the quality of scaling between the dotted lines, as we don't expect scaling to hold much outside this range. The exponent $\phi = -\alpha/\nu$ in this case. We determined errors in ϕ similarly to those in Fig. 20, also discussed in the text. The scaling analysis yielded a value of $\alpha/\nu = 0.4(1)$, which together with the value $\nu = 0.6$ gives $\alpha = 0.24(6)$.

XIX. COMMENTS ON THE NONEQUILIBRIUM DATA

The errors in the static exponents for the two–temperature model are generous, but the result is still clear. They are quite different than in the equilibrium case; the symmetry–breaking field imposed on the model has had a definite effect on the physics. We have not determined the phase diagram for the model, but the critical point we have discovered is actually connected by a continuous phase boundary to the Ising critical point. [37] The

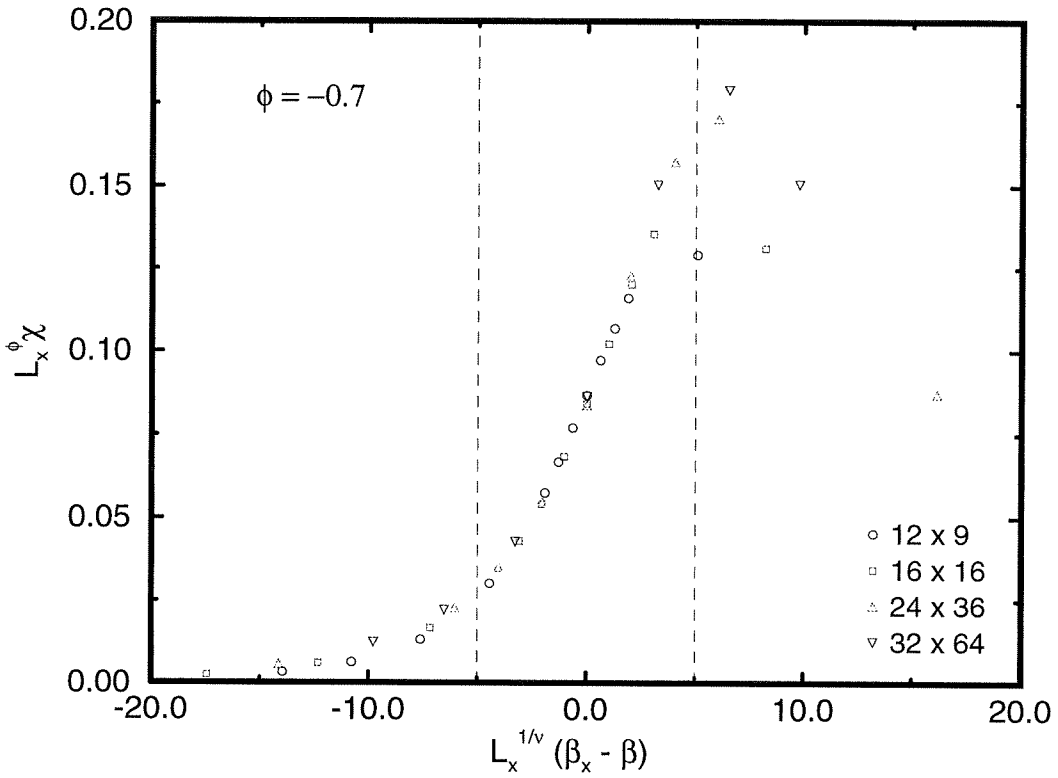

FIG. 23. Finite–size scaling plot for the susceptibility χ as defined in Eqn. (28). This plot is the counterpart to Fig. 22, and comments made in that figure's caption apply here as well. In this case, $\phi = -\gamma/\nu$, and we determined $\gamma/\nu = 0.7(1)$. Together with the value $\nu = 0.6$ we found $\gamma = 0.48(6)$.

scaling for Ψ, C, and χ is quite good in the regime where we expect scaling to hold; the scaling of Ψ is the best of the three. This should immediately raise the question: what has happened to the order parameter? We chastised the order parameter in the equilibrium case, and blamed its poor behavior near T_c for inconsistent results. Now, for lack of any better idea, we come crawling back to Ψ in the nonequilibrium case, and it works wonderfully! Indeed, a glance at Fig. 24 shows much better order parameter behavior near β_{xc}. Of course, as we reduce the asymmetry in β_x and β_y, we come closer to the equilibrium Kawasaki model, and we know that Ψ fails miserably there. This is the key to understanding what is happening.

When $\beta_x \neq \beta_y$, there is one "fast" degree of freedom and one "slow" degree of freedom; namely the hotter and cooler temperatures, respectively. The fast degree of freedom can more easily perform significant reorganization of domains, leading to an altered value of Ψ. We considered the $\beta_y = 0$ case, which makes the anisotropy as dramatic as possible; the y axis of the lattice undergoes fast diffusion — exchanges are always made because the Boltzmann factor equals one. Carrying this argument further, we would expect Ψ to characterize the transition more and more poorly as we reduce the asymmetry in β_x and β_y, which corresponds to studying a smaller and smaller perturbation from Ising criticality. It would be interesting to test this conjecture.

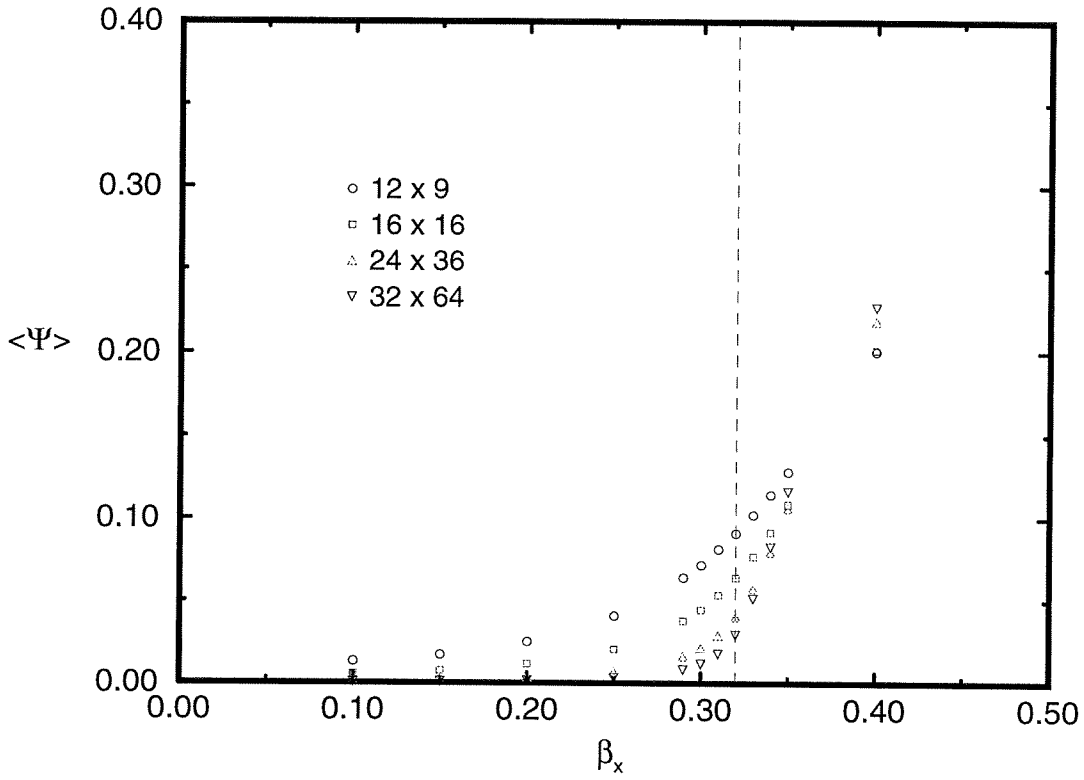

FIG. 24. Behavior of Ψ for the nonequilibrium two–temperature Ising model with changes in β_x. β_y is fixed at zero for this data; the critical value of β_x is indicated with the dotted line. System sizes are as denoted in the legend. Compare with Fig. 14, and we see in this case Ψ is behaving more in accord with what we want from an order parameter. Note especially the fast drop in Ψ near β_{xc} as $L_x L_y$ increases; this is in marked contrast to Fig. 14.

XX. CONCLUSION

When we review the work covered in this paper, one can easily lose sight of the forest for the trees. There is a great deal of space devoted to mean–field calculations, discussions of the Monte Carlo method, and the extraction of critical exponents and determination of errors

from simulation. We studied the equilibrium Ising model from multiple angles, including both spin–flip and spin–exchange dynamics. We then introduced a nonequilibrium variant of the Ising model and studied its behavior, finding very different results when compared to the equilibrium case. We should here reflect upon the physics that all of this analysis has taught us.

We started with a mean–field theory. As all mean–field theories do, it tries to describe the behavior of each object in the system in terms of the average behavior of the other objects. Mean–field theory did give us some qualitatively correct behavior; for example, χ diverged at T_c as we expected it should. However, when we began to push the theory farther, by calculating T_c and the critical exponents, mean–field theory was a resounding failure. From this we began to realize that correlations between our objects (spins for the Ising model) are not only *important* near critical points but *dominant*. After all, recall that the exponents we predicted were nowhere near their correct values, and that mean–field theory (at the level of the Weiss approximation) did not predict any divergence in the specific heat.

Seeing that mean–field theory is not adequate for describing critical behavior, we proceeded to perform a series of numerical "experiments," aimed at trying to understand the Ising model's behavior near the critical point. We calculated the exponents for the classic Glauber dynamics case, and then moved on to Kawasaki (spin–exchange) dynamics. The goal here was to see what kind of relevance the *dynamics* has on the *statics*. After all, we earlier mentioned conservation laws can affect the critical behavior of a system, and the effect of the Kawasaki dynamics is to add a strict conservation law to the model. To a good extent we found no difference in the critical behavior under this new dynamics, and when there were differences we argued that it was our technique and not the physics that caused these differences.

This comparison of the two types of dynamics led us to realize the importance of a fundamental feature of equilibrium physics, that of the principle of detailed balance. Detailed balance in effect tells us we can generate states in any number of ways, but we end up with Boltzmann distributed, and hence equilibrium, states if the transition probabilities satisfy a simple relationship. To use a thermodynamic analogy, we can adiabatically cool first and then isothermally change the pressure, or we can do things the other way around. While what happens *during* the relaxation is quite different, the end (static) product of the process is the same.

Having tried to show that the static behavior of the Ising model is similar for dynamics that generate equilibrium states, we introduced a version of the model that cannot come to thermal equilibrium. When we studied the behavior of the two–temperature Ising model, we saw some qualitative features that remind us of the equilibrium model. In the regime we studied, the model has a transition from an ordered steady–state to a disordered state. The order parameter, and its fluctuations and those of the energy, followed power law behavior near the transition, similar to what we found in the equilibrium case. However, the static exponents that characterize this behavior assumed completely different values from what we measured in the equilibrium case. We thus demonstrate that our perturbation from equilibrium, and hence violation of detailed balance, gives rise to completely new behavior. This occurs even near a critical point, where we expect the system to be less and less sensitive to microscopic details.

The results of the study of the nonequilibrium Ising model should be striking to the

reader. Consider that a binary fluid, the mixture of two hydrocarbons insoluble below a critical temperature, and the 3D equilibrium Ising model have the same critical exponents. The two–temperature model and the 2D equilibrium model do not, and the equilibrium model is even a special case ($\beta_x = \beta_y$) of the two–temperature model! Though not all nonequilibrium perturbations to equilibrium systems result in new critical behavior [46,47,14], it is clear that in this case the two competing heat baths have had a significant effect on the model's static critical properties.

This paper only scratches the surface, both of equilibrium phase behavior and nonequilibrium transitions. As we stated earlier, the nonequilibrium variants of the Ising model are nearly endless [12]. There are equilibrium continuous–spin models, like the Heisenberg and XY models, that develop complicated structures like spin waves, spin solitons, and vortices. Even the equilibrium Ising model with Glauber dynamics, known since 1925, is experiencing a resurgence via the study of first passage and early–time critical exponents. [51–54] The field of statistical mechanics is rife with interesting and elegant phenomena; the very notion of universality is one of the most striking and beautiful results in all of physics. We hope that this work can serve as an introduction to some of this elegance.

APPENDIX A: THE BETHE APPROXIMATION

In this appendix we give a short description of the Bethe approximation [23]. The Bethe approximation adds some local information that the WMF and BW theories lack. We write the Hamiltonian for one spin σ_0 as

$$\mathcal{H} = -h\sigma_0 - J\sum_{i=1}^{z}\sigma_0\sigma_i - h_e\sum_{i=1}^{N}\sigma_i,$$

where the first two terms represent the exact interaction of σ_0 with the applied field h and its z nearest neighbors, while the second term represents an effective interaction between the z neighbors and *their* surrounding spins. We thus write the partition function for one spin as

$$Z_1 = \sum_{\sigma_0=\pm1}\cdots\sum_{\sigma_z=\pm1} e^{\beta h\sigma_0}e^{\beta J\sigma_0\sum_{i=1}^{z}\sigma_i}e^{\beta h_e\sum_{i=1}^{z}\sigma_i}.$$

The self consistency condition

$$\langle\sigma_0\rangle = \frac{1}{z}\left\langle\sum_{i=1}^{z}\sigma_i\right\rangle$$

is imposed, since σ_0 was selected arbitrarily. Both of these averages can be computed from Z. This, and a lot of algebra, eventually leads to the equation of state

$$\tanh\beta\left(\frac{h_e-h}{z-1}\right) = \tanh\beta J\tanh\beta h_e,$$

from which one can derive the zero–field transition temperature

$$T_c = \frac{2}{\ln\left(\frac{z}{z-2}\right)},$$

expressed in units of J/k_B. Notice that the Bethe approximation does predict the correct transition temperature in the $d=1$ model where $z=2$; the WMF and BW theories predict $T_c = 2.0$ in one dimension. Thus the transition temperature is moving closer to the correct value, but one can show (though we do not) that the critical exponents as calculated by the Bethe approximation assume identical values to exponents from BW and WMF. Thus we see that the amount of algebra at each level of approximation seems to be diverging to infinity more quickly than our answers are converging to their correct values!

APPENDIX B: CORRELATION FUNCTIONS AND FOURIER TRANSFORMS

In this appendix we derive the relationship

$$C_M(0,t) = \frac{\mathcal{F}\left(\left|\widetilde{M}[k]\right|^2\right)}{\sum_k \left|\widetilde{M}[k]\right|^2},$$

which is of tremendous use in calculating correlation functions from steady–state time series obtained from MC simulation. The requirement that the data be in a steady–state situation amounts to a requirement of ergodicity.[18] Assuming we do have such a stationary time series of N elements, we can view it as a function on $\mathbf{P}_N$, the cyclic group of N elements. In $\mathbf{P}_N$,

$$M[n] = M[n+N].$$

We write the Fourier series of M on $\mathbf{P}_N$ as follows:

$$M[n] = \frac{1}{N}\sum_{k=0}^{N-1} \widetilde{M}[k] e^{2\pi i k n/N}.$$

We are interested in the quantity

$$\frac{1}{N}\sum_{k=0}^{N-1} \delta M[k]\delta M[k+T], \tag{B1}$$

where T is some increment taking us to a later "time" in the data, or equivalently another state. We have here used

$$\delta M[k] = M[k] - \overline{M} = \frac{1}{N}\sum_{k=0}^{N-1} \widetilde{M}[k] e^{2\pi i k n/N} - \frac{1}{N}\widetilde{M}[0],$$

which is just M minus its mean value. We are interested in Eqn. (B1) because it is the numerator of Eqn. (16). We continue and compute

$$\delta M[n]\delta M[n+T] = \left(\frac{1}{N}\sum_{k=1}^{N-1} \widetilde{M}[k] e^{2\pi i k n/N} \frac{1}{N}\sum_{l=1}^{N-1} \widetilde{M}[l] e^{2\pi i l (n+T)/N}\right).$$

Notice the sums start at $k, l = 1$; this is because of the subtraction of the mean of M, represented by the zeroth Fourier component. We thus write for the numerator of $C_M(t,t')$

$$\langle \delta M[n]\delta M[n+T]\rangle = \frac{1}{N}\sum_{n=0}^{N-1}\left(\frac{1}{N}\sum_{k=1}^{N-1} \widetilde{M}[k] e^{2\pi i k n/N} \frac{1}{N}\sum_{l=1}^{N-1} \widetilde{M}[l] e^{2\pi i l (n+T)/N}\right), \tag{B2}$$

---

[18] Thus, this technique does *not* work when one wants to study relaxation to equilibrium, as is the case for a critical quench or a zero–temperature quench to study coarsening.

which implies for the denominator of $C_M(t,t')$, which is just the equal time correlator, we have

$$\langle \delta M[n] \delta M[n] \rangle = \frac{1}{N} \sum_{n=0}^{N-1} \left(\frac{1}{N} \sum_{k=1}^{N-1} \widetilde{M}[k] e^{2\pi i k n/N} \frac{1}{N} \sum_{l=1}^{N-1} \widetilde{M}[l] e^{2\pi i l n/N} \right). \tag{B3}$$

Consider each term in the sum over n in Eqn. (B3); each summand is of the form

$$\widetilde{M}[k] \widetilde{M}[l] e^{2\pi i (k+l) n/N}.$$

Focus on the exponential. It will have a value

$$e^{2\pi i (k+l) n/N} = \begin{cases} 1 & \text{if } k+l = mN \\ 0 & \text{otherwise} \end{cases} \tag{B4}$$

since if $k+l = mN$, the argument of the exponential is an integer multiple of $2\pi i$; otherwise we are summing the N^{th} roots of unity. Since all multiples of N are equivalent in $\mathbf{P}_N$, the first condition amounts to $k+l = N$. We can rewrite Eqn. (B3) as

$$\langle \delta M[n] \delta M[n] \rangle = \frac{1}{N^3} \sum_{k=1}^{N-1} \widetilde{M}[k] \sum_{l=1}^{N-1} \widetilde{M}[l] \left(\sum_{n=0}^{N-1} e^{2\pi i (k+l) n/N} \right).$$

If we apply the condition on the exponential (Eqn. (B4)) we see

$$\sum_{n=0}^{N-1} e^{2\pi i (k+l) n/N} = \begin{cases} N & \text{if } k+l = N \\ 0 & \text{otherwise,} \end{cases}$$

which can be more concisely written as

$$\sum_{n=0}^{N-1} e^{2\pi i (k+l) n/N} = N \delta_{k+l,N}.$$

Thus

$$\begin{aligned} \langle \delta M[n] \delta M[n] \rangle &= \tfrac{1}{N^3} \textstyle\sum_{k=1}^{N-1} \widetilde{M}[k] \sum_{l=1}^{N-1} \widetilde{M}[l] N \delta_{k+l,N} \\ &= \tfrac{1}{N^2} \textstyle\sum_{k=1}^{N-1} \widetilde{M}[k] \widetilde{M}[N-k], \end{aligned}$$

which we can even further simplify by realizing

$$\widetilde{M}[N-k] = \widetilde{M}^*[k].$$

The validity of the previous statement rests both on the fact that we are dealing with an N-periodic group and that M is real–valued, i.e.

$$M^*[k] = M[k], \qquad \forall k.$$

Thus we finally see for the equal time correlator

$$\begin{aligned}\langle \delta M[n]\delta M[n]\rangle &= \tfrac{1}{N^2}\textstyle\sum_{k=1}^{N-1} \widetilde{M}[k]\widetilde{M}^*[k] \\ &= \tfrac{1}{N^2}\textstyle\sum_{k=1}^{N-1} \left|\widetilde{M}[k]\right|^2.\end{aligned}$$

This equation suggests a simple flowchart that we can use when we want to compute the correlator in practice:

$$M[n] \xrightarrow{\text{FFT}} \widetilde{M}[n] \xrightarrow{\text{sum}} \langle \delta M[n]\delta M[n]\rangle.$$

We still need the numerator to compute the entire correlation function; what we in fact have already done is determine only the normalizing factor! All the real information is in Eqn. (B2). Luckily, the procedure is exactly the same as before, but we are carrying around an extra exponential factor, due to the time increment T. Following the previous line of reasoning leads us to

$$\langle \delta M[n]\delta M[n+T]\rangle = \frac{1}{N^2}\sum_{k=1}^{N-1} \left|\widetilde{M}[k]\right|^2 e^{-2\pi i k T/N}.$$

We immediately recognize this expression as

$$\mathcal{F}\left(\left|\widetilde{M}[T]\right|^2\right) - \left|\widetilde{M}[0]\right|^2,$$

which we can in practice compute with another FFT! Thus we combine the equal and unequal time correlators to determine

$$C_M(t,t') = \frac{\mathcal{F}\left(\left|\widetilde{M}[k]\right|^2\right)}{\sum_{k=1}^{N-1}\left|\widetilde{M}[k]\right|^2}.$$

Changing the arguments (for example to 0 and t) introduces a phase factor when we take the Fourier transform, but this factor is irrelevant to the modulus of the coefficients, which is what determines C_M.

REFERENCES

[1] W. Lenz, *Phys. Zeitschrift* **20**, 613 (1920).
[2] E. Ising, *Z. der Physik* **31**, 253 (1925).
[3] L. Onsager, Phys. Rev. **65**, 117 (1944).
[4] R. M. Ziff, E. Gulari, and Y. Barshad, Phys. Rev. Lett. **56**, 2553 (1986).
[5] J. Zhou, S. Redner, and H. Park, J. Phys. A **26**, 4197 (1993).
[6] C. Henley, Phys. Rev. Lett. **71**, 2741 (1993).
[7] P. Bak, K. Chen, and C. Tang, Phys. Lett. A **147**, 297 (1990).
[8] P. Brassberger and H. Krantz, J. Stat. Phys. **63**, 683 (1991).
[9] S. Ispolatov, P. L. Krapivsky, and S. Redner, Euro. J. of Phys. B, xxx (1998).
[10] E. Ben-Naim, P. L. Krapivsky, and S. Redner, Phys. Rev. E **50**, 822 (1994).
[11] I. Ispolatov, P. L. Krapivsky, and S. Redner, Phys. Rev. E **54**, 1274 (1996).
[12] *Nonequilibrium Statistical Mechanics in One Dimension*, ed. V. Privman (Cambridge University Press, Cambridge, U. K., 1997).
[13] P. L. Garrido, A. Labarta, and J. Marro, J. Stat. Phys. **49**, 551 (1987).
[14] H. W. Blöte, J. R. Heringa, A. Hoogland, and R. K. P. Zia, J. Phys. A **23**, 3799 (1990).
[15] T. L. Hill, J. Chem. Phys. **76**, 1122 (1982).
[16] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, U. K., 1995).
[17] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
[18] K. Kawasaki, Phys. Rev. **145**, 224 (1966).
[19] L. P. Kadanoff, Physics **2**, 263 (1966).
[20] K. G. Wilson and J. Kogut, Phys. Rep. **12C**, 75 (1974).
[21] P. Weiss, Comptes Rendus **143**, 1136 (1906); J. Phys. (France) **6**, 661 (1907).
[22] W. L. Bragg and E. J. Williams, Proc. Roy. Soc. London A **145**, 699 (1934); **151**, 540 (1935); **152**, 231 (1935).
[23] H. A. Bethe, Proc. Roy. Soc. London A **150**, 552 (1935).
[24] R. Kikuchi, Phys. Rev. **81**, 988 (1951).
[25] R. Dickman, Phys. Rev. A **34**, 4246 (1986).
[26] R. Dickman, Phys. Lett. A **122**, 463 (1987).
[27] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York, NY, 1971).
[28] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group* (Oxford University Press, Oxford, U. K., 1992).
[29] N. Metropolis, et. al., J. Chem. Phys. **21**, 1087 (1953).
[30] K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics* (Springer–Verlag, Heidelberg, Germany, 1988).
[31] M. E. Fisher and M. N. Barber, Phys. Rev. Lett. **28**, 1516 (1972).
[32] M. E. Fisher, in *Proceedings of the International Summer School Enrico Fermi 1970, Course 51, Varenna, Italy*, ed. by M. S. Green (Academic, New York, NY, 1971).
[33] K. Binder, *Z. Phys. B* - Condensed Matter **34**, 119 (1981).
[34] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[35] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[36] H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods: Applications to Physical Systems* (Addison-Wesley, Menlo Park, CA, 1988).
[37] K. E. Bassler and Z. Rácz, Phys. Rev. Lett. **73**, 1320 (1994).
[38] P. Tamayo and W. Klein, Phys. Rev. Lett. **63**, 2757 (1989).
[39] B. Schmittmann and R. K. P. Zia, Phys. Rev. Lett. **66**, 357 (1991).
[40] P. L. Garrido, J. L. Lebowitz, C. Maes, and H. Spohn, Phys. Rev. A **42**, 1954 (1990); Z. Cheng, P. L. Garrido, J. L. Lebowitz, and J. L. Valles, Europhys. Lett. **14**, 507 (1991).
[41] K.-t. Leung, Phys. Rev. Lett. **66**. 453 (1991).
[42] D. P. Landau, Phys. Rev. B **13**, 2997 (1976).
[43] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical Recipes in C, Second Edition* (Cambridge University Press, 1996).
[44] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).
[45] E. L. Praestgaard, H. Larsen, and R. K. P. Zia, Europhys. Lett. **25**, 447 (1994).
[46] G. Grinstein, C. Jayaprakash, and Y. He, Phys. Rev. Lett. **55**, 2527 (1985).
[47] C. H. Bennet and G. Grinstein, Phys. Rev. Lett. **55**, 657 (1985); P. L. Garrido, A. Labarta, and J. Marro, J. Stat. Phys. **49**, 551 (1987).
[48] A. DeMasi, P. A. Ferrari, and J. L. Lebowitz, Phys. Rev. Lett. **55**, 1947 (1985).
[49] J. M. Gonzalez-Miranda, P. L. Garrido, J. Marro, and J. L. Lebowitz, Phys. Rev. Lett. **59**, 1934 (1987).
[50] P. L. Garrido, J. Marro, and J. M. Gonzalez-Miranda, Phys. Rev. A **40**, 5802 (1989).
[51] S. N. Majumdar, A. J. Bray, S. J. Cornell, and C. Sire, Phys. Rev. Lett. **77**, 3704 (1996).
[52] S. N. Majumdar and C. Sire, Phys. Rev. Lett. **77**, 1420 (1996).
[53] B. Derrida, V. Hakim, and V. Pasquier, J. Stat. Phys. **85**, 763 (1996).
[54] H. K. Janssen, H. Schaub, and B. Schmittmann, *Z. Phys. B* - Condensed Matter **73**, 539 (1989).