Quantum Monte Carlo calculations

applied to magnetic molecules

by

Larry Paul Engelhardt

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Condensed Matter Physics

Program of Study Committee:
Marshall Luban, Major Professor
Bruce Harmon
David Johnston
John Lajoie
Glenn Luecke

Iowa State University

Ames, Iowa

2006

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Larry Paul Engelhardt

has met the dissertation requirements of Iowa State University

---

Committee Member

Committee Member

Committee Member

Committee Member

[signature]

Major Professor

For the Major Program

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.   Overview and introduction

We have calculated the equilibrium thermodynamic properties of Heisenberg spin systems using a quantum Monte Carlo (QMC) method. We have used some of these systems as models to describe recently synthesized magnetic molecules, and—upon comparing the results of these calculations with experimental data—have obtained accurate estimates for the basic parameters of these models. We have also performed calculations for other systems that are of more general interest, being relevant both for existing experimental data and for future experiments. Utilizing the concept of importance sampling, these calculations can be carried out in an arbitrarily large quantum Hilbert space, while still avoiding any approximations that would introduce systematic errors. The only errors are statistical in nature, and as such, their magnitudes are accurately estimated during the course of a simulation.[1] Frustrated spin systems present a major challenge to the QMC method, nevertheless, in many instances progress can be made.

1.1   Thesis organization

In this chapter, the field of magnetic molecules is introduced, paying particular attention to the characteristics that distinguish magnetic molecules from other systems that are studied in condensed matter physics. We briefly outline the typical path by which we learn about magnetic molecules, which requires a close relationship between experiments and theoretical calculations. The typical experiments are introduced here, while the theoretical methods are discussed in the next chapter. Each of these theoretical methods has a considerable limita-

---

[1] Although these statistical errors have been estimated for all of our calculations, the errors associated with the QMC data that are presented in this thesis are so small that error bars would not be visible on the plots, and have thus not been included.

tion, also described in Chapter 2, which together serve to motivate the present work. As is shown throughout the later chapters, the present QMC method is often able to provide useful information where other methods fail.

In Chapter 3, the use of Monte Carlo methods in statistical physics is reviewed, building up the fundamental ideas that are necessary in order to understand the method that has been used in this work. With these ideas in hand, we then provide a detailed explanation of the current QMC method in Chapter 4. The remainder of the thesis is devoted to presenting specific results: Chapters 5 and 6 contain articles in which this method has been used to answer general questions that are relevant to broad classes of systems. Then, in Chapter 7, we provide an analysis of four different species of magnetic molecules that have recently been synthesized and studied. In all cases, comparisons between QMC calculations and experimental data allow us to distinguish a viable microscopic model and make predictions for future experiments. In Chapter 8, the infamous "negative sign problem" is described in detail, and we clearly indicate the limitations on QMC that are imposed by this obstacle. Finally, Chapter 9 contains a summary of the present work and the expected directions for future research.

1.2 Introduction to magnetic molecules

The field of magnetic molecules exploded in the early 1990s,[1–4] and has continued to grow rapidly in the years since then.[5–7] As will be described in this chapter, magnetic molecules are experimental realizations of zero–dimensional, nanoscale systems of interacting quantum spins. As such, they produce a number of interesting quantum mechanical effects, including quantum entanglement[8] and, in the presence of an external magnetic field, level–crossings.[2] Since these systems exhibit magnetism that is of molecular origin, they are also (not surprisingly) of interest for quantum computing.[10] Magnetic molecules have received recent attention for other applications as well, including refrigeration[11] and use in the biomedical industry as MRI agents.[12] The focus of the present work is, however, not on applications. Instead, we seek to develop, employ, and assess a new theoretical tool that has not been previously used

---

[2]Level–crossings are described in detail in Chapter 5. For experimental results, see, for example, Refs. [2] and [9].

to study magnetic molecule systems.

The theory of magnetic molecules is typically quite complex. To appreciate this complexity, we must first have some understanding of the physical systems to which the term "magnetic molecules" refers. Magnetic molecules are synthesized by chemists in the form of macro-scopic (~ 100 milligram) crystalline or powder samples. Each sample contains a huge number ($\sim 10^{20}$) of identical molecular units which, in turn, are each composed of ~ 100 atoms. Most of these atoms are non–magnetic, leaving a smaller number N (typically ranging from 2 to 30) of magnetic ions in the core of each magnetic molecule. The non–magnetic ions contribute directly to the magnetism only in the form of weak diamagnetism, but they nonetheless play two very important roles: They (i) mediate strong superexchange interactions[3] between neigh-boring magnetic *ions*; and (ii) serve to magnetically isolate the *molecules* from one another by increasing the separation between their magnetic cores.

Since the molecules are well isolated, the interactions between molecules are orders of mag-nitude weaker than the intramolecular interactions, and are thus negligible.[4] This leads to a very important consequence: Experimental data measure the superposition of $\sim 10^{20}$ identical, *independent* molecules. The sample, when held at a fixed temperature, can hence be regarded as a canonical ensemble of such systems, and theoretical calculations are thus concerned with calculating the thermodynamic properties of a *single* magnetic molecule. The problem, how-ever, is that first–principles calculations are still formidable, involving the consideration of hundreds of interacting electrons, requiring considerable approximations.[5] Instead, we choose to approach magnetic molecules in terms of a simplified (but often very accurate) model for which we obtain numerically exact results by using a QMC method.

Throughout this work, we use the isotropic Heisenberg model,[15] which has long been used to model these types of interactions. This model relies on two central assumptions: First, each

---

[3]When two magnetic ions are separated by a non–magnetic ligand (atom or atoms), electrons from both magnetic ions can, by interacting with the nearby ligand electrons, have an effect on one another, resulting in a so-called "superexchange" interaction.

[4]The typical energy scales of intramolecular interactions are on the order of tens of degrees Kelvin, while the intermolecular interactions are on the order of milli-Kelvin. These weak intermolecular interactions can, however, be relevant for certain types of systems that are not considered here.[13]

[5]For a review of first–principles calculations for magnetic molecules, see [14] and the references therein.

magnetic ion possesses a well defined magnetic moment or "spin". These spins result from the strong coupling of valence electrons within a given ion, and hence their values s depend upon the element and ionization in question. For example, the ions studied in Chapter 7 include Cu^{2+} and V^{4+} with $s = 1/2$, Ni^{2+} with $s = 1$, and Cr^{3+} with $s = 3/2$; and Chapter 5 includes an analysis of the magnetic molecule $\{Fe_{12}\}$, which is composed of Fe^{3+} ions with $s = 5/2$. (Note that all of these ions belong to the Fe group, having partially filled 3d orbitals.) The second assumption is that the spins of neighboring magnetic ions within the same magnetic molecule interact isotropically (i.e., without a preferred direction based upon the local environment of the ion) with an interaction that depends on the dot-product between any pair of interacting spin vectors. This is believed to be a good approximation for these ions, while for other ions, e.g. Mn^{2+} and Mn^{3+}, additional anisotropic terms are known to be of importance.[16]

The problem thus reduces to that of calculating the properties of a relatively small number[6] N of interacting spins. However, even with these assumptions, the calculations are *still* very complex, as we describe in the next chapter. This final complexity arises from the fact that the number of ways in which the spins can be coupled together grows exponentially with N, forbidding the application of theoretical methods that require such enumeration. Instead, we use QMC for the task of calculating thermodynamic properties; the calculation of other properties, such as the non-equilibrium behavior of molecules, necessitates the use of totally different methods and is therefore the focus of other work.[18] In the following section, the typical experiments are introduced which motivate the development of these various theoretical methods.

1.3 Roles of experiment and theory

The study of magnetic molecules begins with the chemical synthesis of a particular sample. Then, with the sample in hand, one must next perform a series of experiments to determine the microscopic nature of the sample, both in terms of the structure and the physical interactions. Early on in this process, the geometric (crystal) structure is probed by x-ray diffraction,[19]

---

[6]The "standard" size of magnetic molecules is $2 \leq N \leq 30$, but G. Christou's group has also synthesized a $\{Mn_{84}\}$ cluster[17] which is much larger than other existing magnetic molecules.

whereby the details of the observed diffraction patterns provide information regarding the various bonds, including the types of ligands, interatomic distances, and bond angles. Based on this information, and previous experience with similar structures, rough predictions can then be made for the types of magnetic interactions that are likely to be present within the molecules.

As the term "*magnetic* molecules" suggests, we are centrally interested in magnetic properties. There are a variety of experiments which can be used to probe these properties, but the first to be carried out is generally the measurement of the weak-field magnetic susceptibility χ_0 as a function of temperature. This is because χ_0 is readily accessible experimentally, and there are usually a few pieces of information that are easily extracted from the data. These experiments are conducted using a SQUID magnetometer,[20] within which the magnetization of the sample is measured in a weak (~ 0.1 Tesla) magnetic field as the temperature is slowly lowered from room temperature (300 K) down to 2 K. The theory of this and other properties will be described in greater detail in the following chapter.

For reasons that are described in the next chapter, the computation and analysis of χ_0 data is often very challenging, but it can also offer a great deal of information. (As we describe in Chapter 7, we have successfully performed such analyses using QMC.) To analyze these data, we begin with the simplest model that is consistent with the rough predictions offered by the x-ray diffraction data. From such a model, theoretical values of χ_0 are calculated and compared with the data. Then, the parameters that describe the interactions in the Heisenberg model are adjusted, seeking to answer two questions: Is there a set of microscopic parameters that will adequately reproduce the experimental data? If so, is this result unique, or will a different set of parameters provide essentially the same result? If there is indeed a unique set of acceptable parameters, we are then prepared to make predictions for additional experiments. If not, the theoretical model is apparently too simple and must be further developed.[7]

Based on the model that emerges, we can then make detailed numerical predictions for other experiments. One such experiment is the measurement of low-temperature (presently down to

---

[7]For example, one could consider the inclusion of additional bonds, anisotropies, or more complicated interactions such as those described in Ref. [21] and the references therein.

as low as 20 mK) magnetization. This can be carried out either in static or pulsed fields, and both methods have several advantages and disadvantages. The maximum magnitudes of the static fields are currently ≈ 30 T, while the pulsed fields can somewhat exceed 50 T. Hence, if we were to predict a so-called "level-crossing field" (which are associated with the intersection of discrete energy levels, and are described in detail in Chapter 5) to occur at 40 T, a pulsed field measurement would be necessary in order to test this prediction. However, during the course of a measurement, the pulsed fields increase and decrease at rates of several Teslas per milli-second. Clearly then, if the equilibriation of the magnetic molecules were to take place on a time scale of ms, the experiment would *not* provide the equilibrium magnetization.[8] The static measurements do provide equilibrium data, but as a result they will not allow one to study dynamical behavior including hysteresis effects.

Inelastic neutron scattering (INS) is another technique that allows one to study similar features to those probed with low-temperature magnetization, but in a totally different way.[23] A neutron has a magnetic moment, which facilitates interactions between an incident beam of neutrons and the magnetic molecules contained within a target sample, thereby causing neutrons to scatter. A great deal of this scattering is elastic, i.e., without energy transfer, but the inelastic contribution is also readily measurable. At low temperatures, this occurs when the neutron transfers energy to a molecule, leaving the molecule in an excited state, while the neutron is detected with measurably lower energy than it initially had. This loss of neutron energy hence provides information regarding the energy gaps between the ground state and certain excited states within the magnetic molecules.[9] At higher temperatures, energy can also be transferred *from* the molecules *to* the neutrons, and as a result both energy gain and loss are observed.

Resonance experiments are also important, and these include nuclear magnetic resonance (NMR), and electron spin resonance (ESR). In NMR experiments, the nuclei within the molecules are used to probe the spins of the nearby ions. By exposing a sample to an alternating

---

[8]A description of the equilibriation process in magnetic molecules and the implications for pulsed field measurements can be found in Ref. [22].

[9]The energy spectra of magnetic molecules and the effect of temperature are discussed in greater detail in Chapters 2 and 5.

magnetic field and measuring the time that it takes these spins to respond to the changing fields, time-dependent correlations are obtained;[24] and by subjecting magnetic molecules to a large external magnetic field, one can study how level-crossings affect this behavior.[25] The spins can also be studied directly using ESR,[26] whereby one looks for transitions between spin states, which are shown as peaks in the spectrum of radiation that is absorbed (or emitted) by the magnetic molecules. These data are particularly useful for determining the g-value a given magnetic ion, including anisotropies that might be present.[10] Given this great variety of experimental tools, magnetic molecules can be thoroughly studied and understood, provided that one is able to construct a microscopic model that is able to link these various experiments, hence producing a more complete picture.

Bibliography

[1] R. Sessoli, D. Gatteschi, A. Caneschi, and M. Novak, Nature **365** (1993).

[2] D. Gatteschi, A. Caneschi, L. Pardi, and R. Sessoli, Science **265**, 1054 (1994).

[3] O. Kahn, *Molecular Magnetism* (John Wiley and Sons, 1993).

[4] D. D. Awschalom and D. P. DiVincenzo, Physics Today **48** (4), 43 (1995).

[5] R. E. P. Winpenny, Adv. Inorg. Chem. **52**, 1 (2001).

[6] R. E. P. Winpenny, Dalton Trans. **1**, 1 (2002).

[7] R. E. P. Winpenny, Comp. Coord. Chem. II **7**, 125 (2004).

[8] X. Wang and P. Zanardi, Phys. Lett. A **301**, 1 (2002).

[9] Y. Inagaki, T. Asano, Y. Ajiro, Y. Narumi, K. Kindo, A. Cornia, and D. Gatteschi, J. Phys. Soc. Jap. **75**, 1178 (2003).

[10] M. N. Leuenberger and D. Loss, Nature **410**, 789 (2001).

---

[10]In some cases (not studied here) local anisotropies can induce a magnetization that is not parallel to the external magnetic field, requiring one to adopt a g-tensor.

[11] F. Torres, J. M. Hernandez, X. Bohigas, and J. Tejada, Appl. Phys. Lett. **77**, 3248 (2000).

[12] E. Rodríguez, A. Roig, E. Molins, C. Arús, M. R. Quintero, M. E. C. nas, S. Cerdán, P. Lopez-Larrubia, and C. Sanfeliu, NMR in Biomedicine **18**, 300 (2005).

[13] A. Morello, F. L. Mettes, F. Luis, J. F. Fernáandez, J. Krzystek, G. Aromí, G. Christou, and L. J. de Jongh, Phys. Rev. Lett. **90**, 017206 (2003).

[14] A. V. Postnokiv, J. Kortus, M. Pederson, *Density functional studies of molecular magnets*, Scientific Highlight of the month, February 2004, Newsletter **61**, of the Ψ_k–Network. URL http://psi-k.dl.ac.uk./newsletters/News_61/Highlight_61.pdf

[15] W. Heisenberg, Z. Phys. **49**, 619 (1928).

[16] A. Cornia, R. Sessoli, L. Sorace, D. Gatteschi, A. L. Barra, and C. Daiguebonne, Phys. Rev. Lett. **89**, 257201 (2002).

[17] A. J. Tasiopoulos, A. Vinslava, W. Wernsdorfer, K. A. Abboud, and G. Christou, Angew. Chem. Int. Ed. **43**, 2117 (2004).

[18] I. Rousochatzakis and M. Luban, Phys. Rev. B **72**, 134424 (2005).

[19] J. P. Glusker, M. Lewis, and M. Rossi, *Crystal Structure Analysis for Chemists and Biologists* (Wiley, 1994).

[20] J. Clarke, Scientific American **271** (2), 46 (1994).

[21] J. R. Pilbrow and M. R. Lowrey, Rep. Prog. Phys. **43**, 433 (1980).

[22] I. Rousochatzakis, Ph.D. thesis, Iowa State University (2005).

[23] G. L. Squires, *Introduction to the Theory of Thermal Neutron Scattering* (Dover Publications, Inc., 1996).

[24] P. J. Hore, J. A. Jones, and S. Wimperis, *NMR: The Toolkit* (Oxford University Press, 2000).

[25] M.-H. Julien, Z. H. Jang, A. Lascialfari, F. Borsa, M. Horvatić, A. Caneschi, and D. Gatteschi, Phys. Rev. Lett. **83**, 227 (1999).

[26] F. E. Mabbs and D. Collison, *Electron Paramagnetic Resonance of d Transition Metal Compounds* (Elsevier Science, 1992).

CHAPTER 2.   Thermodynamic theory of finite quantum spin systems

In this chapter we review the theory of quantum statistical mechanics, specifically as it pertains to quantum spin systems and the present QMC method. Thermodynamic properties are initially introduced in the context of classical statistical mechanics, but are readily generalized to quantum systems. The quantum Heisenberg model is also introduced, as well as the challenges that one confronts when attempting to calculate thermodynamic properties for the Heisenberg Hamiltonian. These challenges serve to motivate the use of the present QMC method (described in Chapter 4) for systems of magnetic molecules. For additional background regarding the material that is contained in this chapter, see for example Refs. [1], [2], and [3].

2.1   Thermodynamic properties

There are four (equilibrium) thermodynamic properties that are simultaneously calculated during the course of every QMC simulation that we have performed: The internal energy U, the specific heat C, the magnetization M, and the differential susceptibility χ. The internal energy is not directly accessible in experiments, and is therefore not of primary interest for this work. Furthermore, the specific heat, though measurable, is experimentally inconvenient for most magnetic molecules. This is because the non-magnetic portion of a sample (which comprises the majority of the mass) provides a large contribution to the specific heat which is difficult to separate from the magnetic contribution.[4] The remaining properties, M and χ, are the primary focus of this work; however, as we describe in Chapters 3 and 4, U and C always "come along for the ride" during the course of a Monte Carlo calculation. Therefore, all four properties are introduced in this chapter.

The theory of these properties begins with Boltzmann's idea[1] that, in equilibrium, a microscopic system will occupy a state (which we will denote with an index i) with a *probability* that depends on the energy E_i of the state and the temperature T. More specifically, that probability $\mathcal{P}_i$ is proportional to $\exp(-E_i/k_BT)$, where k_B is Boltzmann's constant. The probability must be normalized such that the total probability of occupying any state is unity, thus

$$\mathcal{P}_i = \frac{e^{-E_i/k_BT}}{Z}, \tag{2.1}$$

where the normalization factor Z is called the "partition function" and is given by[2]

$$Z = \sum_i e^{-E_i/k_BT}, \tag{2.2}$$

with the summation extending over all states, i. The thermodynamic properties listed above are then weighted (thermodynamic) averages and the derivatives thereof, with the weights given by the probabilities $\mathcal{P}_i$. Very generally, this means that the thermodynamic average (henceforth denoted $\langle \cdots \rangle_T$) of a quantity Q is given by

$$\langle Q \rangle_T \equiv \sum_i Q_i \mathcal{P}_i \tag{2.3a}$$

$$= \frac{\sum_i Q_i e^{-E_i/k_BT}}{\sum_i e^{-E_i/k_BT}}, \tag{2.3b}$$

and it depends only on T, the energy spectrum $\{E_i\}$, and the value of Q_i that is associated with each state i.

Using this basic theory, let us now explicitly write down the formulas for U, M, C, and χ in terms of T and $\{E_i\}$. The internal energy U is the thermodynamic average of the (microscopic) energy, and is thus given by

$$U \equiv \langle E \rangle_T \tag{2.4a}$$

$$= \sum_i E_i \mathcal{P}_i \tag{2.4b}$$

$$= \frac{\sum_i E_i e^{-E_i/k_BT}}{\sum_i e^{-E_i/k_BT}}. \tag{2.4c}$$

---

[1] For a review of Boltzmann's work in statistical mechanics, see e.g. Ref. [5].

[2] For classical systems that involve continuous degrees of freedom, this sum is replaced by an integral over all phase-space; for the purposes of the present work, the summation is much more relevant.

Similarly, the magnetization M is the thermodynamic average of the individual (microscopic) magnetic moments M_i, so it is given by[3]

$$M = \sum_i M_i \mathcal{P}_i \tag{2.5a}$$

$$= \frac{\sum_i M_i e^{-E_i/k_B T}}{\sum_i e^{-E_i/k_B T}}. \tag{2.5b}$$

Inspecting Eq. (2.4c), U can alternatively be written in terms of a derivative of Z. At this point it is convenient to define the inverse temperature $\beta \equiv 1/k_B T$ in order to compactly write

$$U = -\frac{1}{Z}\frac{\partial Z}{\partial \beta}. \tag{2.6}$$

Since, in a magnetic field[3] H, the energy spectrum is shifted ($E_i \rightarrow E_i - HM_i$ for all i), the magnetization from Eq. (2.5b) can also be written as a derivative of Z,

$$M = \frac{1}{\beta Z}\frac{\partial Z}{\partial H}. \tag{2.7}$$

Eqs. (2.6) and (2.7) turn out to be useful representations for QMC, as we show in Chapter 4.

The specific heat and magnetic susceptibility are defined as derivatives of U and M, respectively. Differentiating Eq. (2.4c), C can be written

$$C \equiv \frac{\partial U}{\partial T} \tag{2.8a}$$

$$= -k_B \beta^2 \frac{\partial U}{\partial \beta} \tag{2.8b}$$

$$= k_B \beta^2 \left(\langle E^2 \rangle_T - \langle E \rangle_T^2 \right), \tag{2.8c}$$

while the differentiation of Eq. (2.5b) gives a similar relation for χ,

$$\chi \equiv \frac{\partial M}{\partial H} \tag{2.9a}$$

$$= \beta \left(\langle M^2 \rangle_T - \langle M \rangle_T^2 \right). \tag{2.9b}$$

It is important to note that all four of these quantities (U, C, M, and χ) can be written directly in terms of thermodynamic averages, without the need for further differentiation. In

---

[3]More precisely, M is the component of the magnetization vector $\vec{M}$ along a particular axis, and M_i is the component of an individual magnetic moment vector $\vec{M}_i$ along the *same* axis. However, since we deal exclusively with isotropic interactions, it is sufficient to note that this must be the same axis as that of an external magnetic field $\vec{H}$, and disregard the vanishing transverse components.

particular, the zero-field susceptibility, $\chi_0 \equiv \lim_{H \to 0} (\chi)$, is given by the fluctuations of the magnetic moments in Eq. (2.9b), and it can thus be calculated in the absence of a field. Experimentally, this quantity is obtained by measuring M and dividing by H, and as long as H is sufficiently weak (~ 0.1 T), M/H is essentially identical to χ_0.

In order to perform calculations for a sample of magnetic molecules, the preceding formulas are immediately applicable, and (using the methods that are described in detail in Chapters 3 and 4) they are employed extensively throughout the later chapters. To do this, instead of thinking in terms of a probability $\mathcal{P}_i$ for a *single* molecule, one can instead recognize that the number $\mathcal{N}$ of individual magnetic molecules is so large ($> 10^{20}$) that the number of magnetic molecules in state i will be accurately given by $\mathcal{P}_i \mathcal{N}$; and the magnetization that is measured for such a sample is then given by the product $\mathcal{N}M$. Similarly, the purpose of the Monte Carlo calculations is to generate a large ensemble of simulated (individual) magnetic molecules, where each member of the ensemble is chosen so as to correctly reproduce the partition function. The one remaining issue to be addressed is the precise definition of the words "states" and "energies" with regard to quantum spin systems. This is clarified in the following section.

2.2 Representing quantum spin systems

For classical systems, the calculation of E_i for any given state i is typically simple. The challenge of calculating Z (and hence the thermodynamic properties) instead lies in properly accounting for the contribution that each state makes to Z, i.e., performing the necessary sums or integrals.[4] For quantum systems this is not the case. In addition to being faced with unwieldy sums, the calculation of any given E_i can be an insurmountable task in and of itself. This difference arises because a classical "state" is easily constructed by simply specifying all of the degrees of freedom. In contrast, an energy E_i is *only* defined for a quantum state $|\phi_i\rangle$ if it satisfies the time-independent Schrödinger equation,[6, 7]

$$\mathcal{H}|\phi_i\rangle = E_i|\phi_i\rangle, \tag{2.10}$$

---

[4]In the next chapter the application of the Monte Carlo method to this challenge is described in detail.

where $\mathcal{H}$ is the Hamiltonian operator[5] which defines the model system and its interactions. The task of evaluating the energy spectrum hence amounts to solving this eigenvalue equation [Eq. (2.10)] for the eigenvalues E_i and eigenstates $|\phi_i\rangle$. As we discuss below, this is straightforward for small spin systems, but becomes very complex with increasing system size.

Throughout this work calculations are performed only for the Heisenberg model, described by the Hamiltonian

$$\mathcal{H} = \sum_{\langle j,k \rangle} J_{j,k}\vec{\underline{s}}_j \cdot \vec{\underline{s}}_k + g\mu_B \vec{H} \cdot \sum_{j=1}^{N} \vec{\underline{s}}_j, \tag{2.11}$$

where the summation $\langle j, k \rangle$ is over all distinct pairs of interacting spins, the spin operators $\vec{\underline{s}}_j$ are given in units of $\hbar$, g is the spectroscopic splitting factor,[6] and μ_B is the Bohr magneton.[6] The quantity $J_{j,k}$ is often called the "exchange constant";[8] its magnitude determines the strength of the interaction between spins j and k, and its sign determines whether the interaction is antiferromagetic (AFM) or ferromagnetic (FM). For $J_{j,k} > 0$, the interaction is AFM and (roughly speaking) this tends to align spins j and k anti-parallel to one another, while $J_{j,k} < 0$ describes a FM interaction, causing the spins to tend to point in the same direction.

For this particular Hamiltonian, the square of the total spin operator $\vec{\underline{S}}^2$ and the z-component of the total spin operator $\underline{S}^z$ both commute with $\mathcal{H}$. This implies that—in the basis that consists of $|\phi_i\rangle$, which are the eigenstates of $\mathcal{H}$—the eigenvalues that are associated with $\vec{\underline{S}}^2$ and $\underline{S}^z$ are good quantum numbers. They are given by the solutions of the eigenvalue equations,

$$\vec{\underline{S}}^2|\phi_i\rangle = S(S+1)|\phi_i\rangle, \tag{2.12a}$$

$$\underline{S}^z|\phi_i\rangle = M_S|\phi_i\rangle, \tag{2.12b}$$

where S is referred to as the total spin of the state $|\phi_i\rangle$, and M_S, the z-component of the total spin, takes on values ranging in integer steps from $-S$ to S. Specifying S and M_S does

---

[5]Note that the tilde beneath the character $\mathcal{H}$ is meant to distinguish the character as representing a quantum operator. This notation is used throughout this work.

[6]Note that Eq. (2.11) assumes g to be a scalar (i.e., fully isotropic) quantity. Although it is usually small, some amount of anisotropy is expected for any physical system. Unfortunately though, realistic anisotropic interactions are outside the scope of this thesis, as we describe in Sec. 9.1. Additionally, for certain magnetic molecules more than one species of magnetic ions is present, so g varies from site to site. This is described in Chapter 7.

not, however, uniquely specify the state $|\phi_i\rangle$. For example, there are typically very many eigenstates with the quantum numbers $S = 0$, $M_S = 0$. Also, it is usually[7] not obvious how one can evaluate $\mathcal{H}|\phi_i\rangle$ without first beginning with a different basis, and then constructing $|\phi_i\rangle$ as a linear combination of those basis states. These two issues are both avoided by choosing the basis described below.

For the sake of convenience[8] we choose to represent the spin state (upon which $\mathcal{H}$ acts) in terms of the basis states for which the z-component of each spin is a good quantum number. These states will be represented $|\psi_\alpha\rangle$, where the index α denotes a particular N-tuple of quantum numbers, and each of these N quantum numbers will be denoted by m_j ($1 \leq j \leq N$). This choice can be stated very compactly with the eigenvalue equation,

$$\underset{\sim}{s}_j^z|\psi_\alpha\rangle = m_j|\psi_\alpha\rangle, \qquad (-s_j \leq m_j \leq s_j), \qquad (2.13)$$

where the operator $\underset{\sim}{s}_j^z$ is the z-component of $\underset{\sim}{\vec{s}}_j$, and m_j is the corresponding eigenvalue that is associated with $|\psi_\alpha\rangle$. (These basis states $|\psi_\alpha\rangle$ are henceforth referred to simply as the "z-states" for brevity.) In this basis, M_S is still a good quantum number, given by

$$M_S = \sum_{j=1}^{N} m_j, \qquad (2.14)$$

but S is not a good quantum number. For a simple example of z-states, consider a system composed of $N = 2$ spins, each having $s = 1$. There would be 9 z-states, and they can be written in terms of the quantum numbers m_1 and m_2 using the notation $|\psi_\alpha\rangle = |m_1, m_2\rangle$. In this notation they are:

$$|-1,-1\rangle \qquad (M_S = -2);$$
$$|-1,0\rangle, |0,-1\rangle \qquad (M_S = -1);$$
$$|-1,1\rangle, |0,0\rangle, |1,-1\rangle \qquad (M_S = 0);$$
$$|0,1\rangle, |1,0\rangle \qquad (M_S = 1);$$
$$|1,1\rangle \qquad (M_S = 2).$$

---

[7]In Sec. 6.2.3 we describe a few special cases where this is not the case.

[8]This choice is important for the QMC method as we describe in Chapter 4.

Since there are $2s+1$ possible quantum numbers for each of the N spins, the number of z-states is always[9] equal to $(2s+1)^N$, so they quickly become tedious to write out. Very importantly though, the z-states are generally *not* the same as $|\phi_i\rangle$ which are the eigenstates of $\mathcal{H}$. Instead, each $|\phi_i\rangle$ is a linear combination of the z-states, and can be written

$$|\phi_i\rangle = \sum_{\alpha=1}^{(2s+1)^N} c_{i,\alpha}|\psi_\alpha\rangle. \tag{2.15}$$

Each $|\phi_i\rangle$ can therefore be defined by a $(2s+1)^N$-component vector, $(c_{i,1}, c_{i,2}, \dots)$, so we say that the system has a Hilbert space of dimension $D = (2s+1)^N$.

When adopting this basis, it is essential to know what happens when the Hamiltonian [Eq. (2.11)] acts on the z-states. Recalling Eq. (2.13), the Zeeman term,[6] $g\mu_B \vec{H} \cdot \sum_j \vec{s}_j$, can be trivially replaced with its eigenvalue $g\mu_B H M_S$, provided that we choose $\vec{H}$ to be directed along the z-axis. The dot products, $\vec{s}_j \cdot \vec{s}_k$, do not simply provide eigenvalues, but can be written[9]

$$\vec{s}_j \cdot \vec{s}_k = s_j^z s_k^z + \frac{1}{2}\left(s_j^+ s_k^- + s_j^- s_k^+\right), \tag{2.16}$$

where the operators $s_j^\pm$—when acting on a spin that is initially in a state with quantum numbers s_j and m_j (represented $|s_j, m_j\rangle$)—alter the state such that[7]

$$s_j^\pm |s_j, m_j\rangle = \sqrt{s_j(s_j+1) - m_j(m_j \pm 1)}|s_j, m_j \pm 1\rangle. \tag{2.17}$$

This means that the quantity $\langle \psi_\alpha|\mathcal{H}|\psi_\gamma\rangle$ is non-zero for many combinations of $\alpha \neq \gamma$, which has important consequences for the calculation of thermodynamic properties. This is discussed in the next section.

2.3 Quantum statistical mechanics

Assuming for a moment that one had already determined the energy eigenvalues E_i, the formulas from Sec. 2.1 could be immediately applied to the (quantum) Heisenberg Hamiltonian to calculate U, C, M, and χ. To do this, one would simply insert these energies (for each E_i) and replace the classical magnetic moments with the appropriate quantum numbers, $M_i \rightarrow$

---

[9]For a system composed of spins with different values of s, the number of basis states is more generally $\prod_{j=1}^{N}(2s_j + 1)$. Such cases are considered in Chapter 7.

$g\mu_B M_S$; the partition function from Eq. (2.2) would then still apply. There is, however, no way to know these eigenvalues from the start, and their calculation is often difficult or impossible. Instead, we now review how the equations from Sec. 2.1 can be generalized to an arbitrary basis.

To accomplish this generalization, consider the matrix-representation[3] of $\mathcal{H}$. In the basis that is composed of the $|\phi_i\rangle$ states, this matrix is diagonal, i.e., the matrix elements $\langle\phi_i|\mathcal{H}|\phi_j\rangle$ are zero for off-diagonal terms $(i \neq j)$, while the diagonal terms $(i = j)$ give $\langle\phi_i|\mathcal{H}|\phi_i\rangle = E_i$. The "density matrix", which is the matrix representation of the operator $\exp(-\beta\mathcal{H})$, is therefore also diagonal in this basis, since $\langle\phi_i|\exp(-\beta\mathcal{H})|\phi_j\rangle$ is equal to $\exp(-\beta E_i)$ for $i = j$, and is zero for $i \neq j$. Clearly then, the partition function, as it is written in Eq. (2.2), is a sum over all of the diagonal elements of the density matrix; or in the language of linear algebra,[10]

$$Z = \mathrm{Tr}\left(e^{-\beta\mathcal{H}}\right). \tag{2.18}$$

Finally, recall that the trace of an operator is independent of the basis, so Eq. (2.18) can be applied for any choice of basis, and in particular the z-basis. This is a very important point which will be exploited, as we describe in Chapter 4, to calculate the thermodynamic properties *without* calculating the energy eigenvalues E_i, but by instead numerically evaluating traces in terms of the z-states.

2.4   Motivation for using quantum Monte Carlo

Using the formulas that have been reviewed in the previous sections, the calculation of thermodynamic properties for the Heisenberg Hamiltonian is (in principle) a straightforward task: One first calculates the matrix elements of $\mathcal{H}$ in some basis. Choosing the z-states for example, these matrix elements, $\langle\psi_\alpha|\mathcal{H}|\psi_\gamma\rangle$, can be calculated using Eqs. (2.16) and (2.17); Using methods that are well-known from linear algebra,[10] the Hamiltonian matrix would then be diagonalized to obtain E_i and $|\phi_i\rangle$ as linear combinations of $\langle\psi_\alpha|\mathcal{H}|\psi_\gamma\rangle$ and $|\psi_\gamma\rangle$, respectively.

Although straightforward, this calculation would however be extremely impractical for large systems. The diagonalization of a matrix (which in this case has $D \times D$ total elements)

has a computational complexity that grows as the cube of the number of matrix elements. Consequently, even when taking full advantage of symmetries,[12] and only diagonalizing the necessary sub-spaces of $\mathcal{H}$, matrix diagonalization methods are still currently limited to $D \lesssim 10^6$ or 10^7. For example, the $\{Ni_{12}\}$ cluster ($N = 12$ spins, each with $s = 1$) that is studied in Sec. 7.3 has $D \approx 5.31 \times 10^5$, so matrix diagonalization is feasible, but it requires much more computation time than our QMC calculations.[10] On the other hand, some of the molecules described in Chapter 7 have $D > 10^9$, so matrix diagonalization is not even an option. For the study of sufficiently small spin systems (e.g., the study of a $\{Cr_8\}$ molecule with $D = 65,536$ is described in Ref. [11]), computer codes are currently available which can carry out these matrix calculations. In particular, we have used the MAGPACK (Magnetic Properties Analysis Package for Spin Clusters) package, described in Refs. [12] and [13], to rigorously test our QMC calculations on small systems before approaching the large systems that are described in Chapters 5–8.

Other methods, such as the Lanczos[14, 15] and Davidson[16] methods, are able to handle larger D by only attempting to calculate a smaller number ν of low-energy states, where $\nu \ll D$. These methods begin with states that are linear combinations of all of the D basis states, but then project them onto a smaller subspace (composed of only ν states) within which $\nu \times \nu$ matrices are diagonalized. The obvious advantage of such an approach is that one does not need to diagonalize huge matrices. However, it is still necessary to record the linear combinations of the D basis states, so practical limitations are still imposed by the value of D. Specifically, these calculations are currently restricted to $D \lesssim 10^9$. Furthermore, since only low-energy states are considered, Eq. (2.1) implies that these methods will only provide valid results for thermodynamic properties at low temperatures.

We would like to finally mention one other theoretical method that has been commonly used when applying the Heisenberg Hamiltonian to the study of magnetic molecules: The classical spin model. This approximation is described in detail in Sec. 6.2.1, but the general

---

[10]For this particular molecule, we were able to consider roughly a thousand different sets of exchange parameters over the course of a weekend, allowing a thorough search of the complex parameter space, while the exact diagonalization required several hours for each set of exchange parameters. See Sec. 7.3 for more detail.

idea is to replace the quantum spin operators in Eq. (2.11) by *classical* unit vectors, $\vec{s}_j \rightarrow \vec{e}_j$, and use classical Monte Carlo methods to sample the resulting partition function. Since this is a classical approach, it is not limited by D, but it suffers from a different limitation. Namely, below a certain temperature the results of a classical spin model will differ substantially from those of the quantum spins that it is attempting to represent. An accurate estimate of when (i.e., at what temperature for a given value of s) this difference occurs was not previously known, so in Chapter 6 we systematically compare the results of the quantum and classical Heisenberg models for many systems in order to establish clear guidelines. These results are provided in Sec. 6.3.

To summarize, every method that exists for calculating the thermodynamic properties of spin systems has a crucial limitation, shown schematically in Fig. 2.1. Exact analytical results are often not possible, so we instead resort to numerical methods. Exact (numerical) diagonalization can provide results for arbitrary geometries and interactions, but is currently limited to systems with $D \lesssim 10^6$. Lanczos and Davidson methods allow one to study a restricted subset of eigenvalues for larger systems, but are still limited to $D \lesssim 10^9$. One could instead use a classical model, and this will provide accurate results, unhampered by D, as long as s and T are sufficiently large,[11] but at lower temperatures the results are not reliable.

The combination of these limitations result in the large "hole" that appears in Fig. 2.1, suggesting that an additional tool (or tools) is needed in order to successfully study the full variety of magnetic molecules; this provided the motivation for the present work. The QMC method described in Chapter 4 is immune to the size of D (unlike the diagonalization methods), and does not involve systematic approximations (unlike the classical approximation). As a result, we are indeed able to fill in much of the hole, and provide results for systems that were previously untouchable.

Despite this success however, the QMC method has a different set of limitations. Recall that the advantage of the QMC approach is that we do not calculate energy eigenvalues and eigenstates. We instead use Monte Carlo methods (described in the next chapter) to sample

---

[11]See Chapter 6 for further details.
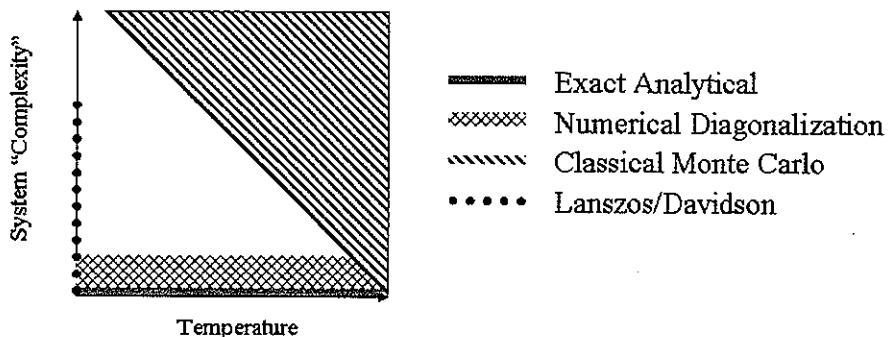
Figure 2.1 Schematic diagram which includes the theoretical methods that have typically been used to obtain the thermodynamic properties of magnetic molecules, and shows the regimes for which each of the methods can (and cannot) be applied.

the partition function, avoiding the obstacle imposed by large values of D. The cost of this is that our QMC results do not provide information about matrix elements, and hence cannot be used to calculate time-dependent phenomena. Furthermore, the QMC method suffers from the well-known "negative sign problem" which, for frustrated systems, restricts our calculations to high temperatures. This is a serious limitation, which is discussed in detail in Chapter 8.

Bibliography

[1] L. E. Reichl, *A Modern Course in Statistical Physics* (John Wiley and Sons, 1998), 2nd ed.

[2] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, vol. 5 of *Course of Theoretical Physics* (Pergamon Press, 1980), 3rd ed.

[3] R. Shankar, *Principles of Quantum Mechanics* (Springer, 1994), 2nd ed.

[4] M. Affronte, J. C. Lasjaunias, and A. Cornia, Eur. Phys. J. B **15**, 633 (2000).

[5] E. G. D. Cohen, cond-mat/9608054.

[6] R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles* (John Wiley and Sons, 1985), 2nd ed.

[7] D. J. Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, 1995).

[8] D. C. Mattis, *The Theory of Magnetism I: Statics and Dynamics*, vol. 17 of *Springer Series in Solid-State Sciences* (Springer-Verlag, 1988), 2nd ed.

[9] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, 1996), 4th ed.

[10] A. D. Martin and V. J. Mizel, *Introduction to Linear Algebra* (McGraw-Hill, 1966).

[11] M. Luban, P. Kögerler, L. L. Miller, and R. E. P. Winpenny, J. Appl. Phys. Soc. **93**, 7083 (2003).

[12] J. J. Borrás-Almenar, J. M. Clemente-Juan, E. Coronado, and B. S. Tsukerblat, Inorg. Chem. **38**, 6081 (1999).

[13] J. J. Borrás-Almenar, J. M. Clemente-Juan, E. Coronado, and B. S. Tsukerblat, J. Comp. Chem. **22**, 985 (2001).

[14] C. Lanczos, J. Res. Nat. Bur. Stand. **45**, 255 (1950).

[15] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations* (Birkhäuser, 1985).

[16] E. R. Davidson, J. Comput. Phys **17**, 87 (1975).

CHAPTER 3.   The Monte Carlo method in classical statistical physics

The term "Monte Carlo" (MC) refers to any method that relies on random numbers to perform a calculation. This is an extremely broad definition, which includes calculations in all areas of science and engineering, and even economics. In statistical mechanics alone, the use of MC methods is extremely varied, including the determination of ground states with simulated annealing[1] and the study of dynamics with kinetic Monte Carlo.[2] A broad review of the uses of MC methods in statistical physics can be found in Refs. [3] and [4]. In this chapter, we instead restrict our discussion to the issues that are directly relevant to the calculations that are described in the later chapters.

As we described in the previous chapter, the evaluation of thermodynamic properties amounts to the computation of averages, weighted by the terms that comprise the partition function. For classical models, these terms are written $\exp(-\beta E_i)$, yielding averages that take the form of Eq. (2.3) for a general quantity Q. More specifically, for the calculation of U, C, M, and χ, one simply replaces Q_i in Eq. (2.3b) with E_i, E_i^2, M_i, and M_i^2, as we have shown in Eqs. (2.4)–(2.9). Each of these quantities are trivial to calculate for any *single* term, but the summation over all of the states is often impossible. However, one can avoid this obstacle by instead using the MC strategy that is reviewed in this chapter.[1] As we describe below, these thermodynamic averages can thereby be obtained to arbitrary accuracy, where higher accuracy simply requires more computation time. Furthermore, the statistical errors are reliably estimated during the course of a MC simulation, as we describe in Sec. 3.5.

---

[1] This approach also applies to systems with continuous degrees of freedom in *exactly* the same way. For these cases however, it is the evaluation of complicated multi-dimensional integrals, $Z = \int d\Omega \exp(-\beta E)$, that one wishes to avoid.

3.1   Monte Carlo steps based on detailed balance

In order to use a MC method to calculate thermodynamic properties, one need only recall the principle of "detailed balance", which relates the probability of occupying a microscopic state [given by Eq. (2.1)] to the probability of making a transition between states. Specifically, *if* these transition probabilities satisfy the detailed balance condition, *then* the distribution of states thus obtained will correspond to the equilibrium distribution. The detailed balance condition to be satisfied is extremely simple to write down, and is given by

$$\frac{P_{j \to i}}{P_{i \to j}} = \frac{\mathcal{P}_i}{\mathcal{P}_j}, \tag{3.1}$$

where $\mathcal{P}_i$ and $\mathcal{P}_j$ represent the probabilities (in equilibrium at a temperature T) of occupying state i and state j, respectively [see Eq. (2.1)]; and $P_{j \to i}$ and $P_{i \to j}$ represent the probabilities of making a transition from state j to state i, and from state i to state j, respectively. A Monte Carlo calculation then proceeds by making transitions in such a way that Eq. (3.1) is satisfied.

To see that the detailed balance condition truly will produce the desired equilibrium distribution, consider first a system having only two states, i and j, with $p \equiv \mathcal{P}_i/\mathcal{P}_j > 0$. As transitions are made back and forth between the two states, the transitions from j to i will occur p times as often as those from i to j. As a result, the system will spend p times as much time in state i, so the correct distribution is indeed produced. Now consider a third state, k, with a probability $\mathcal{P}_k$ that satisfies $q \equiv \mathcal{P}_j/\mathcal{P}_k > 0$. We already showed that states i and j are occupied according to the correct distribution; similarly, the system will spend q times as much time in state j as it does in state k. Therefore, the equilibrium distribution is again obtained. In fact, by continuing this argument, Eq. (3.1) will *always* produce the correct equilibrium distribution, as long as it is possible to get from any state to any other state in a finite number of transitions—a condition known as "ergodicity".

To show how the detailed balance condition can be used, let us now consider a concrete example. Suppose that our system consists of one person, who we will name Carl, and the state of that person is specified solely by the city in which he lives. He can live in any one of

N_c cities, represented $i = 1, 2, 3, \ldots, N_c$, with populations $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, $\ldots$, $\mathcal{P}_{N_c}$, respectively. One could then perform a Monte Carlo simulation in the following manner:

(1) Begin with Carl living in one of the N_c cities, chosen at random.

(2) Propose that Carl move from his current city to a neighboring city.

(3) Choose to accept or reject this proposed move based on a probability that is consistent with Eq. (3.1).

(4) Repeat steps 2 and 3 many times. (The details of steps 2 and 3 will be made more explicit in the following section.)

After many proposed moves (each of which is referred to as a "Monte Carlo step", or "MCS"), the number of times that Carl has chosen to live in city i will be approximately proportional to $\mathcal{P}_i$. Of course, for any finite number of MCS, represented N_{MC}, this proportionality is not exact, but the exact distribution is *approached* in the limit $N_{MC} \to \infty$. Suppose now that there is some quantity Q—e.g., Carl's salary—which is determined entirely by the city in which he lives. Furthermore, assume that there are many such people who are identical to one another and do not influence one another.[2] The average salary that these people earn is then easily calculated by recording the value of Q_i at each MCS and averaging, i.e.,

$$\langle Q \rangle \approx \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} Q_i. \tag{3.2}$$

For any finite number of MCS, Eq. (3.2) is also only approximate, with the exact equality being approached in the limit $N_{MC} \to \infty$. This is perfectly analogous to the rolling of a 6-sided die. A small number of rolls will not provide accurate information about the die, but after *many* rolls, each number (1–6) would be encountered 1/6 of the time, so the average value rolled would be $21/6 = 3.5$. Of course if the die were "weighted", then we would not have $\mathcal{P}_i = 1/6$ for all i, and a different distribution would be obtained.

---

[2]Economically, this assumption would of course be ridiculous, but it can be valid for physical systems. For a very relevant example, there are many identical, non-interacting magnetic molecules in a macroscopic sample.

3.2   From detailed balance to importance sampling

The simple four-step algorithm (and the subsequent averaging) that was described in the previous section very accurately describes the "big picture" of a Monte Carlo calculation. However, there are many additional—often subtle—details that must be elucidated. First, the discussion of the previous section involved making transitions with probabilities that depended on $\mathcal{P}_i$. This would appear to be problematic, since $\mathcal{P}_i$, which is defined in Eq. (2.1), depends on Z. In fact, the whole motivation for using MC is that one typically cannot calculate Z. Therefore we cannot have a useful algorithm that requires a knowledge of Z. However, if one proceeds to insert $\mathcal{P}_i$ and $\mathcal{P}_j$ into Eq. (3.1), the resulting form of the detailed balance equation,

$$\frac{P_{j\rightarrow i}}{P_{i\rightarrow j}} = \frac{\exp(-\beta E_i)/Z}{\exp(-\beta E_j)/Z} \tag{3.3a}$$

$$= \exp\left(-\beta(E_i - E_j)\right), \tag{3.3b}$$

is *independent* of Z. In fact, it depends on only two energies—those of the current state and the proposed state—and it is hence easily evaluated.

Secondly, note that the total number of (microscopic) states is often incomprehensibly large. Perhaps the simplest such example is that of classical Ising spins, i.e., spins that can point only "up" or "down".[5–7] A *small* 3-dimensional lattice of Ising spins, having only 10 spins along each edge (for a total of $10 \times 10 \times 10$ spins), would give $2^{1000} \approx 10^{301}$ different possible states! Obviously one could never hope to separately count the contribution that each state makes to the partition function. We can, however, still perform MC calculations, and actually obtain useful results, by applying a very important principle: The vast majority of these states can often be *ignored*, because the contribution that they make to the partition function is utterly negligible. Furthermore, the relatively small subset of remaining, "important" states can be "sampled", just as voters are sampled in a political poll. As is the case in polls, a relatively small sample will yield accurate results, provided that important demographics are not excluded. (In the language of Monte Carlo simulations, this condition of non-exclusion is enforced by ergodicity: It must be possible to access all of the important regions of phase-space

in fewer than N_{MC} Monte Carlo steps.)

The application of this principle is generally known as "importance sampling".[3] It was first used in 1953 by Metropolis, *et al.* in Ref. [8], and has since been applied to numerous systems, some of which are discussed in Ref. [3]. Their method, now known as the "Metropolis algorithm", consisted of simply taking the principle of detailed balance; writing it as it appears in Eq. (3.3b); and choosing to always accept a transition from a higher energy state to a lower energy state. For any given transition, if one defines state i to be the state with higher energy (i.e., $E_i \geq E_j$), then this choice is equivalent to setting the denominator of Eq. (3.3b) equal to one. Therefore, in the Metropolis algorithm, the probability of accepting a proposed transition is given by

$$P_{j \to i} = \begin{cases} \exp\left(-\beta(E_i - E_j)\right) & \text{if } E_i > E_j, \\ 1 & \text{if } E_i \leq E_j. \end{cases} \tag{3.4}$$

3.3 Proposing and accepting transitions

We can now begin to clarify steps (2) and (3) of the 4-step algorithm that was presented in the previous section. First, recall that $P_{j \to i}$ is the probability of *making* a particular transition from the current state j to a different state i, while step (2) involves *proposing* a transition, and step (3) involves *accepting* the transition. For lack of a better symbol, we will represent the probability of proposing this transition by $R_{j \to i}$, and the probability of accepting the proposed transition will be represented $A_{j \to i}$. In order to make a transition, it must first be proposed [in step (2)], and then be accepted [in step (3)], giving a total probability of $P_{j \to i} = R_{j \to i} A_{j \to i}$.

Often $R_{j \to i}$ can be trivially set equal to $R_{i \to j}$ (as explained below), in which case both $R_{j \to i}$ and $R_{i \to j}$ can be eliminated from the detailed balance equations, leaving only $A_{j \to i}$ and $A_{i \to j}$. Clearly, in this situation, Eqs. (3.3) and (3.4) would describe the probability of *accepting* a proposed transition, so Eq. (3.4) could be used in step (3). For example, again consider a system of Ising spins. Given a current state—specified by which spins point up, and which point down—the next state can be proposed by simply flipping random spins. Since each spin

---

[3]The alternative to importance sampling is called "simple sampling". It is described in Ref. [3] and the references therein.
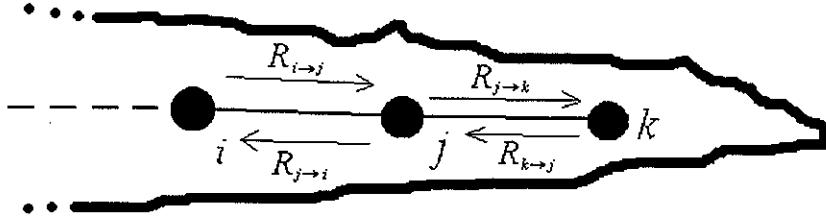
Figure 3.1   Diagram illustrating the probabilities of proposing transitions
($R_{j\to i}$, etc.) for the example of a person moving from city to
city along a peninsula.

has only two possible states, up or down, $R_{j\to i}$ is then automatically independent of i, and
can hence be ignored.[4]

Actually, it turns out that $R_{j\to i}$ and $R_{i\to j}$ can always be set equal to one another and
eliminated from the detailed balance equations. Therefore Eq. (3.4) can always be used to
determine $A_{j\to i}$ in step (3). Sometimes, however, how to accomplish this [in step (2)] requires
careful thought. For instance, recall the example of Carl moving from city to city from the
previous section. Suppose three cities—i, j, and k—are all on a peninsula, with k at the tip.
This situation is illustrated in Fig. 3.1. If Carl currently lives in j, then he can move to either i
or k; and in step (2) of the algorithm, we can set $R_{j\to i} = R_{j\to k} = 1/2$. This means that a move
will definitely be proposed. If, however, Carl currently lives in k, then the only city to which
a move can be proposed is j. Since we already set $R_{j\to k} = 1/2$, we must also set $R_{k\to j} = 1/2$
in order for Eq. (3.4) to describe the acceptance probability. This means that there is a 50%
chance that a move will not even be proposed.

This issue of dealing with $R_{j\to i}$ can often be ignored (as it is with the Ising model),
so frequently Eq. (3.4) is introduced as the acceptance probability without mentioning the
corresponding subtleties of step (2). The treatment of $R_{j\to i}$ is, however, very relevant to the

---

[4]Specifically, with 1000 Ising spins, one could propose a new state by flipping 1 spin. There would be
1000 different spins from which to choose; so by choosing the spin at random, we would have $R_{j\to i} = 1/1000$,
independent of i.

QMC method that we have employed. In that context, each MCS involves spin raising and lowering operators acting on quantum z-states. For $s = 1/2$, the spins are simply "flipped" (much like Ising spins); but for $s > 1/2$, when $m_j = s_j$ the spin state can be lowered, but cannot be raised (much like the city at the end of the peninsula). This is described in greater detail in Chapter 4.

3.4   Using random numbers to attain probabilities

Finally, we are ready to precisely state steps (2) and (3) of the 4-step algorithm that was presented in the previous section. In step (2), a transition from the current state j to a different state i is proposed with a probability $R_{j \to i}$, the details of which depend on the the specific problem under consideration. For some problems (e.g., the Ising model) a transition is always proposed, while for other problems (e.g., Carl moving between cities) it is sometimes necessary to *not* propose a transition. The decision of which transition to propose (or not propose) can be made in the following way: First, select an interval between 0 and 1 of size $R_{j \to i}$. Next, generate a random number r, uniformly distributed between 0 and 1. The transition from j to i is then proposed if and only if r falls in the selected interval.[5]

In step (3), the proposed transition is accepted (or rejected) with a probability given by Eq. (3.4). If the proposed state has a lower energy than the current state, the proposed transition is accepted. If not, we can make the necessary decision by generating another random number r, again uniformly distributed between 0 and 1. If r is less than $P_{j \to i}$, then the proposed transition is accepted. If $r > P_{j \to i}$, the transition is rejected.

One can equivalently view this decision in terms of the curves that appear in Fig. 3.2. These curves represent Eq. (3.4), plotted as a function of the energy difference, $E_i - E_j > 0$, for multiple fixed temperatures. If the energy of the proposed state E_i is greater than the energy of the current state E_j, then one generates the random number r. Whether or not to accept the proposed transition depends on three values: the temperature, the energy difference, and r. The temperature defines a *single* curve which is relevant for all N_{MC} of the MC steps.

---

[5] Note that in Fig. 3.1 the two transitions from j ($R_{j \to i} = R_{j \to k} = 0.5$) will together span the entire interval from 0 to 1, while $R_{k \to j}$ leaves half of the interval unoccupied.
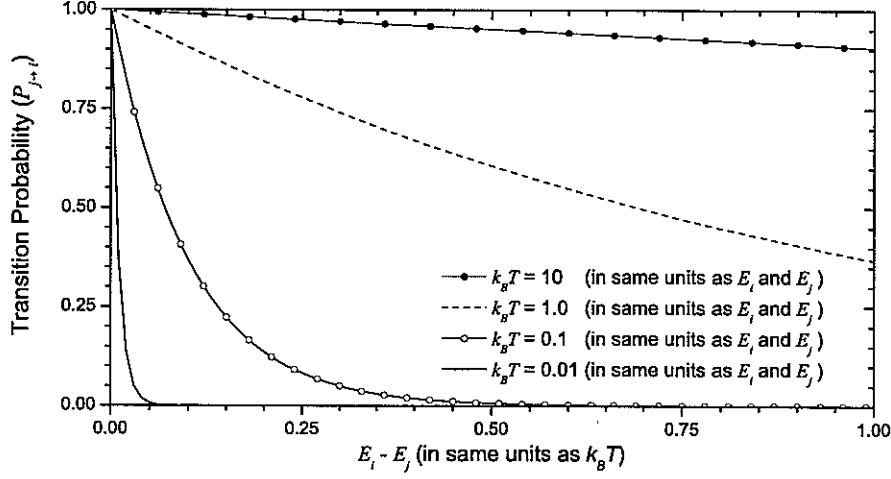
Figure 3.2   The probability of accepting a proposed transition (from state j to state i) is shown as a function of the energy difference, $E_i - E_j > 0$, for four different temperatures. Transitions from higher energies to lower energies ($E_i \leq E_j$) are accepted automatically.

(Four examples of such curves have been included in Fig. 3.2.) Along this curve, there is furthermore only a single point that is relevant for making the decision [in step (3)] for any particular MC step; it is given by the intersection of the relevant curve with a vertical line, drawn at the current value of $E_i - E_j$. The random number, $0 \leq r \leq 1$, can then be pictured as a point lying somewhere along this line. If r lies *below* the point of intersection, the transition is accepted. If r lies above the curve, the transition is rejected.

Fig. 3.2 also provides important additional insight into the physics that is involved, and in particular, how it is that many of the states come to be ignored. Clearly, the higher the energy of the proposed state, the less likely it is that the transition will be accepted; and as a result, fewer Monte Carlo steps will end up at the high energy states. As the temperature is lowered, it becomes even less likely to accept transitions to high energy states. In fact, at low temperatures, transitions to the high energy states will *never* be accepted. Suppose, for example, that we have chosen to use a million MC steps ($N_{MC} = 10^6$). If the proposed state (i) were to have an energy that gives $E_i - E_j = 23 \times k_B T$, then the corresponding acceptance probability would be $P_{j \to i} \approx 10^{-10}$, and it would be extremely improbable for such a transition

to be accepted in the 10^6 MCS. States such as these are therefore ignored, and it is the proposed states for which $P_{j \to i}$ is close to unity (and $\mathcal{P}_i$ is large) that detailed importance sampling takes place.

3.5   Ensemble averaging and error-bars

From the discussion of the previous sections, it should be clear how—based on the principle of detailed balance—one can perform a Monte Carlo calculation by making transitions back and forth among the important states. During each MCS of this process, one can record information about the current state, and use this information to calculate averages using Eq. (3.2). There are, however, important issues that deal with the statistical analysis of these data that have not yet been addressed. This is the purpose of the present section.

3.5.1   Reaching equilibrium

First, note that step (1) of the 4-step algorithm begins with a *random* initial state. It is likely that this state (i) will have a very small value of $\mathcal{P}_i$, in which case it should not be counted among the "important" states. The second state (j) will be more likely to be an important state, since such transitions (with $\mathcal{P}_j > \mathcal{P}_i$) are accepted preferentially. It might, however, require several MCS before the important region of the phase space (which should actually be sampled) is reached. This is a noteworthy problem, but has a very easy solution: simply choose not to count a certain number N_{eq} of the initial MCS, since the states being encountered are probably not yet representative of the equilibrium distribution.

The value of N_{eq} that should be used depends on the specific model being studied, but its exact value is fairly arbitrary. It can be chosen by starting a simulation and observing how the quantities of interest (e.g., E_i and M_i) evolve. For the QMC calculations that are described in Chapters 5–8, we have always disregarded the first $N_{eq} = 50,000$ MCS. For these simulations, equilibrium was always reached within about 10,000 MCS, so it was important to choose $N_{eq} > 10,000$. Since 50,000 MCS required a relatively small amount of additional computation time, this number was chosen to be somewhat larger than necessary.

After this initial equilibriation is complete, the subsequent N_{MC} states that are encountered, and in particular the frequency with which they are sampled, will indeed accurately reflect the equilibrium distribution. As a result, Eq. (3.2) can be used to estimate the necessary thermodynamic averages, and the value thus obtained will approach the exact value in the limit $N_{MC} \to \infty$. Naturally, however, we want to get results using a *finite* number of MCS, and a correspondingly finite amount of computation time. To know whether or not these results (with a finite sample size) are "good enough", we need to know the size of the corresponding statistical errors.

3.5.2 Accounting for correlations

Recall from statistics[9] that for a finite number N_{si} of statistically independent observations of a quantity Q, the statistical error in the resulting estimate of Q is given by

$$\delta Q = \sqrt{\frac{\langle Q^2 \rangle - \langle Q \rangle^2}{N_{si}}}. \tag{3.5}$$

The quantities $\langle Q^2 \rangle$ and $\langle Q \rangle^2$ can each be estimated using Eq. (3.2); and their difference gives the "width" of the distribution being sampled, which is a well-defined, finite number.[6] Since the numerator is finite, we have the very important result: $\delta Q \propto 1/\sqrt{N_{si}}$, decreasing toward zero with increasing N_{si}. The errors in the averages that we calculate can therefore be made *arbitrarily small* by choosing N_{si} to be large,[7] and the calculations are subsequently referred to as "numerically exact".

There is however a remaining complication. The values of Q_i that we obtain will *not* be independent from one another. This is because, as we described in the previous section, each new state is proposed by making some change to the current state.[8] In fact, in order to stay in the important region of phase space—and hence have a high acceptance probability—it is actually advantageous to propose relatively *small* changes to the current state, which

---

[6]From Eqs. (2.8c) and (2.9b), the numerator of Eq. (3.5) is actually proportional to $\sqrt{C}$ or $\sqrt{\chi}$ if Q represents the energy or magnetic moment, respectively.

[7]When the negative sign problem is present, the amount of useful statistics that are accumulated decrease *exponentially* with decreasing temperature, dominating the $1/\sqrt{N_{si}}$ advantage that is provided by increased computation time, and restricting us to relatively high temperatures. See Chapter 8 for details.

[8]A process such as this (where each new state is selected based on the current state) is called "Markovian", and the series of states thus obtained is called a "Markov chain".

would cause successive states to be strongly *dependent* on one another. (Methods that avoid this apparent "Catch-22" are mentioned in the following section.) As it stands, Eq. (3.5) is therefore insufficient.

Fortunately, it is relatively easy to account for this complication as well. During the course of a simulation, we can estimate the "auto-correlation function" of each quantity Q, given by

$$\phi(t) = \frac{\langle Q_i Q_{i+t} \rangle - \langle Q \rangle^2}{\langle Q^2 \rangle - \langle Q \rangle^2}, \tag{3.6}$$

which is a measure of how strongly correlated the sampled values are when they are separated by t MC steps. To calculate the quantity $\langle Q_i Q_{i+t} \rangle$, we simply multiply Q_i (from the ith MCS) by Q_{i+t} [from the $(i+t)$th MCS], and average over all $N_{MC} - t$ of these products; i.e.,

$$\langle Q_i Q_{i+t} \rangle \equiv \frac{1}{N_{MC} - t} \sum_{i=1}^{N_{MC}-t} Q_i Q_{i+t}. \tag{3.7}$$

If the values were perfectly correlated (i.e., $Q_i = Q_{i+t}$ for all i), then clearly one would find $\langle Q_i Q_{i+t} \rangle = \langle Q^2 \rangle$, such that $\phi(t) = 1$. In the other limit, if the steps were totally uncorrelated, we would have $\langle Q_i Q_{i+t} \rangle = \langle Q_i \rangle \langle Q_{i+t} \rangle = \langle Q \rangle^2$, such that $\phi(t) = 0$. By definition then, $\phi(0) = 1$, and as the states become less correlated, $\phi(t)$ will decrease to zero with increasing t. To determine when the states are no longer correlated, we use the "integrated auto-correlation time",

$$\tau_{AC} \equiv \sum_{t=1}^{\infty} \phi(t), \tag{3.8}$$

which will have a value of 1 if there are no correlations; otherwise, τ_{AC} represents the approximate number of MCS that are necessary in order for correlations to vanish. [Note that the summation in Eq. (3.8) can be safely truncated at any $t \gg \tau_{AC}$, since these values of $\phi(t)$ are small, random fluctuations that do not significantly affect the value of τ_{AC}.]

The simplest solution to the problem of correlations in Eq. (3.5)—and the solution that we have implemented in our QMC calculations—is to group the data into a large number N_{si} of "bins" that have negligible correlations with one another. To ensure negligible inter-bin correlations, we simply choose the number of MCS per bin N_{MC} to be much larger than τ_{AC}. (For our QMC calculations, we typically find $\tau_{AC} < 10$, and in all cases, $\tau_{AC} \ll 100$; so, we safely used $N_{MC} \geq 1,000$ for the calculations that are described in Chapters 5–8.) The data

within each bin are used to calculate N_{si} different estimates of the thermodynamic properties (U, C, M, and χ) using Eq. (3.2). It is then these N_{si} estimates that are used to calculate the error-bars for each of the thermodynamic properties using Eq. (3.5).

The main idea is that even though the individual MCS are correlated, each of these N_{si} estimates truly are statistically independent, as long as we have chosen $N_{MC} \gg \tau_{AC}$. Note also that this approach in no way affects the averages that are obtained, but does affect the estimates of the statistical errors. By choosing $N_{MC} \gg \tau_{AC}$ we are slightly overestimating our statistical errors, while choosing $N_{MC} < \tau_{AC}$ (or totally ignoring correlations) would results in considerable *underestimation* of the statistical errors. Finally, note that the statistical errors, when estimated in this manner, still decrease inversely proportional to the square-root of the *total* number of MCS: Clearly from Eq. (3.5) we still have $\delta Q \propto 1/\sqrt{N_{si}}$. In addition, as N_{MC} increases, the distribution of the (individual bin) estimates becomes narrower, with a width proportional to $1/\sqrt{N_{MC}}$, such that $\delta Q \propto 1/\sqrt{N_{MC}N_{si}}$.

3.6 Summary

In addition to the Metropolis algorithm, there are many other ways in which the principle of detailed balance can be used to perform importance sampling which rely on many (if not all) of the principles that have been expressed in this chapter. Firstly, the choice to set the denominator equal to unity in Eq. (3.3b) is not unique. One could instead choose the probabilities to be symmetric (i.e., $P_{i \to j} + P_{j \to i} = 1$),[10] or could choose the probabilities in some other, more general, way.[11] More substantial modifications involve choosing the proposed transitions in such a way that they are always accepted.[12, 13] Versions of these so-called "cluster methods" have also been implemented in QMC algorithms,[14, 15] and are predecessors of the "directed-loop"[16] method which we use, and have described in the next chapter. Recently, there has also been considerable interest in developing MC methods that do not simulate a canonical (fixed temperature) ensemble, but instead consider multiple types of ensembles simultaneously.[17, 18] One such method can be used to estimate the density of states directly,[19] and a version of this method has recently been applied to QMC calculations

as well.[20]

To conclude, we would like to now present an algorithm which uses all of the ideas from this chapter, but is sufficiently general to be immediately applied in the following chapter. In particular, we will now use the term "configuration" instead of "state" to refer to a term of the partition function that is encountered in a MCS. This is to avoid confusion with quantum states in the next chapter. Also, each MCS will involve multiple proposed changes. This is reflected in step (2) of the following algorithm.

(1) Begin with a randomly chosen configuration.

(2) Change configurations by taking a Monte Carlo step. (See Secs. 3.2–3.4.)

    (2a) Propose changes to the configuration, each with some probability R.

    (2b) Choose whether or not to accept each of the proposed changes with some probability A, such that each of the total probabilities $P = RA$ satisfy detailed balance.

(3) Repeat step (2) N_{eq} times. (See Sec. 3.5.1.)

(4) Collect data for a bin. (See Sec. 3.5.2.)

    (4a) Take a Monte Carlo step [described in step (2)].

    (4b) Record the quantities of interest. (See Sec. 2.1.)

    (4c) Repeat (4a) and (4b) N_{MC} times.

    (4d) For each property, average the N_{MC} values using Eq. (3.2).

(5) Repeat step (4) N_{si} times.

(6) For each of the thermodynamic properties, average the N_{si} values to obtain a final estimate.

(7) Estimate the uncertainty in each of the estimates from step (6) using Eq. (3.5). (See Sec. 3.5.2.)

Bibliography

[1] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, Science **220**, 671 (1983).

[2] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comp. Phys. **17**, 10 (1975).

[3] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2000).

[4] K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction*, vol. 80 of *Springer Series in Solid-State Sciences* (Springer, 2002), 4th ed.

[5] E. Ising, Z. Phys. **31**, 253 (1925).

[6] L. Onsager, Phys. Rev. **65**, 117 (1944).

[7] B. A. Cipra, Amer. Math. Monthly **94**, 937 (1987).

[8] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[9] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1 (John Wiley and Sons, 1968).

[10] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).

[11] H. Müller-Krumbhaar and K. Binder, J. Stat. Phys. **8**, 1 (1973).

[12] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[13] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).

[14] H. G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).

[15] B. B. Beard and U.-J. Wiese, Phys. Rev. Lett. **77**, 5130 (1996).

[16] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002).

[17] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[18] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).

[19] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).

[20] M. Troyer, S. Wessel, and F. Alet, Phys. Rev. Lett. **90**, 120201 (2003).

CHAPTER 4.   Description of the quantum Monte Carlo method

In this chapter we provide a detailed description of the quantum Monte Carlo (QMC) method that we have implemented, relying heavily upon the ideas that have been reviewed in Chapters 2 and 3. This particular QMC method—which we have used to produce the results that are presented in Chapters 5-8—was introduced in Ref. [1], and is known as the "Stochastic Series Expansion method with directed-loops". Note that, although highly effective, this is by no means the *only* QMC method that could have been used to carry out these calculations. In fact, earlier versions of this method exist that are based on many of the same ideas,[2–4] and alternative methods exist as well. For instance, one could use the Trotter-Suzuki method[6, 7] which is briefly described in Sec. 4.2.1. However, this has the disadvantage that one must either extrapolate approximate results,[6] or deal more extensively with floating point (decimal) numbers rather than integers.[7] One could also choose to use the stochastic series expansion (SSE) method *without* directed loops,[3] but this has disadvantages as well, which are mentioned at the beginning of Sec. 4.6.2. For a recent review of the chronological development of the various QMC methods, as well as their relative advantages and disadvantages, see Ref. [5].

4.1   Introduction

The key premise of our calculations is that we wish to utilize the principles of detailed balance and importance sampling, as they have been described in the previous chapter, to calculate thermodynamic properties for *quantum* spins systems. In order to proceed, we must first recall the central requirement that (if satisfied) allows the strategies of Chapter 3 to be useful: *Each individual term of the partition function must be easy to evaluate.* As long as this

condition is satisfied—i.e., as long as we can write

$$Z = \sum_i W_i, \tag{4.1}$$

and easily evaluate each term W_i—the principle of detailed balance can be written

$$\frac{P_{j \to i}}{P_{i \to j}} = \frac{W_i}{W_j}, \tag{4.2}$$

without any reference to the partition function. This is important because (as was the case in Chapter 3) when the partition function is not needed, calculations can be performed *regardless* of the complexity of the sum that appears in Eq. (4.1).

For classical systems, this requirement is trivially satisfied. Each term is given by $\exp(-\beta E_i)$, so Eq. (4.2) reduces to Eq. (3.3b), and each transition probability depends on only two energies and the temperature. For quantum systems on the other hand, recall from Sec. 2.3 that the partition function is given by

$$Z = \mathrm{Tr}\left(e^{-\beta \mathcal{H}}\right), \tag{4.3}$$

which is a sum over all of the diagonal elements of the density operator. As we described in Sec. 2.4, these matrix elements are often extremely difficult (and sometimes impossible) to evaluate. Therefore, the first step in developing a QMC method is to recast the partition function, Eq. (4.3), in the form of Eq. (4.1) in such a way that each term W_i is easily evaluated. The details of this chapter become rather involved, so the following two paragraphs provide a brief "road map" to the material that is covered.

The following section (and in particular Sec. 4.2.2) includes a description of exactly *how* Eq. (4.3) can be written in the form of Eq. (4.1). This leads (after some algebra) to Eq. (4.15) which—for the sake of QMC updating and sampling—turns out to be problematic for two reasons: (1) As we show in Sec. 4.5, when antiferromagnetic interactions are present, Eq. (4.15) *always* produces some terms W_i that have negative values, regardless of the geometry. This would be problematic for sampling, but is easily rectified by adjusting Eq. (4.9a) to instead take the form of Eq. (4.23). (2) The sizes of the $(d+1)$-dimensional configurations (described in Sec. 4.4) vary among the terms in Eq. (4.15), and would therefore be inconvenient to represent

and update in computer memory. For this reason, Eq. (4.15) is finally rewritten in the more useful form of Eq. (4.21) in Sec. 4.4.

In Sec. 4.3, the explicit form of the thermodynamic averages are derived, which [recalling Eqs. (2.3) and (3.2)] can be used to sample the partition function, as it is written in Eq. (4.21). This sampling is carried out using Monte Carlo updates that are described in detail in Sec. 4.6. Each update consists of two stages, the first of which is performed with probabilities that are given by Eq. (4.29), which follow directly from the detailed balance condition [Eq. (4.2)]. The second stage of the update is considerably more involved. The detailed balance condition for this stage is (eventually) given by Eq. (4.37), where the specific probabilities that we have employed are provided in Eqs. (4.45), (4.46), (4.48), and (4.49). Finally, in Sec. 4.7 we address a number of specific details of our implementation.

4.2   Representing the partition function

4.2.1   Trotter-Suzuki method

There are two main avenues by which people proceed in recasting the partition function. One method (which we have *not* implemented, so will describe only very briefly) is to use the Trotter-Suzuki decomposition.[6] This method begins by introducing an integer m, and trivially writing

$$Z = \text{Tr}\left(\left[\exp\left(-\frac{\beta}{m}\mathcal{H}\right)\right]^m\right). \tag{4.4}$$

Recalling that the Hamiltonian depends on pairs of interacting spins, the partition function can be written[1]

$$Z = \text{Tr}\left(\left[\exp\left(\frac{\beta}{m}\sum_{b=1}^{N_b}\mathcal{H}_b\right)\right]^m\right) \tag{4.5a}$$

$$\approx \text{Tr}\left(\left[\prod_{b=1}^{N_b}\exp\left(\frac{\beta}{m}\mathcal{H}_b\right)\right]^m\right), \tag{4.5b}$$

where the approximate equality is very poor for small m, but becomes exact in the limit $m \to \infty$. (This is because the commutators, $[\mathcal{H}_{b_1}, \mathcal{H}_{b_2}] \propto \beta^2/m^2$, vanish for large m.) Rewriting

---

[1]Additional details regarding the "bond operators" $\mathcal{H}_b$ are provided in the following subsection.

the partition function in this way is useful because Eq. (4.5b) is a product of 2-spin operators, whose matrix elements are easily evaluated. This approach has been used both by choosing many finite m and extrapolating to the $m \to \infty$ limit, and by starting with $m = \infty$ and writing Eq. (4.5b) as a path-integral.[7]

4.2.2 Handscomb's method

The method that we have implemented begins in a totally different manner from the Trotter-Suzuki method, but ends up with a very similar result.[8] Specifically, the desired result is to write the partition function in terms of 2-spin operators, in such a way that their matrix elements are easily evaluated. This can be accomplished by first writing the series expansion of the density operator,

$$Z = \sum_{\psi} \langle \psi | e^{-\beta \mathcal{H}} | \psi \rangle \tag{4.6a}$$

$$= \sum_{\psi} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \langle \psi | \left(-\mathcal{H} \right)^n | \psi \rangle, \tag{4.6b}$$

where the trace has been written explicitly as a summation over all of the basis states $|\psi\rangle$, which we choose to be the z-states. (See Secs. 2.2 and 2.3.)

Next, the Hamiltonian is expressed as a summation over bond operators,[2]

$$-\mathcal{H} = \sum_{b=1}^{N_b} \mathcal{H}_b, \tag{4.7}$$

where N_b represents the total number of distinct bonds in the system, and the index b is used to distinguish between these bonds. With the Hamiltonian expressed in this way, inspection of Eqs. (2.11) and (2.16) reveals that each bond operator is given by

$$-\mathcal{H}_b = J_b \left(s_{1(b)}^z + s_{2(b)}^z \right) + \frac{J_b}{2} \left(s_{1(b)}^+ s_{2(b)}^- + s_{1(b)}^- s_{2(b)}^+ \right) + g\mu_B H \left(\frac{s_{1(b)}^z}{\xi_{1(b)}} + \frac{s_{2(b)}^z}{\xi_{2(b)}} \right), \tag{4.8}$$

where, since each bond connects two spins, the subscripts $1(b)$ and $2(b)$ refer to the two spins that are connected by bond b. Also, note that in Eq. (2.11) the field-dependent (Zeeman) portion of the Hamiltonian involves a sum over all *spins*, while in Eq. (4.7) the Hamiltonian

---

[2]The operators $\mathcal{H}_b$ are defined so as to give a negative sign which conveniently disappears when they are substituted into Eq. (4.6b).
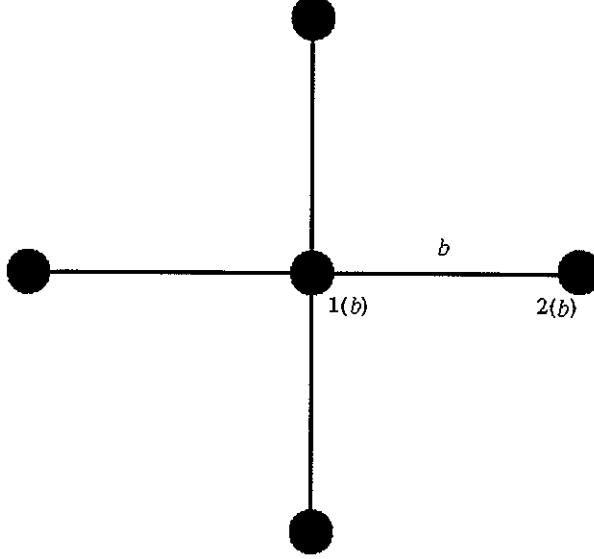
Figure 4.1   Diagram showing a system of $N = 5$ spins, connected by $N_b = 4$ bonds. Bond b connects the central spin [which has $\xi_{1(b)} = 4$ neighbors] and the right-most spin [which has $\xi_{2(b)} = 1$ neighbor].

is written as a sum over all *bonds*. In order to not over-count these terms, we must therefore divide the Zeeman terms by the number of bonds, $\xi_{1(b)}$ and $\xi_{2(b)}$, that connect to spins $1(b)$ and $2(b)$, respectively. To illustrate this point, consider the example system shown in Fig. 4.1. Four bonds connect to the central spin, so the corresponding Zeeman term would be counted $\xi_{1(b)} = 4$ times in Eq. (4.7), while only $\xi_{2(b)} = 1$ bond connects to the right-most spin. In Eq. (4.8), one therefore divides these two terms by 4 and 1, respectively.

For the purpose of the QMC algorithm, it will also be important to differentiate between diagonal and off-diagonal operators. This can be done by writing $\mathcal{H}_b = \mathcal{H}_{b,1} + \mathcal{H}_{b,2}$, where $\mathcal{H}_{b,1}$ and $\mathcal{H}_{b,2}$ represent the diagonal and off-diagonal portions of $\mathcal{H}_b$, respectively.[3] They can be read directly from Eq. (4.8), giving

$$\mathcal{H}_{b,1} \;=\; -J_b \left(s^z_{1(b)} + s^z_{2(b)} \right) - g\mu_B H \left(\frac{s^z_{1(b)}}{\xi_{1(b)}} + \frac{s^z_{2(b)}}{\xi_{2(b)}} \right) \tag{4.9a}$$

$$\mathcal{H}_{b,2} \;=\; -\frac{J_b}{2} \left(s^+_{1(b)} s^-_{2(b)} + s^-_{1(b)} s^+_{2(b)} \right). \tag{4.9b}$$

---

[3]The terms "diagonal" and "off-diagonal" are referring to the matrix elements of $\mathcal{H}_b$, when specifically represented using the z-states (introduced in Sec. 2.2) as the basis.

The next (*crucial*) step is to recognize that the partition function,

$$Z = \sum_{\psi} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \langle \psi | \left(\sum_{b=1}^{N_b} \sum_{o=1}^{2} \mathcal{H}_{b,o} \right)^n | \psi \rangle, \tag{4.10}$$

which involves summations to the nth power (i.e. products of sums), can instead be written in terms of sums of products, where each product consists of a sequence of n operators. To make this clearer, consider the simplest possible example: one bond connecting two spins. For this example, the sums of products for the four lowest orders of the series expansion are

$$\left(\sum_{b=1}^{N_b} \sum_{o=1}^{2} \mathcal{H}_{b,o} \right)^n = \begin{cases} 1 & \text{for } n = 0, \\ \mathcal{H}_{1,1} + \mathcal{H}_{1,2} & \text{for } n = 1, \\ \mathcal{H}_{1,1}^2 + \mathcal{H}_{1,2}^2 + \mathcal{H}_{1,1}\mathcal{H}_{1,2} + \mathcal{H}_{1,2}\mathcal{H}_{1,1} & \text{for } n = 2, \\ \mathcal{H}_{1,1}^3 + \mathcal{H}_{1,1}^2\mathcal{H}_{1,2} + \mathcal{H}_{1,1}\mathcal{H}_{1,2}\mathcal{H}_{1,1} + \mathcal{H}_{1,1}\mathcal{H}_{1,2}^2 \\ \quad + \mathcal{H}_{1,2}^3 + \mathcal{H}_{1,2}^2\mathcal{H}_{1,1} + \mathcal{H}_{1,2}\mathcal{H}_{1,1}\mathcal{H}_{1,2} + \mathcal{H}_{1,2}\mathcal{H}_{1,1}^2 & \text{for } n = 3. \end{cases} \tag{4.11}$$

For a given n, each sequence of n operators, $\mathcal{H}_{b_1,o_1}\mathcal{H}_{b_2,o_2} \cdots \mathcal{H}_{b_n,o_n}$, will be represented S_n, such that

$$\left(\sum_{b=1}^{N_b} \sum_{o=1}^{2} \mathcal{H}_{b,o} \right)^n = \sum_{S_n} \mathcal{H}_{b_1,o_1}\mathcal{H}_{b_2,o_2} \cdots \mathcal{H}_{b_n,o_n}. \tag{4.12}$$

Substituting Eq. (4.12) into Eq. (4.10), one arrives at

$$Z = \sum_{\psi} \sum_{n=0}^{\infty} \sum_{S_n} \frac{\beta^n}{n!} \langle \psi | \mathcal{H}_{b_1,o_1}\mathcal{H}_{b_2,o_2} \cdots \mathcal{H}_{b_n,o_n} | \psi \rangle. \tag{4.13}$$

Eq. (4.13) is now just one step away from achieving the goal of writing Z in terms of 2-spin matrix elements. This final step is to insert an identity operator $\underline{1}$, written as a summation over all z-states, in between each of the operators that appear in Eq. (4.13). For example, this insertion between the second and third operator in the sequence gives

$$\mathcal{H}_{b_2,o_2}\mathcal{H}_{b_3,o_3} = \mathcal{H}_{b_2,o_2}\underline{1}\mathcal{H}_{b_3,o_3} \tag{4.14a}$$

$$= \sum_{\psi_3} \mathcal{H}_{b_2,o_2}|\psi_3\rangle\langle\psi_3|\mathcal{H}_{b_3,o_3}. \tag{4.14b}$$

Repeating this for every operator in the sequence, Eq. (4.13) becomes

$$Z = \sum_{n=0}^{\infty} \sum_{S_n} \sum_{\psi_1} \sum_{\psi_2} \cdots \sum_{\psi_n} \frac{\beta^n}{n!} \langle\psi_1|\mathcal{H}_{b_1,o_1}|\psi_2\rangle\langle\psi_2|\mathcal{H}_{b_2,o_2}|\psi_3\rangle \cdots \langle\psi_n|\mathcal{H}_{b_n,o_n}|\psi_1\rangle \tag{4.15a}$$

$$= \sum_{i} W_i, \tag{4.15b}$$

where the original trace is now written as the sum over ψ_1 in Eq. (4.15a), and Eq. (4.15b) stresses how this process has finally led to a partition function that has the form of Eq. (4.1).

The partition function—as it appears in Eq. (4.15)—is now ideally suited for using the Monte Carlo method to perform importance sampling in the following two senses:[4] (1) Each individual term W_i is easy to evaluate, involving only products of two spin matrix elements.[5] Recall from Sec. 4.1 that this is *precisely* that requirement that—if satisfied—allows us to use Eq. (4.2) to calculate thermodynamic properties without calculating the (often unwieldy) partition function. (2) The summation in Eq. (4.15) is *extremely* complex. There are n sums $(\psi_1, \psi_2, \ldots, \psi_n)$ which each extend over all $(2s + 1)^N$ of the z-states, so it is clear that importance sampling will be necessary. This approach was first used by R. Handscomb in 1964 to perform calculations for the $s = 1/2$ Heisenberg ferromagnet.[4] However, an efficient, general method of performing Monte Carlo updates was lacking for decades to follow. This was remedied with the introduction of the stochastic series expansion (SSE) method[3]—and, more recently, directed loops[1]—as we describe in Sec. 4.6. First though, we address the issue of extracting thermodynamic properties from Eq. (4.15) in Sec. 4.3, and in Secs. 4.4 and 4.5 we discuss how to interpret the terms W_i that appear in Eq. (4.15).

4.3   Thermodynamic properties revisited

Recalling the material that was discussed in Chapter 2—in particular Secs. 2.1 and 2.2— Eq. (4.15) can be straightforwardly used to calculate thermodynamic properties of a form similar to that of Eq. (2.3b). The only difference is that the exponential terms, $\exp(-\beta E_i)$, must be replaced in Eq. (2.3b) with the more complex terms W_i from Eq. (4.15) [which will be modified slightly to give Eq. (4.21) in the next section]. To determine the quantities Q_i that must be sampled in order to calculate U, C, M, and χ, we now recall the derivatives of the partition function, Eqs. (2.6), (2.7), (2.8), and (2.9), and apply them to the *quantum* partition function.

---

[4]In a different sense, Eq. (4.15) is actually not quite ideal, and the terms W_i are therefore rewritten in the more useful form of Eq. (4.21) in Sec. 4.4.

[5]These 2-spin matrix elements are evaluated using Eqs. (2.13) and (2.17).

To calculate the magnetization, one can differentiate Eq. (4.3) using Eq. (2.7). This produces a factor of $g\mu_B \sum_{j=1}^{N} \underline{s}_j^z$ from the Heisenberg Hamiltonian[6] [Eq. (2.11)], which, from Eqs. (2.13) and (2.14), commutes with the Hamiltonian. (Note, any single operator $\underline{s}_j^z$ does not commute with $\mathcal{H}$, but since M_S is a good quantum number, the sum does commute.) By representing the quantum number M_S that corresponds to the term W_i with an additional index, $M_{S,i}$, the magnetization can be written

$$M \; = \; g\mu_B \frac{\sum_i M_{S,i} W_i}{\sum_i W_i} \tag{4.16a}$$

$$= \; g\mu_B \langle M_S \rangle, \tag{4.16b}$$

where—in analogy to Eq. (2.3b)—we now define

$$\langle Q \rangle \equiv \frac{\sum_i Q_i W_i}{\sum_i W_i} \tag{4.17}$$

for a general quantity Q. Recalling Chapter 3, M can then be calculated by sampling (i.e., recording and later averaging) the quantum number $M_{S,i}$ for each term W_i that is encountered during a MC simulation.[7] As we show below, χ, U, and C can straightforwardly be written in terms similar averages.

From Eq. (2.9), the magnetic susceptibility can be obtained by differentiating the partition function a second time. This yields

$$\chi \; = \; \beta g^2 \mu_B^2 \left[\left(\frac{\sum_i M_{S,i}^2 W_i}{\sum_i W_i} \right) - \left(\frac{\sum_i M_{S,i} W_i}{\sum_i W_i} \right)^2 \right] \tag{4.18a}$$

$$= \; \beta g^2 \mu_B^2 \left[\langle M_S^2 \rangle - \langle M_S \rangle^2 \right], \tag{4.18b}$$

which depends on the averages of both M_S and M_S^2. The internal energy is obtained by differentiating the partition function with respect to β as was shown in Eq. (2.6). Note however that the *only* temperature dependence in Eq. (4.15) is the factor of β^n that appears in each

---

[6]This assumes that all spins share a common value of g. This is often not the case for magnetic molecules, so we have derived more general forms of the magnetization and magnetic susceptibility, which are also applicable to systems with different values of g. This derivation has been included in Appendix A.

[7]This assumes (of course) that one has some strategy for taking a Monte Carlo step. This is described in great detail in Sec. 4.6.

term W_i. Clearly then, $\frac{\partial W_i}{\partial \beta} = n\beta^{-1}W_i$, so

$$U = -k_B T \frac{\sum_i n_i W_i}{\sum_i W_i} \tag{4.19a}$$

$$= -k_B T \langle n \rangle, \tag{4.19b}$$

where the subscript i on n_i stresses that this value corresponds to the term W_i. Differentiating again, one obtains the specific heat from Eq. (2.8),

$$C = k_B \left[\left(\frac{\sum_i n_i^2 W_i}{\sum_i W_i} \right) - \left(\frac{\sum_i n_i W_i}{\sum_i W_i} \right)^2 - \left(\frac{\sum_i n_i W_i}{\sum_i W_i} \right) \right] \tag{4.20a}$$

$$= k_B \left[\langle n^2 \rangle - \langle n \rangle^2 - \langle n \rangle \right]. \tag{4.20b}$$

From these equations, the calculation of the four thermodynamic quantities requires one to record two (and only two) quantities for each Monte Carlo step: $M_{S,i}$ and n_i. Then, from these numbers, we are able to calculate estimates of M, χ, U, and C in the manner discussed in Sec. 3.5. Furthermore, by recording these data in separate "bins", we are able to properly account for correlations in the data, and hence obtain accurate estimates for the uncertainties. This has also been explained in Sec. 3.5, and the process of dealing with this data is summarized in Sec. 3.6.

4.4 Understanding the terms in the partition function

Before proceeding to describe how the terms that comprise the partition function are sampled in a QMC simulation, it is necessary to make a slight additional adjustment to Eq. (4.15), and to further clarify the meaning of the individual terms W_i. Following the factor of $\beta^n/n!$, each term in Eq. (4.15) begins with a state, ψ_1, which is followed by an operator, $\mathcal{H}_{b_1,o_1}$. If that operator is diagonal (i.e., if $o_1 = 1$), then the state ψ_2 is identical to ψ_1. If, on the other hand, the operator is off-diagonal ($o_1 = 2$), then ψ_2 is different than ψ_1. Specifically, spins $b_1(1)$ and $b_1(2)$ will have different m_j quantum numbers in the state ψ_1 than they do in state ψ_2. This is due to the action of the raising and lowering operators in Eq. (4.9b), as described in Sec. 4.2.2.[8] After the second state, ψ_2, the second operator, $\mathcal{H}_{b_2,o_2}$, leads to the third state,

---

[8]Note that, given a particular state ψ_1, there are potentially *two* states ψ_2 that will yield non-zero matrix elements, $\langle \psi_1 | \mathcal{H}_{b_1,o_1} | \psi_2 \rangle$, corresponding to the two terms in Eq. (4.9b). Each of these states ψ_2 belong to different terms W_i; so, in a sense, the operator $\mathcal{H}_{b_1,o_1}$ is determined by the states, ψ_1 and ψ_2, and not vice-versa.

ψ_3, which again is either identical to ψ_2 (if $o_2 = 1$) or differs from ψ_2 in two quantum numbers (if $o_2 = 2$). This process continues for all n operators, and eventually leads back to the state ψ_1.

The number of bond operators n and the number of states n clearly vary from term to term. This turns out to be inconvenient, both in terms of storing these terms in computer memory, and for taking Monte Carlo steps. This is avoided in the SSE method by considering a *fixed* number of operators in the sequence, and a correspondingly *fixed* number of states, ψ_k. Fixing this number to be L, the series will then be truncated at $n = L$, which will have *no* effect on the results that are obtained, provided that L is chosen to be large enough that no important terms are neglected. (This is discussed in greater detail in Sec. 4.7.1.)

The process by which one converts Eq. (4.15) into an equation with a fixed number of operators proceeds as follows: First, a sufficiently large value of L is chosen, such that $L > n$ for every important term. (See Sec. 4.7.1.) Then, $L - n$ *identity* operators, $\underline{1}$, are "filled in", such that there are L *total* operators, even though there are only n *bond* operators. Consider, for example, a term that consists of three bond operators, $\mathcal{H}_{b_1,o_1} \mathcal{H}_{b_2,o_2} \mathcal{H}_{b_3,o_3}$. If we were to truncate the series at $L = 4$ total operators,[9] there would be four choices for where to insert the additional (identity) operator: Before $\mathcal{H}_{b_1,o_1}$, before $\mathcal{H}_{b_2,o_2}$, before $\mathcal{H}_{b_3,o_3}$, or after $\mathcal{H}_{b_3,o_3}$. (One such choice—before $\mathcal{H}_{b_3,o_3}$—is shown diagrammatically in Fig. 4.2.) There are then *four* terms, each with the same value of W_i, instead of the original *one* term from Eq. (4.15). To account for this, each of the terms must be divided by four in order to not change the value of the partition function Z. If the example had instead consisted of $n = 2$ bond operators and $L = 4$ total operators, there would have been six choices for where to insert the $L - n = 2$ identity operators. In general, when $L - n$ identity operators are inserted, there will be $\frac{L!}{n!(L-n)!}$ terms spawned from the initial term that had only n operators. Therefore, the partition function is preserved by dividing each term of Eq. (4.15) by $\frac{L!}{n!(L-n)!}$, giving

$$W_i = \frac{\beta^n (L - n)!}{L!} \langle \psi_1 | \mathcal{H}_{b_1,o_1} | \psi_2 \rangle \langle \psi_2 | \mathcal{H}_{b_2,o_2} | \psi_3 \rangle \cdots \langle \psi_L | \mathcal{H}_{b_L,o_L} | \psi_1 \rangle, \qquad (4.21)$$

[9]This example has been specifically chosen for its simplicity. Truncating a series at $L = 4$ could only be possible at *very* high temperatures. In fact, we always choose $L \geq 10$, and often $L > 1,000$ operators are necessary.
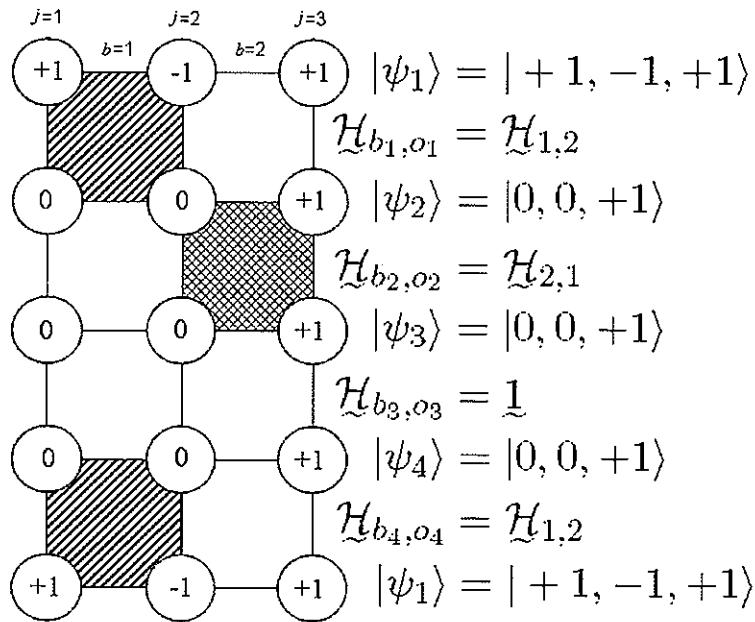
Figure 4.2   Diagram representing one (simple) term of the partition function for a chain of $N = 3$ spins, connected by $N_b = 2$ bonds. The three intrinsic spins s_j are arbitrary, but have been assumed to be integers for this example, such that each m_j quantum number is also an integer. This term has $n = 3$ bond operators and one identity operator, $\underline{1}$, for a total of $L = 4$ operators and $L = 4$ states (ψ_1, ψ_2, ψ_3, ψ_4). The quantum numbers m_j ($j = 1$, 2, 3) that correspond to each of these states are shown inside of the circles, where each circle appears below the corresponding value of j, and to the left of the corresponding state $|\psi_k\rangle$. The off-diagonal bond operators are represented with diagonal shading, the diagonal bond operator is represented with cross-hatched shading, and no shading is shown for the identity operator.

where the value of n varies from term to term, but the value of L is fixed.[10]

The diagram shown in Fig. 4.2 corresponds the particularly simple example of $N = 3$ spins, connected by $N_b = 2$ bonds, with $n = 3$ and $L = 4$. However, despite its simplicity, it illustrates a means of visualizing these terms which can *always* be used, even for much larger, more complex systems. Specifically, for a d-dimensional quantum system, each term W_i can always be represented in $(d+1)$-dimensional space, just as this one-dimensional system has been represented in two-dimensional space. The "extra" dimension (proceeding from the top to the bottom of Fig. 4.2) extends through the L operators and L states, always arriving back at the state ψ_1. The process of taking a Monte Carlo step then consists of nothing more than changing some or all of the L operators $\mathcal{H}_{b_k, o_k}$, and the L states ψ_k, in a diagram such as the one shown in Fig. 4.2. This process is described in detail in Sec. 4.6, continuing to make use of (and build upon) this same example that has been shown in Fig. 4.2.

4.5 Attempting to ensure that terms are positive

Before describing the details of how a Monte Carlo step is taken, there is one remaining issue that must be addressed. We wish to sample all of the important terms W_i, as they are written in Eq. (4.21), using the methods described in Chapter 3. In particular, we wish to use Eq. (4.2) to determine the *probabilities* with which transitions should occur, which depend on the values of W_i. By definition, a probability must be non-negative ($0 \le P_{i \to j} \le 1$), which was obviously the case for classical systems, since $\exp(-\beta E_i) > 0$ for all E_i. However, for the terms W_i that are given by Eq. (4.21), it is not immediately obvious whether these terms will be positive or negative; and in fact, terms sometimes are negative! The purpose of the present section is to address this issue.

Each term W_i from Eq. (4.21) involves a factor of $\beta^n \frac{L!}{n!(L-n)!}$ (which is clearly positive) multiplied by a product of L matrix elements. However, $L - n$ of these matrix elements correspond to identity operators, and are therefore equal to unity. The question is then whether the product of the remaining n matrix elements is positive or negative. To answer this question,

---

[10]Note that we henceforth adopt the notation used in Eq. (4.21). Namely, each of the L operators $\mathcal{H}_{b_k, o_k}$ in the sequence has a distinct index, $1 \le k \le L$, including the identity operators.

recall the definitions of the bond operators, given in Eq. (4.9).

The diagonal bond operators, defined in Eq. (4.9a), depend on the operators $s_{1(b)}^z$ and $s_{2(b)}^z$. Recalling Eq. (2.13), these operators yield the quantum numbers $m_{1(b)}$ and $m_{2(b)}$, corresponding to spins $1(b)$ and $2(b)$, respectively. These quantum numbers can be either positive or negative, so the resulting matrix elements—as they appear in Eq. (4.9a)—can also be either positive or negative. However, since these are *diagonal* operators, one can very easily adjust their definition to *ensure* that the matrix elements are positive. This is accomplished by simply adding a constant,

$$c = \sum_{b=1}^{N_b} c_b, \tag{4.22}$$

to the Hamiltonian, such that Eq. (4.9a) becomes

$$\mathcal{H}_{b,1} = c_b - J_b \left(s_{1(b)}^z + s_{2(b)}^z \right) - g\mu_B H \left(\frac{s_{1(b)}^z}{\xi_{1(b)}} + \frac{s_{2(b)}^z}{\xi_{2(b)}} \right). \tag{4.23}$$

Changing the Hamiltonian in this way will have no effect on the resulting magnetization, magnetic susceptibility, or specific heat, but it *will* change the internal energy. Specifically, the resulting internal energy will be larger than that of the original Hamiltonian by an amount c, which can be trivially subtracted at the end of a calculation.

It is perhaps worthwhile to clarify the meaning of the word "constant" when it is used in the context of Eq. (4.22), referring to values of c and c_b. Every value of c_b must be held fixed throughout a simulation, while the thermodynamic quantities are calculated for a particular set of parameters. (In other words, the Hamiltonian cannot be altered during the course of a calculation.) However, when repeating a calculation with different parameters, different values of c_b might be necessary. For example, as an external field H becomes large, large values of c_b become necessary. This is due to the Zeeman term in Eq. (4.9a), which *decreases* proportional to H for positive values of $m_{1(b)}$ and $m_{2(b)}$.

For the calculations that we have performed, each value of c_b is chosen at the beginning of a calculation, so that the smallest matrix element for each bond b has some small, positive value (e.g., 0.01). Calculations can then be performed for many different temperatures without changing the values of c_b, since the Hamiltonian—and in particular Eq. (4.9a)—does not depend

on T. This is what happens when we calculate $\chi(T)$. For calculating $M(H)$, different values of c_b must be determined for each H. Naturally though, this is not done by hand, yet a negligible amount of computation time is spent in determining valid values of c_b; the corresponding diagonal operators, Eq. (4.23); and the resulting matrix elements.

As we have described above, any diagonal operator $\mathcal{H}_{b,1}$ can be defined such that the corresponding matrix elements are positive. The only remaining factors to be considered are the off-diagonal matrix elements that appear in Eq. (4.21), whose operators are defined in Eq. (4.9b). The matrix elements for the raising and lowering operators are calculated using Eq. (2.17), and are always positive. Therefore, the sign of an off-diagonal matrix element is determined entirely by the sign of the exchange constant J_b. If a bond describes a ferromagnetic (FM) interaction (i.e., $J_b < 0$), then the corresponding off-diagonal matrix elements are positive; while antiferromagnetic (AFM) interactions ($J_b > 0$) produce *negative* matrix elements.

This issue of dealing with the "sign" of matrix elements can now be summarized very concisely: (1) Diagonal matrix elements can easily be made positive by shifting the Hamiltonian [and the constituent bond operators in Eq. (4.9a)] by an appropriate constant, yielding Eq. (4.23). (2) FM bonds have $J_b < 0$, so the corresponding matrix elements, calculated from Eqs. (4.9b) and (2.17), are positive. (3) AFM bonds have $J_b > 0$, so the resulting matrix elements [from Eqs. (4.9b) and (2.17)] are negative. These three results give two very simple rules regarding the signs of the terms W_i:

(1) If all of the bonds are FM (i.e., if $J_b < 0$ for all b), then all of the terms W_i from Eq. (4.21) are positive.

(2) If AFM bonds are present, then a term W_i from Eq. (4.21) is negative (and hence problematic) if and only if in contains an *odd* number of off-diagonal matrix elements that correspond to AFM bonds.

It turns out that, for many geometries, all terms W_i are positive, even for AFM interactions. This is because, for these geometries, it is impossible to begin with a state ψ_1, and return to the same state ψ_1 via the operation of an odd number of off-diagonal operators. These geometrical

issues are closely related to both classical frustration, and the "negative sign problem". This is discussed in detail in Chapter 8, and in particular Sec. 8.2.

4.6 Monte Carlo updating

We are now finally prepared to describe how Monte Carlo updating proceeds for the QMC algorithm that has been implemented. Again, the goal is to sample the terms W_i that comprise the partition function, as they have been written in Eq. (4.21), using the principles of detailed balance [Eq. (4.2)] and importance sampling. As we described in Sec. 4.4, this task is equivalent to altering *diagrams*, such as the one that was shown in Fig. 4.2, in such a way that all of the important configurations are reached. There is, however, an important challenge that exists in executing such a QMC update; and this challenge is *not* present in classical Monte Carlo simulations. Namely, most changes—if proposed in an arbitrary manner—would lead to configurations for which W_i is identically zero, and should therefore not be sampled. For example, choose one of the 12 distinct m_j quantum numbers shown in Fig. 4.2, and try raising or lowering its value by one unit. The resulting term would automatically have $W_i = 0$. This is because the total angular momentum would no longer be conserved from one of the z-states to the next, and all bond operators $\mathcal{H}_{b_k, o_k}$ conserve M_S.

Instead of proposing arbitrary changes to individual quantum numbers, the method that has been implemented proceeds by making global changes to many quantum numbers, or equivalently, many states ψ_k. This update occurs in two separate stages. The first stage, called the "diagonal update" involves changing the order of the series expansion by replacing identity operators with diagonal bond operators, and vice-versa. This is described in the following subsection. The second stage of the update is considerably more involved. It involves constructing "directed-loops" through the $(d+1)$-dimensional space (such as that shown in Fig. 4.2) and altering the m_j quantum numbers of all spin states along the loop. In addition to altering spin states, this process also has the effect of changing diagonal operators to off-diagonal operators, and vice-versa, such that all operators eventually become amenable to the diagonal update. These directed-loop updates are described in detail in Secs. 4.6.2 and 4.6.3, and the specific

probabilities that we have used in our implementation are presented in Sec. 4.7.3.

4.6.1 Stochastic series expansion—diagonal update

There are a few key features of the stochastic series expansion (SSE) method, most of which have already been described. First, one expands the density operator, just as Handscomb did in 1964,[4] to obtain Eq. (4.15). Next, additional identity operators are added, and the series is truncated at L total operators,[3] yielding Eq. (4.21). By doing this, one gains the advantage that the size of the $(d+1)$-dimensional space (such as that shown in Fig. 4.2) does not change from term to term. Finally, by adding a constant to the Hamiltonian, one can often ensure that the terms W_i are positive, as described in Sec. 4.5. With these adjustments in place, it is now relatively straightforward to perform the diagonal update of the SSE method.

This update proceeds by separately adding and removing diagonal operators $\mathcal{H}_{b_k,1}$ ($1 \leq k \leq L$) with probabilities P_a and P_r, respectively. Recalling the discussion of Chapter 3, the new configurations that result from these changes will be encountered with frequencies that are proportional to the values of the terms [from Eq. (4.21)] that they represent—and can hence be used for importance sampling—provided that detailed balance is satisfied. In this context, the detailed balance condition can be restated,

$$\frac{P_a}{P_r} = \frac{W_a}{W_r}, \tag{4.24}$$

where W_a and W_r are the values of terms [from Eq. (4.21)] that differ *only* by the addition or removal of a single matrix element, so P_a and P_r are *exactly* analogous to $P_{j \to i}$ and $P_{i \to j}$ from Eq. (4.2). Specifically, W_r represents the value of a term that is *lacking* a bond operator $\mathcal{H}_{b_k,1}$ (and the corresponding matrix element $w_k \equiv \langle \psi_k | \mathcal{H}_{b_k,1} | \psi_{k+1} \rangle$) relative to the term whose value is W_a. By representing the entire product of n matrix elements[11] in the term W_r more compactly as π_n, these terms can be written—from Eq. (4.21)—as

$$W_r = \frac{\beta^n (L-n)!}{L!} \pi_n \tag{4.25a}$$

$$W_a = \frac{\beta^{n+1} [L-(n+1)]!}{L!} \pi_n w_k, \tag{4.25b}$$

---

[11]There are actually L matrix elements, but $L - n$ of them are identically equal to unity, so can be omitted from the product.

where n is the number of bond operators in the term W_r, and $n+1$ is the number of bond operators in the term W_a. Substituting Eq. (4.25) into Eq. (4.24), the detailed balance condition becomes

$$\frac{P_a}{P_r} = \frac{\beta w_k}{L-n},$$ (4.26)

which depends on only the *single* matrix element w_k whose addition or removal is being considered.

In order to now understand how the probabilities in Eq. (4.26) can actually be *used*, recall again Fig. 4.2 from Sec. 4.4. This diagram represents a term that has four operators. Two of them ($\mathcal{H}_{b_1,o_1}$ and $\mathcal{H}_{b_4,o_4}$) are off-diagonal; one ($\mathcal{H}_{b_2,o_2}$) is a diagonal bond operator; and one ($\mathcal{H}_{b_3,o_3}$) is an identity operator. Given this initial configuration, the diagonal update would proceed by proposing to remove[12] the bond operator $\mathcal{H}_{b_2,o_2}$ with a probability P_r that depends on the value of w_2. We would also propose to add a bond operator $\mathcal{H}_{b_3,o_3}$ (replacing the identity operator) with a probability P_a that depends on the value of the proposed matrix element w_3. The two off-diagonal operators are left unchanged in this stage of the update, because removing either of them (and replacing it with an identity operator) would produce a term that is identically zero.

More generally, for any given configuration we simply step through all of the operators $\mathcal{H}_{b_k,o_k}$ ($1 \le k \le L$), changing the diagonal bond operators into identity operators with probabilities P_r; changing the identity operators into diagonal bond operators with probabilities P_a; and leaving the off-diagonal operators unchanged. The only remaining task for this subsection is then to extract individual probabilities from the ratio that was given in Eq. (4.26). To do so, we must be careful to follow the rules that were discussed in Sec. 3.3 regarding proposing and accepting changes. Specifically, recall that *making* a change involves first *proposing* a change, and then *accepting* a change. Using the notation of Sec. 3.3, we have

$$P_r = R_r A_r$$ (4.27a)

$$P_a = R_a A_a.$$ (4.27b)

---

[12]Note that when we say to "remove" this operator, what is actually meant is to "replace the bond operator $\mathcal{H}_{b_2,o_2} = \mathcal{H}_{2,1}$ with an identity operator, such that $\mathcal{H}_{b_2,o_2} = \underset{\sim}{1}$".

The simplest way to propose the removal of an operator is to always let $R_r = 1$, meaning that we definitely propose the removal of any diagonal bond operator. However, when proposing the *addition* of a diagonal bond operator, there are N_b different bonds from which to choose. Therefore, if we definitely propose the addition of an operator $\mathcal{H}_{b_k,o_k}$, with a randomly chosen bond b_k ($1 \leq b_k \leq N_b$), this gives $R_a = 1/N_b$. Inserting R_r and R_a into Eqs. (4.27a) and (4.27b), respectively; and further substituting these into Eq. (4.26) yields

$$\frac{A_a}{A_r} = N_b \frac{P_a}{P_r} \tag{4.28a}$$

$$= \frac{N_b \beta w_k}{L - n}. \tag{4.28b}$$

Choosing to use Metropolis-like acceptance rates (which maximize A_a and A_r), one arrives at the final acceptance probabilities,

$$A_r = \begin{cases} \frac{L-n}{N_b \beta w_k} & \text{if } \frac{L-n}{N_b \beta w_k} < 1, \\ 1 & \text{if } \frac{L-n}{N_b \beta w_k} \geq 1, \end{cases} \tag{4.29a}$$

$$A_a = \begin{cases} \frac{N_b \beta w_k}{L-n} & \text{if } \frac{N_b \beta w_k}{L-n} < 1, \\ 1 & \text{if } \frac{N_b \beta w_k}{L-n} \geq 1. \end{cases} \tag{4.29b}$$

(Note again that the value of n used here does *not* include the operator that is being potentially added or removed.) By generating random numbers (as we have described in Sec. 3.4), it is these probabilities [Eqs. (4.29a) and (4.29b)] that are used to determine whether or not to accept a proposed removal or addition of a bond operator.

4.6.2 Constructing loops

The diagonal updates that were described in the previous subsection allow one to sample various terms of the partition function by adding and removing diagonal bond operators, and hence varying the order n of the expansion. However, that update alone is clearly not sufficient to access all of the possible terms W_i that appear in Eq. (4.21)—and all of the corresponding configurations that could occur in a diagram such the one shown in Fig. 4.2—and must therefore be supplemented with a separate update. In particular, it is necessary to change the quantum numbers m_j, and equivalently, to change between diagonal and off-diagonal bond operators. In

the original formulation of the SSE method,[3] this was accomplished through a series of local updates, each of which changed a small number of m_j quantum numbers. This updating was later replaced by the "operator-loop" update,[2] which produces large changes in the $(d+1)$-dimensional configuration with a single "loop". This approach has two main advantages over the initial SSE method. (1) The resulting integrated auto-correlation times[13] are much smaller, especially at low temperatures, and (2) the algorithm is less complicated to implement. These operator-loops were initially introduced in the context of $s = 1/2$ models, where spin states are changed by simply "flipping" individual spins (from "up" to "down", or vice-versa), but were later generalized to larger values of s, with the introduction of the so-called "directed loop equations".[1] The purpose of the present subsection is to describe the process of constructing such a loop in a manner that is consistent with detailed balance. This process leads naturally to the directed loop equations which are described in detail in the following subsection, and in Sec. 4.7.3 we present the precise probabilities that we have used in our implementation.

The general idea of this update is that a loop is constructed through the $(d+1)$-dimensional configuration, and all of the quantum numbers along this loop are changed by one unit. This idea is very simple, but in order to describe the details of how it is accomplished—and in particular, how detailed balance is fulfilled—it will be necessary to introduce some additional notation and terminology. First, the initial configuration, such as the one shown in Fig. 4.2, will be denoted C_i, and the corresponding term that it represents from Eq. (4.21) is represented W_i; while the final configuration, that is reached by constructing a loop, will be denoted C_f, and its term [from Eq. (4.21)] is W_f. The detailed balance condition it then given by

$$\frac{P_{i \to f}}{P_{f \to i}} = \frac{W_f}{W_i},\qquad (4.30)$$

where $P_{i \to f}$ and $P_{f \to i}$ are the probability of changing from configuration C_i to C_f and from C_f to C_i, respectively.

The challenge in fulfilling Eq. (4.30) is that there are typically *many* choices that must be made in order to change $C_i \to C_f$, where each choice must be made with an appropriate

---

[13]See Sec. 3.5.2 for a description of correlations between successive Monte Carlo steps.

probability, such that the *total* probabilities satisfy Eq. (4.30). Specifically, the construction of a single loop proceeds according to the following algorithm:[14]

(1) Choose where and how the loop will begin.

(2) Choose where the loop goes next.

(3) Repeat step (2) until the loop returns back to where it began in step (1).

An example of one particular path that a loop could follow is shown in Fig. 4.3(a) by a thick line, where the initial configuration C_i was shown in Fig. 4.2, and the resulting (post-loop) configuration C_f is shown in Fig. 4.3(b). With the aid of this example, we are now prepared to introduce some additional terminology that will be useful in clarifying the details of the 3-step algorithm that is given above, in particular, how—for each step—the probabilities can be chosen in such a way that Eq. (4.30) is satisfied.

First, note that the *path* shown in Fig. 4.3(a) could be describing any one of several *loops*. This is because we have not distinguished where the loop began on the diagram, or the direction in which it was constructed (clockwise or counterclockwise). We stress this point because, for *any* change $C_i \rightarrow C_f$, there are multiple loops that would each (separately) accomplish the same change. If we distinguish each of these loops with an index l, and represent the probability of constructing loop l by P_l, then the total probability of making this change ($C_i \rightarrow C_f$) is given by

$$P_{i \rightarrow f} = \sum_l P_l, \tag{4.31}$$

where the summation extends over all loops l that accomplish the given change.

A natural question that one might ask is, "How many loops are there that could accomplish this change?" There are 6 quantum numbers that lie along this path (and are therefore changed by the loop), and each loop could be constructed either clockwise or counterclockwise. A reasonable answer might therefore seem to be that there are 12 such loops. However, some of these quantum numbers must be the same as that of the next or previous spin state, because there is no bond operator in between the states. For example, in Fig. 4.2, consider the three

---

[14]The details of this algorithm are described below in greater detail.
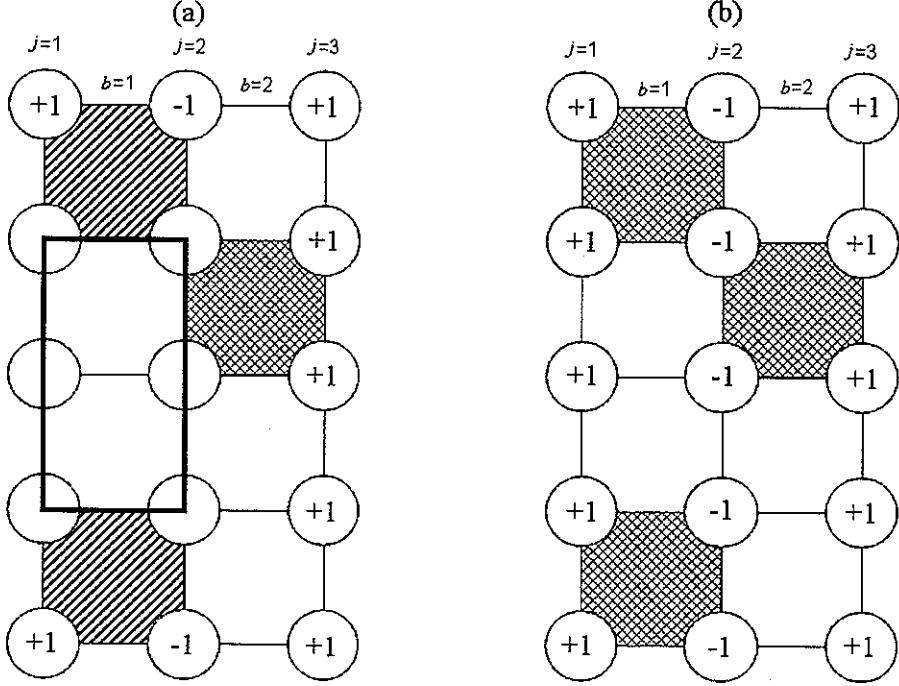
Figure 4.3   Diagrams representing (a) the path of a loop (thick line), which
has been constructed in the initial configuration $\mathcal{C}_i$ that was
represented in Fig. 4.2, and (b) the final configuration $\mathcal{C}_f$ that
results. As was the case in Fig. 4.2, off-diagonal bond opera-
tors are represented with diagonal shading, and diagonal bond
operators are represented with cross-hatched shading. Again,
the system represented here involves $N = 3$ spins, or arbitrary
(integer) spins s_j.

quantum numbers that have $m_j = 0$ in the $j = 1$ (left-most) column. There are no bond
operators between these three states (ψ_2, ψ_3, and ψ_4) that affect the $j = 1$ spin. Furthermore,
no bond operators are added or removed during the loop update. Therefore we have the
constraint that these three quantum numbers must be identical to one another; i.e., if one of
them is changed, all three must be changed. There is a similar constraint (affecting 2 spins)
in the lower right corner of the path, so there are actually 6 loops that could accomplish the
change shown in Fig. 4.3.

In order to take such constraints into account—and not preferentially propose certain loops
more often than other loops—these loops are generally described in terms of "legs".[1, 9] Each
leg represents a possible location (and direction) for a loop to begin, so there are 6 legs along
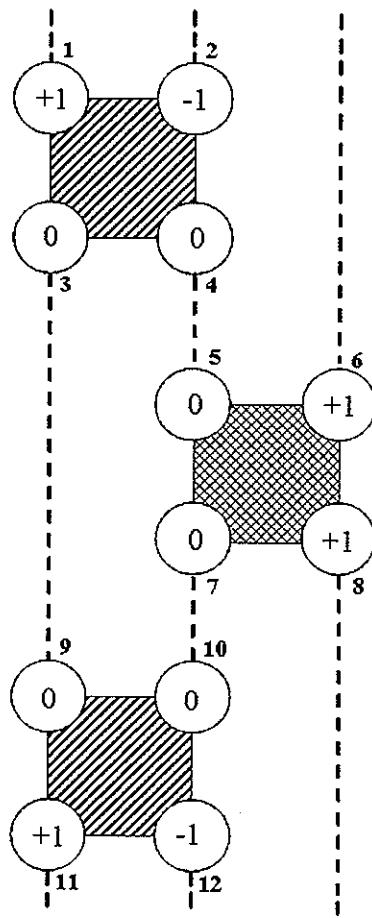
Figure 4.4   Diagram showing the 12 "legs" for the configuration shown in
Fig. 4.2. Any quantum numbers that are not adjacent to an
operator are irrelevant to the construction of a loop, and have
thus been omitted from this representation.

the path that was shown in Fig. 4.3. In the entire diagram, there are a total of 12 legs for

this example configuration, which are shown in Fig. 4.4, numbered 1–12. There are two other

features of these legs which should also be noted. First, every bond operator is attached to 4

legs, so the total number of legs is always $4n$. (For our example $n = 3$, so $4n = 12$.) Secondly,

every leg is connected (via a dashed line in Fig. 4.4) to exactly one other leg. In other words,

the legs come in *pairs* that must share the same quantum numbers. For our example, the

pairs are: (1,11), (2,12), (3,9), (4,5), (6,8), and (7,10). Finally, note that three of these leg

pairs "wind around" the top/bottom of the configuration. In fact, it is always the case that

N pairs of legs wind around in this way, because ψ_1 appears both at the top and bottom of a

configuration.

Having carefully introduced the meaning of "loops" and "legs", we are now finally in a position to more precisely state the three step algorithm that was given above. In step (1), one of the legs is chosen at random, and is proposed as the first leg of the loop. The number of legs remains fixed throughout the loop update, so the probability of choosing any leg is $R_l = \frac{1}{4n}$. Next, we must choose whether to raise or lower the m_j quantum number of the leg that has been proposed. If this first quantum number has a value that is initially $m_j \neq |s_j|$, then it could be *either* raised *or* lowered. In this situation, the proposed leg is accepted with probability 1, and a loop begins by either raising or lowering the value, each with an equal probability $P_m = \frac{1}{2}$. If, on the other hand, $m_j = s_j$, then the value could *only* be lowered; and if $m_j = -s_j$, the value could *only* be raised. We again set $P_m = \frac{1}{2}$, so there is a 50% chance that the proposed leg will not be accepted. (Recalling Sec. 3.3, this is perfectly analogous to the situation of Carl moving along a peninsula.) After changing the quantum number of the first leg, step (1) is complete, and we proceed to step (2).

For the sake of concreteness, consider one particular loop which could have been responsible for the update $(\mathcal{C}_i \rightarrow \mathcal{C}_f)$ that took place in Fig. 4.3: Suppose the loop began[15] at leg 3 (as labeled in Fig. 4.4) by raising the quantum number. Next, in step (2), we chose to proceed to leg 4 (and lower its quantum number) with some probability[16] p_1. We then *automatically* followed the dashed line to leg 5 whose quantum number was also changed. (It is in fact the *same* quantum number as the one that was changed at leg 4.) Step (2) was then repeated, and leg 7 was chosen (and its quantum number lowered) with a probability[16] p_2. Leg 7 led to leg 10 (lowering both quantum numbers), and we returned to step (2). Finally, leg 9 was chosen with a probability[16] p_3, which connects back to leg 3. We raised the necessary quantum numbers between legs 9 and 3, and—since we returned to the initial leg—the loop closed.

The construction of the above loop involved 5 individual choices, which (chronologically) had probabilities $R_l = \frac{1}{4n}$, $P_m = \frac{1}{2}$, p_1, p_2, and p_3. The total probability of constructing the

---

[15]Note that choosing to begin at either leg 3 or leg 9 would involve changing the *same* quantum numbers along the dashed line that connects this pair of legs. The difference is that choosing to begin at leg 3 results in a *clockwise* loop around the path, while choosing to begin at leg 9 would have given a *counterclockwise* loop.

[16]How these probabilities are chosen is the subject of the next subsection.

loop is therefore equal the the product of these five factors, $R_l P_m p_1 p_2 p_3$. For a general loop l, if we define N_l to be the number of times step 2 is executed, then the probability of constructing any loop is

$$P_l = \frac{1}{8n} \prod_{\mu=1}^{N_l} p_\mu. \tag{4.32}$$

Substituting this product into Eq. (4.31), we have

$$P_{i \to f} = \frac{1}{8n} \sum_l \prod_{\mu=1}^{N_l} p_\mu; \tag{4.33}$$

and if we represent the loops that "undo" these changes (i.e., that change $C_f \to C_i$) with an index l', then one also has

$$P_{f \to i} = \frac{1}{8n} \sum_{l'} \prod_{\nu=1}^{N_{l'}} p'_\nu. \tag{4.34}$$

Substituting Eqs. (4.33) and (4.34) into Eq. (4.30), the detailed balance condition becomes

$$\frac{\sum_l \prod_{\mu=1}^{N_l} p_\mu}{\sum_{l'} \prod_{\nu=1}^{N_{l'}} p'_\nu} = \frac{W_f}{W_i}, \tag{4.35}$$

where the probabilities from step (1) have conveniently canceled out. The values of W_i and W_f can be inserted from Eq. (4.21), giving

$$\frac{\sum_l \prod_{\mu=1}^{N_l} p_\mu}{\sum_{l'} \prod_{\nu=1}^{N_{l'}} p'_\nu} = \frac{\pi'_n}{\pi_n}, \tag{4.36}$$

where π_n and π'_n represent the product of the n initial and final matrix elements, respectively. The purpose of the following subsection is now to show how the individual probabilities p_μ and p'_ν can be chosen in such a way that Eq. (4.36) is satisfied, in order to be used in step (2) of the three-step loop-construction algorithm.

Before delving into the details of the probabilities, p_μ and p'_ν, there is one last point that should be made regarding the construction of loops. The purpose of a loop is to change the initial configuration C_i as much as possible, such that the resulting configuration C_f is different

from—and *hopefully* not strongly correlated with—C_i. However, it is often the case that a *single* loop is not sufficient to produce uncorrelated configurations. To remedy this, one can instead construct a number $\mathcal{N}$ of loops for each Monte Carlo step, where each loop separately satisfies detailed balance. To choose the ideal value of $\mathcal{N}$, we count the number of legs that are encountered in each loop during the initial equilibration: For a given loop l, each of the N_l times step (2) occurs, two new legs are altered,[17] so the number of number of legs altered by the loop is $2N_l$. Since there were a *total* of $4n$ legs, but $2N_l$ legs were actually *altered*, the ideal value of $\mathcal{N}$ *would* have been $2n/N_l$, in order to have the opportunity to alter each leg once. Therefore, to determine $\mathcal{N}$, we simply average the values of $2n/N_l$ that are found during equilibration, and round to the to the nearest integer (≥ 1).[18] It should finally be stressed that $\mathcal{N}$, whatever it is chosen to be, must remain *fixed* during a simulation, to ensure that detailed balance continues to be satisfied.

4.6.3 Directed loop equations

The key to determining the unknown probabilities p_μ and p'_ν in Eq. (4.36) lies in determining how the loops l (that can cause the change $C_i \to C_f$) are related to the loops l' (that can cause the change $C_f \to C_i$). After carefully examining Figs. 4.2–4.4, it should be clear that the following (very important) statement is true: For every loop l, there is a corresponding loop l' that follows the *same path*, but *begins* at the *last leg* of loop l, and *ends* at the *first leg* of loop l, raising the quantum numbers that were lowered in loop l, and lowering the quantum numbers that were raised in loop l.[19]

For the example loop l from the previous subsection, the corresponding loop l' would proceed through the legs (as numbered in Fig. 4.4) in the following order: $[9\ (-)] \to [10\ (+)]$ $\to [7\ (+)] \to [5\ (+)] \to [4\ (+)] \to [3\ (-)] \to 9$, where $(+)$ or $(-)$ refer to quantum numbers

[17]To be more precise, two new legs are *potentially* altered. It is entirely possible that the same leg could be altered multiple times by a single loop.

[18]The precise value of $\mathcal{N}$ that is used can be chosen somewhat arbitrarily. However, if one were to choose $\mathcal{N} \ll 2n/N_l$, then the simulation would suffer from a large integrated auto-correlation time [Eq. (3.8)]. If, on the other hand, one were to use $\mathcal{N} \gg 2n/N_l$, then the later loops would be undoing the changes that were already made by the earlier loops, hence wasting computation time.

[19]Since the loops l and l' follow the same path, it is clear that $N_l = N_{l'}$, so one can replace $N_{l'} \to N_l$ in Eq. (4.36).

being raised or lowered, respectively. To be even more specific, consider both the first ($\mu = 1$) choice of our loop l (to proceed from leg 3 to leg 4), which was made with probability p_1; and the last ($\nu = 3$) choice of the corresponding loop l' (to proceed from leg 4 to leg 3), which was made with probability p_3'. The former change ($\mu = 1$) is shown in Fig. 4.5(a), with the result shown in Fig. 4.5(b); while the latter change ($\nu = 3$) is shown in Fig. 4.5(c), with its result shown in Fig. 4.5(d). These two changes—and, in fact, *every* μ^{th} change of loop l and $(N_l - \mu + 1)^{\text{th}}$ change of loop l'—are related to one another in a very important way, allowing one to satisfy Eq. (4.36).

In order to formulate this crucial relationship between p_μ and $p'_{N_l - \mu + 1}$, first note that the strange looking index, $N_l - \mu + 1$, simply indicates that the pair of related choices—in which we are interested—occur at the *same* location on the path: If we count up from $\mu = 1$ for loop l, then (because l and l' are constructed in opposite directions) we must count *down* from $\nu = N_l$ for loop l' to arrive at the same location. For this reason, we will henceforth define $\nu \equiv N_l - \mu + 1$ for the sake of convenience. Next, let w_μ represent the *initial* value of the matrix element that will be changed—with probability p_μ—by the μth decision of loop l. Similarly, w'_ν will represent the initial value of the matrix element that will be changed—with probability p'_ν—by the νth decision of loop l'.[20] Using this notation, Eq. (4.36) will then be satisfied (as we will show shortly) if we choose

$$\frac{p_\mu}{p'_\nu} = \frac{w'_\nu}{w_\mu} \tag{4.37}$$

for *each* of the N_l decisions that are made during the loop construction.

In order to see that Eq. (4.37) will indeed satisfy Eq. (4.36), first recall that the loops extend over the *same* paths in both the numerator and the denominator, encountering the *same* matrix elements. Therefore, the summations simply produce a single factor that is common to both the numerator and the denominator. (For our example path, a factor of 6 would appear in both the numerator and denominator, since there were 6 possible loops that would accomplish

---

[20]It is important to observe that in the μth step of loop l, the change that occurs is $w_\mu \to w'_\nu$; while in the νth step of l', we have the change $w'_\nu \to w_\mu$.
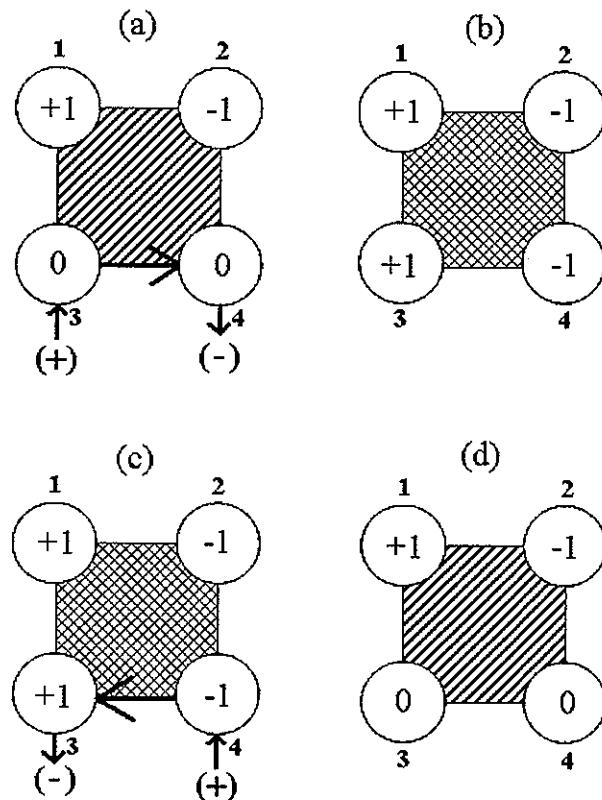
Figure 4.5   Diagrams illustrating (a) the first choice that was made in the construction of the example loop l (from the end of Sec. 4.6.2), which results in the matrix element shown in (b); and (c) the final choice that would be made in the construction of the corresponding loop l', which results in the matrix element shown in (d). In (a) and (c), (+) or (−) indicates a quantum number being raised or lowered, respectively; and the leg numbers (1–4) have been included in (a)–(d). These diagrams could be describing any integer spins.

the change $\mathcal{C}_i \rightarrow \mathcal{C}_f$.) The summations can hence be eliminated from Eq. (4.36), giving

$$\frac{\displaystyle\prod_{\mu=1}^{N_l} p_\mu}{\displaystyle\prod_{\nu=1}^{N_l} p'_\nu} = \frac{\pi'_n}{\pi_n}. \tag{4.38}$$

Next, since p'_ν are c-numbers, the order of the product in the denominator can be reversed, such that the product extends from $\nu = N_l$ to 1, which—still using the definition $\nu \equiv N_l - \mu + 1$—is equivalent to a product extending from $\mu = 1$ to N_l. Therefore, the left-hand side of Eq. (4.38) can be written as a single product,

$$\prod_{\mu=1}^{N_l} \frac{p_\mu}{p'_\nu} = \frac{\pi'_n}{\pi_n}, \tag{4.39}$$

where Eq. (4.39) is now clearly a product of the ratios from Eq. (4.37). Any matrix element that is not encountered by the loop will be unchanged, and will thus cancel from both the numerator and denominator the right-hand side of Eq. (4.39), leaving only the N_l final (updated) matrix elements w'_ν in the numerator and the N_l initial matrix elements w_μ in the denominator. Eq. (4.37) therefore *does* satisfy detailed balance,[21] and *can* be used for step (2) of the loop construction algorithm.

The only task that remains in order to obtain the individual probabilities—to be used in step (2) of the loop construction—is to *normalize* the probabilities such that the total probability of continuing the the loop (by choosing one of the four legs connected to the current matrix element) is equal to unity. For example, consider again the first ($\mu = 1$) choice [shown in Fig. 4.5(a)] that was made in the construction of the example loop l; and, for the sake of this example, let us now assume that $s_j = 1$ for the $j = 1$ (left) spin. With $s_1 = 1$, there would have been *three* (not four) possible choices that could have been made for where to proceed from leg 3. Leg 1 would was not an option, because raising its $m_j = 1$ value is impossible for $s_j = 1$, and lowering the value would have violated angular momentum conservation from state ψ_1 to state ψ_2. (See Fig. 4.2.) The three other choices {[leg 2 (+)], [leg 4 (−)], or [leg 3

---

[21]One could also imagine a situation in which the same matrix element is encountered more than once in the construction of a loop. In that case, the substitution of Eq. (4.37) into Eq. (4.39) would cause any *intermediate* matrix elements to appear in both the numerator and the denominator. They would hence cancel out, and detailed balance would still be preserved.
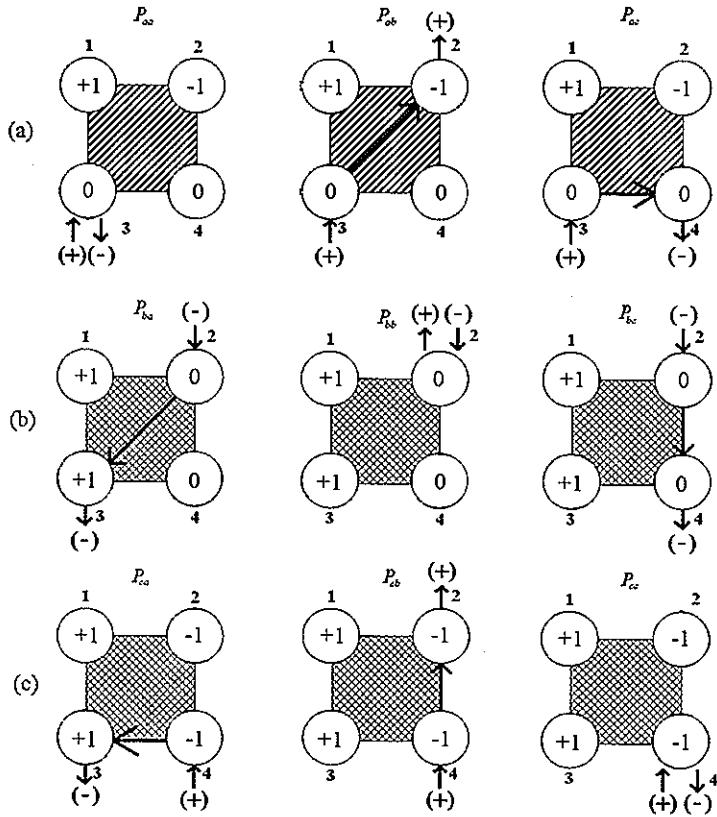
Figure 4.6 Diagram showing the three possible legs that could be chosen—
given an initial leg, and an initial matrix element—for three
different different situations [(a), (b), and (c)]. These 9 proba-
bilities are all related to one another via Eqs. (4.40) and (4.41),
and could describe any integer spins.

$(-)]$} all *are* possible, so the sum of their probabilities must equal 1. These three options are
shown in Fig. 4.6(a).

We now must adopt a different notation[22] in order to write down the final form of the
directed loop equations. In particular, let P_{aa}, P_{ab}, and P_{ac} represent the probabilities of
making the first, second, and third choices in Fig. 4.6(a), respectively; and let the initial
value of this matrix element be represented w_a. Additionally, we will define w_b and w_c to
be the matrix elements that result from the choices whose probabilities are P_{ab} and P_{ac},

---

[22]This is necessary because the derivation of Eq. (4.37) requires us to keep track of where we are in the *loops*
(with the indices μ and ν), while the final form of the detailed balance contains absolutely no reference to the
rest of the loop.

respectively. The choices that "undo" P_{ab} and P_{ac} are labeled P_{ba} and P_{ca}, and are shown in Figs. 4.6(b) and 4.6(c), respectively, along with the other possible choices that could be made when changing *from* the matrix elements w_b and w_c.[23]

Diagrams such as the one shown in Fig. 4.6 were first created in Ref. [1], and are extremely helpful in determining the probabilities, P_{aa}, P_{ab}, etc. From the normalization condition, in each row of Fig. 4.6 the sum of the three probabilities is equal to one. Therefore, if the probabilities in Figs. 4.6(a), 4.6(b), and 4.6(c) are multiplied by w_a, w_b, and w_c, respectively; then one obviously arrives at

$$P_{aa}w_a + P_{ab}w_a + P_{ac}w_a = w_a \qquad (4.40a)$$

$$P_{ba}w_b + P_{bb}w_b + P_{bc}w_b = w_b \qquad (4.40b)$$

$$P_{ca}w_c + P_{cb}w_c + P_{cc}w_c = w_c, \qquad (4.40c)$$

where, from Eq. (4.37), the *products* satisfy

$$P_{ab}w_a = P_{ba}w_b \qquad (4.41a)$$

$$P_{ac}w_a = P_{ca}w_c \qquad (4.41b)$$

$$P_{bc}w_b = P_{cb}w_c. \qquad (4.41c)$$

Therefore, the products on the left-hand sides of Eqs. (4.40a)–(4.40c) form a *symmetric* 3-by-3 matrix, where the sum of each row is equal to w_a, w_b, or w_c; and dividing the elements of this matrix by the appropriate matrix elements (w_a, w_b, or w_c), we obtain the desired probabilities. Note that there is a great deal of freedom in choosing the specific probabilities, since there are 6 equations and 9 unknown probabilities. The specific values that we have used in our calculations are provided in Eq. (4.46) of Sec. 4.7.3.

Finally, note that in some situations any of the four legs are possible choices for how to continue a loop, and in other situations only two choices are possible. An example of each of these scenarios is provided in Fig. 4.7. In Fig. 4.7(a), none of the four choices would lead to the

---

[23]Note that the choices along the diagonal of Fig. 4.6 cause the loop to "turn around" and *undo* the last change that was made by the loop. Choices such as these were termed "bounces" in Ref. [1], and are discussed again in Sec. 4.7.3.
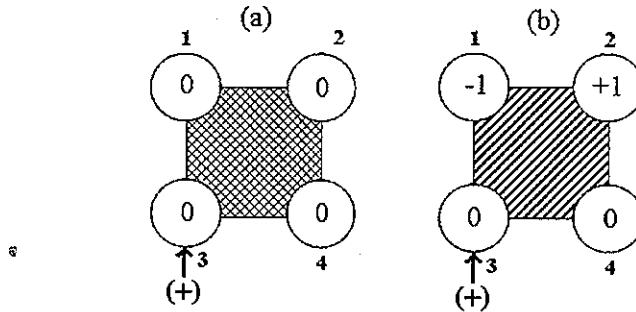
Figure 4.7   Examples of two situations that could be encountered during the construction of a loop. In (a) there are four possible choices for where to proceed next, and in (b) there are only two possible choices. These scenarios would be encountered for any pairs of integer spins.

(forbidden) result $m_j > |s_j|$, and none of the choices would lead to a (forbidden) difference of two units between the m_j quantum numbers in successive states, ψ_k and ψ_{k+1}. Therefore, all choices are possible. In Fig. 4.7(b), two of the choices (leg 2 or leg 4) would cause a difference of two units between the values of m_j between successive states. Therefore, there are only two possible choices (leg 1 or leg 3) in Fig. 4.7(b). Following the same recipe that was used to create Fig. 4.6, the situations in Fig. 4.7(a) and Fig. 4.7(b) would lead to a 2-by-2 symmetric matrix and a 4-by-4 symmetric matrix, respectively. These probabilities are provided in Sec. 4.7.3 as well.

4.7   Practical issues

The material that has been discussed in the previous sections of this chapter has—in some form, although often in a very terse, sometimes incomplete, fashion—already been published in one or more of the references that are cited at the end of this chapter. Rather than producing original *results*, the intended purpose of those sections has instead been to provide an original *description* of this previously known material, which is *hopefully* more complete and easier to understand than the published articles. The current section is somewhat different in the sense that the information presented here focuses more on the details of our particular

68

implementation. The method of determining the truncation L, which we describe in Sec. 4.7.1, was introduced in Refs. [3] and [11], and we provide data to illustrate why the truncation does not affect our results. In Secs. 4.7.2 and 4.7.3 we indicate how the issues of dealing with matrix elements and directed loop equations depend on the intrinsic spins s. This has not (to our knowledge) been previously published, but follows straightforwardly from the material that was discussed earlier in this chapter. In Sec. 4.7.3 we also describe the details of how we have implemented the directed loop equations that were derived in Secs. 4.6.2 and 4.6.3. The probabilities that we use [given in Eqs. (4.45), (4.46), (4.48), and (4.49)] were arrived at independently of Ref. [9] (at roughly the same time as that article's publication), but are entirely consistent with the probabilities suggested therein. For further discussion of the efficiency of various directed loop choices, see Refs. [9] and [10].

4.7.1 Truncating the series expansion

In Sec. 4.4 we claimed that truncating the series expansion at some order L does not influence the results, provided that L is chosen to be sufficiently large. In this subsection, we now show *why* this is true, and we review the simple method that we have used to select L. First, the claim that truncation does not affect our results can easily be understood with the inspection of Fig. 4.8, where—for three different temperatures—we have plotted histograms of the number of times each order n was encountered during a calculation consisting of $100,000$ MCS. Recalling the discussion of Secs. 3.1 and 3.2, these data can be interpreted as the "importance" of the various terms W_i,[24] plotted as a function of n. In all cases, for values of n that are greater than the "peak" value, the data clearly decrease exponentially with increasing n. Therefore, by choosing L even *slightly* larger than the largest n that was encountered, we could have safely used an order of magnitude more Monte Carlo steps, and still would never have reached $n = L$.

To select a value of L that is sufficiently large, we have used the method that was described

---

[24]These data are somewhat analogous to Fig. 3.2 from Chapter 3. In that case, terms with high *energies* are automatically ignored, whereas in Fig. 4.8 the terms with large n are automatically ignored. In both cases, the temperature determines which terms are (and are not) important.
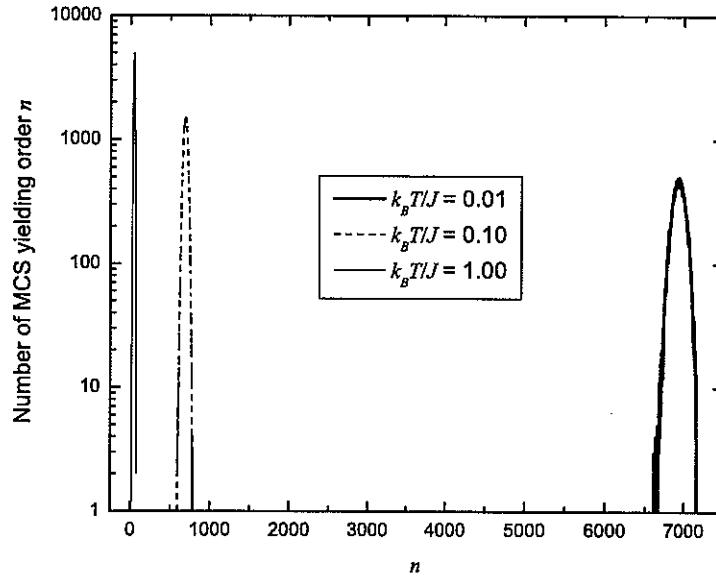
Figure 4.8   Histogram of how many times the nth order of the expansion was encountered during a simulation consisting of $100,000$ Monte Carlo steps. The system being simulated was a ring of $N = 100$ spins $s = 1/2$, interacting via identical AFM bonds of strength J, with $H = 0$.

in Ref. [11], and relies upon the exponential decrease that occurs at the right edge of each distribution in Fig. 4.8. Namely, we begin each simulation with a small value of L,[25] and whenever the diagonal update produces a term with $n = L$, we increase the value of L. For our calculations, we have chosen to increase L by 20% whenever $n = L$, which (for our purposes) has produced a sufficiently large value of L when using $50,000$ MCS to reach equilibrium. Then, after this initial equilibriation, L is fixed for the remainder of the MCS. To ensure that L has indeed been chosen to be large enough, we simply check to make sure that $n < L$ throughout the subsequent data collection, which—when satisfied—implies that the terms with $n \geq L$ are not important enough to be sampled in the given number of MCS.

---

[25]We always begin each simulation with $L = 10$.

4.7.2 Dealing with matrix elements

As we described in Sec. 4.6, both stages of the Monte Carlo update (the diagonal update and the directed loops) require us to know the values of two-spin matrix elements in order to proceed. The corresponding two-spin bond operators were given in Eqs. (4.23) and (4.9b), and their matrix elements are easily evaluated using Eqs. (2.13) and (2.17). The resulting values are

$$\langle s_{1(b)}, m_{1(b)}|\langle s_{2(b)}, m_{2(b)}|\mathcal{H}_{b,1}|s_{1(b)}, m_{1(b)}\rangle|s_{2(b)}, m_{2(b)}\rangle$$
$$= c_b - J_b\left(m_{1(b)} + m_{2(b)}\right) - g\mu_B H\left(\frac{m_{1(b)}}{\xi_{1(b)}} + \frac{m_{2(b)}}{\xi_{2(b)}}\right) \tag{4.42}$$

and

$$\langle s_{1(b)}, m_{1(b)} + 1|\langle s_{2(b)}, m_{2(b)} - 1|\mathcal{H}_{b,2}|s_{1(b)}m_{1(b)}\rangle|s_{2(b)}m_{2(b)}\rangle$$
$$= -\frac{J_b}{2}\sqrt{\left[s(s+1) - m_{1(b)}(m_{1(b)} + 1)\right]\left[s(s+1) - m_{1(b)}(m_{1(b)} - 1)\right]} \tag{4.43}$$

$$\langle s_{1(b)}, m_{1(b)} - 1|\langle s_{2(b)}, m_{2(b)} + 1|\mathcal{H}_{b,2}|s_{1(b)}m_{1(b)}\rangle|s_{2(b)}m_{2(b)}\rangle$$
$$= -\frac{J_b}{2}\sqrt{\left[s(s+1) - m_{1(b)}(m_{1(b)} - 1)\right]\left[s(s+1) - m_{1(b)}(m_{1(b)} + 1)\right]}, \tag{4.44}$$

for the diagonal and off-diagonal matrix elements, respectively. (All other matrix elements are identically zero.)

Although easy to evaluate, there are potentially *many* matrix elements that might be needed. The number of *total* matrix elements for two spins s is $(2s + 1)^4$ (the square of the dimension of the Hilbert space), but most of these matrix elements are clearly zero. First, The states must have the same total spin; i.e., the sum $m_{1(b_k)} + m_{2(b_k)}$ must be the same for the state ψ_k as it is for the state ψ_{k+1} in order for the matrix element $\langle\psi_k|\mathcal{H}_{b_k,o_k}|\psi_{k+1}\rangle$ to be non-zero. Secondly, each individual quantum number can only change by one unit.[26] The number of matrix elements that satisfy these conditions (and are hence non-zero) are shown in the right-most column of Table 4.1.

We have implemented our calculations for values of s up to $\frac{9}{2}$, and have thus only shown these values in the table. However, the number of nonzero matrix elements for *any* s can

---

[26]It is for this reason that only two choices are possible for some of the choices that are made during the construction of directed loop. [See the example in Fig. 4.7(b).]

s	Total matrix elements	Non-zero matrix elements
$\frac{1}{2}$	16	6
1	81	17
$\frac{3}{2}$	256	34
2	625	57
$\frac{5}{2}$	1296	86
3	2401	121
$\frac{7}{2}$	4096	162
4	6561	209
$\frac{9}{2}$	10,000	262

Table 4.1   The total number of 2-spin matrix elements, and the number of non-zero non-zero 2-spin matrix elements for intrinsic spins s.

be obtained iteratively. This is accomplished by adding $5 + 12s$ to the number of non-zero matrix elements for spins s, to determine the number for $s + \frac{1}{2}$. This number may seem very random, but the *difference* (between the values for any s and $s + \frac{1}{2}$) simply increases by six each time.[27] Evaluating hundreds of matrix elements (as shown in Table 4.1) is actually not at all problematic, especially since the their calculation only needs to be performed *once* for each value of the external field H. We perform these calculations before any Monte Carlo steps take place, and store them in a table, to be looked up later when they are needed during the MC update.

4.7.3   Solutions to the directed loop equations

The sole purpose of this subsection is to present the specific probabilities that we have used for the construction of directed loops. First, consider the diagrams [such as the one shown in Fig. 4.7(b)] for which there are only two possible choices for how to proceed. For these situations, there are two normalization equations [analogous to Eq. (4.40)], and one detailed balance equation [analogous to Eq. (4.41)]. Given three equations for the four unknown probabilities, there is only one free parameter that needs to be fixed. To fix this parameter—as well as all other parameters that are encountered in this subsection—we always strive to

---

[27]This is because—when going from spin s to spin $s + \frac{1}{2}$—the number of matrix elements with $M_S = 0$ increases by 3, and the number of matrix elements that *had* a value of $|M_S|$ turns into the number of matrix elements that now have $|M_S| = \pm 1$. To understand this more clearly, the reader is encouraged to sketch the matrix representation of the $\underline{S}^z$ operator for a few choices of s.

eliminate "bounces" (defined in footnote 23) whenever possible. Since bounces undo changes that have already been made (and computation time that has already been invested), this is clearly a reasonable choice, and the subsequent effect on efficiency has been previously investigated.[1, 9, 10] Setting one of the bounce probabilities (P_{aa}) equal to zero, one easily obtains the following probabilities for the situations when there are only two choices for how to continue loop construction.

$$\begin{pmatrix} P_{aa} & P_{ab} \\ P_{ba} & P_{bb} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{w_a}{w_b} & 1 - \frac{w_a}{w_b} \end{pmatrix} \tag{4.45}$$

In Eq. (4.45) it is clearly necessary that $w_b > w_a$ in order that $P_{bb} > 0$, and similar requirements are imposed by the other probabilities below. Therefore, for every calculation, after the matrix elements have been calculated, but before taking any MC steps, we compare the relevant matrix elements with one another in order to define the indices (a, b, etc.) appropriately. We are then careful to *define* the larger matrix element to be w_b and the smaller matrix to be w_a in Eq. (4.45).

For the cases when a loop can continue in any one of three possible ways (as was the case in Fig. 4.6), we are able to set both P_{aa} and P_{bb} equal to zero (when we define $w_c > w_a, w_b$), leaving one remaining parameter. The probabilities [from Eqs. (4.40) and (4.41)] are then given by

$$\begin{pmatrix} P_{aa} & P_{ab} & P_{ac} \\ P_{ba} & P_{bb} & P_{bc} \\ P_{ca} & P_{cb} & P_{cc} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\alpha_3}{w_a} & \frac{\beta_3}{w_a} \\ \frac{\alpha_3}{w_b} & 0 & \frac{\gamma_3}{w_b} \\ \frac{\beta_3}{w_c} & \frac{\gamma_3}{w_c} & \frac{b}{w_c} \end{pmatrix}, \tag{4.46}$$

where

$$\alpha_3 = \frac{1}{2}(b + w_a + w_b - w_c) \tag{4.47a}$$

$$\beta_3 = \frac{1}{2}(-b + w_a - w_b + w_c) \tag{4.47b}$$

$$\gamma_3 = \frac{1}{2}(-b - w_a + w_b + w_c). \tag{4.47c}$$

Note that if $w_c > w_a + w_b$, then we must set $b > w_c - w_a - w_b$ in order for α_3 (and hence P_{ab} and P_{ba}) to be positive; whereas if $w_c < w_a + w_b$, then b can be chosen to be some small positive

number.[28] For our calculations, we have chosen $b \equiv \epsilon + w_c - w_a - w_b$ when $w_c < w_a + w_b$; and $b \equiv \epsilon$ otherwise, and we have typically chosen $\epsilon = 0.01$.

Finally, when all four legs are allowed choices for the continuation of a loop [as was the case in Fig. 4.7(a)], there is even more freedom in choosing the probabilities. When $w_d > w_a + w_b + w_c$, all but one of the bounces can be eliminated, allowing the very simple solution,

$$
\begin{pmatrix}
P_{aa} & P_{ab} & P_{ac} & P_{ad} \\
P_{ba} & P_{bb} & P_{bc} & P_{bd} \\
P_{ca} & P_{cb} & P_{cc} & P_{cd} \\
P_{da} & P_{db} & P_{dc} & P_{dd}
\end{pmatrix}
=
\begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
\frac{w_a}{w_d} & \frac{w_b}{w_d} & \frac{w_c}{w_d} & 1 - \frac{w_a+w_b+w_c}{w_d}
\end{pmatrix}.
\tag{4.48}
$$

Note, however, that if $w_d < w_a + w_b + w_c$, then this solution would give $P_{44} < 0$, so a different solution is necessary. Actually, when $w_d < w_a + w_b + w_c$, all four bounces can be set equal to zero, and we use the probabilities,

$$
\begin{pmatrix}
P_{aa} & P_{ab} & P_{ac} & P_{ad} \\
P_{ba} & P_{bb} & P_{bc} & P_{bd} \\
P_{ca} & P_{cb} & P_{cc} & P_{cd} \\
P_{da} & P_{db} & P_{dc} & P_{dd}
\end{pmatrix}
=
\begin{pmatrix}
0 & \frac{\delta}{w_a} & \frac{\delta}{w_a} & 1 - \frac{2\delta}{w_a} \\
\frac{\delta}{w_b} & 0 & \frac{\alpha_4}{w_b} & \frac{\beta_4}{w_b} \\
\frac{\delta}{w_c} & \frac{\alpha_4}{w_c} & 0 & \frac{\gamma_4}{w_c} \\
\frac{w_a-2\delta}{w_d} & \frac{\beta_4}{w_d} & \frac{\gamma_4}{w_d} & 0
\end{pmatrix},
\tag{4.49}
$$

where

$$
\alpha_4 = \frac{1}{2}(w_a + w_b + w_c - w_d) - 2\delta
\tag{4.50a}
$$

$$
\beta_4 = \frac{1}{2}(-w_a + w_b - w_c + w_d) + \delta
\tag{4.50b}
$$

$$
\gamma_4 = \frac{1}{2}(-w_a - w_b + w_c + w_d) + \delta.
\tag{4.50c}
$$

We must (of course) be careful to choose the remaining parameter δ such that all probabilities are positive. In particular, if $w_a + w_b + w_c - w_d < 2w_a$, then we must choose δ in the range $0 < \delta < (w_a + w_b + w_c - w_d)/4$; and otherwise we must choose $0 < \delta < w_a/2$. For the sake of simplicity, we have simply chosen δ to always lie in the center of whichever range is relevant.

[28]Note that if $w_c \leq w_a + w_b$, one could in fact set $b = 0$ and eliminate this bounce as well. This turn out however to cause the directed loops to not be ergodic. In particular, when the negative sign problem is present (see Chapter 8) this choice makes it impossible to make transitions between the terms for which $W_i > 0$ and those with $W_i < 0$.

s	2×2 diagrams	3×3 diagrams	4×4 diagrams
$\frac{1}{2}$	0	8	0
1	8	16	8
$\frac{3}{2}$	24	24	24
2	48	32	48
$\frac{5}{2}$	80	40	80
3	120	48	120
$\frac{7}{2}$	168	56	168
4	224	64	224
$\frac{9}{2}$	288	72	288

Table 4.2   The number of diagrams (such as the one shown in Fig. 4.6) that involve: four probabilities, and are given by the 2×2 matrix in Eq. (4.45); nine probabilities, and are given by the 3×3 matrix in Eq. (4.46); 16 probabilities, and are given by either the 4×4 matrix in Eq. (4.48), or the 4×4 matrix in Eq. (4.49).

We should stress once again that—in *all* cases—it is necessary to correctly order the probabilities according to the values of the corresponding matrix elements. This ordering process must occur for *every* diagram, such as (for example) the one diagram that was shown in Fig. 4.6. In Table 4.2 we have provided the total number of diagrams of each type (i.e., 2×2, 3×3, and 4×4) that exist for a given spin s. There are often several hundred diagrams to be considered, which might seem imposing. However, the ordering of the matrix elements (and the corresponding probabilities) only needs to happen *once* for each simulation. Before any Monte Carlo steps occur, the probabilities are calculated and written into a table. They are then accessed later, when they are needed during the construction of loops (much like the matrix elements are only calculated once and then stored for later use).

4.7.4   Computation

In this final subsection, we provide a description of the computational issues that are involved in our calculations, including (among other things) the dependence of the computation time on the relevant parameters: the strength of the exchange interaction J, the temperature T, the external magnetic field H, the number of spins N, and the intrinsic spin s. In all cases, these calculations were performed using a linux cluster, were coded using Fortran 90, and

MPI was used for parallelization. Almost all of the calculations were carried out specifically using the "Hal2004" cluster, which belongs to the Ames Laboratory Condensed Matter Physics group, and is a 64 node cluster, with 128 3.06 GHz Intel Xeon processors. However, a few calculations were also performed using (1) the Ames Laboratory "Medusa" cluster, which has 11 nodes, with 22 2.4 GHz Intel Xeon processors; and (2) the "HPC-class" cluster, which is operated by Iowa State University's High Performance Computing group, and consists of 88 2.8 GHz Intel Xeon processors processors on 44 nodes.

The sampling method that was described at the end of Chapter 3 is extremely straightforward to parallelize. When using p processors to calculate N_{si} statistically independent estimates—or N_{si} "bins"[29]—each processor can simply be assigned N_{si}/p of the bins. Each processor then does the same calculation, but using a different sequence of random numbers; and there is no need for communication between the processors until after all Monte Carlo (MC) steps have been completed. The process of taking MC steps is by far the most time consuming part of a calculation (since many steps are usually needed), so the computation time scales very efficiently with p. The one detriment to the efficiency of the scaling is that the initial equilibration—described in Sec. 3.5.1—must take place before the collection of any useful data, and this is not easily parallelized. Instead, we usually choose the simplest strategy, i.e., reaching equilibrium independently on *every* processor. This does not require much extra computation, provided that (on each processor) the number of MC steps that take place *after* equilibriation is considerably larger than the number of MC steps used to reach equilibrium. We have tried various other schemes to avoid this issue, such as having each of the processors deal with different temperatures or fields. In some cases this can be somewhat more efficient, but in other cases it can leave a large number of processors idle. We generally only use a few processors at a time, so the former (simplest) method works quite well.

For a given calculation, we typically use 2–8 processors, and $\chi(T)$ results—spanning the experimentally relevant range (2–300 K)—are obtained in a few minutes; whereas $M(H)$ results— from $H = 0$ to a few tens of Teslas—typically require a few hours. However, for large N and

---

[29]See Sec. 3.5.2.

very low T, $M(H)$ calculations can take up to a few days. (For example, the large N data that is reported in Chapter 5 required multiple days of computation.) The reason for the difference in computation time between the different types of data [$\chi(T)$ vs. $M(H)$], as well as the dependence on N and s, are described below. First though, it is helpful to note that the parameters J, T, and H do not appear independently in the partition function, since Z depends only on the *ratio*, $\mathcal{H}/(k_B T)$. Therefore, if we assume (for the sake of simplicity) that all bonds have the same strength J, and all spins share the the same g, then only *two* parameters are necessary in order to describe how the computation time depends on J, T, and H. These parameters can be defined

$$t \equiv \frac{k_B T}{J} \tag{4.51a}$$

$$h \equiv \frac{g \mu_B H}{J}, \tag{4.51b}$$

and the partition function becomes

$$Z = \text{Tr} \left[\exp \left(-\frac{1}{t} \sum_{\langle j,k \rangle} \vec{s}_j \cdot \vec{s}_k - \frac{h}{t} \sum_{j=1}^{N} s_j^z \right) \right]. \tag{4.52}$$

It is important to note that low-temperature calculations are always more time consuming, such that low-temperature $M(H)$ calculations take longer than those for $\chi(T)$. More specifically, at low temperatures, the computation time grows proportional to $1/t$ as the temperature decreases (with fixed J). This dependence can be understood from the inspection of Eq. (4.19). In the limit $t \to 0$, the internal energy U approaches a constant, temperature-independent value. In order for the right-hand side of Eq. (4.19) to also be independent of temperature, $\langle n \rangle$ clearly must grow proportional to $1/t$. Recalling Sec. 4.6, the amount of work to be done for each MC step—and hence the resulting computation time—is proportional to n for both stages of the update.[30] Therefore, the computation must grow proportional to $1/t$, exactly as we have observed.

As for the dependence on h, computation time increases somewhat with with increasing external field, but this dependence is much less drastic. Specifically, it is typical for the

---

[30] Actually, for the diagonal update, the computation time is proportional to L; but L is always chosen to be slightly larger than the largest n, so $L \propto \langle n \rangle$.

computation time (per data point) to grow by roughly a factor of two as h increases from zero up to the saturation field.[31] It should not be a surprise that computation time increases with h, given the discussion of Sec. 4.5, and Eq. 4.23 in particular. As h increases, larger values of c_b become necessary in order for the matrix elements associated with $\mathcal{H}_{b,1}$ to remain positive for positive values of $m_{1(b)}$ and $m_{2(b)}$. Larger values of c_b are equivalent lower energy eigenvalues, so U will also be lowered. Therefore, Eq. (4.19) again tells us that $\langle n \rangle$ increases, hence resulting in increased computation time. However, the dependence of the computation time on h depends on the details of the energy spectrum, so there is not a simple relation like there was for t.

As larger systems are considered, the computation time grows linearly with N, which is far more desirable than the exponential growth with N that is encountered when attempting to diagonalize the full Hamiltonian matrix. The reason for this growth can again be understood in terms of Eq. (4.19). The precise value of U will naturally depend on the details of the structure. However, for similar geometries, the internal energy *per spin* will be approximately independent of N. Therefore, $U \propto N$, so $\langle n \rangle$—and hence the computation time—scale linearly with N.

Finally, the computation time increases with increasing s, which can be understood in exactly the same way as the other results. As s increases (with fixed t), each Heisenberg term in the Hamiltonian can be "more satisfied", so the internal energy is again lowered, leading to longer computations. Specifically, we have found that the increase in computation time is roughly proportional to $s(s+1)$ for a fixed value of t. In the language of Chapter 6,[32] this means that the computation time actually scales proportional to $J_c/(k_B T)$, which can replace the $1/t = J/(k_B T)$ dependence which we already showed for fixed s.

Bibliography

[1] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002).

---

[31]The "saturation field" is the value of H at which the state with $S = -Ns$ becomes the ground state. This will be clarified in the next chapter.

[32]$J_c \equiv Js(s+1)$ which, as we describe in Chapter 6, is the quantity that should be fixed in order to describe the approach to the classical limit.

[2] A. W. Sandvik, Phys. Rev. B **59**, 14157(R) (1999).

[3] A. W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991).

[4] D. C. Handscomb, Proc. Cambridge Philos. Soc. **60**, 115 (1964).

[5] M. Troyer, F. Alet, S. Trebst, and S. Wessel, in *The Monte Carlo Method in the Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm*, edited by J. E. Gubernatis (American Institute of Physics, 2003), pp. 156–169.

[6] M. Suzuki, Prog. Theor. Phys. **56**, 1454 (1976).

[7] B. B. Beard and U.-J. Wiese, Phys. Rev. Lett. **77**, 5130 (1996).

[8] A. W. Sandvik, R. R. P. Singh, and D. K. Campbell, Phys. Rev. B **56**, 14510 (1997).

[9] O. F. Syljuåsen, Phys. Rev. E **67**, 046701 (2003).

[10] F. Alet, S. Wessel, and M. Troyer, Phys. Rev. E **71**, 036706 (2005).

[11] A. W. Sandvik, J. Phys. A **25**, 3667 (1992).

CHAPTER 5. Low temperature magnetization and the excitation spectrum of antiferromagnetic Heisenberg spin rings

A paper published in Physical Review B[1]

L. Engelhardt and M. Luban

Department of Physics and Astronomy and Ames Laboratory,

Iowa State University, Ames, Iowa 50011, USA

Abstract

Accurate results are obtained for the low-temperature magnetization vs magnetic field of Heisenberg spin rings consisting of an even number N of intrinsic spins $s = 1/2, 1, 3/2, 2, 5/2, 3, 7/2$ with nearest-neighbor antiferromagnetic (AF) exchange by employing a numerically exact quantum Monte Carlo method. A straightforward analysis of this data, in particular the values of the level-crossing fields, provides accurate results for the lowest energy eigenvalue $E_N(S, s)$ for each value of the total spin quantum number S. In particular, the results are substantially more accurate than those provided by the rotational band approximation. For $s \leq 5/2$, data are presented for all even $N \leq 20$, which are particularly relevant for experiments on finite magnetic rings. Furthermore, we find that for $s \geq 3/2$ the dependence of $E_N(S, s)$ on s can be described by a scaling relation, and this relation is shown to hold well for ring sizes up to $N = 80$ for all intrinsic spins in the range $3/2 \leq s \leq 7/2$. Considering ring sizes in the interval $8 \leq N \leq 50$, we find that the energy gap between the ground state and the first excited state approaches zero proportional to $1/N^\alpha$, where $\alpha \approx 0.76$ for $s = 3/2$ and $\alpha \approx 0.84$ for $s = 5/2$. Finally, we demonstrate the usefulness of our present results for $E_N(S, s)$ by

examining the Fe_{12} ring-type magnetic molecule, leading to a more accurate estimate of the exchange constant for this system than has been obtained heretofore.

5.1 Introduction

Since the early 1990s, the field of magnetic molecules has blossomed and the number of different species that exist is increasing rapidly.[1–5] In particular, there is a large family of so-called ring-type magnetic molecules[3, 6–16] that we focus on in the present work. Within such molecules there are embedded an even number N of identical paramagnetic ions of intrinsic spin s occupying N equally spaced sites defining a ring. Each such ion ("spin") is coupled to its two nearest neighbors via an AF exchange interaction, resulting in systems that can often[6, 7, 10–12, 14] be well represented by an isotropic Heisenberg model with a single exchange energy, $J > 0$, of the form

$$\mathcal{H} = J\sum_{i=1}^{N} \vec{s}_i \cdot \vec{s}_{i+1} + g\mu_B \vec{H} \cdot \sum_{i=1}^{N} \vec{s}_i, \tag{5.1}$$

where the spin operators $\vec{s}_i$ are given in units of $\hbar$, g is the spectroscopic splitting factor, and μ_B is the Bohr magneton. In the first term of Eq. (5.1), the cyclic character of the system is maintained by requiring that $\vec{s}_{N+1} \equiv \vec{s}_1$. The second term describes the standard Zeeman effect, where the external field $\vec{H}$ is typically defined to be directed along the z-axis. The total spin operators S^2 and S_z then commute with $\mathcal{H}$, and the eigenstates are described by quantum numbers S and M_S whose values range from 0 to Ns and from $-S$ to S, respectively. In Fig. 5.1 we display the zero-field energy spectrum corresponding to Eq. (5.1) for a particular example, $s = 3/2$ and $N = 6$, with the subset of minimal energies (SME) shown in red (gray). The SME are closely related to what are called level-crossing fields, quantities which are used to study the SME in great detail throughout this work for many choices of s and N.

In an external magnetic field, the $(2S + 1)$-fold degeneracy of each field-free multiplet is lifted due to a shift, $g\mu_B H M_S$, originating in the Zeeman term. As the external field is increased from $H = 0$, the ground state will change (among the members of the zero-field SME) successively from $S = 0$, $M_S = 0$ to $S = 1$, $M_S = -1$, etc., in integer steps of S and M_S until $S = -M_S = Ns$, corresponding to saturation of the magnetization. Each of the
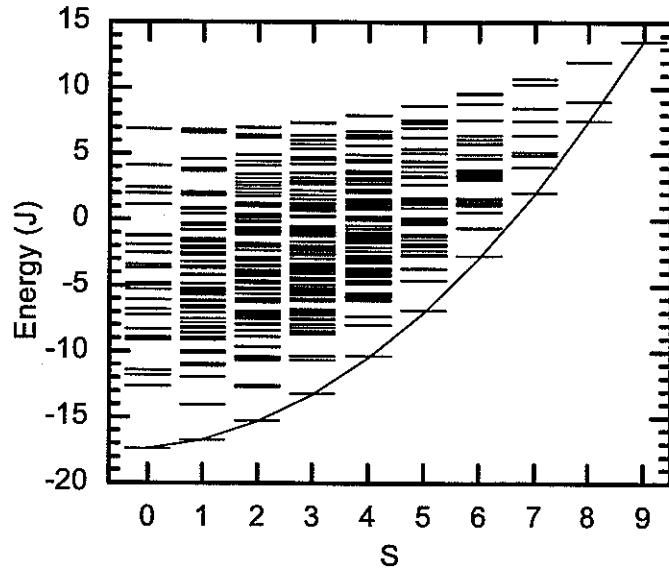
Figure 5.1   Complete energy spectrum for $N = 6$ and $s = 3/2$ obtained by diagonalizing the Hamiltonain in Eq. (5.1) for $H = 0$. The energy levels shown in gray are the subset of minimal energies (SME) for this system as explained in the text. The solid line is a parabola, given in the text, that gives an excellent fit to the SME.

Ns changes of the ground state quantum numbers is referred to as a level-crossing, and the field at which the nth level-crossing occurs is denoted in the following by H_n. By determining these fields, we seek to record the characteristics of the SME as a function of s and N. This is accomplished using the difference equation,

$$E_N(S, s) = E_N(S - 1, s) + g\mu_B H_n \qquad (5.2)$$

for $S = n$, where n extends from 1 to Ns. We elaborate on this connection between the SME and the H_n in detail in the following section.

In order to appreciate the details of the SME, we first review some generic features of the spectra and, in particular, the SME, that are already known. It has been noted,[12, 17] and is clearly evident in Fig. 5.1, that the SME are accurately approximated by a quadratic dependence on S of the form $E(S) \propto S(S + 1)$, as for a quantum rotor. The solid curve in

Fig. 5.1 describes the parabola $E(S) = cJS(S+1)/(2N) + E_G$, where $c = 4.14$ gives the best fit to this SME which has a ground state energy $E_G = -17.393J$. (The reason for the inclusion of the factor of $2N$ in this equation will become clear in Sec. 5.2.) If the SME were strictly parabolic in S, this would give rise to uniformly spaced level-crossing fields. Although uniform spacing is approximately realized in Fig. 5.2 for our example, we find that the accuracy of such an approximation deteriorates for larger values of N. This is explored in detail in Sec. 5.2.

Above the SME there exists a large forest of energy levels. Although many of these levels lie very close to one another, there is a relatively large energy separation between the SME and the higher energy levels, which has been previously observed.[17, 18] Since at a low temperature T only the lowest levels can be thermally occupied, and all other levels lie well above the SME, the magnetization as a function of field $M(H)$ consists of a series of thermally broadened steps that arise at the level-crossing fields and are determined solely by the SME. [The magnetization is also a function of T, but we will write $M(H)$ for the sake of brevity.] This step-like property is illustrated in Fig. 5.2, where $M(H)$ and dM/dH, the differential susceptibility, are shown for the $s = 3/2$, $N = 6$ example with $k_BT/J = 0.1$. The data in Fig. 5.2(a) were calculated in three different ways: using the partition function that includes the complete energy spectrum; using the quantum Monte Carlo (QMC) method employed in this work; and using a modified partition function that includes only the states belonging to the SME. The sharp peaks that appear in Fig. 5.2(b) were calculated using QMC and the susceptibility fluctuation formula to give dM/dH directly, not by differentiating $M(H)$.

The three data sets shown in Fig. 5.2(a) are all identical to 4 significant figures, supporting the assertion that the SME are sufficient for analyzing low-temperature experimental data of this type. For larger values of N, lower temperatures are needed in order to obtain this degree of agreement, especially in the vicinity of the saturation field. For this reason, we have carefully checked that as the temperature is lowered, the level-crossing fields, have indeed converged to their limiting, temperature-independent values. As shown in Sec. 5.2 it is these fields that are then used to calculate the SME function $E_N(S, s)$.

Despite the very simple appearance of the Hamiltonian of Eq. (5.1), the evaluation of the
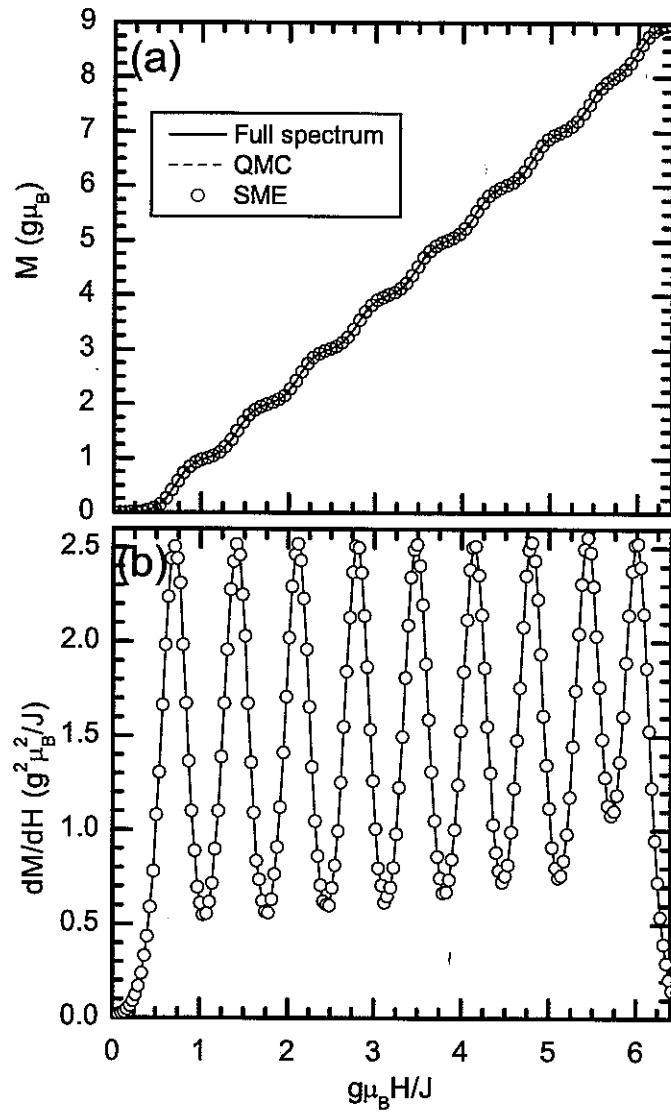
Figure 5.2 (a) $M(H)$ and (b) dM/dH for the example system ($N = 6$ and $s = 3/2$) at a fixed temperature, $k_BT/J = 0.1$. The data in (b) are obtained by quantum Monte Carlo (QMC) calculations, and as a guide to the eye, the line connects successive data points.

corresponding energy eigenvalues and resulting thermodynamic properties frequently presents a major challenge. The most straightforward way to deal with this Hamiltonian, and the method that is usually employed when analyzing magnetic molecules, is to numerically diagonalize the Hamiltonian matrix. This yields energy spectra such as that shown in Fig. 5.1. However, even for relatively small rings, the dimensionality of the Hilbert space, given by $D = (2s + 1)^N$, is so large that the exact diagonalization of the Hamiltonian matrix becomes totally impractical. For the small ring that has been considered as an example, $s = 3/2$ and $N = 6$ gives $D = 4096$. If we consider a larger ring, for example $s = 5/2$ and $N = 12$, which will be analyzed in Sec. 5.4, we already have $D \approx 2.2 \times 10^9$, which is well beyond the practical limit of existing computers. For $s = 5/2$ and $N = 20$, D is a staggering 3.6×10^{15}.

We can entirely avoid the obstacles confronting matrix diagonalization by using a QMC method that is not restricted by the dimensionality of the Hilbert space. Here, we only focus on low-temperature $M(H)$ and $dM/dH(H)$, which are used to determine the SME, but other thermodynamic quantities (such as the temperature-dependent susceptibility, specific heat, and internal energy) are also readily attainable using this method for all temperatures and fields and are, in fact, computationally much less demanding than the present low-temperature studies.

As seen above, knowledge of the SME enables one to obtain accurate values of low-temperature $M(H)$ and $dM/dH(H)$ data. To this end, the SME are calculated in Sec. 5.2 for all s from 1/2 to 5/2 and all even N from 4 to 20. These data are presented in the form of convenient, dimensionless "spectral coefficients" that will be introduced in Sec. 5.2. The spectral coefficients are also presented for larger rings, $N = 40$, 80, and larger intrinsic spins, $s = 3$ and $s = 7/2$. Such large values of N and s are useful for studying the approach to the classical limit ($s \to \infty$).

In Sec. 5.3 the energy gap $\Delta_s(N)$ between the $S = 0$ ground state and the lowest $S = 1$ state, which can be inferred from the first level-crossing field, is analyzed in greater detail for successively larger values of N, up to $N = 50$, for $s = 3/2, 2, 5/2$. This gap is experimentally relevant for NMR and inelastic neutron scattering (INS) experiments and is also important for analyzing low temperature, low-field susceptibility data. Finally, as an illustration of the use-

fulness of the present results, in Sec. 5.4 we analyze an existing ring-type magnetic molecule[11] composed of 12 Fe^{3+} ions ($s = 5/2$), leading to an improved estimate for the exchange constant. With the experimental advancements that are being made both in the synthesis of molecules and in high field magnetization studies, we anticipate that the use of the theoretical data presented in this work will complement future experiments in a much needed way, providing more accurate estimates of microscopic parameters for future ring-type molecules.

5.2 Spectral Coefficients

Since the Hilbert space associated with $\mathcal{H}$ is often too large to allow diagonalization of the Hamiltonian matrix, other theoretical methods must be found. To analyze low-field susceptibility data, classical spin models and scaled-up data from smaller systems can sometimes be useful.[6, 11, 12, 14] However, the level-crossings that are observed in high-field experiments have no classical analog and cannot be easily scaled up. For this reason, reliable theoretical data have previously been lacking, and a main goal of the current work is to remedy this situation through detailed QMC calculations.

In order to learn about the nature of the SME, we used the stochastic series expansion method[19] with directed loops as described in Ref. [20]. This method is based on a series expansion of the partition function, whereby different orders of the expansion are sampled by adding and/or removing the diagonal portion of two-spin (bond) operators according to detailed balance. To ensure that off-diagonal operators are sampled ergodically, "loops" are constructed in an extended phase space. For the case of $s = 1/2$, updating proceeds by simply flipping all spins along a loop from spin down (up) to spin up (down). For $s > 1/2$, each loop involves the evaluation of matrix elements of bond operators, acting on spin states that have been altered by spin raising and/or lowering operators. These loops will fulfill the detailed balance condition provided that one is careful to incorporate the appropriate "directed-loop equations"[20] for the given choice of s.

Using this method, we simultaneously calculate both M and dM/dH versus H at fixed temperatures, an example of which was shown in Fig. 5.2. From these data, we can very

accurately infer the level-crossing fields and thereby reconstruct the SME. This follows from Eq. (5.2) which gives

$$E_N(S, s) = g\mu_B \sum_{n=1}^{S} H_n + E_G, \qquad (1 \le S \le Ns), \qquad (5.3)$$

where $E_G \equiv E_N(0, s)$ is the ground state energy. It is convenient to define the quantities

$$h_n \equiv \frac{g\mu_B H_n}{J} = \frac{c_n(N, s)n}{N}, \qquad (n = 1, \ldots, Ns), \qquad (5.4)$$

where the dimensionless numbers $c_n(N, s)$ will be referred to as "spectral coefficients". The energy spectrum of the SME may thus be written as

$$E_N(S, s) = \frac{J}{N} \sum_{n=1}^{S} n c_n(N, s) + E_G, \qquad (1 \le S \le Ns). \qquad (5.5)$$

Note that if $c_n(N, s)$ were independent of n and given by $c(N, s)$, Eq. (5.5) would reduce to

$$E_N(S, s) = \frac{c(N, s)J}{N} \frac{S(S+1)}{2} + E_G, \qquad (0 \le S \le Ns), \qquad (5.6)$$

the so-called rotational band model, that has often been employed to analyze magnetization data.[6, 17, 21, 22] Inspecting Eq. (5.4), the rotational band model immediately implies that the level-crossing fields are equally spaced, which, as we will demonstrate in the Secs. 5.2.1–5.2.3 subsequent subsections, is hardly the case. Instead, Eq. (5.5), in conjunction with the spectral coefficients presented in Secs. 5.2.1–5.2.3, provides a highly accurate, yet convenient means of representing $E_N(S, s)$ and, thus, for analyzing low-temperature magnetization data.

Based on previously known properties of Heisenberg rings, it is easy to show that $c_n(N, s)$ is exactly 4 for a very few special cases. These are listed here and will be useful in discussing the results of our calculations in subsequent subsections: (i) In the case of the $N = 4$ ring, $c_n(4, s) = 4$, independent of n and s. This is easily derived by describing this system in terms of two sublattices, each consisting of two spins. As a result, the SME is given exactly by $E_4(S, s) = JS(S + 1)/2 + E_G$. (ii) In the limit of classical spins,[2] $\lim_{s \to \infty} c_n(N, s) = 4$, for all n and N. (iii) In all cases $c_{n=Ns}(N, s) = 4$, independent of N and s. This follows from

---

[2]For rings of classical spins with N even, the SME can be described by the continuous function $E_N(S_c) = 2J_c S_c^2/N + E_G$, given in Eq. (80) of Ref. [41]. Replacing the classical exchange constant $J_c = s(s+1)J$, and the classical spin S_c by their quantum analogs, we obtain $E_N(S, s \to \infty) = 2JS(S+1)/N + E_G$ from which item II follows.

Eq. (5.5) and the fact that the state with $S = Ns$ has energy $E_N(Ns, s) = JNs^2$, while the SME energy with $S = Ns - 1$ is $E_N(Ns - 1, s) = Js(Ns - 4)$.[23]

Since the spectral coefficients have a value of exactly 4, both in the limit of very small rings [item (i)] and in the limit of very large intrinsic spins [item ii)], one might expect that the replacement, $c_n(N, s) = c(N, s) \approx 4$, independent of n, would provide a very good approximation, for example, for Fe^{3+} ions ($s = 5/2$) in small rings ($N \leq 20$). However, as shown in Secs. 5.2.1–5.2.3 for different choices of s and N, the spectral coefficients do vary significantly with n.

5.2.1 $s = 1/2$, and $s = 1$

Rings of $s = 1/2$ spins have been studied using many methods, and a great deal is known about their spectra. In the 1960s, the lowest energies, $E_{N=\infty}(S, s = 1/2)$, were calculated[24] in the thermodynamic limit using the Bethe ansatz,[25, 26] while numerical diagonalization[27] was carried out on finite rings. More recently, work has continued for finite N using methods that include the quantum Monte Carlo[28], renormalization group[29, 30], Lanczos[18, 31], and conformal field theory methods.[32, 33]

The lowest eigenvalues for small $s = 1/2$ rings can be easily obtained from straightforward matrix diagonalization, but are included here both for completeness and to assess the usefulness of Eqs. (5.5) and (5.6). The spectral coefficients that are shown in Fig. 5.3(a) as a function of $n/(Ns)$ define the SME for small $s = 1/2$ rings. One can immediately notice that the c_n vary with n, and most are much larger than 4, implying that a rotational band approximation provides a relatively poor approximation to these spectra.

Also included in Fig. 5.3(a) are the spectral coefficients corresponding to Griffith's original $M(H)$ result for the infinite $s = 1/2$ chain,[24] which is shown as a solid curve in Fig. 5.3(b). In the thermodynamic limit, the transformation from magnetization to spectral coefficients can be accomplished by making the replacement, $n/N \to m_0$, where m_0 is the zero temperature magnetization per spin in units of $g\mu_B$. Eq. (5.4) can then be rewritten,

$$c_n(N = \infty, s) = \frac{h_n}{m_0(h_n)}.$$

(5.7)
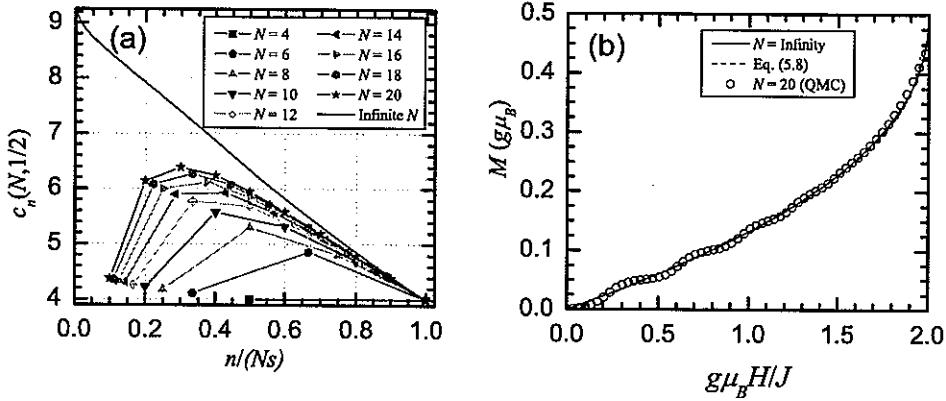
Figure 5.3 (a) Spectral coefficients for small $s = 1/2$ rings. The solid lines are included to guide the eye, and the continuous curve corresponds to the $N = \infty$ magnetization data of Ref. [24]. (b) $M(H)$ for $N = \infty$ (from Ref. [24]), for $N = 20$ (QMC) and an approximation based on (5.8).

As can be seen in Fig. 5.3(a), for $N = \infty$ the spectral coefficients form a nearly linear function of $n/(Ns)$ over a very large range of this variable. Approximating these data as a linear function,

$$c_n(N = \infty, 1/2) \approx \alpha - \beta \frac{n}{Ns}, \tag{5.8}$$

and substituting Eq. (5.8) into Eq. (5.7), again replacing $n/N \to m_0$, the resulting approximate magnetization is

$$m_0 = \frac{\alpha s}{2\beta} \left(1 - \sqrt{1 - \frac{4\beta}{s\alpha^2}h}\right), \tag{5.9}$$

where $s = 1/2$. Fitting the $c_n(\infty, 1/2)$ data to the linear function, we find $\alpha = 8.9$ and $\beta = 5.07$. The corresponding curve terminates at the point ($h = 1.953$, $m_0 = 0.439$), rather than at (2, 0.5), but otherwise is virtually indistinguishable in Fig. 5.3(b) from the exact magnetization (solid curve). This deviation of the terminus is due to the fact that the linear approximation of Eq. (5.8) does not incorporate a small positive curvature of the cluster coefficients as a function of $n/(Ns)$ as n approaches Ns. Also included in Fig. 5.3(b) is $M(H)$ for the $s = 1/2$, $N = 20$ ring at a temperature $k_B T/J = 0.05$. These data are nearly identical to that of the infinite ring, except for the existence of thermally broadened steps associated with level-crossings.
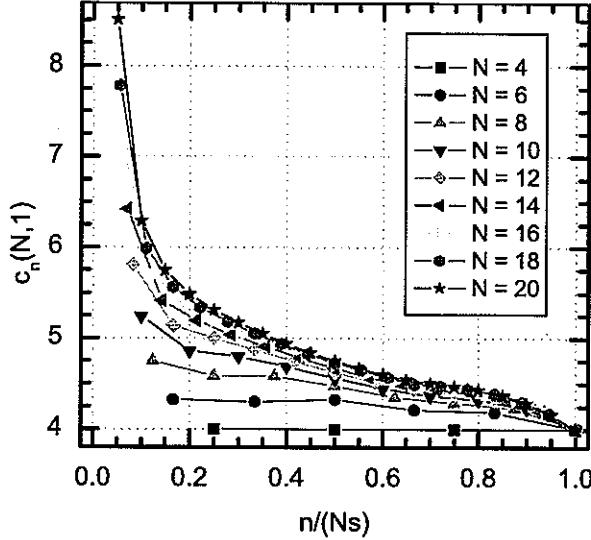
Figure 5.4   Spectral coefficients for small rings of intrinsic spins $s = 1$. The solid lines are included to guide the eye.

Heisenberg rings of $s = 1$ spins have received a great deal of attention since Haldane's prediction[34] that a finite gap separates the ground state from the first excited state in infinite rings of integer spins s.[35] In the notation of the present work, this gap is given by $\Delta_s(N) \equiv E_N(1, s) - E_G = \frac{J}{N} c_1(N, s)$, and the values of $c_1(N, 1)$, seen as the left-most points in Fig. 5.4, are in good agreement with published values[36] of $\Delta_1(N)$. Values of $\Delta_s(N)$ for all s in the range $1/2 \leq s \leq 5/2$ will be discussed in Sec. 5.3.

Note, however, that the data presented here and in Sec. 5.2.2 include not only the first energy gap [associated with $c_1(N, s)$], but *all* energy levels that belong to the SME. Studying the details of of the SME, we find a very rich structure. For instance, it is evident in Fig. 5.4 that $c_n(N, 1)$ decreases rapidly with increasing n, unlike the corresponding data for $s = 1/2$. For $n \geq 0.4Ns$ the value of $c_n(N, 1)$ has already fallen below 5 for $N \leq 20$, whereas for $s = 1/2$ this value is not reached until $n > 0.75Ns$. In this sense, increasing s from $1/2$ to 1 is a significant step on our way toward the classical limit, stated in item (ii) earlier in Sec. 5.2.
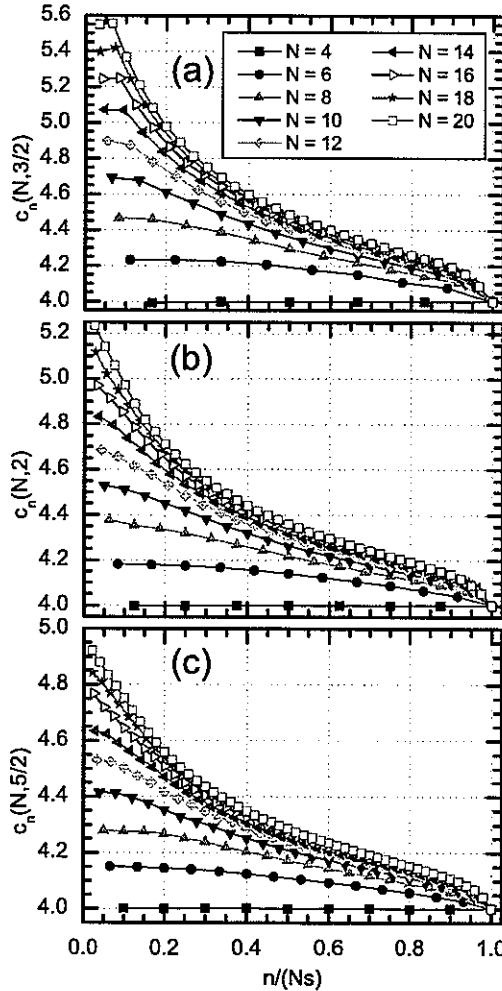
Figure 5.5   Spectral   coefficients   for   small   rings   of   intrinsic   spins
$s = 3/2, 2, 5/2$. The solid lines are included to guide the eye.

5.2.2 $s = 3/2$, $s = 2$, and $s = 5/2$

Systems of larger intrinsic spins have also been studied in recent years[17, 18, 37–39] but
with less frequency than $s = 1/2$ and $s = 1$ systems. Since a knowledge of the spectral
coefficients for $s = 3/2$, 2, and 5/2 is important for a number of molecular rings, these data
are presented in Fig. 5.5 for all $N \leq 20$. The values of $c_1(N, s)$ that appear in Figs. 5.5(a)
and 5.5(b) agree with the values of $\Delta_s(N)$ that have been published[39] ($N \leq 10$). Again,
besides the first energy gap, the SME exhibit several interesting characteristics, which are
reflected in the spectral coefficients.
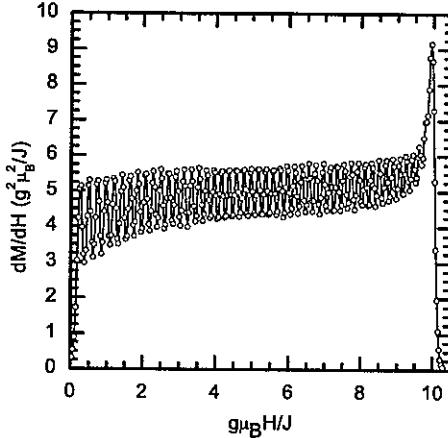
Figure 5.6   Differential susceptibility for a ring of $N = 20$ intrinsic spins $s = 5/2$ at a temperature, $k_B T/J = 0.05$. The large peak immediately before saturation $(g\mu_B H/J = 10)$ is discussed in the text.

Of course, the spectral coefficients for $N = 4$ are all equal to 4 as required by item (i). As N increases with fixed s and $n/(Ns)$, the corresponding spectral coefficients increase from 4 monotonically, resulting in the series of nonintersecting curves seen in Fig. 5.5. This is consistent with Waldmann's observation[18] that the rotational band model becomes poorer for larger rings.

Anchored at 4 for $n = Ns$ [item (iii)], and always approaching 4 from above, the values of the spectral coefficients decrease sharply as n approaches Ns. This ubiquitous drop can be discussed in a number of contexts. Recalling Eq. (5.4), this is clearly equivalent to a compression of the level-crossing fields as saturation is approached. At low temperatures, this results in a large slope of $M(H)$, as can be seen in Fig. 5.6 for $N = 20$ and $s = 5/2$. In terms of the energy spectrum, this implies that the curvature of the SME decreases for large S.

Finally, note that as s is increased with fixed N and $n/(Ns)$, the spectral coefficients descend toward 4 [item (ii)], but only very slowly. Even for $s = 5/2$, most of the spectral coefficients shown in Fig. 5.5(c) are considerably larger than 4, indicating that one is still far from the classical limit that is stated in item II. This behavior is explored in Sec. 5.2.3 with the inclusion of larger values of intrinsic spin.

5.2.3 Scaling relation for large s

Thus far we have presented the spectral coefficients that define the SME as a function of three variables, s, N, and $n/(Ns)$, and some general trends have emerged. Now, considering larger values of s and N, we would like to make more quantitative statements regarding the functional dependence of $c_n(N, s)$ on these variables. To that end, we have calculated the spectral coefficients for values of s up to $7/2$ and present those data for $3/2 \leq s \leq 7/2$.
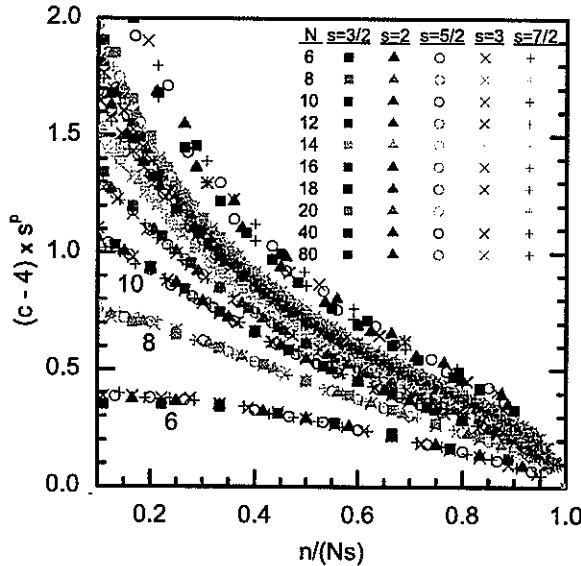
Figure 5.7   Spectral coefficients adjusted to investigate the scaling behavior of Eq. (5.10) for all s in the range $3/2 \leq s \leq 7/2$. The data shown correspond to $p = 1.05$.

In Fig. 5.7, we plot the quantity $[c_n(N, s) - 4] \times s^p$ as a function of $n/(Ns)$ for the choice $p = 1.05$. From these data the s dependence of the spectral coefficients is immediately evident. For each value of N, the data for all s lie on a single curve, implying that the spectral coefficients scale according to

$$c_n(N, s) = 4 + s^{-p} f(N, n/(Ns)). \tag{5.10}$$

In particular, for $s \to \infty$ Eq. (5.10) will be in accord with item (ii). The slow approach to 4 as s is increased is noteworthy, as even $s = 7/2$ is still far away from the classical limit. Choosing

a slightly different value for the scaling exponent p, such as 1.03 or 1.07, yields visibly inferior results, so we conclude that $p = 1.05 \pm 0.01$.

A few of the spectral coefficients are also calculated for larger rings, $N = 40$ and $N = 80$. The inclusion of these data in Fig. 5.7 serves two purposes. First, these data suggest that $f(N, n/(Ns))$ is, indeed, converging to a finite limiting curve in the limit $N \rightarrow \infty$, which defines the zero temperature $M(H)$ of an infinite chain of spins s. Secondly, the larger N data strengthen our belief that the scaling relation (5.10) is valid for all N.

Note that in Fig. 5.7 data are only included for $n/(Ns) > 0.1$. The data for small $n/(Ns)$ have not been included because the error in calculating c_n using the QMC method rapidly increases as $n/(Ns)$ decreases towards zero. The $n = 1$ (gap) behavior is considered in Sec. 5.3.

5.3 Energy Gap

We now explore the energy gap $\Delta_s(N)$ between the ground state and the first excited SME level. Values of this gap are shown in Fig. 5.8(a) for rings of $N \leq 20$ spins $s \leq 5/2$. Much like the behavior of the full SME discussed in Sec. 5.2, this gap systematically approaches the limiting $s = \infty$ form as s increases from $3/2$, while the $s = 1/2$ and $s = 1$ data exhibit distinctly different trends.

Specifically, the energy gap for $s = 1/2$ rings is very similar to the energy gap that would be obtained for a ring with the same value of N but very large s. This large s limit, indicated in Fig. 5.8(a) as the "classical rotational band," follows from item (ii) and is given by $\Delta_\infty(N) = 4J/N$. By contrast, $s = 1$ rings have much larger gaps. Note also that these are already within 3.5% of the limiting, $N = \infty$ value even for $N = 20$. The known limiting value,[35] $\Delta_1(\infty) \approx 0.4105J$, is indicated by an arrow on the right side of Fig. 5.8(a).

Recall that for any s, $\Delta_s(4) = J$ from item (i). Considering $N = 6$, the classical result $\Delta_\infty(6) = 2J/3$ is still a reasonable approximation to $\Delta_s(6)$, with a relative error of only a few percent for any s. However, as N increases further this error continues to grow, and with $N = 20$ it is nearly 25% for $s = 5/2$ and nearly 40% for $s = 3/2$.

Although the classical result is not sufficient, we find that the energy gaps for $s > 1$ are
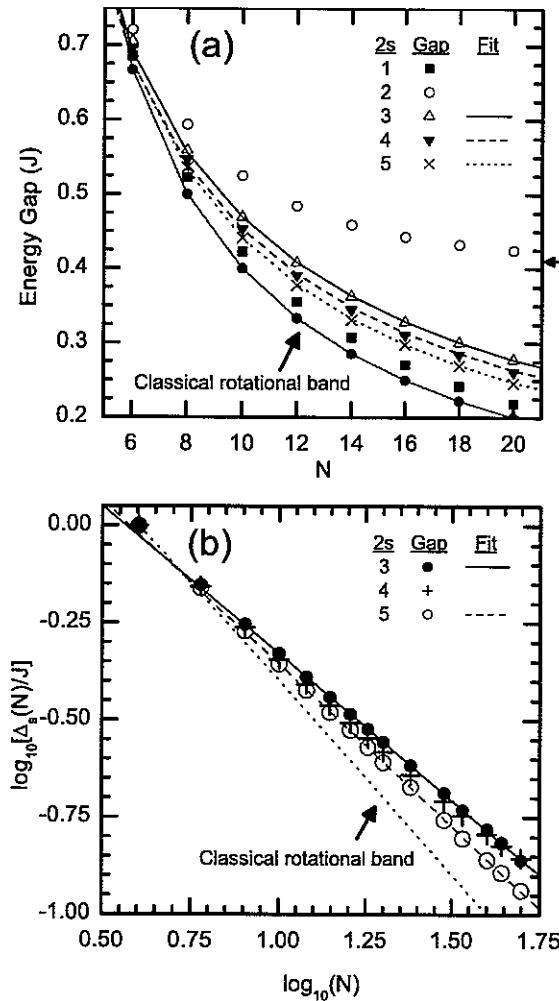
Figure 5.8   Energy gap $\Delta_s(N)$ for rings of N spins (a) for all $s \leq 5/2$ (symbols) along with the best fits to Eq. (5.11) (curves) as described in the text and (b) with N varying from 4 to 50 for $s = 3/2$, 2, 5/2 (symbols) along with the best fits to Eq. (5.11) (lines) as described in the text.

well described by a slightly more general power-law dependence on N of the form

$$\Delta_s(N) \sim \Omega N^{-\alpha}. \tag{5.11}$$

The curves in Fig. 5.8(a) were obtained by choosing: $\Omega = 2.68J$ and $\alpha = 0.757$ for $s = 3/2$; $\Omega = 2.73J$ and $\alpha = 0.781$ for $s = 2$; $\Omega = 3.03J$ and $\alpha = 0.837$ for $s = 5/2$; while, of course, $\Omega = 4J$ and $\alpha = 1$ corresponds to $s = \infty$. With these choices of Ω and α, excellent agreement with the QMC data is obtained in the range $8 \leq N \leq 20$ and it is clear that the classical limit is, indeed, being approached with increasing s for both Ω and α.

For the half odd integer spins, $s = 3/2$ and $5/2$, the agreement with Eq. (5.11) continues for larger values of N. The same data are shown in Fig. 5.8(b), now including $N \leq 50$, and the QMC data agree with the power-law formulas to within a fraction of a percent for all ring sizes in the range $8 \leq N \leq 50$, which is comparable to our uncertainties in $\Delta_s(N)$. The values of $\Delta_2(N)$ begin to diverge from the power-law dependence for $N \gtrsim 30$, which is to be expected since they must eventually converge to a nonzero value. This gap for $s = 2$ chains has been previously studied in great detail, and density matrix renormalization group calculations have yielded a value[40] of $\Delta_2(N) = 0.0876J \pm 0.0013J$ in the limit as $N \to \infty$. One can see in Fig. 5.8(b) that $\Delta_2(N)$ is beginning to approach its limiting value, having become larger than $\Delta_{3/2}(N)$ for $N \geq 50$, but data for much larger rings would be necessary in order to obtain an accurate estimate for the limit $N \to \infty$.

The rotational band result, $\Delta_s(N) = 4J/N$, has been used in the past[11, 12, 21] as an estimate of $\Delta_s(N)$. Although this provides a reasonable approximation for $N < 10$, as we have seen it quickly diverges from the correct result with increasing N. As such, it would be prudent to use the more accurate results presented here when attempting to relate J to the experimentally measured lowest energy gap, e.g., by using INS, NMR, low-field susceptibility, or magnetization data.

5.4   An application: Fe_{12}

In this section we apply our results to a known magnetic molecule,[11] whose analysis has been challenged by a Hilbert space of dimension $D = 6^{12} \approx 2.2 \times 10^9$. The molecule is comprised

of 12 Fe^{3+} ions ($s = 5/2$), whose interaction was first investigated[11] by measuring the low-field susceptibility $\chi_0(T)$ as a function of temperature and fitting these data to an approximation of the $s = 5/2$, $N = \infty$ chain. The exchange energy thus obtained was $J/k_B = 31.9$ K for $g = 2.00$. The field-dependent magnetization of the molecule has also been measured and analyzed, and the first four level-crossing fields at low temperatures were[21] $H_1 = 10.1 \pm 0.2T$, $H_2 = 19.6 \pm 0.2T$, $H_3 = 29.6 \pm 0.4T$, $H_4 = 39.1 \pm 0.8T$. An analysis of the magnetization was given in Ref. [21] using the classical rotational band $c(N, s) = 4$, and this yielded the estimate $J/k_B = 40.7$ K with $g = 2.02$. Note that the latter estimate is more than 25% larger than the former[11] estimate. Given the results of Sec. 5.2, one can expect that the estimate $J = 40.7$ K will be considerably larger than what will result from an accurate treatment of the Heisenberg model. This is borne out in the following.

In Fig. 5.9, we compare the four measured level-crossing fields with our low-temperature. ($k_B T/J = 0.01$) QMC results. At this low temperature, each level-crossing of the theoretical $s = 5/2$, $N = 12$ Heisenberg ring is clearly indicated by a narrow peak in dM/dH. Note that the peaks in the QMC data arising from the parameters $J = 31.9$ K and $g = 2.00$ occur at fields that are considerably below the experimental level-crossings indicated by the dashed vertical lines. On the other hand, the QMC peaks that correspond to $J = 40.7$ K and $g = 2.02$ are all at fields significantly greater than the measured values. Particularly at high fields, these discrepancies become quite pronounced, suggesting that neither choice of parameters is consistent with the experimental data. However, we find that the predictions of the Heisenberg model agree very well with the experimental data if we use $J = 35.2$ K and $g = 2.0$. With this choice of parameters, each of the four theoretical peaks clearly coincides with a measured level-crossing shown in Fig. 5.9.

Without using the dM/dH level-crossing field data directly, one can easily arrive at the same estimate based on the spectral coefficients of Sec. 5.2. Recalling Eq. (5.4), the ratio of J to g is given by $J/g = N\mu_B H_n/[nc_n(N, s)]$. An estimate of this ratio for a given molecule is then obtained by simply inserting a measured value of H_n and the corresponding $c_n(N, s)$ from Fig. 5.5.
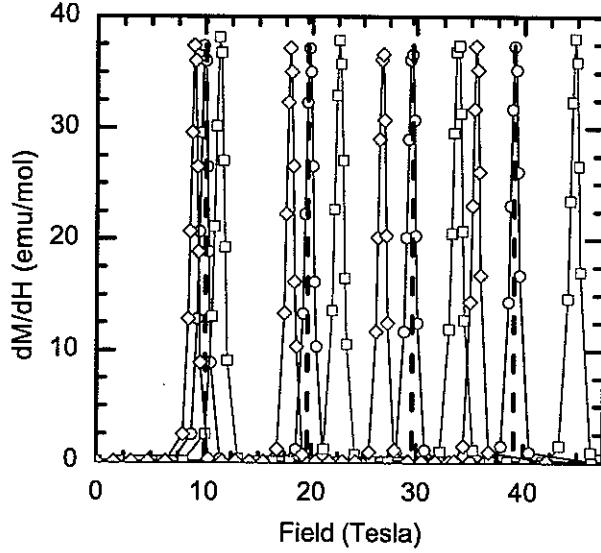
Figure 5.9   The four measured level-crossing fields[21] (dashed vertical
              lines) are compared with the theoretical dM/dH that result
              from $N = 12$, $s = 5/2$ Heisenberg rings with $k_B T/J = 0.01$.
              The theoretical data are shown for the following three choices
              of J and g: $J = 31.9$ K and $g = 2.00$ (diamonds) from Ref. [11];
              $J = 40.7$ K and $g = 2.02$ (squares) from Ref. [21]; $J = 35.2$ K
              and $g = 2.0$ (circles) are our best estimates.

Alternatively, from the measured H_n we can construct an experimental analog of the spec-
tral coefficients by fixing the ratio J/g in Eq. (5.4). In Fig. 5.10, we display those results for
the four measured H_n (and their uncertainties). These data are in good agreement with the
spectral coefficients if we choose the ratio $J/(k_B g) = 17.6$ K, consistent with our previously
stated estimate.

A small decrease with increasing n is observable in the spectral coefficients derived from
the experimental values of the level-crossing fields. This is expected from the data presented
in Sec. 5.2 but more level-crossings and/or smaller experimental error bars are needed in order
to clarify this point. These data are also useful for getting a sense of the typical errors in the
spectral coefficients that were presented in Sec. 5.2. As shown in Fig. 5.10, the error bars of
the QMC data decrease very rapidly with increasing n and are in fact not visible in Figs. 5.5
and 5.7.

Our conclusion is that the existing data for the Fe_{12} molecule are best fit by the choice
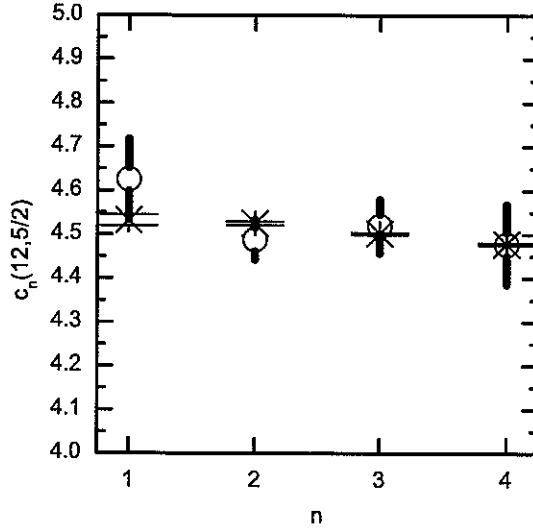
Figure 5.10   The first four spectral coefficients for the $N = 12$, $s = 5/2$ ring
($\times$), along with their errors bars, compared with the results
of inserting the first four level-crossing fields and uncertainties
from Ref. [21] into Eq. (5.4) with the choice, $J/(k_B g) = 17.6$
K ($\circ$).

$g = 2.0$, $J = 35.2$ K. This value of J is 13.5% smaller than the value that resulted[21] from analyzing the experimental level-crossing fields using $c(N,s) = 4$. This reflects the fact that the spectral coefficients, although not constant, exceed 4 by approximately 13%. A similar analysis would be equally straightforward for any other rings whose spectral coefficients are shown in Sec. 5.2.

5.5   Summary

In this paper, we have utilized a quantum Monte Carlo (QMC) method[19, 20] to calculate detailed properties of the general quantum Heisenberg ring. This system consists of an even number N of equally spaced spins s mounted on a ring, where the spins interact via nearest-neighbor antiferromagnetic isotropic exchange, with a single exchange constant J. As this system does not exhibit magnetic frustration, it was possible to calculate thermodynamic quantities down to very low temperatures. In this work our primary focus has been on the accu-

rate determination of the level-crossing fields, which, in turn, directly provide the lowest-energy eigenvalue $E_N(S, s)$ for each value of the total spin quantum number S. By introducing the notation of spectral coefficients [see Eq. (5.4)], denoted by $c_n(N, s)$, we obtained an especially convenient representation of $E_N(S, s)$, given by Eq. (5.5). As the QMC method operates without diagonalizing the Hamiltonian matrix, we were able to obtain results for spins $s = 1/2$, 1, 3/2, 2, 5/2, 3, 7/2, focusing mostly on $N \leq 20$ as these are experimentally relevant, although $N \leq 80$ were also considered. Among our principal results, we have found that the set of level-crossing fields are not uniformly spaced, and thus the spectral coefficients $c_n(N, s)$ vary significantly with n. Equivalently, $E_N(S, s)$ departs from the strictly quadratic dependence on S, referred to as the rotational band approximation[17] (equivalently, the Landé interval rule). These deviations from uniform spacing are fairly small for $N \leq 8$, however, they become increasingly severe with increasing N. Similarly, the ground state energy gap, which may be written as $\Delta_s(N) = E_N(1, s) - E_N(0, s) = c_1(N, s)J/N$, varies significantly with N and s. For $s = 1$, 2 we find that, consistent with the Haldane result,[34] $\Delta_s(N)$ is, indeed, converging to a nonzero limiting gap for large N and in good agreement with estimates in the literature[35, 40] for these two choices of s. By contrast, for $s = 3/2$, 5/2, we find that $\Delta_s(N)$ appears to decrease to zero for large N according to a power law, [see Eq. (5.11)], where the exponent is given by $\alpha \approx 0.76$ for $s = 3/2$ and $\alpha \approx 0.84$ for $s = 5/2$. The increase of α towards unity with increasing s is consistent with the known rigorous result,[41] $\Delta_\infty(N) = 4J/N$, for the classical Heisenberg ring, which may be pictured as the quantum Heisenberg ring in the limit $s \to \infty$. We also find that the departure of the general spectral coefficient $c_n(N, s)$ from the classical result $c_n(N, \infty) = 4$ is characterized by power-law behavior [see Eq. (5.10)]. Finally, we have illustrated the practical utility of our present results for the level-crossing fields and $E_N(S, s)$ by considering the ring-type magnetic molecule known[11] as Fe_{12}. In particular, our analysis of the existing[21] experimental data for level-crossing fields shows that this system can be very well described by the nearest-neighbor Heisenberg model with antiferromagnetic exchange, and we are able to provide a new and improved estimate of the exchange constant. Although only rings with even N have been considered in this work, we suspect that similar

scaling relations may hold for other structures as well.

Acknowledgements

Bibliography

[1] R. Sessoli, D. Gatteschi, A. Caneschi, and M. Novak, Nature **365** (1993).

[2] D. Gatteschi, A. Caneschi, L. Pardi, and R. Sessoli, Science **265**, 1054 (1994).

[3] R. E. P. Winpenny, Adv. Inorg. Chem. **52**, 1 (2001).

[4] R. E. P. Winpenny, Dalton Trans. **1**, 1 (2002).

[5] R. E. P. Winpenny, Comp. Coord. Chem. II **7**, 125 (2004).

[6] K. Taft, C. Delfs, G. Papaefthymiou, S. Foner, D. Gatteschi, and J. Lippard, J. Am. Chem. Soc. **116**, 823 (1994).

[7] G. L. Abbati, A. Cornia, A. C. Fabretti, W. Malavasi, L. Schenetti, A. Caneschi, and D. Gatteschi, Inorg. Chem. **36**, 6443 (1997).

[8] R. W. Saalfrank, I. Bernt, E. Uller, and F. Hampel, Angew. Chem. Int. Ed. **36**, 2482 (1997).

[9] S. P. Watton, P. Fuhrmann, L. E. Pence, A. Cornia, G. L. Abbati, and S. J. Lippard, Angew. Chem. **36**, 2774 (1997).

[10] I. M. Atkinson, C. Benelli, M. Murrie, S. Parsons, and R. E. P. Winpenny, Chem. Comm. pp. 285–286 (1999).

[11] A. Caneschi, A. Cornia, A. Fabretti, and D. Gatteschi, Angew. Chem. Int. Ed. **38**, 1295 (1999).

[12] G. L. Abbati, A. Caneschi, A. Cornia, A. C. Fabretti, and D. Gatteschi, Inorg. Chim. Acta **297**, 291 (2000).

[13] S.-X. Liu, S. Lin, B.-Z.Lin, C.-C. Lin, and J.-Q. Huang, Angew. Chem. **113**, 1118 (2001).

[14] E. J. L. McInnes, C. Anson, A. K. Powell, A. J. Thomson, and S. P. amd R. Sessoli, Chem. Comm. pp. 89–90 (2001).

[15] C.-H. Chang, K. C. Hwang, C.-S. Liu, Y. Chi, A. J. Carty, L. Scoles, S.-M. Peng, G.-H. Lee, and J. Reedijk, Angew. Chem. **113**, 4787 (2001).

[16] A. L. Dearden, S. Parsons, and R. E. P. Winpenny, Angew. Chem. Int. Ed. **40**, 151 (2001).

[17] J. Schnack and M. Luban, Phys. Rev. B **63**, 14418 (2000).

[18] O. Waldmann, Phys. Rev. B **65**, 024424 (2001).

[19] A. W. Sandvik and J. Kurkijarvi, Phys. Rev. B **43**, 5950 (1991).

[20] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002).

[21] Y. Inagaki, T. Asano, Y. Ajiro, Y. Narumi, K. Kindo, A. Cornia, and D. Gatteschi, J. Phys. Soc. Jap. **75**, 1178 (2003).

[22] M. Julien, Z. Jang, A. Lascialfari, F. Borsa, M. Horvatic, A. Caneschi, and D. Gatteschi, Phys. Rev. Lett. **83**, 227 (1999).

[23] H. J. Schmidt, J. Phys. A: Math. Gen. **35**, 6545 (2002).

[24] R. B. Griffiths, Phys. Rev. **133**, A768 (1964).

[25] H. A. Bethe, Z. Phys **71**, 205 (1931).

[26] C. N. Yang and C. P. Yang, Phys. Rev. **150**, 321 (1966).

[27] J. Bonner and M. E. Fisher, Phys. Rev. **135**, A640 (1964).

[28] V. A. Kashurnikov, N. V. Prokof'ev, B. V. Svistunov, and M. Troyer, Phys. Rev. B **59**, 1162 (1999).

[29] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[30] X. Wang and T. Xiang, Phys. Rev. B **56**, 5061 (1997).

[31] M. P. Nightingale, V. S. Viswanath, and G. Müller, Phys. Rev. B **48**, 7696 (1993).

[32] I. Affleck, D. Gepner, H. J. Schulz, and T. Ziman, J. Phys. A: Math. Gen. **22**, 511 (1989).

[33] S. Eggert, I. Affleck, and M. Takahashi, Phys. Rev. Lett. **73**, 332 (1994).

[34] F. D. M. Haldane, Phys. Rev. Lett. **50**, 1153 (1983).

[35] S. Qin, Y.-L. Liu, and L. Yu, Phys. Rev. B **55**, 2721 (1997).

[36] O. Golinelli, T. Jolicoeur, and R. Lacaze, Phys. Rev. B **50**, 3037 (1994).

[37] J. B. Parkinson and J. C. Bonner, Phys. Rev. B **32**, 4703 (1985).

[38] J. Lou, S. Qin, T.-K. Ng, and Z. Su, Phys. Rev. B **65**, 104401 (2002).

[39] J. Schnack, Phys. Rev. B **62**, 14855 (2000).

[40] X. Wang, S. Qin, and L. Yu, Phys. Rev. B **60**, 14529 (1999).

[41] H. J. Schmidt and M. Luban, J. Phys. A: Math. Gen. **36**, 6351 (2003).

CHAPTER 6.   Finite quantum Heisenberg spin models and their approach to the classical limit

L. Engelhardt,[1] C. Schröder,[2] and M. Luban[1]

Abstract

We determine the temperature range over which classical Heisenberg spin models closely reproduce the zero-field susceptibility of the corresponding quantum Heisenberg models for a finite number N of interacting quantum spins s. Using mostly quantum and classical Monte Carlo methods, as well as analytical methods where applicable, we have explored a variety of geometries, including polygons, open chains, and all Platonic and several Archimedean polytopes. These systems range in size from $N = 2$ to 120, and we have considered values of s from 1/2 to 50 for both antiferromagnetic and ferromagnetic exchange. Particular attention is devoted to quantifying the slow convergence of the large s quantum data to the limiting classical data. This is motivated by the desire to define conditions where classical Monte Carlo methods can provide useful predictions for finite quantum Heisenberg spin systems.

6.1   Introduction

In the treatment of magnetic systems, a classical spin model is frequently used, though the microscopic moments are actually quantum in character. For example, this practice is

---

[1] *Department of Physics and Astronomy & Ames Laboratory, Iowa State University, Ames, Iowa 50011, USA*
[2] *Department of Electrical Engineering and Computer Science, University of Applied Sciences Bielefeld, D–33602 Bielefeld, Germany and Ames Laboratory, Ames, Iowa 50011, USA*

sometime followed in the field of magnetic molecules[1] in order to circumvent severe computational difficulties that arise in a quantum treatment. A classical model can, at times, be both conceptually illuminating and provide sufficiently accurate results. However, the accuracy of a classical model's results are generally not known, when compared to those for the quantum spins that are being represented. For that reason, the goal of the present work is to determine the circumstances under which classical models will provide a good approximation to finite systems of interacting quantum spins. In so doing, we are not only able to provide clear "rules of thumb" for specific systems; we also explore the approach to the classical limit for large values of intrinsic spin s. In order to ensure the accuracy of these comparisons, we have compared only quantum *model systems* and their classical counterparts; we have not included comparisons with experimental data, whose underlying Hamiltonian may be in question.

To this end, we have performed calculations using the Heisenberg model, describing a finite number N of interacting spins. Within this model, there are clearly a number of properties that one could calculate when trying to answer the loosely defined question, "When does a classical model accurately simulate the corresponding quantum model?". To select a suitable metric, and hence clarify this question, there are two requirements that we impose. (i) We consider a property that is of relevance to experiments, and (ii) the property, for simplicity reasons, depends on only one physically relevant parameter, such that two regimes exist—In one regime, the classical and quantum models give results that coincide to within some small prescribed error, while in the other regime they produce significantly different results. Both requirements are met by choosing to calculate the zero-field susceptibility χ as a function of temperature T and determining the minimum temperature $T_m(s)$ above which the classical and quantum results for $\chi(T)$ meet a prescribed accuracy criterion for a given choice of s. The details of this correspondence are described in Sec. 6.2.1.

In order to explore the dependence of T_m on s, we have performed calculations for a variety of systems ranging in size from $N = 2$ to 120 sites, with intrinsic spin quantum numbers ranging from $s = 1/2$ to, in some cases, as large as $s = 50$. The classical counterpart to each quantum system is, of course, also considered, as we describe in Sec. 6.2.1. Some of these systems,

specifically with $s \leq 5/2$, are relevant to the study of magnetic molecules.[1] Additionally, the large s calculations allow us to glean valuable information regarding the approach to the classical limit; and by studying many different sizes and geometries, we are able to judge the universality of our conclusions.

These calculations, however, pose a considerable challenge, necessitating the use of a variety of methods. For many existing magnetic molecules, matrix diagonalization of the Hamiltonian is not feasible due to the very large dimensionality of the Hilbert spaces, given by $D = (2s+1)^N$. For example, {Fe$_{12}$} ($s = 5/2$ and $N = 12$)[2] has $D \approx 2.2 \times 10^9$, which is already pushing the current limitations of Lanczos diagonalization, while {Mo$_{72}$Fe$_{30}$} ($s = 5/2$ and $N = 30$)[3] has an associated $D \approx 2.2 \times 10^{23}$. Instead we use classical and quantum Monte Carlo methods that allow us to circumvent the obstacles that large Hilbert spaces and complex multidimensional integrals pose to quantum and classical calculations, respectively. We review these methods in Sec. 6.2.3. This is preceded in Sec. 6.2.2 with a discussion of the few special cases where the spectra of energy eigenvalues can be derived analytically, and matrix diagonalization is unnecessary.

In Sec. 6.3, we present our results for a wide variety of geometries including spin chains, polygons, and a number of Platonic and Archimedean polytopes, where, for each polytope, each vertex represents a spin site and each edge connects a pair of interacting spins. We consider both antiferromagnetic (AFM) and ferromagnetic (FM) interactions, finding that for AFM interactions a classical approximation is typically valid down to a substantially lower temperature T_m than for the corresponding FM case. Furthermore, we present numerical values of T_m and explore its functional dependence on s. Finally, in Sec. 6.4 we summarize our findings.

6.2 Models and methods

6.2.1 Quantum and classical spin models

In order to compare the results of classical and quantum spin models, we use as our starting point the quantum Heisenberg model. We assume all spins to share a common quantum

number s, and further assume that if a distinct pair of spins (represented $\langle i, j \rangle$) interact with one another, they do so with with the same strength J_s as any other pair of interacting spins. This Hamiltonian is then given by[3]

$$\mathcal{H} = J_s \sum_{\langle i,j \rangle} \vec{\tilde{s}}_i \cdot \vec{\tilde{s}}_j + g\mu_B \vec{H} \cdot \sum_{i=1}^{N} \vec{\tilde{s}}_i, \tag{6.1}$$

where the spin operators $\vec{\tilde{s}}_i$ are given in units of $\hbar$, g is the spectroscopic splitting factor, μ_B is the Bohr magneton, and $\vec{H}$ is an external magnetic field. Following Fisher,[4] the classical analog to this quantum Hamiltonian is constructed by first defining rescaled spin operators $\vec{\tilde{e}}_i(s) \equiv \vec{\tilde{s}}_i / \sqrt{s(s+1)}$, in the nature of unit vector operators, whose components satisfy the commutation relation $\tilde{e}_i^x(s)\tilde{e}_i^y(s) - \tilde{e}_i^y(s)\tilde{e}_i^x(s) = i\tilde{e}_i^z(s)/\sqrt{s(s+1)}$ and the cyclic permutations thereof. With this replacement, and introducing new parameters, $J_c \equiv J_s s(s+1)$ and $\mu_c \equiv g\mu_B \sqrt{s(s+1)}$, Eq. (6.1) is rewritten as

$$\mathcal{H} = J_c \sum_{\langle i,j \rangle} \vec{\tilde{e}}_i(s) \cdot \vec{\tilde{e}}_j(s) + \mu_c \vec{H} \cdot \sum_{i=1}^{N} \vec{\tilde{e}}_i(s), \tag{6.2}$$

allowing us to consider a *sequence* of Hamiltonians (6.2) for each geometry, where the members of a sequence differ *only* in s, sharing a common value of J_c. Since the commutators between the $\vec{\tilde{e}}_i(s)$ vanish in the limit of very large s, each of these sequences converge with increasing s toward a classical Hamiltonian,[4]

$$\mathcal{H} = J_c \sum_{\langle i,j \rangle} \vec{e}_i \cdot \vec{e}_j + \mu_c \vec{H} \cdot \sum_{i=1}^{N} \vec{e}_i, \tag{6.3}$$

where all $\vec{\tilde{e}}_i$ operators are replaced by *classical* unit vectors $\vec{e}_i$.

At sufficiently high temperatures, Eqs. (6.2) and (6.3) give the same values of $\chi(T)$, while at low temperatures the values of $\chi(T)$ diverge from one another. Specifically, quantum systems that have an $S = 0$ ground state give $\chi(0) = 0$, with $\chi(T) \propto \exp(-\Delta/k_B T)$ for sufficiently small T, where Δ is the energy gap between the ground state and the lowest $S > 0$ excited state. Classical models, by contrast, do not have a corresponding gap in their energy spectra. As a result, an infinitesimal external magnetic field would induce a proportional (infinitesimal)

---

[3]Tildes, written below symbols, are used to distinguish quantum operators from their classical counterparts.

[4]Operators and c-numbers are of course distinct mathematical objects. However, in the limit $s \to \infty$, Eq. (6.2) is equivalent to Eq. (6.3) in the sense that they both yield the same energy spectra and thermodynamic data.

magnetization, so Eq. (6.3) gives values of $\chi(T)$ that are nonzero in the limit $T \to 0$. For a system whose ground state has $S > 0$, $\chi(T \to 0)$ diverges proportional to $1/T$, as we describe in Sec. 6.3.2.

With this high temperature correspondence (and deviation of classical and quantum results at low temperatures) in mind, we are now prepared to explicitly state the question to be answered in this article: "What is the minimum temperature $T_m(s)$ for which Eqs. (6.2) and (6.3) produce the same zero-field susceptibility $\chi(T)$ to within some predetermined factor, e.g., 2%?" This question is answered in detail in Sec. 6.3 using the methods of Secs. 6.2.2 and 6.2.3, where in all cases $\chi(T)$ is obtained using the fluctuation formula $\chi(T) = (g^2 \mu_B^2 / 3k_B T)\langle S^2 \rangle_T$, where $\langle S^2 \rangle_T$ denotes the ensemble average of S^2.

6.2.2 Analytical methods

For a few special geometries, it is possible to calculate the zero-field energy eigenvalues E analytically by expressing them in terms of the total spin quantum number S and additional spin quantum numbers described below. Given these energies, if one is able to calculate the degeneracy ν_E for each E, it is then straightforward to calculate the zero-field susceptibility. These special geometries can be grouped into two categories. The first category includes systems in which all spins interact equally with all other spins (sometimes referred to as "pantahedra"). In this case, the energies are given by

$$E(S, s) = \frac{J_s}{2}[S(S+1) - Ns(s+1)]. \tag{6.4}$$

It is also possible to write down the energies in the more general case that the spin lattice can be decomposed into two or more sublattices, such that each spin of a given sublattice (whose total spin is labeled S_A, S_B, etc.) interacts with the same strength with *all* spins of the other sublattices, but *no* spins of its own sublattice. To illustrate this, we consider two special cases, whose resulting energy spectra are very similar to Eq. (6.4). For a ring of four spins with nearest-neighbor interactions only, there are two sublattices, each of which consists of a pair of spins that do not interact with one another. In terms of these two sublattices, the energy

spectrum is given by

$$E(S, S_A, S_B) = \frac{J_s}{2}[S(S+1) - S_A(S_A+1) - S_B(S_B+1)], \tag{6.5}$$

while for the octahedron there are three sublattices, resulting in

$$E(S, S_A, S_B, S_C) = \frac{J_s}{2}[S(S+1) - S_A(S_A+1) - S_B(S_B+1) - S_C(S_C+1)]. \tag{6.6}$$

The calculation of $\chi(T)$ for these systems then reduces to enumerating all of the ways in which the individual spins of each sublattice can couple together to yield a given energy, hence providing the values of ν_E. Using these data, we are able to calculate $\chi(T)$ for very large values of s and D, which are well beyond the current limitations of matrix diagonalization. To compare with classical data, we use the results of Ref. [5], as well as additional results that have been derived for the classical $N = 5$ pantahedron.[6]

6.2.3   Monte Carlo methods

For both the quantum and the classical model Hamiltonians that were introduced in Sec. 6.2.1, the analytic calculation of $\chi(T)$ is not feasible for general geometries of spins. In both cases, this calculation involves an ensemble average which becomes very challenging with increasing N. For quantum spins, the value of the Hilbert space dimensionality D can become so large ($\gtrsim 10^7$) that it is impractical to attempt to compute all of the energy eigenvalues, while the classical versions often involve intractable $3N$-dimensional integrals.

Both classical and quantum Monte Carlo methods exploit the following idea. Instead of seeing it as a hinderance that one cannot include the contributions from all of the quantum eigenstates and all of the classical phase space, one can instead recognize that it is often *unnecessary* to include all of the states and use this to our advantage. When the excitation energy of a state is large relative to the thermal energy $k_B T$, the state will not contribute significantly to the ensemble average, and can hence be ignored. In this spirit, one can use the method of importance sampling to perform the necessary averaging, whereby the states that make large contributions are accurately sampled, and the states that make negligible contributions are ignored. (See, for example, Ref. [7].) While there are statistical errors involved in

this approach, the errors can be made arbitrarily small by performing longer sampling runs. Very importantly, no systematic errors are introduced, and the statistical errors are accurately estimated during the course of a calculation.

For the quantum calculations, we use the method that was introduced in Ref. [8] and has been recently used to study similar finite systems.[9, 10] The idea of this method is that a high-temperature expansion can be used to express the partition function in terms of two-spin matrix elements, which are easily evaluated.[11] The trade-off is that, in order to consider the full range of temperatures, many complicated terms must be sampled, and the Monte Carlo updating becomes very involved. However, this updating can be efficiently performed using the so-called "directed-loop" equations[8] to simultaneously satisfy detailed balance, and ergodically sample the relevant phase space. For the classical Monte Carlo calculations, importance sampling is carried out using the standard Metropolis method.[12]

Using these methods we have considered a great variety of geometries in Sec. 6.3, with sizes ranging up to $N = 120$. For this largest system, we considered quantum spins up to $s = 9/2$, with dimension $D = 10^{120}$. However, as impressive as this is, the quantum Monte Carlo (QMC) method also has a serious limitation. For classical systems that have frustrated ground states,[13] the quantum analogs suffer from the infamous negative sign problem,[14] and results can only be obtained for relatively high temperatures.[9] As a result, for such frustrated systems QMC calculations are *sometimes* able to provide a more complete description of experimental data than one would obtain from classical Monte Carlo. In other situations, QMC calculations are limited to temperatures that are greater than T_m, in which case QMC offers no additional information beyond that which is given by classical Monte Carlo. This issue is addressed in greater detail in the Appendix.

6.3   Results

6.3.1   AFM interactions

In this section we present and discuss $\chi(T)$ and $T_m(s)$ data for systems of spins that interact via AFM interactions. As was described in Sec. 6.2.1, the quantity $T_m(s)$ is defined *only* in

terms of some prescribed discrepancy, which we choose here to be a 2% difference between the values of χ for quantum spins s and classical spins. This value is chosen simply because a 2% difference is just visible to the eye; choosing a different number, such as 5%, does in fact lead to the same conclusions, only with a different numerical prefactor.
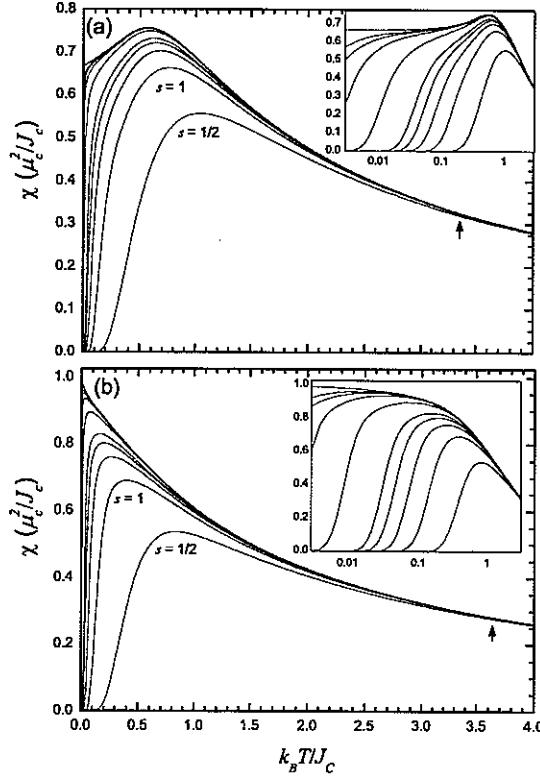

Figure 6.1   Zero-field susceptibility in units of μ_c^2/J_c for the AFM (a) square and (b) tetrahedron. In both cases, results are shown for intrinsic spins $s = 1/2$, 1, $3/2$, 2, $5/2$, 5, 15, 20 and classical spins. The $s = 1/2$ and $s = 1$ curves are labeled, and the larger values of s proceed systematically toward the classical results, which are the uppermost curves. In the inset the same data are plotted with a logarithmic temperature scale, and, as described in the text, the arrows indicate where a 2% difference is reached between the $s = 1/2$ and $s = \infty$ data.

In Fig. 6.1 we show $\chi(T)$ for the square and tetrahedron with quantum spins ranging from $s = 1/2$ to $s = 20$ as well as classical spins. For values of $s > 5/2$ we display only the results for $s = 5$, 10, 15 and 20, but calculations were performed for all $s \leq 50$, and we observe the

same trends for all values of s. Note that in both plots an arrow indicates the temperature at which a 2% difference occurs between the classical and the $s = 1/2$ data, hence providing the corresponding values of $T_m(s = \frac{1}{2})$ in terms of J_c. For both geometries we find that $k_B T_m(\frac{1}{2})/J_c \approx 3.5$, and the values of $k_B T_m(s)/J_c$ clearly decrease toward zero with increasing s. These same trends were also found for the other structures described in Sec. 6.2.2, and thus the corresponding $\chi(T)$ data are not shown.
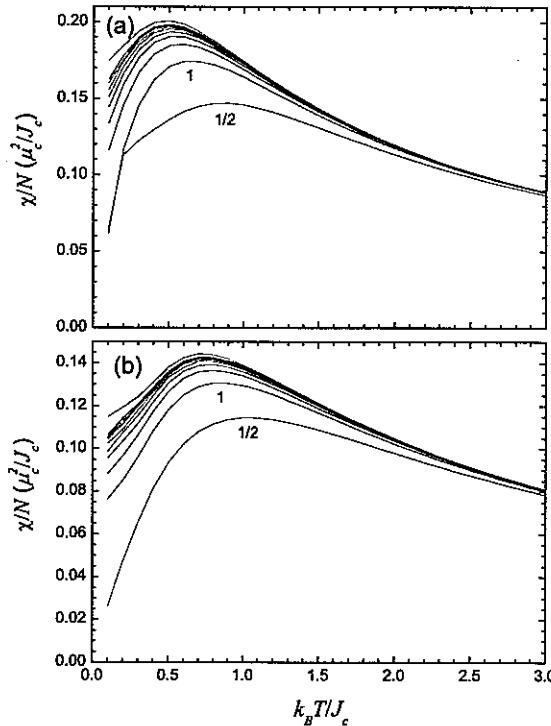
Figure 6.2   Zero-field susceptibility per spin in units of μ_c^2/J_c for the AFM (a) $N = 20$ ring and (b) $N = 120$ Archimedean solid (great rhombicosidodecahedron). In both cases, results are shown for intrinsic spins $s = 1/2$, 1, $\ldots$, $9/2$ and classical spins. The $s = 1/2$ and $s = 1$ curves are labeled, and the larger values of s proceed systematically toward the classical results, which are the uppermost curves.

In Fig. 6.2 is shown the susceptibility per spin, $\chi(T)/N$, obtained using Monte Carlo methods, for a ring of $N = 20$ spins and an Archimedean solid of $N = 120$ spins, with s ranging from $1/2$ to $9/2$ as well as classical data. In both Figs. 6.1 and 6.2, the classical limit is indeed

being approached with increasing s, but only slowly. Specifically, for $s \approx 5$ the quantum and classical results differ noticeably when $k_B T \lesssim J_c/2$; and in Fig. 6.1 the $s = 20$ results clearly diverge from the classical curves for $k_B T \lesssim 0.01 J_c$. We seek to quantify this approach to the classical limit, and have thus determined $k_B T_m(s)/J_s$, which equals $s(s+1)k_B T_m/J_c$, for many geometries.[5] One could instead analyze $k_B T_m(s)/J_c$, but, as we will show, by plotting $k_B T_m(s)/J_s$ the functional dependence of T_m on s is easily extracted.
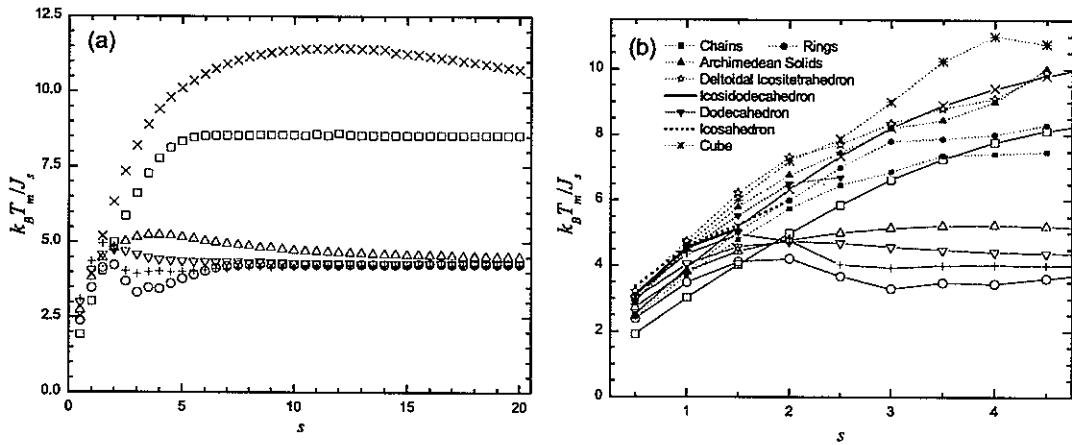

Figure 6.3 T_m vs. s for systems with AFM interactions. In (a) data are included up to $s = 20$ for the dimer ($\square$), triangle ($\circ$), square ($\times$), tetrahedron ($\triangle$), $N = 5$ pantahedron (∇), and octahedron ($+$). The same symbols also apply in (b), where data are plotted for all of the geometries that have been considered. As a guide to the eye, lines connect successive data points. For additional details regarding these data, see the Appendix.

In Fig. 6.3(a), we plot $k_B T_m(s)/J_s$ for the geometries described in Sec. 6.2.2, including values of s extending up to $s = 20$. For all of these structures, there are clearly two distinct regimes of s values. For small s ($\lesssim 5/2$) the data increase linearly with s, and for large s ($\gtrsim 5$) the $k_B T_m(s)/J_s$ data saturate. In the intermediate range, $5/2 \lesssim s \lesssim 5$, the behavior is crossing-over between the two limiting cases. [For the square, larger values of s ($\gtrsim 40$) are

---

[5]Descriptions of the geometries, as well as some technical details regarding the results for different geometries, are contained in the Appendix.

needed to reach saturation, but $k_B T_m(s)/J_s$ eventually approaches a constant value, ≈ 9.0.] These data immediately imply that, although $k_B T_m(s)/J_c$ decreases monotonically with s, its functional dependence is different for large and small s: For small s, $k_B T_m(s)/J_s \propto s$, or equivalently, $k_B T_m(s)/J_c \propto 1/(s+1)$; while for large s, $k_B T_m(s)/J_c \propto 1/s^2$.

In order to formulate a useful rule of thumb, we must include not only the s-dependence, but also the proportionality constants. For large s there are clearly two proportionality constants that are relevant for the geometries shown in Fig. 6.3(a). For the dimer and square $k_B T_m/J_s \approx 9$, while $k_B T_m/J_s \approx 4.5$ for the pantahedra and octahedron. In Fig. 6.3(b) we have included data for $s < 5$ for many additional structures, calculated using the Monte Carlo methods described in Sec. 6.2.3. Again, we find that in *all* cases, $k_B T_m/J_s$ increases linearly with s for small s, and then approaches saturation for larger s. For the pantahedra and octahedron this saturation is already reached by $s \approx 2$, while $k_B T_m/J_s$ continues to increase for the other structures. In fact, other than the pantahedra and octahedron, all of this data for $s \leq 3$ can be described to within 30% by the function $k_B T_m(s)/J_s \approx 2(s+1)$, giving the very simple rule of thumb, $k_B T_m/J_c \approx 2/s$ for this interval of s.

6.3.2 FM interactions

The same systems (described in the Appendix) but with FM interactions were also studied in the same manner as was described for AFM systems. There are, however, some distinct differences between the AFM and FM results. When all interactions are FM, the ground states of the classical systems are realized by aligning all spins; similarly, the ground states of the quantum systems have $S = Ns$. As a result, $\chi(T)$ diverges at low T proportional to $1/T$ in all cases, so instead of plotting $\chi(T)$, it is more instructive to plot $T\chi(T)$ as is shown in Fig. 6.4. Moreover, the limiting $T \to 0$ values of $T\chi$, henceforth referred to as $(T\chi)_0$, converge with increasing s to the classical result.[6] Consequently, there will be some value of $s \equiv s^*$ which, when exceeded, will give results that differ by less than 2% even at $T = 0$. Therefore, $T_m = 0$

---

[6]For FM exchange, the ground state has $S = Ns$, and the values of $(T\chi)_0$ are easily calculated from the fluctuation formula for the cases of both classical and quantum spins. For quantum spins, $(T\chi)_0 = \frac{\mu_c^2}{3k_B} \frac{N(Ns+1)}{s+1}$; and for classical spins, $(T\chi)_0 = \frac{\mu_c^2}{3k_B} N^2$.

for $s > s^*$, so for large s we do not obtain the AFM result, i.e., $k_B T_m(s)/J_c \propto 1/s^2$. Note however that the value of s^* is typically quite large (e.g., $s^* = 45$ for $N = 10$), and there is still a large region of s values, shown in Fig. 6.5(a), for which $k_B T_m/J_s$ is independent of s, or equivalently, $k_B T_m(s)/J_c \propto 1/s^2$.
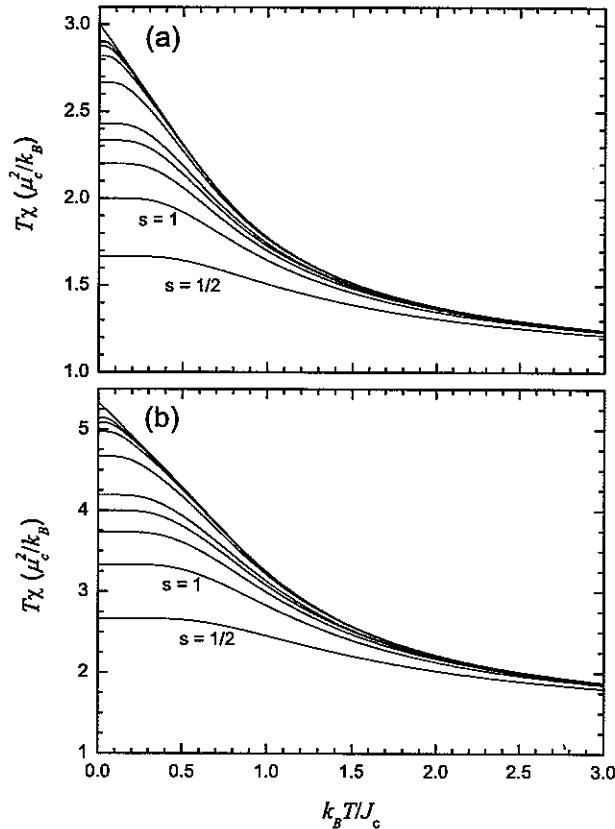
Figure 6.4   The product of temperature times zero-field susceptibility per spin in units of μ_c^2/k_B for the FM (a) icosahedron and (b) icosi-dodecahedron. In both cases, results are shown for intrinsic spins $s = 1/2, 1, \ldots, 9/2$ and classical spins. The $s = 1/2$ and $s = 1$ curves are labeled, and the larger values of s proceed systematically toward the classical result, which are the uppermost curves.

As shown Fig. 6.5, $k_B T_m/J_s$ indeed decreases for large s as s^* is approached, but the data are nonetheless quite similar to those obtained for the AFM systems. For $s \lesssim 5$, $k_B T_m/J_s$ is increasing, and this increase is again linear with s. The most striking feature of Fig. 6.5 is

perhaps that $k_B T_m / J_s$ can be quite large, exceeding 30 for the case of the $N = 5$ pantahedron and the octahedron, and exceeding 10 for almost all of the systems when $s \gtrsim 5/2$. There is a very large variation in the data between the geometries, but the median values of T_m can be roughly described by $k_B T_m / J_s \approx 5s$ for $s \lesssim 5$. Comparing with the AFM rule of thumb, for small s these values of T_m are similar to those found for AFM interactions, but for large s, T_m is much larger when the interactions are FM.
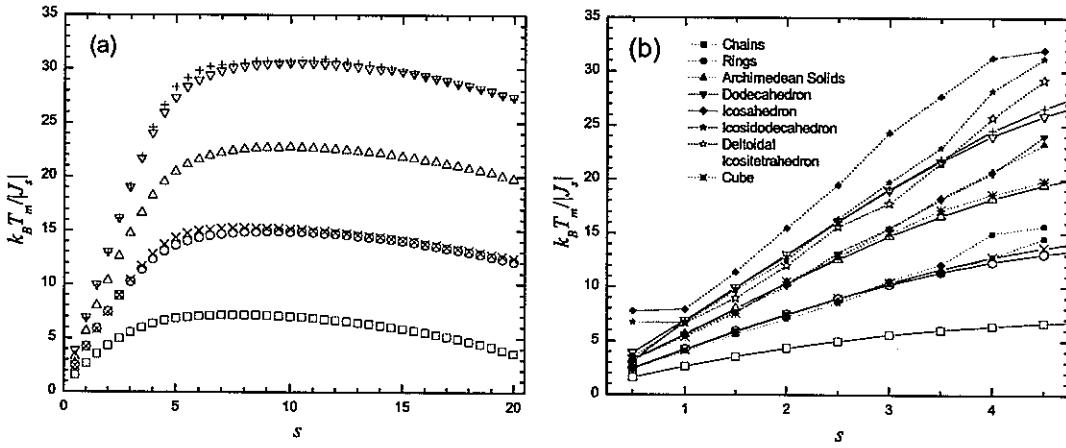

Figure 6.5 T_m vs. s for systems with FM interactions. In (a) data are included up to $s = 20$ for the dimer ($\square$), triangle ($\circ$), square ($\times$), tetrahedron ($\triangle$), $N = 5$ pantahedron (∇), and octahedron ($+$). The same symbols also apply in (b), where data are plotted for all of the geometries that have been considered. As a guide to the eye, lines connect successive data points. For additional details regarding these data, see the Appendix.

6.4   Summary

In this article we have utilized a combination of quantum Monte Carlo, classical Monte Carlo, and analytical methods to study and quantify how, with increasing intrinsic spin s, the results of quantum Heisenberg model systems approach the results of the corresponding classical Heisenberg models. To this end, the zero-field susceptibility was calculated for many

geometries, and from these data we have extracted the minimum temperature $T_m(s)$ at which the quantum (spin s) results differ by less that 2% from the corresponding classical results. In terms of the exchange constants J_s and J_c (defined in Sec. 6.2.1) we have found for small s that $k_B T_m/J_s \propto s$, or equivalently $k_B T_m/J_c \propto 1/(s+1)$, while for large s we found that $k_B T_m/J_s$ saturates, and thus $k_B T_m/J_c \propto 1/s^2$. In particular, for almost all of the AFM systems that we have studied, T_m can be described to within 30% by the rule-of-thumb $k_B T_m/J_c \approx 2/s$, which is valid for $s \lesssim 3$ and should therefore be relevant to the analysis of experimental susceptibility data. When FM interactions are present we have found that, to within a factor of 2, the rule-of-thumb $k_B T_m/J_c \approx 5/(s+1)$ applies for $s \lesssim 5$. More precise results have been obtained for many specific systems and can be found by inspecting Figs. 6.3 and 6.5.

The application of these results to the analysis of any real system of interacting quantum spins would be straightforward, provided one had some estimate of the relevant energy scale J_s. For example, inspection of Fig. 6.3(b) shows that a classical model will accurately reproduce the results of quantum spins $s = 3/2$ for temperatures $T \gtrsim (5 \pm 1)J_s/k_B$. Additionally, the present large s results [shown in Fig. 6.3(a)] also underscore the fact that a huge value of s does *not* automatically imply that a classical spin model is valid. For example, a spin triangle with an exchange constant $J_s/k_B = 10$ K could accurately be described by a classical model only for $T \gtrsim 40$ K, even if $s = 50$. Finally, we remark that these results are only meant to provide information about *static* properties. Obtaining classical and quantum time-correlation functions and comparing them in an analogous manner should provide interesting and useful results and would be a worthy avenue for future study.

Acknowledgements

Appendix: Description of geometries

The various geometries studied here can almost all be placed into one of three general categories: Open chains, closed rings (polygons), or three-dimensional polytopes. For the chains and rings, we find that with increasing N the results for $T_m(s)$ rapidly converge to a single curve that is essentially independent of N and is valid for both even and odd N. The data labeled "chains" and "rings" in Figs. 6.3(b) and 6.5(b) are hence averages over these data with N varying from 10 to 20. (For the classical open chains, the exact susceptibility has long been known,[4] and these exact results are used here.) The category of chains also includes the simplest possible case, which is one pair of interacting spins (dimer), while the rings include the triangle and square. Data for these three systems each appear separately in Figs. 6.3 and 6.5.

The polytopes that were studied include four of the 13 Archimedean solids (AS) and all five Platonic solids (PS). (A complete description of all of these polytopes can be found in Ref. [15].) Of these AS, three of the structures were chosen because they are the only non-frustrated AS that exist (i.e., they are bipartite lattices), and hence allow QMC calculations to proceed down to arbitrarily low T, even for AFM exchange. They are the truncated octahedron ($N = 24$), great rhombicuboctahedron ($N = 48$), and great rhombicosidodecahedron ($N = 120$). They also share an additional property: As is the case with large N chains and rings, these three AS produce the same values of $T_m(s)$, independent of N, so their averages also appear in Figs. 6.3(b) and 6.5(b), labeled "Archimedean Solids".

Recall from Sec. 6.3.2 that s^* depends on N, which implies for $s \lesssim s^*$ that T_m must also depend on N. For $s < 5$ however, we find that the variation of T_m with N is always smaller than the associated error-bars for the chains, rings, and non-frustrated AS. For small s, these error-bars are smaller than the symbols that appear in Figs. 6.3(b) and 6.5(b), while for $s = 9/2$ the uncertainties in $k_B T_m / J_s$ are roughly ± 1. In Figs. 6.3(a) and 6.5(a), the errors are considerably smaller than the associated symbols for all s.

The remaining AS that we studied is the icosidodecahedron, whose structure is adopted by multiple species of magnetic molecules.[3, 9, 16, 17] For this geometry, AFM interactions give rise to a classical ground state configuration that is frustrated,[18] so the QMC calculations

are restricted to relatively high T.[14] Specifically, these QMC calculations for $T < T_m$ were possible only with $s = 1/2$, 1, and 3/2. Frustration also occurs for four of the five PS, the cube being the one exception. Two of the PS, the tetrahedron and octahedron, although being frustrated, were calculated using the method of Sec. 6.2.2; while for the other two PS, the icosahedron and dodecahedron, QMC calculations were used. For the icosahedron, we were able to proceed to temperatures below T_m for $s \leq 2$, and, for the dodecahedron, could handle all $s \leq 5/2$. Consequently, data for larger s do not appear in 6.3(b) for the icosidodecahedron, icosahedron, and dodecahedron.

There are two other geometries for which $\chi(T)$ and T_m have also been calculated: The $N = 5$ pantahedron and the deltoidal icositetrahedron.[15] The pantahedra were described in Sec. 6.2.2, and the deltoidal icositetrahedron is a polytope which is unfrustrated and of lower symmetry than the PS or AS. For both of these structures, the resulting $k_B T_m / J_s$ data appear in Figs. 6.3(b) and 6.5(b). In order to obtain the T_m data that have been presented here, a great deal of additional $\chi(T)$ data were obviously calculated which have not been shown; these data are available from the authors for all of the structures.

Bibliography

[1] A recent useful survey of magnetic molecules is given in R. E. P. Winpenny, Comp. Coord. Chem. II **7**, 125 (2004).

[2] A. Caneschi, A. Cornia, A. Fabretti, and D. Gatteschi, Angew. Chem. Int. Ed. **38**, 1295 (1999).

[3] A. Müller, M. Luban, C. Schröder, R. Modler, P. Kögerler, M. Axenovich, J. Schnack, P. Canfield, S. Bud'ko, and N. Harrison, ChemPhysChem **2**, 517 (2001).

[4] M. E. Fisher, Am. J. Phys. **32**, 343 (1964).

[5] O. Ciftja, M. Luban, M. Auslender, and J. H. Luscombe, Phys. Rev. B **60**, 10122 (1999).

[6] M. Luban, unpublished.

[7] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2000).

[8] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002).

[9] A. Müller, A. Todea, J. van Slageren, M. Dressel, H. Bögge, M. Schmidtmann, M. Luban, L. Engelhardt, and M. Rusu, Angew. Chem. Int. Ed. **44**, 3857 (2005).

[10] L. Engelhardt and M. Luban, Phys. Rev. B **73**, 054430 (2006).

[11] D. C. Handscomb, Proc. Cambridge Philos. Soc. **60**, 115 (1964).

[12] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[13] R. Moessner and A. P. Ramirez, Physics Today **59(2)**, 24 (2006).

[14] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[15] URL http://mathworld.wolfram.com.

[16] B. Botar, P. Kögerler, and C. L. Hill, Chem. Commun. pp. 3138–3140 (2005).

[17] B. Botar, P. Kögerler, and C. L. Hill, J. Am. Chem. Soc. Comm. **128**, 5336–5337 (2006).

[18] M. Axenovich and M. Luban, Phys. Rev. B **63**, 100407(R) (2001).

CHAPTER 7. Applications to theoretically challenging magnetic molecules

In this chapter we present our analysis of four different species of recently synthesized magnetic molecule systems. Each of these systems is theoretically challenging for methods that rely on matrix diagonalization, due to the large sizes of their Hilbert spaces D. Specifically, D ranges from $531,441$ for $\{Ni_{12}\}$, up to over a billion for $\{V_{30}\}$, making these ideal systems for us to study. In all four of these cases, we have used the QMC method that was described in Chapter 4, and have determined a microscopic model that accurately reproduces the measured data.[1] Furthermore, in most cases (the exception being $\{V_{30}\}$) we have also calculated the *predicted* low-temperature magnetization versus magnetic field, which can be tested with pulsed-field measurements, and also provides information about the underlying energy spectra that are associated with these magnetic molecules.

Before coming to our analysis, first note that two of the magnetic molecule systems studied in this chapter—$\{Cr_{10}Cu_2\}$ and $\{Cr_{12}Ni_3\}$—contain two *different* magnetic ions within each molecule. The QMC method that we described in Chapter 4 is perfectly applicable to any combination of intrinsic spins s, but we did *not* address the matter of different values of g. In fact, Eq. (4.16) requires that all spins share the *same* g, and QMC calculations subsequently involve sampling the quantum number M_S (defined in Sec. 2.2). By contrast, when different values of g are present, the magnetic moments associated with the states ψ_k ($1 \le k \le L$) vary from state to state within a given term W_i [of Eq. (4.21)], even though the quantum number M_S is fixed. Therefore, we have derived more general formulas for the magnetization and

---

[1]For all of the magnetic molecules that are studied in this chapter, the value of g associated with each ion has been suggested by the synthesis chemists, based on their analysis of the system. In the future, we hope to also include anisotropic interactions, as described in Chapter 9.

susceptibility which are still applicable when multiple values of g are present. This derivation is included in Appendix A, where we show that

$$M = \mu_B \left\langle \frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right\rangle \qquad (7.1a)$$

$$\chi = \beta \mu_B^2 \left[\left\langle \left(\frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right)^2 \right\rangle - \left\langle \frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right\rangle^2 \right], \qquad (7.1b)$$

where g_j is the g-value for the jth spin site ($1 \leq j \leq N$), and $m_{j,k}$ is the quantum number corresponding to the operator $\underline{s}_j^z$, acting on the state $|\psi_k\rangle$, of the term W_i.[2]

7.1 Analysis of the $\{Cr_{10}Cu_2\}$ magnetic molecule

The first magnetic molecule that we describe in this chapter is the $\{Cr_{10}Cu_2\}$ ring that has been synthesized by R. Winpenny's group at the University of Manchester, and was initially reported in Ref. [1]. The system is comprised of 10 chromium (Cr) ions, which each have an ionization of $+3$ and a resulting spin $s = 3/2$; and two copper (Cu) ions, which each have an ionization of $+2$ and a resulting spin $s = 1/2$.[3] The structure of this system was studied using x-ray crystallography, and these data were used to create the representation shown in Fig. 7.1(a). The two Cu ions are situated on the opposite sides of the ring to one another, with five Cr ions on each side of a Cu ion. These data also reveal that the ligands (O and F atoms) that connect each pair of (crystallographically equivalent) Cr ions are the same for all Cr-Cr bonds, so the Cr-Cr interactions are modeled with a single exchange interaction J_1. Furthermore, the molecules possess $180°$ rotational symmetry, so there are no more than two distinct Cr-Cu bonds, which we label J_2 and J_3. This can be seen more easily in Fig. 7.1(b), where we show a further simplified representation, including only the magnetic ions, and distinguishing the bonds (J_1, J_2, and J_3) with different line styles. Note that the two Cr-Cu bonds—represented J_2 and J_3 in Fig. 7.1(b)—correspond to ligands that clearly differ from one another in Fig. 7.1(a), so there is no reason to expect that $J_2 = J_3$.

Given the symmetries of this molecule—and assuming only nearest-neighbor, isotropic

---

[2]Note that M is in fact just the average if the magnetic moments ($\sum_{j=1}^{N} g_j m_{j,k}$) associated with each of the
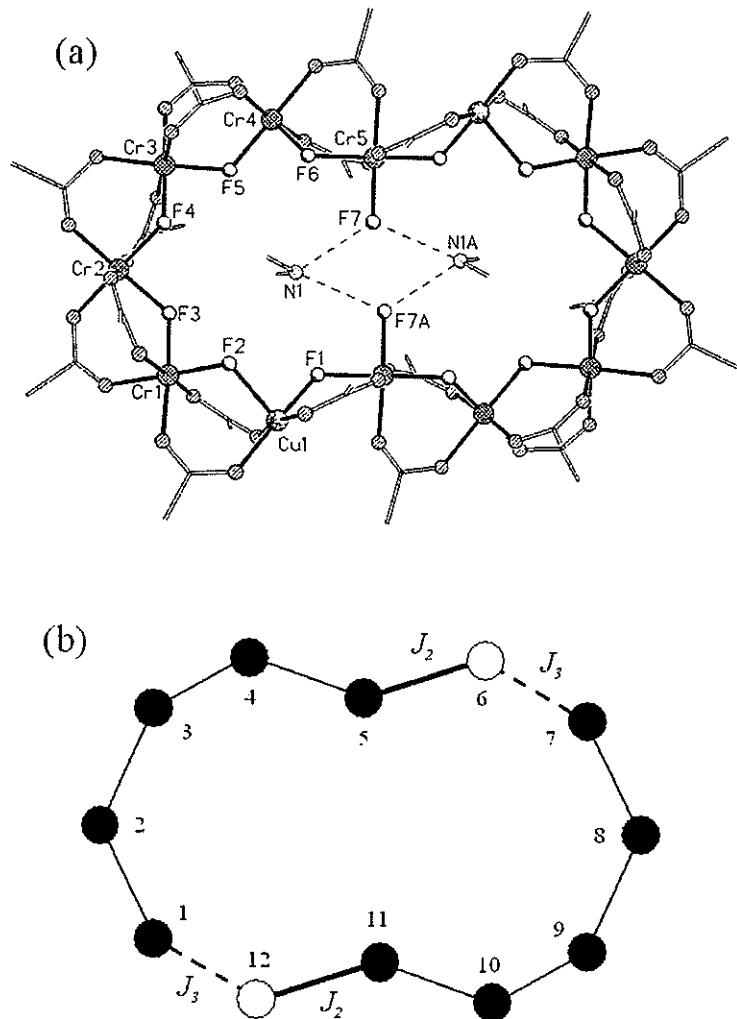
Figure 7.1   Diagrams showing the structure of the $\{Cr_{10}Cu_2\}$ ring. In (a) the magnetic (Cr and Cu) ions are shown, as well as the ligands that link them. This structure will be referred to as **1**. In (b) we show a simplified representation, including only the Cr ions (solid circles numbered 1–5 and 7–11) and Cu ions (open circles numbered 6 and 12), with lines representing the bonds that link them. All Cr-Cr bonds are assumed to be identical, and of strength J_1; whereas two different Cr-Cu bonds, J_2 and J_3, are considered, as shown in (b).

Heisenberg interactions—the Hamiltonian that describes this system is

$$\mathcal{H} = J_1 \left(\sum_{j=1}^{4} \vec{s}_j \cdot \vec{s}_{j+1} + \sum_{j=7}^{10} \vec{s}_j \cdot \vec{s}_{j+1} \right) + J_2 \left(\vec{s}_5 \cdot \vec{s}_6 + \vec{s}_{11} \cdot \vec{s}_{12} \right)$$
$$+ J_3 \left(\vec{s}_6 \cdot \vec{s}_7 + \vec{s}_1 \cdot \vec{s}_{12} \right) + \mu_B H \sum_{j=1}^{12} g_j s_j^z, \tag{7.2}$$

where g_j is different for the Cu sites than it is for the Cr sites. For both of these ions, the typical values of g_j are well known, and we have used $g_j = 2.1$ for Cu ($j = 6, 12$) and $g_j = 1.96$ for Cr (all other j), which were also used in Ref. [1]. There are therefore three parameters (J_1, J_2, and J_3) that must be determined in order to define the theoretical model for this molecule. This is typically done (at least initially) by measuring the weak-field magnetic susceptibility, and varying the free parameters until agreement is obtained with this experimental data.

Following the initial synthesis of this system, the magnetic susceptibility was measured from room temperature down to $T = 2$ K using a SQUID magnetometer with an external magnetic field of $H = 1.0$ Tesla. Due to the molecule's large Hilbert space, $D = \left[2(\frac{3}{2}) + 1 \right]^{10} \left[2(\frac{1}{2}) + 1 \right]^2 > 4 \times 10^6$, matrix diagonalization was not feasible for the *real* molecule. Instead, calculations were performed for a smaller, *imaginary* {Cr$_8$Cu$_2$} molecule, having only four, instead of five, Cr ions on each side of a Cu ion. This data was then "scaled up" in order to approximate the real system. Using this procedure, the best fit was obtained from the values $J_1/k_B = J_2/k_B = 9.8$ K (both AFM) and $J_3/k_B = -2.9$ K (FM).[1] In Fig. 7.2 we show the measured molar susceptibility χ_m times temperature,[4] along with theoretical $T\chi_m$ results that we have obtained by performing QMC calculations—for the *full* Hamiltonian of Eq. (7.2)—using these published values. Clearly there is room for improvement. We therefore sought to answer the following questions: Is there a unique (J_1, J_2, J_3)-triplet that *does* accurately reproduce the experimental susceptibility data? If so, what does this Hamiltonian

---

L states ψ_k. This seems quite reasonable, but is by no means obvious.

[3]For spins $s = 1/2$ and $s = 3/2$, the following information could be added to the tables of Sec. 4.7: There are 14 non-zero two-spin matrix elements; and for the directed loop equations, there are four 2×2 diagrams, 16 3×3 diagrams, and four 4×4 diagrams.

[4]All of the susceptibility data that we present in this chapter is *molar* susceptibility χ_m, in units of "cm^3 per mole formula unit". This quantity, and these units, are obtained by multiplying the (dimensionless) susceptibility χ by the volume per mole (of the complete chemical formula).
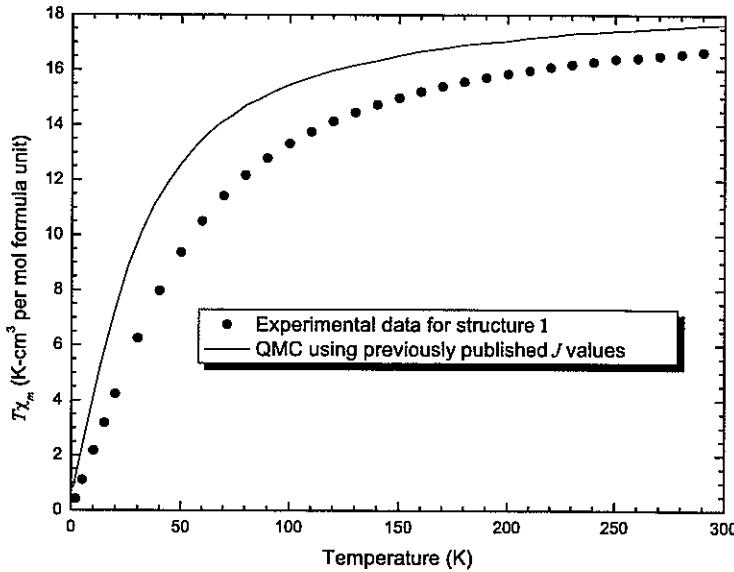
Figure 7.2 $T\chi_m$ as a function of temperature for $\{Cr_{10}Cu_2\}$, including the experimental data for **1**, as well as the QMC data that is obtained using the Heisenberg Hamiltonian of Eq. 7.2 with the previously published values of J_1, J_2, and J_3.

suggest about the energy spectrum (as probed via QMC calculations of M vs. H) that would be relevant to other experiments?

Since the publication of this initial data,[1] the same group has also synthesized another $\{Cr_{10}Cu_2\}$ structure, that is very closely related to **1** [shown in Fig. 7.1(a)]. This newer structure will be referred to as **2**, and is shown in Fig. 7.3. It is nearly identical to **1**, but involves slightly different bonding between the Cr and Cu ions, potentially affecting the bonds labeled J_2 and J_3. The structure of **2** shares the same symmetries with **1**, so the simplified representation shown in Fig. 7.1(b) still applies. Furthermore, the ligands that form the Cr-Cr bonds are very similar in **1** and **2**, so the values of J_1 for **1** and **2** were expected to be similar.

The weak-field magnetic susceptibility has been measured for **2**—just as it was for **1**—and both of these data sets are shown in Fig. 7.4 (symbols), along with our QMC calculations that provide the best fits (lines). In the cases of both **1** and **2**, the high temperature data that we have calculated are determined almost entirely by J_1, the strength of the Cr-Cr interaction, and are relatively insensitive to the values of J_2 and J_3.[5] In fact, by fitting just the $T > 100$ K

[5]This should not be surprising, since there are four times as many J_1 bonds as there are either J_2 or J_3
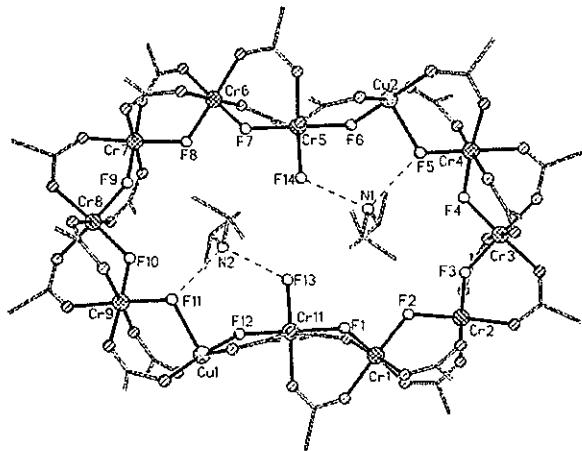
Figure 7.3 The crystal structure for the more recently synthesized $\{Cr_{10}Cu_2\}$ ring. This structure is referred to as **2**.

data, we were able to fix the parameters $J_1/k_B = 17.5 \pm 0.5$ K for **1**, and $J_1/k_B = 20.0 \pm 0.5$ K for **2**. When departing from these narrow ranges of J_1 values, it was subsequently not possible to fit the data by varying J_2 and J_3 in some appropriate way. Therefore, a *unique* value of J_1 exists for each of **1** and **2** which allows us to fit the data. Furthermore, these values are similar but not identical to one another, owing to the fact that the $\chi_m(T)$ data are similar but not identical for $T > 100$ K.

Although J_2 and J_3 have very little effect on the calculated $\chi_m(T)$ curves at high temperatures, they have a rather significant effect at low temperatures. The experimental $\chi_m(T)$ data for **1** show two distinct maxima occurring below $T = 40$ K, which are both accurately reproduced in Fig. 7.4 by choosing $J_1/k_B = 17.25 \pm 0.50$ K, $J_2/k_B = 26 \pm 1$ K, and $J_3/k_B = -18 \pm 1$ K, which are vastly different than the published[1] estimates. The same low-temperature details are not seen in the $\chi_m(T)$ data for **2**. Instead, the data for **2** show a slight "shoulder" around $T = 20$ K, and reach a larger maximum ($\chi_m \approx 0.24$ cm^3/mol) at $T \approx 2$ K. (Note that the initially published estimates of J_1, J_2, and J_3 yield values of χ_m that exceed 0.4 cm^3/mol for $T \lesssim 10$ K, well above the scale shown here.)

These two parameters, J_2 and J_3, together with J_1, turn out to define a unique, relatively small, region of the three-dimensional (J_1, J_2, J_3) parameter space, within which the exper-

bonds.
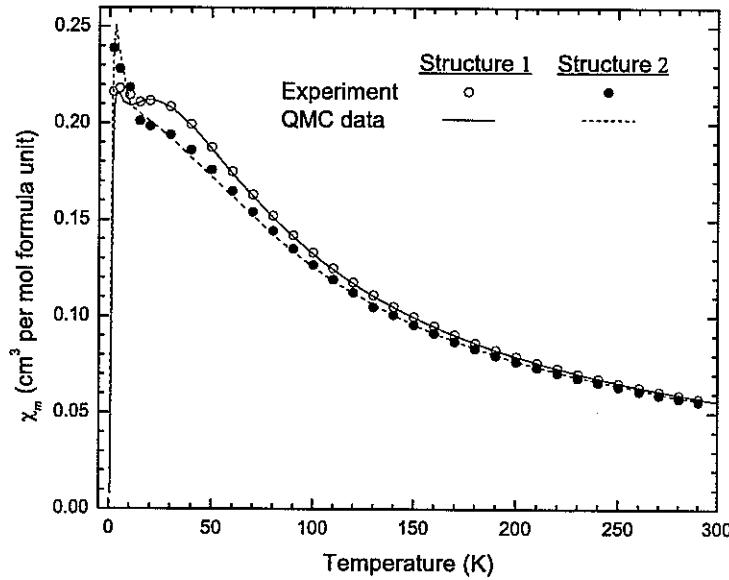
Figure 7.4 χ_m vs. T for both structures (**1** and **2**) of the $\{Cr_{10}Cu_2\}$ ring. The QMC data were obtained by varying J_1, J_2, and J_3 in Eq. (7.2) to find the best fit to the experimental data. The values of these parameters for the QMC data that are plotted here are: $J_1/k_B = 17.25 \pm 0.50$ K (AFM), $J_2/k_B = 26 \pm 1$ K (AFM), and $J_3/k_B = -18 \pm 1$ K (FM) for structure 1; $J_1/k_B = 20.0 \pm 0.5$ K (AFM), $J_2/k_B = 22 \pm 1$ K (AFM), and $J_3/k_B = -13 \pm 1$ K (FM) for structure **2**. For both QMC data sets, we have fixed $g = 1.96$ for the Cr ions, and $g = 2.1$ for the Cu ions.

imental $\chi_m(T)$ data can be accurately reproduced starting from the Heisenberg Hamiltonian of Eq. 7.2. This can be seen in Fig. 7.5, where we present contour plots of the discrepancy between the experimental and theoretical $\chi_m(T)$ data. Each of these two-dimensional plots corresponds to a "cut" through the three-dimensional parameter space, for the *fixed* values of J_1 that were determined by fitting the high temperature data for both 1 and **2**. These discrepancies were calculated by simply summing the squares of the discrepancies between experiment and theory for each of the 32 measured temperatures. The resulting numbers that are shown in the scale on the right edge of Fig. 7.5 are hence in units of $(cm^3/mol)^2$. Note that these contours have been plotted using a logarithmic scale, so the goodness of the fits deteriorates quite rapidly as J_2 or J_3 are varied. Recall that the region of best fit is even narrower (± 0.5

(a)

(b)
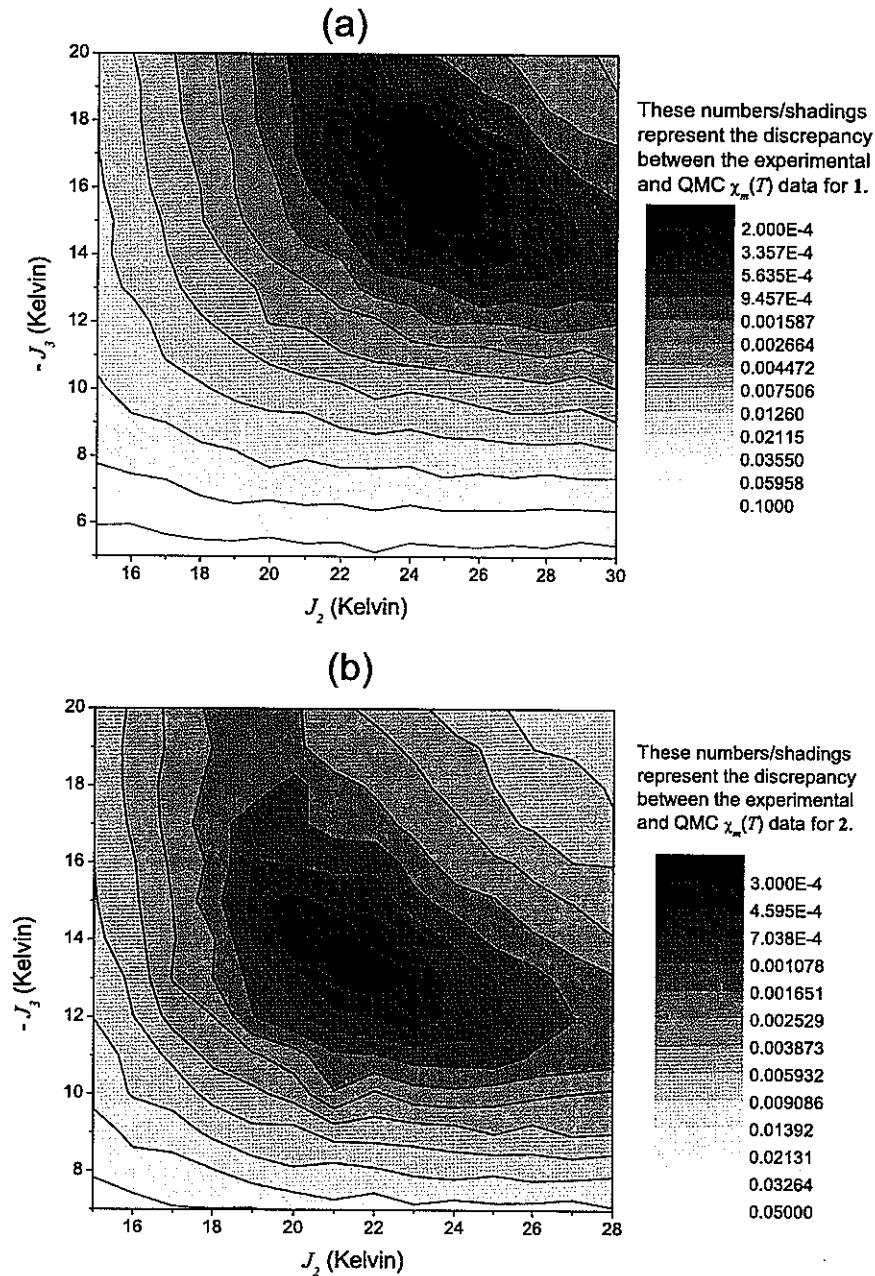
Figure 7.5 Contour maps of the discrepancy (described in the text) between the experimental and QMC susceptibility data $\chi_m(T)$ for (a) structure **1**, and (b) structure **2**. In both cases, J_1 has been fixed to the value that provides a good fit to $\chi(T)$ above $T = 100$ K. These values are $J_1/k_B = 17.5$ K and 20.0 K for **1** and **2**, respectively.

K) for J_1/k_B. The data shown in Fig. 7.5(a) correspond to $J_1/k_B = 17.5$ K, however we have obtained a slightly better fit by altering J_1/k_B to 17.25 K, and varying J_2 and J_3 slightly. This *best* fit data have been shown in Fig. 7.4.

Also note in Fig. 7.4 that the experimental data do not show any signs of decreasing toward zero with decreasing temperature for $T \geq 2$ K (the range that was measured). For this reason, there was initially some question as to whether or not the ground states of **1** and **2** are $S = 0$. However, if the ground states were $S = 1$, the limiting $T \to 0$ value of $T\chi$ would be ≈ 1 K-cm^3/mol, which is more than a factor of two larger than the values that were measured for $T = 2$ K. Instead, the low temperature data from Fig. 7.4 merely suggest an $S = 0$ ground state, separated from the lowest $S > 0$ state by a very small gap in the energy spectrum.

It is interesting to note that the total spin of the ground state can also be deduced by simply placing a *classical* spin vector at the site of each ion in Fig. 7.1(b), and arranging them such that each bond is maximally satisfied (i.e., neighboring spin vectors point antiparallel to one another when they are connected by AFM bonds, and parallel to one another for FM bonds). Choosing the bonds J_1 and J_2 to be AFM with J_3 FM—which is what the data suggest—each spin vector points in the opposite direction as the spin on the opposite side of the molecule. Hence the total spin S is zero. If, on the other hand, *all* bonds were (hypothetically) AFM, then the spins on the opposite sides of the molecule would point in the same direction as one another. Assuming the classical spin vectors for the Cr ($s = 3/2$) sites to have a greater magnitude than the spin vectors for the Cu ($s = 1/2$) sites, the ground state would have $S > 0$. This classical argument has also been confirmed through QMC calculations.

Finally, we present our QMC calculations for the predicted low-temperature magnetization M and differential susceptibility $\frac{dM}{dH}$ in Fig. 7.6 (both calculated per molecule, and plotted versus H) for the parameters that give the best fits to $\chi_m(T)$ for both **1** and **2**. These quantities can be measured directly in pulsed-field experiments (although up to somewhat weaker fields than we have considered), so these data should provide an opportunity to test the validity of the model, i.e. the specific choices of the exchange parameters, at which we have arrived. As we described in Chapter 5, the locations of the "peaks" in Fig. 7.6(b) also
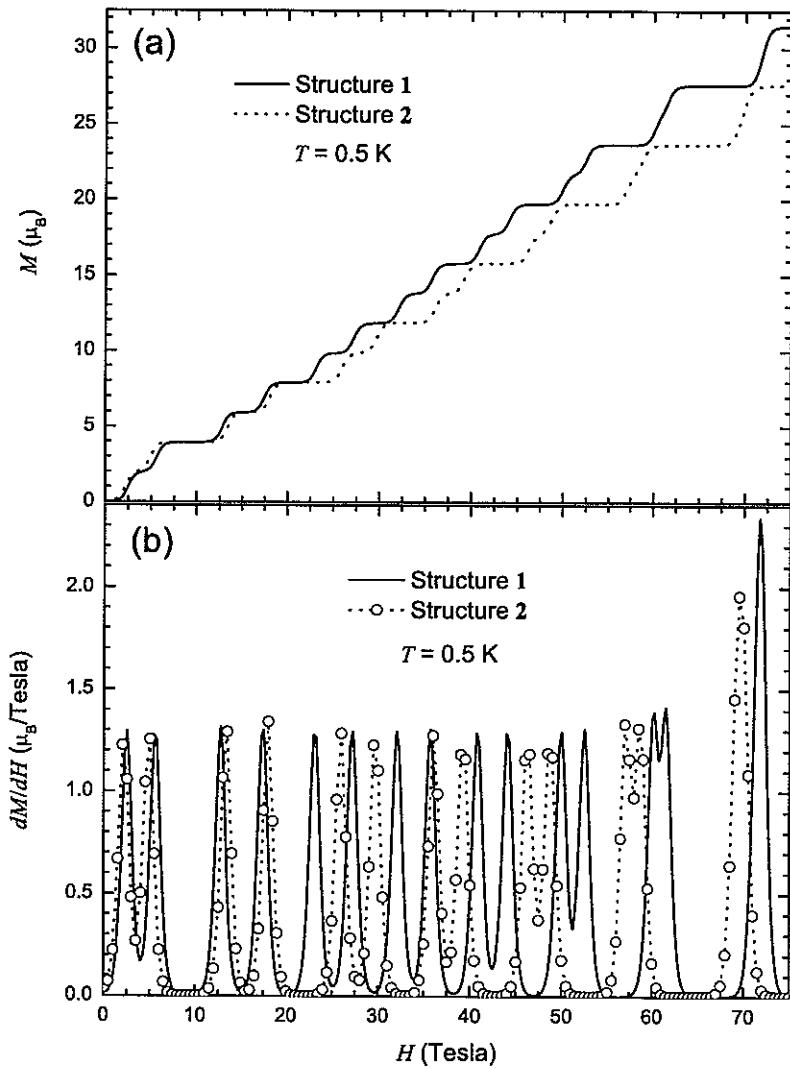
Figure 7.6   (a) M vs. H and (b) $\frac{dM}{dH}$ vs. H, calculated (per $\{Cr_{10}Cu_2\}$ molecule) from Eq. 7.2 for two different sets of the parameters J_1, J_2, and J_3. The solid lines were obtained using $J_1/k_B = 17.25$ K (AFM), $J_2/k_B = 26.0$ K (AFM), and $J_3/k_B = -18.0$ K (FM), which provides the best fit to the $\chi_m(T)$ data for **1**. The dotted lines [with symbols in (b)] correspond to $J_1/k_B = 20.0$ K (AFM), $J_2/k_B = 22.0$ K (AFM), and $J_3/k_B = -13.0$ K (FM), which gives the best fit for **2**. In both cases, a temperature of 0.5 K was used, with $g = 1.96$ for the Cr ions, and $g = 2.1$ for the Cu ions.

reveal information about the energy spectrum of the (model) $\{Cr_{10}Cu_2\}$ magnetic molecule which could be further tested by inelastic neutron scattering or nuclear magnetic resonance experiments. For very large fields [in excess of 70 Tesla for both (**1** and **2**)], the predicted magnetization eventually saturates at a value of 31.5 μ_B per molecule.

This section has focused primarily on the analysis of the existing $\{Cr_{10}Cu_2\}$ susceptibility data, illustrating how the QMC method allows us to determine the unique set of exchange parameters that give agreement between theory and experiment. An additional challenge was provided by the existence of two different, but very *similar* chemical structures, **1** and **2**. As we have shown, it is indeed possible to distinguish two distinct sets of exchange parameters, for which the corresponding Hamiltonians separately reproduce the data for **1** and **2**. These two Hamiltonians then each yield their own predictions for the low-temperature magnetization, as well as the corresponding features of the energy spectra.[6] A manuscript has recently been completed[2] which includes a description of this analysis, as well as a more detailed description of the chemical synthesis, and the analysis of EPR spectroscopy data which suggest that the energies of the lowest $S = 0$, $S = 1$, and $S = 2$ states are entirely consistent with our predicted $\frac{dM}{dH}$ data. This is an ongoing project, and future measurements of the high-field magnetization are expected.

7.2 Analysis of the $\{Cr_{12}Ni_3\}$ magnetic molecule

The same group at the University of Manchester that synthesized the $\{Cr_{10}Cu_2\}$ system, described in the previous section, has also created a $\{Cr_{12}Ni_3\}$ magnetic molecule, whose synthesis was reported in Ref. [3]. This molecule consists of 12 Cr ions, which again each have an ionization of $+3$ and a resulting spin $s = 3/2$; and three Nickel (Ni) ions which each have an ionization of $+2$ and a resulting spin $s = 1$.[7] For this system, x-ray crystallography reveals an open-chain structure, with a Ni ion situated at each end of the chain, and the remaining Ni

---

[6]See Sec. 5.2 for a discussion of the relationship between the low-temperature magnetization and the energy spectrum.

[7]For spins $s = 1$ and $s = 3/2$, the following information could be added to the tables of Sec. 4.7: There are 24 non-zero two-spin matrix elements; and for the directed loop equations, there are 14 2×2 diagrams, 20 3×3 diagrams, and 14 4×4 diagrams.
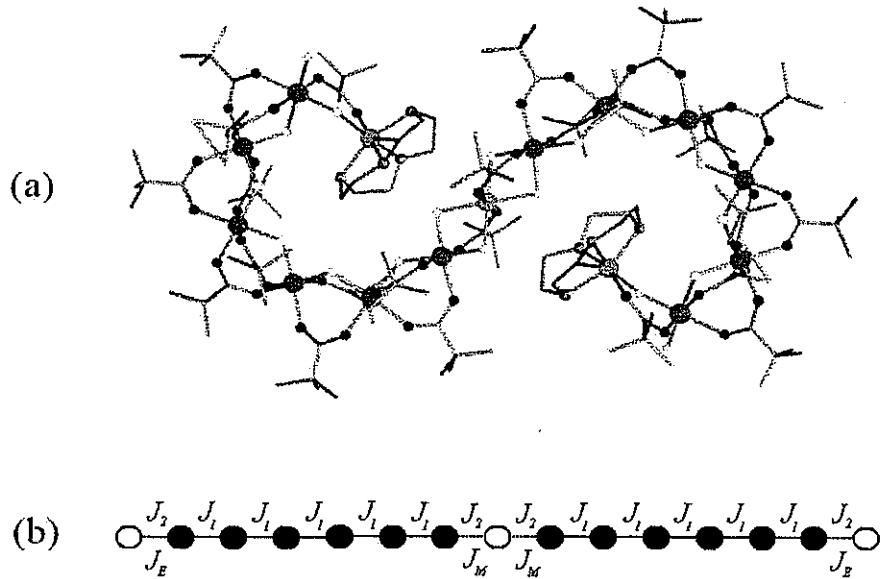
Figure 7.7   Structure of the {Cr$_{12}$Ni$_3$} magnetic molecule. (a) The actual
structure is shown, with the Cr and Ni ions represented by large
circles. Additional details regarding the chemical composition
of this molecule can be found in Ref. [3]. (b) A simplified rep-
resentation is provided, with the Cr and Ni ions represented by
solid and open symbols, respectively. The bonds are labeled J_1,
J_2, J_M, and J_E, as described in the text.

residing in the center. However, these chains do not crystallize in a straight line, but instead
form the interesting "S-shaped" structure shown in Fig. 7.7(a). The same ligands link all
nearest-neighbors Cr-Cr pairs, so a single exchange parameter J_1 can be used to describe all
10 of the Cr-Cr bonds. Also, the ligands that link the central Ni ion to the two neighboring
Cr ions are the same as one another, so these two bonds are described by one parameter J_M
("M" for "Middle"). Similarly, the ligands linking the terminal Ni ions to their nearest Cr
ions are the same at both ends of the molecule, so these two bonds can also be described
by a single parameter J_E ("E" for "End"). The ligands represented by J_M and J_E are not
identical to one another, but for the sake of simplicity, they are (initially) assumed to give rise
to the same interactions, described by $J_2 = J_M = J_E$. This labeling is shown in the simplified
representation that appears in Fig. 7.7(b).

With only two different exchange parameters, J_1 and J_2, the Hamiltonian for this system

is

$$\mathcal{H} = J_1 \left(\sum_{j=2}^{6} \vec{s}_j \cdot \vec{s}_{j+1} + \sum_{j=9}^{13} \vec{s}_j \cdot \vec{s}_{j+1} \right) +$$
$$J_2 \left(\vec{s}_1 \cdot \vec{s}_2 + \vec{s}_7 \cdot \vec{s}_8 + \vec{s}_8 \cdot \vec{s}_9 + \vec{s}_{14} \cdot \vec{s}_{15} \right)$$
$$+ \mu_B H \sum_{j=1}^{15} g_j s_j^z, \tag{7.3}$$

where we assume the same g values that were used[3] in the initial publication: 2.0 for the Cr ions (i.e., $g_j = 2.0$ for $2 \leq j \leq 7$ and $9 \leq j \leq 14$), and 2.2 for the Ni ions ($g_j = 2.2$ for $j = 1$, 8, 15). Following the initial synthesis of this magnetic molecule, $\chi_m(T)$ was measured from $T = 300$ K down to 2 K in a SQUID magnetometer, and these data were used to estimate J_1 and J_2 as reported in Ref. [3]. However, the Hilbert space associated with this Hamiltonian is quite large, given by $D = \left[2(\frac{3}{2}) + 1\right]^{12} \left[2(1) + 1\right]^3 > 4 \times 10^8$. For this reason, Eq. (7.3) was not used directly, but was instead replaced with its classical analog (as we described in Sec. 6.2.1), and this *classical* Hamiltonian was used to fit the experimental data, subsequently yielding the estimates $J_1/k_B = 29.2$ K and $J_2/k_B = 29.9$ K (both AFM). Given the results that have been presented in Chapter 6, we were skeptical of the accuracy of these estimates—as compared with those that would be obtained from Eq. 7.3—and have thus performed QMC calculations using this (quantum) Hamiltonian.

With only 2 free parameters, J_1 and J_2, the process of fitting the experimental $\chi_m(T)$ data was exceedingly straightforward. For a given *ratio*, J_2/J_1, we calculated $T\chi_m$ as a function of $k_B T/J_1$. This is possible because $T\chi_m$ depends only on the ratios $J_1/k_B T$ and $J_2/k_B T$—or equivalently, J_2/J_1 and $k_B T/J_1$—and does not depend on J_1, J_2, and T independently. Therefore, we could plot $T\chi_m(\frac{k_B T}{J_1})$, and by varying the value of J_1 (and hence J_2, since their ratio is fixed) we were able to immediately explore *all* (J_1, J_2)-doublets that share this same ratio. This allowed us to quickly determine whether or not a given J_2/J_1 ratio was feasible. Then, by performing QMC calculations for many different values of J_2/J_1, we were able to find the parameters that give the best fit to the experimental data. Again, as was the case for {Cr$_{10}$Cu$_2$}, we have found that there is only a small, *unique* region of the (J_1, J_2) parameter space which reasonably reproduces the experimental data. The best fit was obtained upon
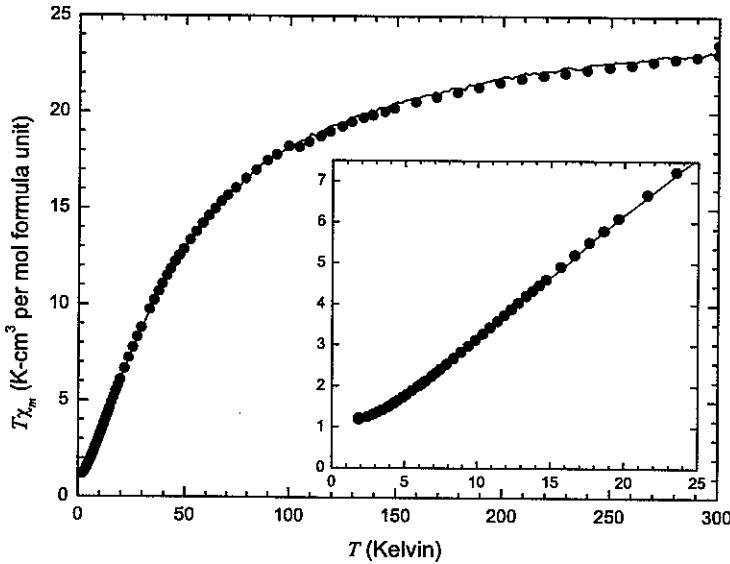
Figure 7.8 $T\chi_m$ vs. T for the $\{Cr_{12}Ni_3\}$ magnetic molecule. The symbols
correspond to the measured data; and the solid line represents
the QMC calculation for $J_1/k_B = 18.0\pm0.5$ K and $J_2/k_B = 9\pm1$
K, with $g = 2.0$ for the Cr ions, and $g = 2.2$ for the Ni ions.

choosing $J_1/k_B = 18.0 \pm 0.5$ K and $J_2/k_B = 9 \pm 1$ K (both AFM), which are quite different

than the previous estimates.[8] The resulting theoretical $T\chi_m$ data are shown in Fig. 7.8, along

with the experimental data.

At low temperatures, the data approach a constant value of $T\chi_m$, slightly larger than 1

K-cm^3/mol, which implies an $S = 1$ ground state. This can also be expected by considering

classical spin vectors (as was the case for $\{Cr_{10}Cu_2\}$), whereby all of the spins cancel one

another out, with the exception of one uncompensated Ni ion. At higher temperatures (in fact

at all temperatures), the experimental data are well fit by the QMC data shown in Fig. 7.8.

However, we wanted to know whether such agreement could also be achieved when allowing

J_M and J_E to differ from one another. To that end, we have additionally considered the 3-

parameter Hamiltonian that results from replacing J_2 with J_M and J_E, as we have shown at

---

[8]As we have described in Chapter 6, for $k_B T/J \gg 1$, a classical Heisenberg model will accurately reproduce
its quantum analog. However, the previous estimates ($J_1/k_B = 29.2$ K and $J_2/k_B = 29.9$ K) produce results that
disagree with the experimental data, even for $T = 300$ K. Therefore, we suspect that there is a typographical
error in Ref. [3], reporting these estimates in units of cm^{-1}, instead of K, hence giving estimates that are a
factor of 1.439 larger than was intended.
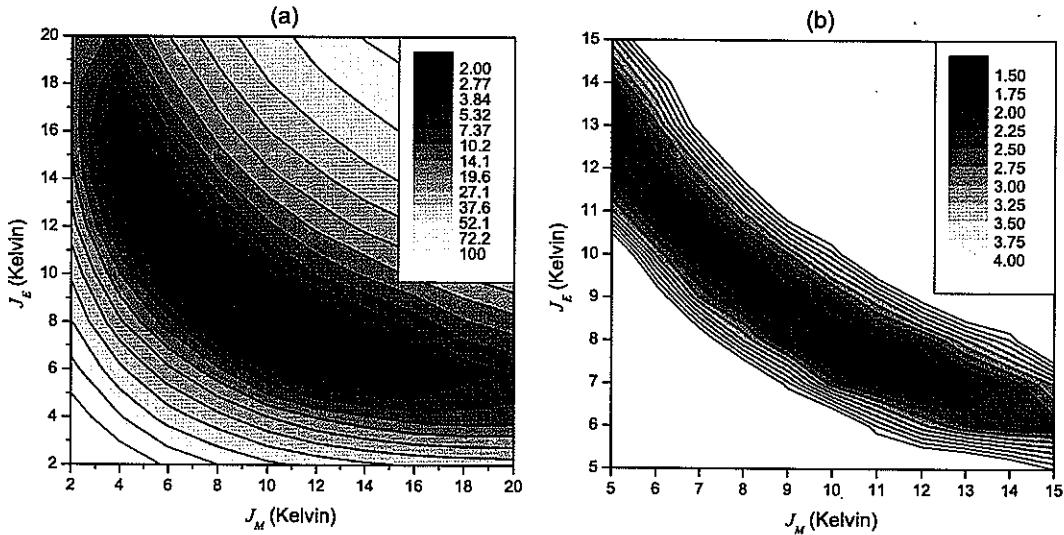
Figure 7.9 Contour plots showing the quality of the fits between theory and experiment for $\{Cr_{12}Ni_3\}$ with a fixed value of $J_1/k_B = 18$ K. The precise meaning of these numbers is described in the text. (a) Data are shown from 2 K to 20 K, which encompass the entire "banana shaped" region of best fit. (b) A close-up of this region is shown using a linear scale.

the bottom of Fig. 7.7(b).

When using three parameters, we have found that J_1 must remain roughly unchanged from its 2-parameter estimate, $J_1/k_B = 18$ K. (This is not surprising, since most of the bonds are of strength J_1.) By subsequently fixing $J_1/k_B = 18$ K, and varying both J_M and J_E, we have explored the two-dimensional (J_M, J_E) plane, searching for points—in addition to $J_2 = J_M = J_E = 9$ K—that produce $T\chi_m$ data which are in agreement with the measured data. Specifically, for each set of parameters, we calculated $T\chi_m$ for each of the 90 measured temperatures; for each value of T, the square of the difference between the experimental and theoretical $T\chi_m$ was calculated; and these squares were summed to determine the goodness of the fit. It is this quantity, in units of $(K\text{-}cm^3/mol)^2$, which is shown in the contour plots
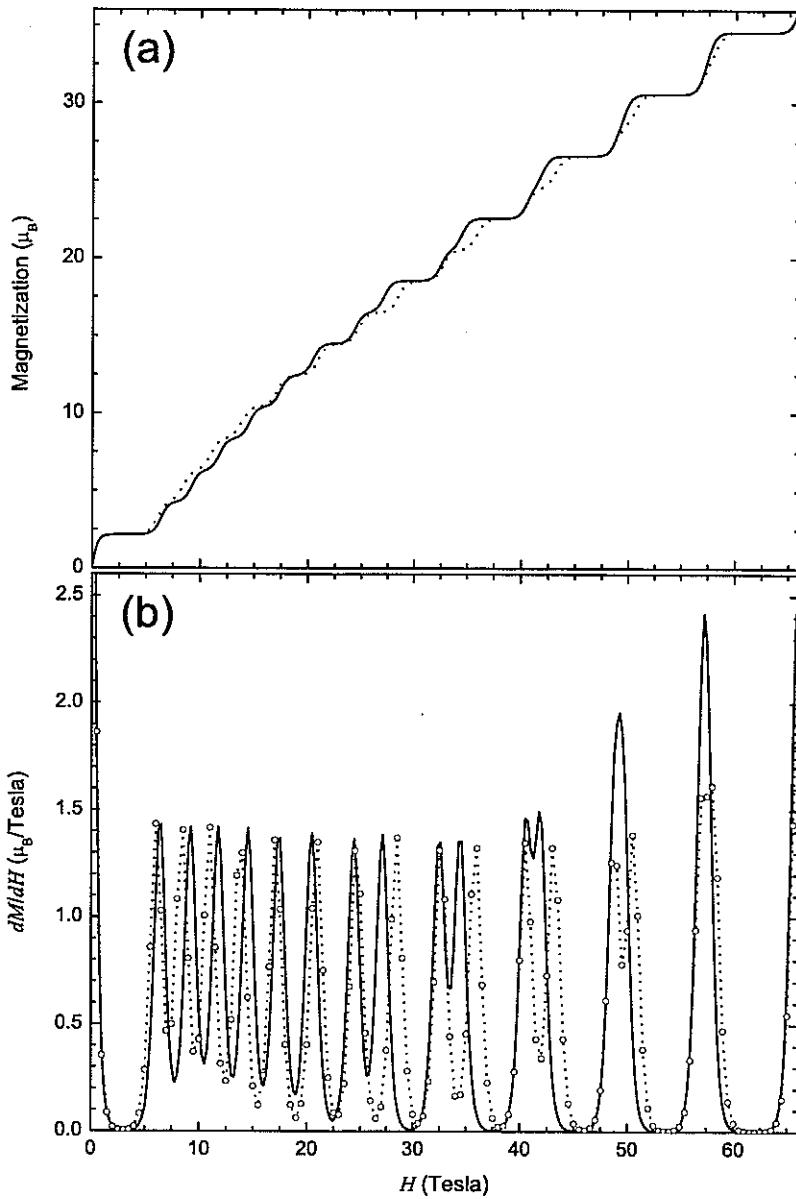
Figure 7.10 Calculated (a) M vs. H and (b) $\frac{dM}{dH}$ vs. H per $\{Cr_{12}Ni_3\}$ magnetic molecule. The solid curves were calculated using $J_1/k_B = 18$ K and $J_2/k_B = 9$ K; and the dotted curves [with symbols in (b)] were calculated using $J_1/k_B = 18$ K, $J_M/k_B = 12$ K, and $J_E/k_B = 7$ K. These calculations were performed with $T = 0.5$ K, assuming $g = 2.0$ for the Cr ions, and $g = 2.2$ for the Ni ions.

of Fig. 7.9. Indeed, along the diagonal line that corresponds to $J_2 = J_M = J_E$, there is a narrow region of good fit centered at $J_2/k_B = 9$ K. However, by simultaneously increasing J_M and decreasing J_E (or vice versa) it is possible to maintain a reasonably good fit throughout a relatively large "banana-shaped" region, which can be seen in Fig. 7.9(a). In Fig. 7.9(b) we show a more detailed plot of this region of best fit. The data shown in Fig. 7.8 (for $J_1/k_B = 18$ K and $J_2/k_B = 9$ K) turn out to provide one of the best possible fits, but a slight improvement to this fit is achieved by choosing $J_1/k_B = 18$ K, $J_M/k_B = 12$ K, and $J_E/k_B = 7$ K.

To summarize, upon fixing $J_M = J_E$, there is only a very small region of the two-dimensional (J_1, J_2) parameter space that will produce a good fit to the experimental data. However, by varying J_M and J_E independently, we have found a somewhat extended region in the three-dimensional parameter space (shown in Fig. 7.9 for $J_1/k_B = 18$ K) which yields fits to the experimental data that are of good quality. We have therefore calculated $M(H)$ and $\frac{dM}{dH}$ for $J_1/k_B = 18$ K and $J_2/k_B = 9$ K (which provides the best fit when using two parameters) and for $J_1 = 18$ K, $J_M = 12$ K, and $J_E = 7$ K (which provides the best fit when using three parameters). These data are plotted in Fig. 7.10, revealing very similar results for both sets of parameters. For both molecules, the (predicted) magnetization eventually saturates to a value of 42.6 μ_B per molecule when $H \approx 75$ Tesla. The low-temperature ($T = 0.5$ K) magnetization for this system has very recently been measured using a pulsed-field for $H \leq 30$ Tesla, and this data will be carefully analyzed in the near future.

7.3 Analysis of the {Ni$_{12}$} magnetic molecule

We have also used QMC calculations to analyze a {Ni$_{12}$} cluster that was synthesized locally by Ames Laboratory chemist, P. Kögerler. For this molecule, x-ray crystallography has revealed the structure shown in Fig. 7.11(a), which includes 12 Ni ions, each of which has an ionization of +2 and a resulting spin $s = 1$. Carefully examining this figure, there are a total of 21 exchange pathways through which the Ni ions can interact with one another. However, even from this relatively complex picture, the magnetic molecule clearly possesses a great deal of symmetry. In particular, there is 120° rotational symmetry, as well as mirror symmetry
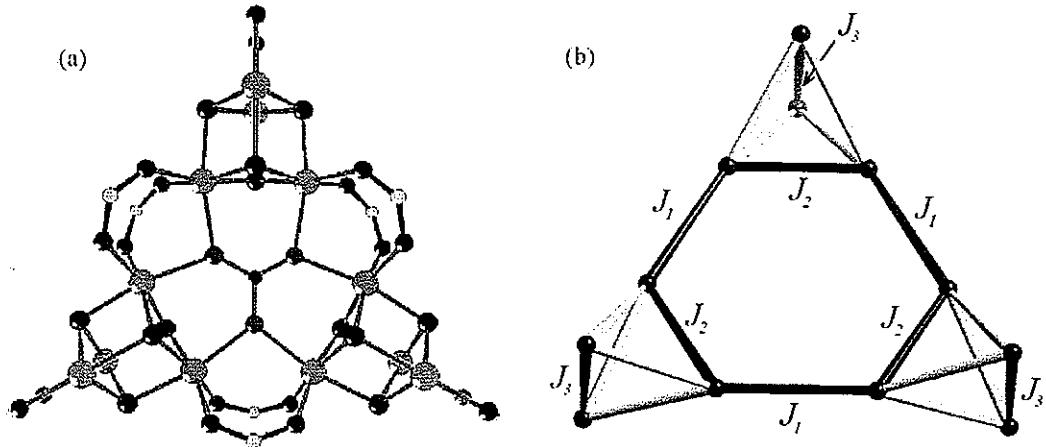
Figure 7.11　Structure of the {Ni$_{12}$} magnetic molecule. (a) The Ni ions are represented by green circles, and the ligand atoms are represented as follows: O=red, C=grey, (O positions of OCH$_3$ groups)=brown, and CO$_3$=(central purple/black). (b) Only the Ni ions are included in this simplified representation, and the J_1, J_2, and J_3 bonds are labeled. The 12 unlabeled (yellow) bonds are of strength J_4.

about the plane that passes vertically through Fig. 7.11(a). These symmetries immediately reduce the number of *distinct* bonds from 21 to (at most) four, which we represent J_1, J_2, J_3, and J_4. These bonds have been labeled accordingly in Fig. 7.11(b), where the 12 unlabeled (yellow) bonds correspond to J_4. The weak-field magnetic susceptibility was also measured for this system, providing useful evidence regarding the nature of these magnetic interactions, as we describe below.

With $N = 12$ Ni ions, each with $s = 1$, the Hilbert space associated with this magnetic molecule is $D = 531,441$, which is small enough to be solved by numerical diagonalization. However, the advantage of using QMC in this situation was the relative speed of the calculations. A QMC calculation of the weak-field susceptibility—from $T = 2$ K to 300 K—required less than 5 minutes for any particular choice of parameters (J_1, J_2, J_3, and J_4), while the numerical diagonalization of the Hamiltonian required several hours of computation. This is therefore an example of how these two methods can effectively *complement* one another: QMC can be used to perform calculations for many different sets of parameters, finding the best fit

to $\chi_m(T)$; then, the resulting Hamiltonian can be diagonalized a single time to determine the full energy spectrum for the optimum set of parameters.

Initially, several assumptions were attempted in order to further reduce the number of parameters from four to (a more manageable) three, or perhaps even two. However, we found that a reasonable fit was *impossible* to obtain with fewer that four distinct bonds. With four parameters though, the exploration of the four-dimensional parameter space is a daunting task, even though each calculation required only a few minutes. For example, sampling a very crude $10 \times 10 \times 10 \times 10$ grid in parameter space would have required multiple days. In order to expedite this process, we thus used the same strategy that was employed for the $\{Cr_{12}Ni_3\}$ system, performing calculations of χ_m as a function of k_BT/J_1 for different values of the ratios: J_2/J_1, J_3/J_1, and J_4/J_1. For each set of three ratios, we then found the value of J_1 that provided the optimum fit, and recorded this set of parameters (J_1,J_2,J_3,J_4) into memory, along with the corresponding goodness of the fit—calculated as a sum of the squares of the differences between experiment[9] and theory.

A thorough exploration of the parameter space then took place over the course of a weekend, finding that the experimental data could be accurately fit by one (and only one) region of the four-dimensional parameter space. This was then followed by a finer search of the region that provided the reasonable fits, yielding that data that are shown in Fig. 7.12. The best fit was subsequently obtained upon choosing $J_1/k_B = 17.5 \pm 0.5$ K, $J_2/k_B = -9.5 \pm 1.5$ K, $J_3/k_B = -22.0 \pm 1.0$ K, and $J_4/k_B = -1.9 \pm 0.1$ K, with $g = 2.21$.[10] (The uncertainties given here represent the approximate size of the region in parameter space in which a good fit is obtained.) Note that these four values are indeed quite different from one another: Two of the bonds, J_1 and J_3, are of similar magnitude, but J_1 is AFM, whereas the other three bonds are FM. By comparison, the other two bonds, J_2 and J_4, are weaker by a factor of roughly 2 and 10, respectively. Also, note that the data are extremely sensitive to the precise value of J_4, which is the weakest FM bond. This could perhaps have been expected, since most of the

---

[9]Note that this procedure required the ability to fit for values of T falling between the measured data points, which was accomplished by replacing the experimental data with a series of piece-wise continuous functions.

[10]The precise value of $g = 2.21$ was determined by allowing g to deviate slightly from 2.2, in order to find the value that allowed the best fit to be obtained.
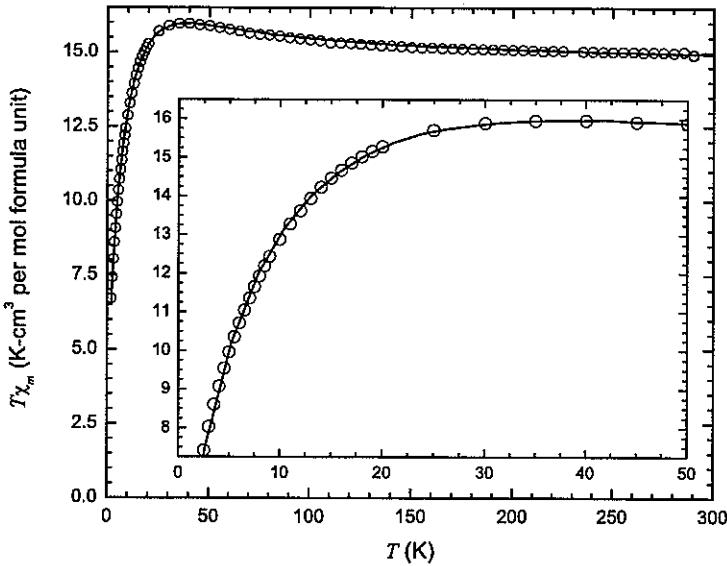
Figure 7.12 $T\chi_m$ vs. T for the $\{Ni_{12}\}$ magnetic molecule. The solid line corresponds to the QMC calculation which used the values of J_1, J_2, J_3, and J_4 that are provided in the text. The symbols represent the experimental data, which has been adjusted for a small diamagnetic contribution of 5×10^{-4} cm^3/mol.

bonds are represented by J_4.

Using these parameters, we have also calculated the predicted low-temperature magnetization as a function of the external magnetic field. These QMC data—calculated using $T = 2$ K—are represented by the solid curve in Fig. 7.13. However, for $T < 2$ K, the QMC method begins to suffer from the "negative sign problem",[11] and data become extremely difficult to obtain as the temperature is lowered toward $T = 1$ K. For this reason, the ability to diagonalize the Hamiltonian was particularly important. Data for $T = 0.5$ K have also been included in Fig. 7.13, which were obtained by numerically diagonalizing the Hamiltonian using the MAGACK[4] package, and reveal a number of "steps" that are not visible in the 2 K QMC data. Furthermore, this diagonalization provided the full energy spectrum, from which we have included a listing (Table 7.1) of all of the energies that lie within 10 K of the $S = 0$ ground state. These should prove valuable for the analysis of future data.

---

[11]The negative sign problem is the subject of the next chapter, and in Sec. 8.2 the emergence of the negative sign problem is described for $\{Ni_{12}\}$ in particular.
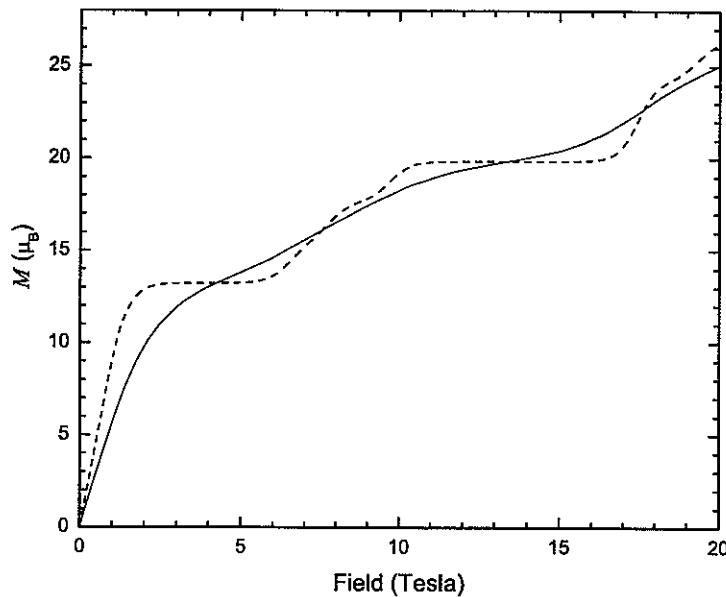
Figure 7.13   Predicted M vs. H per $\{Ni_{12}\}$ molecule. Both curves assume the same Hamiltonian, but the solid curve represents QMC data with $T = 2.0$ K, whereas the dashed curve represents data with $T = 0.5$ K that were obtained by diagonalizing the Hamiltonian matrix.

Following the analysis that has been described above, $M(H)$ was measured for the $\{Ni_{12}\}$ system, using a pulsed field, with a temperature of $T = 0.5$ K. These data were expected to either confirm or refute the prediction contained in Fig. 7.13, however there are problems with these experimental data that are currently unresolved. In particular, SQUID measurements had been made for $H < 5$ T, with $T = 2$ K, giving results that are very similar to the QMC data shown in Fig. 7.13. Upon lowering the temperature from 2 K to 0.5 K, the $M(H)$ curve *should* become more "sharp" as is the case in Fig. 7.13. However, the 0.5 K pulsed-field data actually lie *below* the 2 K SQUID data, a trend which does not make sense thermodynamically. Therefore, a detailed analysis of $M(H)$ is being reserved until reliable data are available. An article describing the synthesis of this system, as well as the results of our QMC analysis was submitted to the Journal of the American Chemical Society, but was not accepted for publication. A revised manuscript is currently being prepared which will be submitted to a different journal in the near future.

S	Excitation energy $(k_B K)$	Multiplicity
0	0	1
1	0.316	2
1	0.379	1
2	0.859	2
2	1.195	2
2	1.234	1
3	1.641	1
3	1.803	1
3	1.969	2
4	2.825	2
4	3.024	1
5	4.245	2
6	5.990	1

Table 7.1   All of the energy eigenvalues (relative to the ground state, $E \equiv 0$) that lie within 10 K of the ground state energy for the $\{Ni_{12}\}$ model Hamiltonian.

7.4   Analysis of the $\{V_{30}\}$ magnetic molecule

The final magnetic molecule to be described in this chapter is the recently synthesized $\{V_{30}\}$ system, which was first synthesized by A. Müller's group at the University of Bielefeld,[12] and has since been synthesized at Ames Laboratory as well.[6, 7] The structure of this molecule is shown in Fig. 7.14(a), where 30 V ions—residing at the centers of the large polyhedra— occupy the 30 vertices of a slightly distorted icosidodecahedron. It should be noted that this same geometrical arrangement [shown in the simplified representation of Fig. 7.14(b)] was also adopted by the previously synthesized $\{Fe_{30}\}$ ($s = 5/2$) magnetic molecule.[8] Each of the 30 V ions have an ionization of +4, and a resulting spin $s = 1/2$; so these systems offer a great deal of potential for comparing the effects of different intrinsic spins in the same geometrical structure. To that end, the weak-field magnetic susceptibility of $\{V_{30}\}$ was measured and analyzed, as we describe below.[13]

The magnetic susceptibility was measured in an applied field of 0.1 Tesla, finding a large

---

[12]We have provided the analysis for this initial report, which was published in Ref. [5], and has also been included as Appendix B.

[13]Note that the expected value of g for the V ions was 1.95, which we have assumed throughout the following.
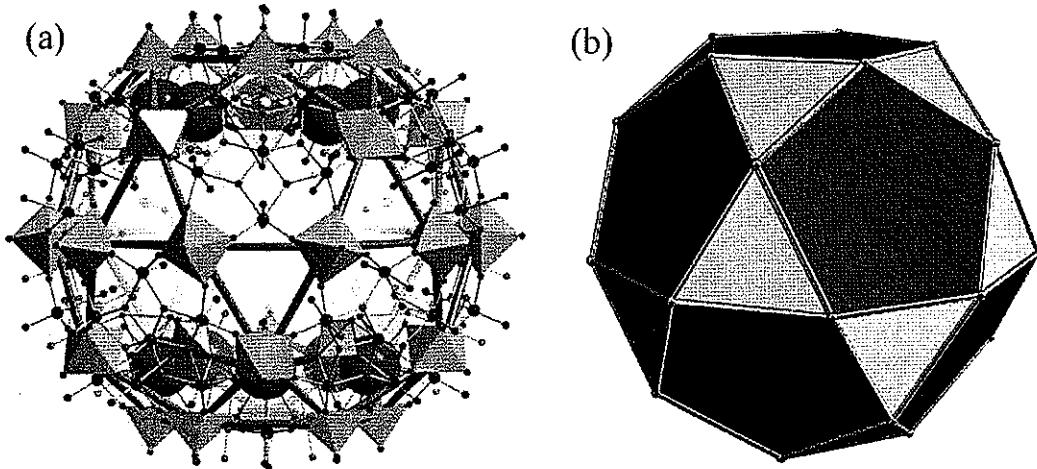
Figure 7.14   Structure of the {V$_{30}$} magnetic molecule. (a) The V ions are
situated at the centers of the large polyhedra, and additional
details regarding the chemical composition can be found in
Appendix B. (b) The V ions reside at the vertices of a nearly
regular icosidodecahedron, such as the one shown here.

magnetization at $T = 2$ K, which was the lowest temperature measured. The value of $T\chi_m$
at 2 K was 0.86 K-cm^3/mol, somewhat smaller than the $T \to 0$ limiting value for an $S = 1$
ground state, which is 0.95 K-cm^3/mol. Furthermore, $T\chi_m$ was increasing with T, even at
$T \approx 2$ K, suggesting an $S = 0$ ground state with some "additional" contribution to the
measured susceptibility. This was attributed to the presence of approximately two detached,
non-interacting V ions per molecule, which would be very difficult to measure experimentally,
and would properly account for the large value of $T\chi_m$ at low temperatures. The data that
would result from these detached $s = 1/2$ V ions were thus subtracted from the measured data,
leaving the "intrinsic" data shown in Fig. 7.15.

For $N = 30$ spins $s = 1/2$, the Hilbert space for this system is $2^{30} > 10^9$, which is far
larger than what is currently solvable by matrix diagonalization. We have therefore used the
QMC method to attempt to fit these data, and have obtained the solid curve that is shown in
Fig. 7.15 for $J = 245 \pm 3$ K and $g = 1.95$. Note that QMC data are conspicuously absent for
$T < 120$ K. This is because the negative sign problem becomes insurmountable for this system
when $k_B T/J \lesssim 1/2$. [The structure of {V$_{30}$}—in particular the icosidodecahedron, defined by
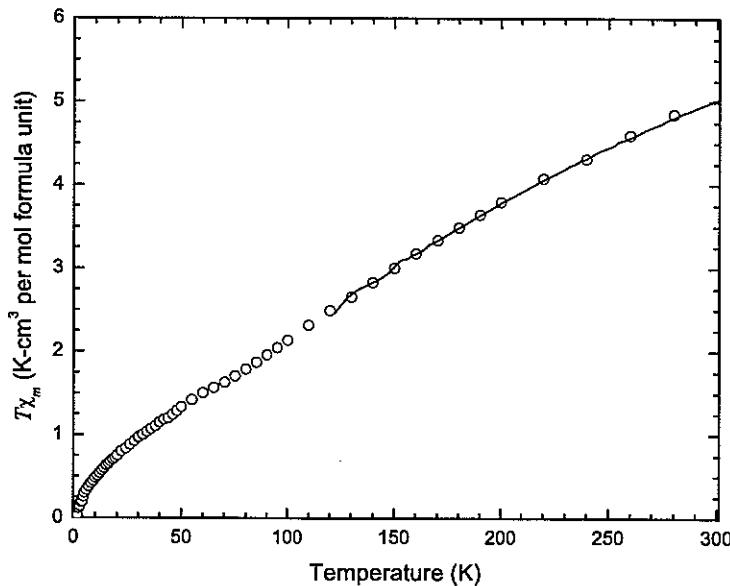
Figure 7.15　$T\chi_m$ vs. T for the $\{V_{30}\}$ magnetic molecule. The "intrin-
sic" experimental data (described in the text) are represented
with symbols, and the solid line represents QMC data for
$J = 245 \pm 3$ K, and $g = 1.95$.

the sites of the magnetic ions and shown in Fig. 7.14(b)—involves many AFM triangles, causing

strong "geometrical frustration" and a severe negative sign problem, as described in the next

chapter.] This is therefore an example of a situation wherein the QMC method is effective in

determining the optimum choice for a *single* J, but can provide no additional information. For

instance, it is entirely possible that $\{V_{30}\}$ should actually be described by multiple J values,

given the slightly distorted structure of the molecule. Furthermore, it would be very interesting

to perform low-temperature QMC calculations to learn about the low-energy excitations, but

this is unfortunately not feasible for the present system.

Bibliography

[1] F. K. Larsen, J. Overgaard, S. Parsons, E. Rentschler, A. A. Smith, G. A. Timco, and
R. E. P. Winpenny, Angew. Chem. Int. Ed. **42**, 5978 (2003).

[2] M. Shanmugam, L. Engelhardt, F. K. Larsen, M. Luban, E. J. L. McInnes, C. A. Muryn,

J. Overgaard, E. Rentschler, G. A. Timco, and R. E. P. Winpenny, Chem. Eur. J. (submitted)

[3] S. L. Heath, R. H. Laye, C. A. Muryn, N. Lima, R. Sessoli, R. Shaw, S. J. Teat, G. A. Timco, and R. E. P. Winpenny, Angew. Chem. Int. Ed. **43**, 6132 (2004).

[4] J. J. Borrás-Almenar, J. M. Clemente-Juan, E. Coronado, and B. S. Tsukerblat, J. Comp. Chem. **22**, 985 (2001).

[5] A. Müller, A. Todea, J. van Slageren, M. Dressel, H. Bögge, M. Schmidtmann, M. Luban, L. Engelhardt, and M. Rusu, Angew. Chem. Int. Ed. **44**, 3857 (2005).

[6] B. Botar, P. Kögerler, and C. L. Hill, Chem. Commun. pp. 3138–3140 (2005).

[7] B. Botar, P. Kögerler, and C. L. Hill, J. Am. Chem. Soc. Comm. **128**, 5336–5337 (2006).

[8] A. Müller, M. Luban, C. Schröder, R. Modler, P. Kögerler, M. Axenovich, J. Schnack, P. Canfield, S. Bud'ko, and N. Harrison, ChemPhysChem **2**, 517 (2001).

CHAPTER 8. The negative sign problem

8.1 Introduction

Recall the following very important message from Secs. 3.1 and 4.1: If a thermodynamic average can be written in the form

$$\langle Q \rangle = \frac{\sum_i Q_i W_i}{\sum_i W_i}, \tag{8.1}$$

then one can proceed[1] with the calculation by using the detailed balance condition,

$$\frac{P_{j \to i}}{P_{i \to j}} = \frac{W_i}{W_j}, \tag{8.2}$$

to sample the terms W_i, *regardless* of the complexity of the summation. However, some of the terms W_i from Eq. (4.21) turn out to be negative for certain types of systems, leading to the so-called "negative sign problem" (NSP). The origin of these negative terms is associated with the presence of negative matrix elements (as we discussed in Sec. 4.5) which depend on whether a bond is ferromagnetic (FM) or antiferromagnetic (AFM). As we describe in Sec. 8.2, the origin of negative terms also depends strongly on the geometry of a system, allowing us to formulate a clear statement as to whether the NSP would, or would not, be present in any given spin system.

From Eq. (8.2), it is clear that these negative values of W_i are potentially problematic, since the left-hand side involves *probabilities*, which of course cannot be negative. However, Eqs. (8.1) and (8.2) can still be applied, with only a slight modification, as we describe in Sec. 8.3. This modification allows one to salvage detailed balance, but at the price of some— sometimes considerable—cancellation among the data which are recorded in the process of a

---

[1]Recall that—in order for this to be useful—each term W_i must be easy to evaluate, as is the case for Eq. (4.21).

QMC calculation. In situations that a great deal of cancellation occurs, the data that remain (after said cancellation) may be barely discernable—or not discernable—from the statistical fluctuations that are inherent in a Monte Carlo calculation, hence causing very large statistical errors. This cancellation is described in greater detail Sec. 8.4, and attempts to overcome the NSP are reviewed in Sec. 8.5. Unfortunately, a general solution to the NSP does not exist,[1] so in Sec. 8.6 we seek to ascertain when the NSP is—and is not—insurmountable for systems that are relevant to the study of magnetic molecules.

8.2   The origin of negative signs

In Sec. 4.5, we showed that the presence of an odd number of off-diagonal, antiferromagnetic (AFM) operators will give rise to a negative value of W_i, defined in Eq. (4.21). We now wish to provide a better appreciation for how this occurs, specifically with regard to the *geometry* of spin systems. The simplest geometry for which a term W_i can be negative is an $N = 3$ ring (i.e., a triangle) with AFM interactions. This very simple geometry is shown in Fig. 8.1(a). For reasons that will soon become clear, it is helpful to represent this triangle as a one-dimensional chain with a periodic boundary condition. This is shown in Fig. 8.1(b), where the $j = 1$ spin is shown on both the left and right ends of the chain to stress the periodicity. With the system represented in one-dimension, we can now construct a two-dimensional diagram—analogous to the one in Fig. 4.2—representing a particular term W_i. Such a diagram[2] is shown in Fig. 8.1(c).

The key feature of Fig. 8.1(c), which is responsible for negative terms, is that it takes an *odd* number of operators—each of which contributes a *negative* matrix element to W_i in Eq. (4.21)—to encircle the ring. In so doing, these operators alter each spin state m_j twice— raising the quantum number once, and lowering the quantum number once—such that the same quantum numbers (0, 0, 0) appear at both the top and the bottom of the diagram; and equivalently, the same spin state, $|\psi_1\rangle = |0,0,0\rangle$, appears at both the left and right sides of a term W_i in Eq. (4.21). This is in contrast to the earlier example of an $N = 3$ open chain, shown in Fig. 4.2, for which each change (raising or lowering a spin state) could *only* be undone by

---

[2]For additional details regarding the meaning of this diagram, see the description of Fig. 4.2, provided both in its caption, and in the text of Sec. 4.4.
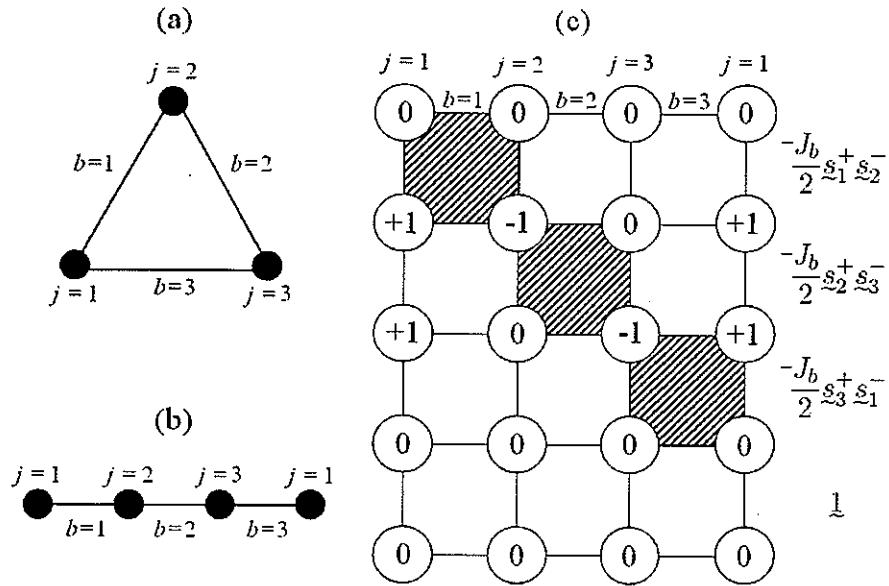
Figure 8.1   Diagrams illustrating the origin of negative terms for an AFM
triangle. (a) A triangle is shown. (b) The triangle is repre-
sented in one dimension. (c) Using this one-dimensional rep-
resentation, a diagram is shown, analogous to Fig. 4.2, which
represents a term for which $W_i < 0$.

another operator corresponding to the *same* bond (i.e., the same value of b in Fig. 4.2) as the
initial change. In other words, for an open chain, all off-diagonal operators must come in *pairs*
in order for the same spin state to appear at both the top and the bottom of the diagram.
Furthermore, each pair of operators involves matrix elements that are either both positive or
both negative (since they correspond to the same bond), so in either case their product is
positive, hence yielding only positive terms W_i.

Now suppose that the triangle consisted of only *two* AFM bonds, and one ferromagnetic
(FM) bond. It is still possible to encircle the triangle with an odd number (three) of off-
diagonal operators, but only two—not three—of the corresponding matrix elements would be
negative. Therefore, the associated term W_i would be positive, and no NSP would be present.
If only one of the triangle's bonds were AFM, and the other two were both FM, encircling
the ring with off-diagonal operators would produce one negative matrix element, so again W_i
would be negative, just as it was for the original scenario (of three AFM bonds).

Rather than a triangle, one could instead consider a ring composed of an *arbitrary* number $N > 2$ of coupled spins. As is the case with the triangle, it would again be possible to return to the same state ψ_1 by encircling the ring with N off-diagonal operators; and whether or not the corresponding term W_i is positive or negative would again depend on the number of AFM bonds—and hence the number or negative matrix elements—that are encountered in the process of going around the ring. If the ring consists of an odd number of AFM bonds, the resulting term W_i would be negative, whereas an even number of AFM bonds would yield $W_i > 0$. This finally provides a very general rule that describes whether or not the NSP is present for an arbitrary geometry:

Theorem 8.1. *A system will suffer from the NSP if and only if that system includes a polygon (or multiple polygons) composed of an odd number of AFM bonds.*

When only AFM bonds are present, it is easy to determine whether or not the NSP is present by simply looking at the geometry. If odd-sided polygons (e.g, triangles or pentagons) are present, so is the NSP.[3] When both AFM and FM bonds are present, the situation is slightly more complicated. To illustrate this, we now present three specific examples in Fig. 8.2.

Consider first the $\{Ni_{12}\}$ magnetic molecule from Sec. 7.3. Its structure is shown in Fig. 8.2(a), and was also shown in Fig. 7.11. Based on our analysis, this system includes three AFM bonds, labeled J_1, and all other bonds are FM. This structure involves many (12) triangles, but they are formed entirely from FM bonds, so they do not produce negative signs. However, the central hexagon, formed by the J_1 and J_2 bonds, involves three AFM (and three FM) bonds, so it *does* cause negative signs. Fortunately though, for this system the NSP only becomes severe for $T < 2$ K.

We have also analyzed another magnetic molecule, synthesized by R. Winpenny's group, composed of 9 Fe ions ($s = 5/2$), which we refer to as $\{Fe_9\}$.[4] This molecule consists of two "fragments"—one with three Fe ions, and one with six Fe ions—which appear to be magnetically uncoupled from one another. The structure of the 6-spin fragment is shown in

---

[3]Examples of systems composed of triangles and pentagons include the dodecahedron (pentagons), icosido-decahedron (both pentagons and triangles), and icosahedron (triangles), all of which are considered in Sec. 8.6.

[4]This analysis did not involve QMC, and has hence not been included in this thesis.
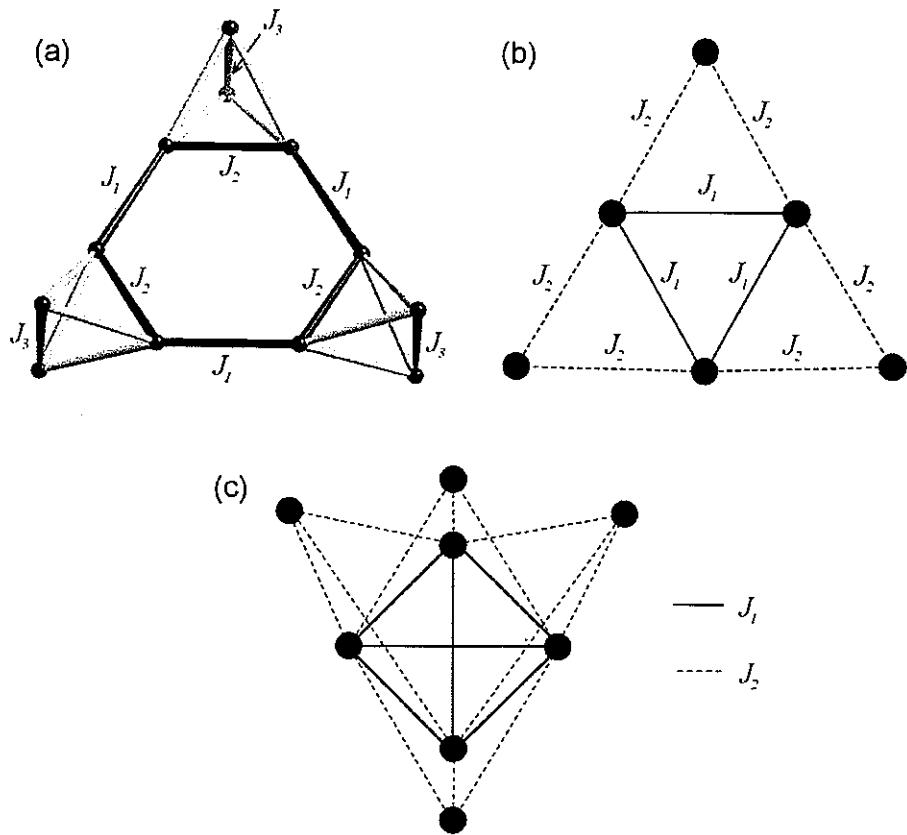
Figure 8.2 Examples of frustrated geometries. (a) The {Ni$_{12}$} magnetic molecule, described in Sec. 7.3, is described by four distinct bond strengths. Three of them (J_1, J_2, and J_3) are shown, and the fourth (J_4) corresponds to the 12 unlabeled bonds. (b) The 6-spin fragment of the {Fe$_9$} magnetic molecule is described (briefly) in the text. (c) The {Fe$_8$} magnetic molecule is also described (briefly) in the text.

Fig. 8.2(b), with the bonds labeled J_1 and J_2.[5] Our analysis has found that both J_1 and J_2 are AFM, such that each of the four triangles contribute to the NSP. For this reason, QMC calculations were not useful for the present system. However, suppose (hypothetically) that only the six bonds labeled J_2 in Fig. 8.2(b) were AFM, and the bonds labeled J_1 were FM. If this were the case, there would have been no polygons with an odd number of AFM bonds, so the NSP would not have been present.

We have also recently studied an {Fe$_8$} magnetic molecule, synthesized by E. Brechin's

---

[5]The symmetry of this molecule allows no more than two different interactions, which we label J_1 and J_2.

group at the University of Edinburgh. This system is composed of 8 Fe ions, four of which couple with one another (with a certain strength, J_1) to form an "inner" tetrahedron. The other four Fe ions each couple (with strength J_2) to three of the inner Fe ions, forming an "outer" tetrahedron. This structure is shown in Fig. 8.2(c). It appears as though both J_1 and J_2 are AFM in the $\{Fe_8\}$ magnetic molecule, such that all 16 triangles would contribute to the NSP. However, if (hypothetically again) J_1 had been FM, then the inner four triangles would have had only FM bonds, and the outer 12 triangles would have each had two (not three) AFM bonds. Therefore, this hypothetical system would not have suffered from the NSP.

For each of the examples shown in Fig. 8.2, one could also construct larger "reducible" polygons by including certain edges—and omitting other edges—of multiple "primitive" polygons. In Fig. 8.2(a), these primitive polygons include the central hexagon, as well as the 12 triangles; and in Figs. 8.2(b) and 8.2(c), the only primitive polygons are triangles. Although one *could* construct reducible polygons, it is not necessary to do so in order to determine whether or not the NSP is present for the following reason: If the NSP is not present (as determined by Theorem 8.1) in any of the primitive polygons, then it will also not be present in any reducible polygons. One can easily be convinced of this by considering a handful of examples, such as those included in Fig. 8.2.

Finally, it is interesting to note that the condition which determines whether or not the NSP is present, stated in Theorem 8.1, is *identical* to the condition which determines whether or not a system is "frustrated". A frustrated spin system can be distinguished by placing a *classical* spin vector at the site of each (quantum) spin, and attempting to simultaneously satisfy (minimize the energy of) all bonds. To do so, spins that interact via FM bonds should be parallel, and spins that interact via AFM bonds should be anti-parallel. If such an arrangement is not possible, the system is said to be frustrated. If, on the other hand, such an arrangement is possible, it can be realized in the following manner:

(1) Pick a starting spin, and let it point in a particular direction.[6]

(2) Next, move to a spin that is connected to the previous spin via a FM (AFM) bond, and

---

[6]Both the spin and its direction can be chosen arbitrarily.

let it be parallel (anti-parallel) to the previous spin.

(3) Repeat step (2) until all spins have been reached.

Note that in step (2), an AFM bond requires the next spin to be "flipped" relative to the direction of its predecessor. However, if an odd number of AFM bonds leads back to the same spin (i.e., if there is a polygon with an odd number of AFM bonds), then an odd number of spin-flips would occur, contradicting a direction that had already been determined, and hence meaning that the system is frustrated. This is perfectly analogous to the previous discussion in which an AFM bond "flipped" the *sign* of the term W_i, hence causing the NSP: An odd number of sign-flips—resulting from an odd number of AFM bonds—cause the NSP; whereas an odd number of spin-flips—resulting from an odd number of AFM bonds—cause frustration. For recent discussions regarding the effects of frustration in classical and quantum systems, see Refs. [2] and [3], respectively.

8.3 Modified sampling strategy

This section describes how it is possible to generalize the thermodynamic averages that take the form of Eq. (8.1)—specifically the quantities discussed in Sec. 4.3—in such a way that Eq. (8.2) involves only *positive* probabilities. This generalization is the topic of the following subsection, and leads to formulas that involve the ratios of multiple QMC estimates. As a result, it was also necessary to estimate our statistical errors in a slightly different manner than that which was presented in Sec. 3.5. This is described in Sec. 8.3.2.

8.3.1 Averaging when negative terms are present

When negative terms W_i [defined in Eq. (4.21)] are present in the partition function, Eq. (8.2) can give the result $\frac{P_{j \to i}}{P_{i \to j}} < 0$, which is clearly inconsistent with our desire to interpret $P_{j \to i}$ and $P_{i \to j}$ as the probabilities of making transitions. However, as we describe in the following, it is possible—and in fact quite straightforward—to rewrite Eq. (8.1) in such a way that these transition probabilities depend upon the *absolute values* of the terms W_i, and

not the terms themselves, some of which have negative values. First, one can trivially replace[7] $W_i = \text{sgn}(W_i)|W_i|$ in Eq. (8.1), giving

$$\langle Q \rangle = \frac{\sum_i Q_i \text{sgn}(W_i)|W_i|}{\sum_i \text{sgn}(W_i)|W_i|}. \qquad (8.3)$$

Now, upon dividing both the numerator and the denominator by $\sum_i |W_i|$, Eq. (8.3) becomes

$$\langle Q \rangle = \left(\frac{\sum_i Q_i \text{sgn}(W_i)|W_i|}{\sum_i |W_i|} \right) \left(\frac{\sum_i \text{sgn}(W_i)|W_i|}{\sum_i |W_i|} \right)^{-1}, \qquad (8.4)$$

such that the two fractions in Eq. (8.4) each have the same form as Eq. (8.1). In fact, the *only* difference between these fractions and the fraction in Eq. (8.1) is that now $|W_i|$ (which is always positive) appears in the both the numerators and denominators, rather than W_i (which is sometimes negative). To clarify this analogy to Eq. (8.1), we now define

$$\langle Q_i \rangle' \equiv \frac{\sum_i Q_i |W_i|}{\sum_i |W_i|} \qquad (8.5)$$

to represent averages that are weighted by $|W_i|$—as opposed to W_i—and Eq. (8.4) becomes

$$\langle Q \rangle = \frac{\langle Q_i \text{sgn}(W_i) \rangle'}{\langle \text{sgn}(W_i) \rangle'}, \qquad (8.6)$$

which is a ratio of two averages. The numerator is the average of the quantity $Q_i \text{sgn}(W_i)$, and the denominator is the average of $\text{sgn}(W_i)$.

The result of the preceding algebra is that Eq. (8.1) can be replaced by Eq. (8.6). This is advantageous because each average $\langle \cdots \rangle'$, defined in Eq. (8.5), can be calculated by taking Monte Carlo steps corresponding to the detailed balance condition,

$$\frac{P_{j \to i}}{P_{i \to j}} = \frac{|W_i|}{|W_j|}, \qquad (8.7)$$

which involves only *positive* values on the right-hand side, and hence positive probabilities on the left-hand side.[8]

In order to calculate thermodynamic properties, we must now rewrite the averages from Sec. 4.3 in the form of Eq. (8.6). This is a straightforward task, but to do so it is helpful

---

[7]The function $\text{sgn}(W_i)$ simply returns the "sign" of W_i: $\text{sgn}(W_i) = 1$ for $W_i > 0$, and $\text{sgn}(W_i) = -1$ for $W_i < 0$.

[8]Although Eq. (8.6) has the desired property that all probabilities are positive, negative terms W_i cause *undesired* cancellation problems in the numerator and denominator, as we describe in Sec. 8.4.

to introduce notation that is more compact. Specifically, for every QMC calculation that we perform, there are five quantities whose averages are calculated for each bin.[9] They are henceforth defined using the following notation:[10]

$$\mathcal{S} \equiv \frac{1}{N_{si}} \sum_{i=1}^{N_{si}} \langle \operatorname{sgn}(W_i) \rangle' \tag{8.8a}$$

$$\mathcal{M} \equiv \frac{1}{N_{si}} \sum_{i=1}^{N_{si}} \langle g\mu_B M_S \operatorname{sgn}(W_i) \rangle' \tag{8.8b}$$

$$\mathcal{M}_{sq} \equiv \frac{1}{N_{si}} \sum_{i=1}^{N_{si}} \langle (g\mu_B M_S)^2 \operatorname{sgn}(W_i) \rangle' \tag{8.8c}$$

$$\mathcal{N} \equiv \frac{1}{N_{si}} \sum_{i=1}^{N_{si}} \langle n_i \operatorname{sgn}(W_i) \rangle' \tag{8.8d}$$

$$\mathcal{N}_{sq} \equiv \frac{1}{N_{si}} \sum_{i=1}^{N_{si}} \langle n_i^2 \operatorname{sgn}(W_i) \rangle' \tag{8.8e}$$

In these definitions, N_{si} represents the number of (statistically independent) bins that are used, and the quantities $\langle \cdots \rangle'$ are averages, calculated using Eq. (8.7), for each of the N_{si} bins. In terms of these newly defined averages, the formulas from Sec. 4.3 become

$$M = \frac{\mathcal{M}}{\mathcal{S}} \tag{8.9a}$$

$$\chi = \frac{1}{k_B T} \left(\frac{\mathcal{M}_{sq}}{\mathcal{S}} - \frac{\mathcal{M}^2}{\mathcal{S}^2} \right) \tag{8.9b}$$

$$U = -k_B T \frac{\mathcal{N}}{\mathcal{S}} - c \tag{8.9c}$$

$$C = k_B \left(\frac{\mathcal{N}_{sq}}{\mathcal{S}} - \frac{\mathcal{N}^2}{\mathcal{S}^2} - \frac{\mathcal{N}}{\mathcal{S}} \right), \tag{8.9d}$$

where the constant c was defined in Sec. 4.5.

We should stress an important distinction between this strategy and the strategy that was described at the end of Chapter 3: Here, we do *not* estimate the thermodynamic quantities— M, χ, U, and C—for each bin. Instead, we estimate the five quantities shown in Eq. (8.8) for each bin; average over the N_{si} bins to obtain $\mathcal{S}$, $\mathcal{M}$, $\mathcal{M}_{sq}$, $\mathcal{N}$, and $\mathcal{N}_{sq}$; and then use these five estimates to calculate M, χ, U, and C from Eq. (8.9). This distinction might seem somewhat

---

[9]The use of "bins" for calculating Monte Carlo averages was discussed in Sec. 3.5.2.

[10]In situations that different values of g are present in a system, one should replace $g\mu_B M_S \rightarrow \frac{\mu_B}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k}$ in Eqs. (8.8b) and (8.8c). This replacement was mentioned at the beginning of Chapter 7, and has been derived in Appendix A.

trivial, but is actually quite important when $\langle \text{sgn}(W_i) \rangle'$ becomes small.[11] Since all four of these thermodynamic quantities involve division by $\mathcal{S}$, it is *crucial* that a precise estimate of this quantity be used, which is accomplished by averaging over all bins in Eq. (8.8a).

The formulas shown in Eq. (8.9), which use the estimates obtained from Eq. (8.8), are what we use for our calculations. Note, however, that each of these formulas depend on either two [in the cases of Eqs. (8.9a) and (8.9c)] or three [in the cases of Eqs. (8.9b) and (8.9d)] estimates, each of which have some associated statistical error. The uncertainties in our final estimates subsequently depend on the statistical errors associated with Eq. (8.8), as well as the covariance between these estimates, as we describe in the following subsection.

8.3.2 Estimating statistical errors with the negative sign problem

Recall from Sec. 3.5 that the error associated with N_{si} independent estimates of a quantity Q is given by

$$\delta Q = \sqrt{\frac{\overline{Q^2} - \overline{Q}^2}{N_{si}}}, \qquad (8.10)$$

where we now use overbars to denote averages. Eq. (8.10) can be used to estimate the uncertainties in our estimates of $\mathcal{S}$, $\mathcal{M}$, $\mathcal{M}_{sq}$, $\mathcal{N}$, and $\mathcal{N}_{sq}$, since each of these quantities are averages over N_{si} statistically independent bins. However, the thermodynamic quantities—as they are written in Eq. (8.9)—are functions of multiple (estimated) variables. Therefore, to estimate the uncertainty in these quantities, we use the following:[4]

Consider a quantity[12] Q, which is a function of multiple other quantities,[13] q_1, q_2, etc. Furthermore, suppose that these other quantities have each been obtained from N_{si} statistically independent estimates, represented $q_{1,i}$, $q_{2,i}$, etc., where $1 \leq i \leq N_{si}$. The error in Q is then given by

$$\delta Q = \sqrt{\sum_{j,k} \frac{\partial Q}{\partial q_j} \frac{\partial Q}{\partial q_k} \sigma_{j,k}}, \qquad (8.11)$$

---

[11]If, for a particular bin, the uncertainty in $\langle \text{sgn}(W_i) \rangle'$ were larger than the quantity itself, division by this quantity could produce nearly infinite estimates of the thermodynamic properties, which would clearly be disastrous to a calculation.

[12]Q could represent M, χ, U, or C.

[13]These quantities could represent $\mathcal{S}$, $\mathcal{M}$, $\mathcal{M}_{sq}$, $\mathcal{N}$, or $\mathcal{N}_{sq}$.

where the summation includes *all* permutations of two quantities (q_j and q_k), from the set $\{q_1, q_2,$ etc.$\}$. For example, when calculating χ, there are three quantities ($\mathcal{S}$, $\mathcal{M}$, and $\mathcal{M}_{sq}$) corresponding to q_j and q_k, so there would be 9 terms in the summation. The symbol $\sigma_{j,k}$ represents the element of the variance-covariance matrix corresponding to the quantities q_j and q_k, and is defined,

$$\sigma_{j,k} \equiv \frac{1}{N_{si}} \sum_{i=1}^{N_{si}} \left(q_{j,i} - \overline{q_j}\right)\left(q_{k,i} - \overline{q_k}\right). \tag{8.12}$$

Finally, to determine the uncertainties that are associated with our estimates in Eq. (8.9), we differentiate each formula in Eq. (8.9), and insert this result—as well as the results of our calculations from Eq. (8.12)—into Eq. (8.11). This yields the following formulas for our uncertainties:

$$\delta M = \left|\frac{\mathcal{M}}{\mathcal{S}}\right| \sqrt{\frac{\sigma_{\mathcal{S},\mathcal{S}}}{\mathcal{S}^2} + \frac{\sigma_{\mathcal{M},\mathcal{M}}}{\mathcal{M}^2} - 2\frac{\sigma_{\mathcal{S},\mathcal{M}}}{\mathcal{S}\mathcal{M}}} \tag{8.13a}$$

$$\delta \chi = \frac{1}{k_B T}\left|\frac{\mathcal{M}}{\mathcal{S}}\right| \sqrt{\begin{array}{c} \xi_m^2 \frac{\sigma_{\mathcal{S},\mathcal{S}}}{\mathcal{S}^2} + \lambda_m^2 \frac{\sigma_{\mathcal{M},\mathcal{M}}}{\mathcal{M}^2} + \frac{\sigma_{\mathcal{M}_{sq},\mathcal{M}_{sq}}}{\mathcal{M}^2} \\ -2\lambda_m \frac{\sigma_{\mathcal{M},\mathcal{M}_{sq}}}{\mathcal{M}^2} + 2\xi_m \frac{\sigma_{\mathcal{S},\mathcal{M}_{sq}}}{\mathcal{S}\mathcal{M}} - 2\lambda_m \xi_m \frac{\sigma_{\mathcal{S},\mathcal{M}}}{\mathcal{S}\mathcal{M}} \end{array}} \tag{8.13b}$$

$$\delta U = k_B T \left|\frac{\mathcal{N}}{\mathcal{S}}\right| \sqrt{\frac{\sigma_{\mathcal{S},\mathcal{S}}}{\mathcal{S}^2} + \frac{\sigma_{\mathcal{N},\mathcal{N}}}{\mathcal{N}^2} - 2\frac{\sigma_{\mathcal{S},\mathcal{N}}}{\mathcal{S}\mathcal{N}}} \tag{8.13c}$$

$$\delta C = k_B \left|\frac{\mathcal{N}}{\mathcal{S}}\right| \sqrt{\begin{array}{c} \xi_n^2 \frac{\sigma_{\mathcal{S},\mathcal{S}}}{\mathcal{S}^2} + \lambda_n^2 \frac{\sigma_{\mathcal{N},\mathcal{N}}}{\mathcal{N}^2} + \frac{\sigma_{\mathcal{N}_{sq},\mathcal{N}_{sq}}}{\mathcal{N}^2} \\ -2\lambda_n \frac{\sigma_{\mathcal{N},\mathcal{N}_{sq}}}{\mathcal{N}^2} + 2\xi_n \frac{\sigma_{\mathcal{S},\mathcal{N}_{sq}}}{\mathcal{S}\mathcal{N}} - 2\lambda_n \xi_n \frac{\sigma_{\mathcal{S},\mathcal{N}}}{\mathcal{S}\mathcal{N}} \end{array}} \tag{8.13d}$$

It is precisely *these* formulas that we have used to calculate our uncertainties in M, χ, U, and C, where λ_m, ξ_m, λ_n, and ξ_n are defined as,

$$\lambda_m \equiv 2\frac{\mathcal{M}}{\mathcal{S}} \tag{8.14a}$$

$$\xi_m \equiv \lambda_m - \frac{\mathcal{M}_{sq}}{\mathcal{M}} \tag{8.14b}$$

$$\lambda_n \equiv 1 + 2\frac{\mathcal{N}}{\mathcal{S}} \tag{8.14c}$$

$$\xi_n \equiv \lambda_n - \frac{\mathcal{N}_{sq}}{\mathcal{N}}. \tag{8.14d}$$

8.4 Cancellation of positive and negative contributions

Thus far in this chapter, we have first described how it is that negative signs emerge. Then—given the existence of negative signs—we have also described how Monte Carlo sampling can still take place, using $|W_i|$ (rather than W_i) to determine transition probabilities. However, we have not yet clarified why the negative sign problem is actually a *problem*. This is the purpose of the present section.

The problem lies in the fact that—especially at low temperatures—almost complete cancellation can take place between positive and negative terms, leading to a poor "signal-to-noise" ratio. This can be most easily understood by examining the frequency with which different types of terms W_i are encountered, and observing how this frequency varies with temperature. For example, let us again consider the icosidodecahedron, where a spin ($s = 1/2$) resides at each vertex, and each edge connects a pair of spins that interacting via an AFM bond of strength J. This model was used in the analysis of the $\{V_{30}\}$ magnetic molecule in Sec. 7.4, and the geometry was shown in Fig. 7.14(b). In Fig. 8.3 we now display histograms for this system (analogous to the histogram shown in Fig. 4.8) which include the relative number of times a given order n of the series expansion was encountered in 10^7 MCS; i.e., a value of 10^{-2} means that this n was encountered 10^5 times. Note that we have further distinguished whether the order was encountered with a positive term ($W_i > 0$) or a negative term ($W_i < 0$), and the *difference* between the positive a negative data is included as well.

Although it is not obvious by looking at Eqs. (8.9c) and (8.9d), the data labeled "difference" in Fig. 8.3 actually determine both U and C. To calculate U, $\frac{\mathcal{N}}{S}$ is nothing more than the average value of the data labeled difference (i.e., the location of the peak). Similarly, to calculate C, $\frac{\mathcal{N}_{sq}}{S} - \frac{\mathcal{N}^2}{S^2}$ is the variance of the data labeled "difference" (i.e., the square of the width of the distribution). Therefore, an accurate estimate of U and C depend on accurate estimates of this distribution, represented by a thick line in Fig. 8.3.

First, consider that data in Fig. 8.3(d), corresponding to a temperature[14] $k_B T/J_c = 1.8$. At this (relatively high) temperature, the terms $W_i < 0$ occur much less frequently than the

---

[14]The parameter $J_c \equiv s(s + 1)J$ was described in Sec. 6.2.1.
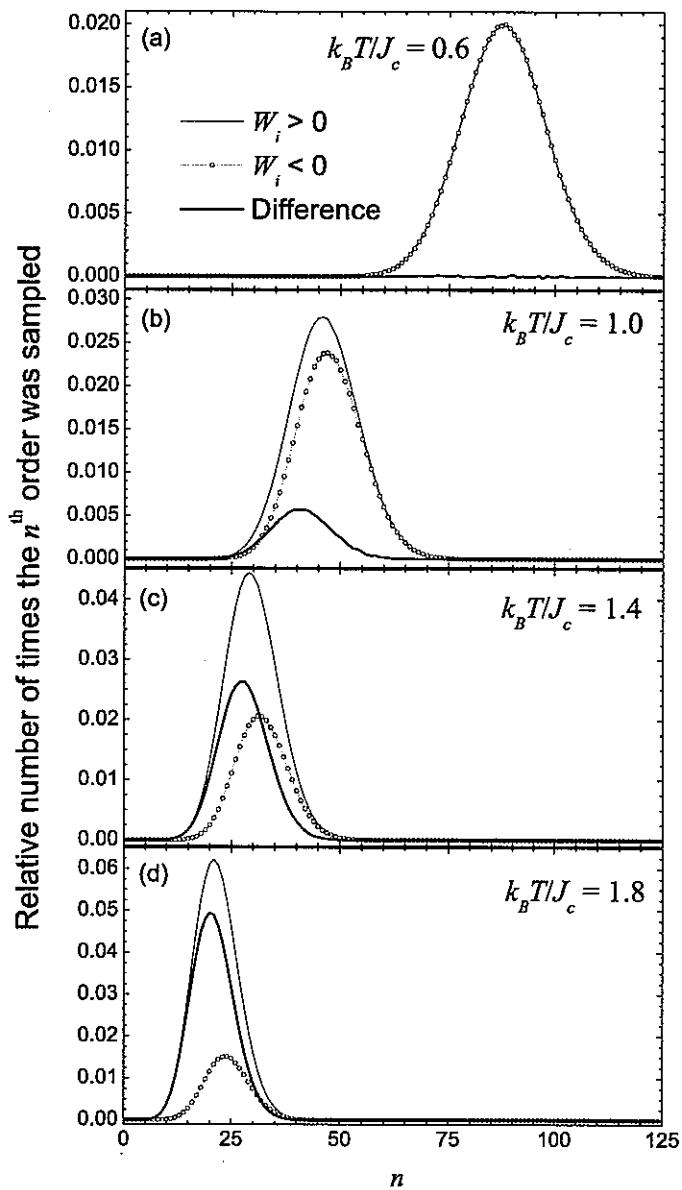
Figure 8.3   Histograms showing the relative number of times that the nth
order was encountered for the $s = 1/2$ AFM icosidodecahedron
(i.e., $\{V_{30}\}$) for four different temperatures. (a) At the lowest
temperature, $k_B T/J_c = 0.6$, virtually complete cancellation oc-
curs between positive and negative terms. (b) At $k_B T/J_c = 1.0$,
large cancellation occurs, but considerable information remains.
(c) At $k_B T/J_c = 1.4$, some cancellation occurs; and (d) at
$k_B T/J_c = 1.8$, the cancellation is relatively small.

terms $W_i > 0$, such that the difference between these data yields the smooth, robust curve shown with the thick line. At $k_B T/J_c = 1.4$, shown in Fig. 8.3(c), the difference between the $W_i > 0$ and $W_i < 0$ data is not as great, but the difference between these data is still quite smooth. As the temperature is lowered to $k_B T/J_c = 1.0$, shown in Fig. 8.3(b), the cancellation between the $W_i > 0$ and $W_i < 0$ data becomes more substantial; and slight "wiggles" become visible in the difference, demonstrating that the statistical errors are no longer small compared with the the data that remains after the cancellation. Finally, in Fig. 8.3(a), we include data for $k_B T/J_c = 0.6$. At this temperature, there is almost complete cancelation between the terms $W_i > 0$ and $W_i < 0$, such that the difference is indistinguishable from the statistical fluctuations. At this temperature and below, useful averages cannot be obtained in a reasonable number of Monte Carlo steps—and a correspondingly reasonable amount of time—so the negative sign problem is indeed a very serious problem.

Inspecting Eqs. (8.9a) and (8.9b), the calculation of M and χ are perfectly analogous to the calculation of U and C, respectively. In particular, we can construct histograms that are analogous to Fig. 8.3, but which show the values of M_S that have been encountered, rather than the values of n. These histograms are shown in Fig. 8.4, corresponding to the same four temperatures that were represented in Fig. 8.3. The data labeled "difference" in Fig. 8.4 now provide the values[15] of M (from the average of the distribution) and χ (from the width of the distribution) for the $s = 1/2$ AFM icosidodecahedron. These data were collected during the same 10^7 Monte Carlo steps that provided the data in Fig. 8.3, and again the cancellation (between the $W_i > 0$ data and the $W_i < 0$ data) becomes more severe with decreasing temperature. At $k_B T/J_c = 0.6$, the remaining data are (again) indistinguishable from the statistical fluctuations, leading to very poor estimates of M and χ, and very large error bars.

---

[15]Since $H = 0$ for this calculation, the data are (of course) symmetrically distributed about $M_S = 0$.
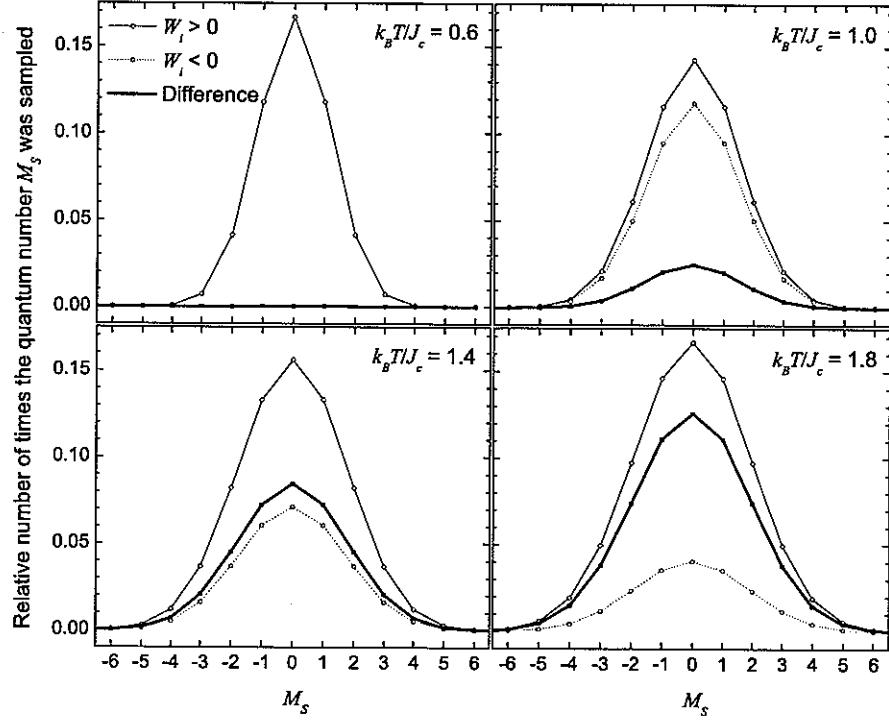
Figure 8.4   Histograms showing the relative number of times that the quan-
tum number M_S was encountered for the $s = 1/2$ AFM icosi-
dodecahedron (i.e., $\{V_{30}\}$) for four different temperatures. (a)
At the lowest temperature, $k_B T/J_c = 0.6$, virtually complete
cancellation occurs between positive and negative terms. (b)
At $k_B T/J_c = 1.0$, large cancellation occurs, but considerable
information remains. (c) At $k_B T/J_c = 1.4$, some cancellation
occurs; and (d) at $k_B T/J_c = 1.8$, the cancellation is relatively
small.

8.5   Attempts to overcome the negative sign problem

The NSP provides a major challenge to QMC calculations for *many* different types of

systems. It is not unique to the particular representation of the partition function that we

have used (i.e., Handscomb's method, described in Sec. 4.2.2), as it also occurs for the Trotter-

Suzuki method[5] (described briefly in Sec. 4.2.1). In addition to spin systems—which have

been the focus of the present work—the NSP also occurs for other types of QMC calculations

as well: When performing calculations that involve interacting fermions on a lattice, the NSP is

present in all but a few special cases.[6, 7] Furthermore, the NSP arises in the QMC calculation

of path integrals for electrons, necessitating certain "fixed-node" approximations.[16]

Since the NSP is so pervasive, a general means of avoiding the problem would be very desirable. One possibility that people have considered is to choose a different basis (rather than the "z-states", introduced in Sec. 2.2) to represent the system of interest. This suggestion can be motivated by observing that *if* the eigenstates of the Hamiltonian had been chosen as the basis states, *then* all matrix elements would have been diagonal, so the NSP would not have been present. However, using the eigenstates cannot be considered a "solution" to the negative sign problem, since their determination has a complexity that grows exponentially with the system size—exactly what one uses QMC to *avoid.*[17] Although the eigenstates of the Hamiltonian would not be a useful basis to choose, in certain (very *specific*) cases, the choice of a different basis has proven to be useful. These include the use of the x-states (analogous to the z-states, defined in Sec. 2.2) for a two-dimensional triangular lattice with certain anisotropic interactions,[9] and the use of a "dimer basis" for a spin-ladder, also with specific types and strengths of anisotropic interactions.[10]

In some other (also very specific) situations, the NSP can be avoided because there are positive portions of the partition function which exactly cancel the negative portions of the partition function. These are referred to as "merons", and their existence was originally recognized and exploited in the context of interacting fermions on a lattice.[11] However, this method can also be applied to XXZ spin models, given by the Hamiltonian

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \left[s_i^z s_j^z + \Delta \left(s_i^x s_j^x + s_i^y s_j^y \right) \right],$$ (8.15)

eliminating the NSP if (and *only* if) $\Delta = -1$.[12] More recently, it has been shown[13] that the meron strategy can also be used for the more general situation of $-1 \leq \Delta < 1$. For $\Delta > -1$, the NSP is reduced, but not eliminated, by the introduction of merons. This reduction is very significant for Δ close to -1, but vanishes continuously as Δ approaches 1. Note that for $\Delta = 1$ the XXZ model becomes the Heisenberg model; hence this strategy is of no help (i.e., merons do not exist) for the Heisenberg model.

---

[16]See, for example, Ref. [8] and the references therein.

[17]This point was stressed in Ref. [1].

To summarize, strategies to overcome the NSP have been found—and successfully implemented—for certain, specific systems for which the NSP is present. However, it was recently shown in Ref. [1] that the NSP is a so-called "NP hard" problem; and since it is generally believed (although unproven) that no polynomial time solution to NP problems exists, a general solution to the NSP also does not exist. Furthermore, the authors of Ref. [1] also argue that a solution to the NSP for the frustrated Heisenberg model—in which we are interested—is "almost certainly impossible" based upon the physics of the problem. Therefore, we are interested in assessing the types of magnetic molecules, and the temperature ranges, for which the NSP does, and does not, allow us to perform calculations in a reasonable amount of time. This is the subject of the following section.

8.6 Assessing the limitations of quantum Monte Carlo for magnetic molecules

From the discussion of Sec. 8.2, we are able to quickly "diagnose" whether or not the NSP is present in any given magnetic molecule, provided we are able to make some initial assumptions as to the sign (AFM or FM) of the interactions. If the NSP is not present, QMC can be used, and accurate results can be obtained, for arbitrarily low temperatures.[18] If, on the other hand, the NSP is present, then results can only be obtained above some minimum temperature, which depends upon both the geometry of the system and the strength of the interactions. In these situations, it is important to have some (albeit approximate) sense of what this minimum temperature might be. Then, given this rough estimate of the minimum temperature, we can assess whether—for the particular system—the use of QMC would be: very useful (as was the case for the $\{Ni_{12}\}$ magnetic molecule described in Sec. 7.3); of some—although limited—usefulness (as was the case for the $\{V_{30}\}$ magnetic molecule described in Sec. 7.4); or not at all useful.

In order to get some sense of the temperature range for which QMC is useful, first consider rings of N spins, coupled via AFM interactions, where N is odd. Recalling Sec. 8.2, such

---

[18]As we described in Sec. 4.7.4, the computation time grows with decreasing temperature proportional to $1/T$. However, in the absence of the NSP, results can be obtained in a reasonable amount of time for $T \ll 1$ K.
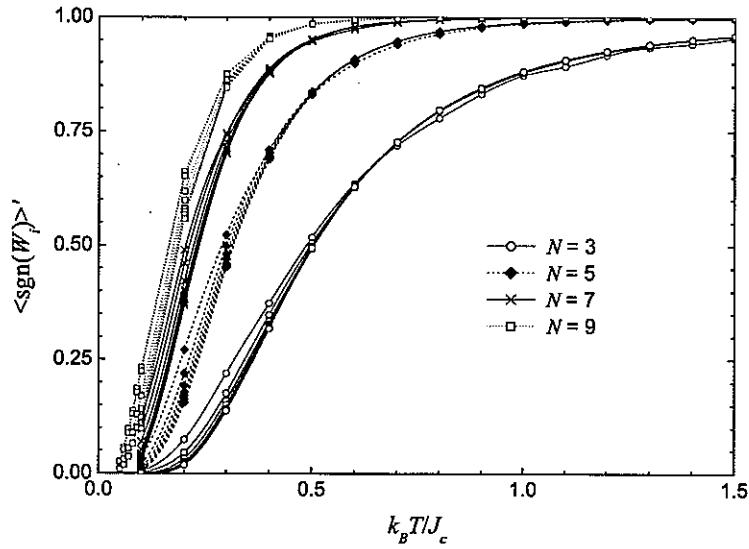
Figure 8.5 $\langle\text{sgn}(W_i)\rangle'$ vs. $k_B T/J_c$ for rings of $N = 3$, 5, 7, and 9 spins. For each N, data are included for $s = 1/2$, 1, ..., 9/2. In all cases, the uppermost data correspond to $s = 1/2$, and the data proceed systematically toward the $s = 9/2$ (lowermost) data. Note that the temperature at which the negative sign problem becomes serious decreases with increasing ring size, N.

systems will suffer from the NSP, so we wish to know the minimum accessible temperatures for systems such as these. To that end, we have calculated the average sign of the terms W_i that are encountered for these systems (i.e. $\langle\text{sgn}(W_i)\rangle'$), and plot that quantity in Fig. 8.5 as a function of temperature for all intrinsic spins from $s = 1/2$ to $s = 9/2$.

There are a number of interesting features to be noted in Fig. 8.5. First, note that in all cases $\langle\text{sgn}(W_i)\rangle'$ decreases with decreasing temperature, eventually dropping toward zero.[19] Although this drop toward zero is ubiquitous, the temperature at which it occurs varies considerably with N. Specifically, for larger values of N, lower temperatures can be achieved before $\langle\text{sgn}(W_i)\rangle'$ drops toward zero; or in other words, for larger N rings, QMC is useful to lower temperatures than it is for small N rings. This can be (at least qualitatively) understood in terms of diagrams such as Fig. 8.1. For the triangle, the sign of a term W_i is determined by a mere three matrix elements; while at low temperatures, the *total* number of matrix elements

---

[19]This is of course consistent with Sec. 8.4.

will be much larger. Hence, the *magnitude* of W_i is affected only a small amount by the three matrix elements that determine the *sign* of W_i, resulting in nearly perfect cancellation when $n \gg 3$. For larger rings, more matrix elements are involved in determining whether or not $W_i > 0$, so larger n (and hence lower T) are necessary in order to achieve the same cancellation.

The other very interesting feature of Fig. 8.5 is that for each N, the data rapidly approach a *single* limiting curve as $s \rightarrow \infty$, when plotted as a function of $k_B T / J_c$. [The quantity $J_c \equiv s(s+1)$ was introduced in Sec. 6.2.1.] In one sense this could be expected, since the physics of the problem scales with[20] J_c as $s \rightarrow \infty$. However, it is interesting that QMC calculations are limited to a minimum temperature (at which $\langle \text{sgn}(W_i) \rangle' \rightarrow 0$) in the limit $s \rightarrow \infty$, even though this limitation—or an analogous limitation—is absent in a classical Monte Carlo calculation of the same system. This point seems to be closely related to the details of how this classical ($s \rightarrow \infty$) limit is achieved, a fundamental issue that is still not sufficiently well understood.

In addition to rings, magnetic molecules also often adopt the structure of polyhedra. We have therefore performed a similar analysis (to that which was just provided for rings) for three ployhedra for which the presence of all AFM bonds gives rise to the NSP. These structure include the dodecahedron, icosidodecahedron [that was shown in Fig. 7.14(b)], and icosahedron; and the corresponding results are shown in Fig. 8.6. Again, calculations have been performed for all s in the range from $1/2$ to $9/2$; and again, for each geometry, a single curve is approached as $s \rightarrow \infty$. However, unlike the data in Fig. 8.5, these data are plotted on semilog scales to underscore how dramatically $\langle \text{sgn}(W_i) \rangle'$ plummets toward zero as the temperature is lowered.

When examining Fig. 8.6, first note that for the dodecahedron [Fig. 8.6(a)] the value $\langle \text{sgn}(W_i) \rangle'$ remains large down to a significantly lower temperature, as compared with the data for the other two structures. From the previous results of the rings, this is to be expected since the dodecahedron is composed of only pentagons, whereas the other two structures include triangles. For all three of these structures though, the value of $k_B T / J_c$ at which the data rapidly plummet toward zero is larger than the corresponding temperature for rings. In other

---

[20]See Sec. 6.2.1.
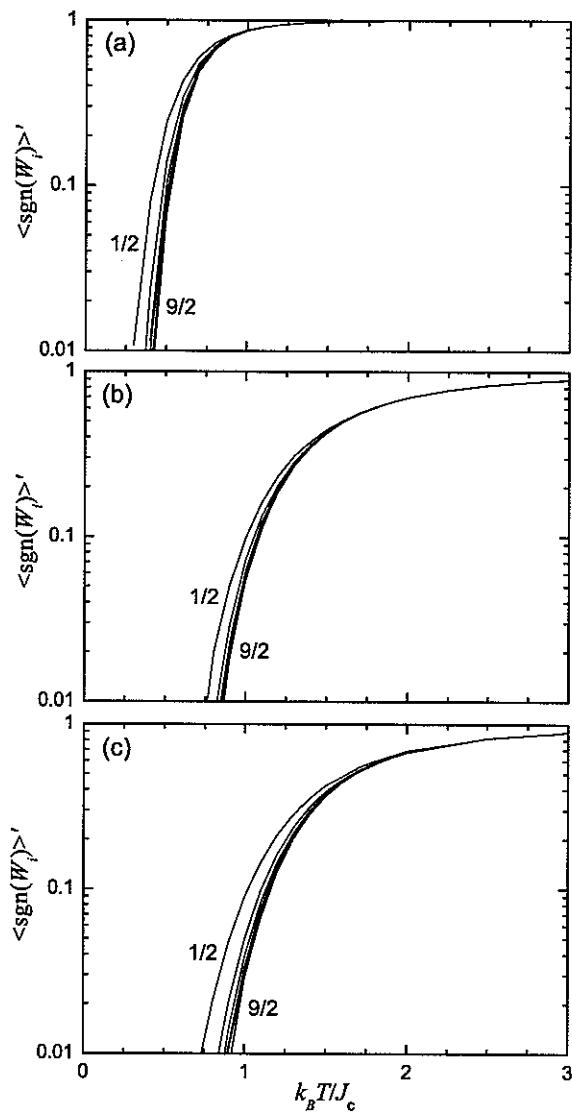
Figure 8.6 $\langle \mathrm{sgn}(W_i) \rangle'$ vs. $k_B T / J_c$ for three frustrated polyhedra: the (a) dodecahedron, (b) icosidodecahedron, and (c) icosahedron. In all cases, results are included for $s = 1/2$, 1, ..., $9/2$. The $s = 1/2$ and $s = 9/2$ curves are labeled, and the other data proceed systematically toward the $s = 9/2$ data with increasing s.

words, a system composed of *multiple* frustrated polygons will stuffer from the NSP at a *higher* temperature than a system with only the single frustrated polygon. In this respect, the icosahedron [Fig. 8.6(c)] can be viewed as a "worst case scenario", since it is composed of entirely triangles, allowing QMC calculations only above $k_B T/J_c \gtrsim 1.0$.

One should also carefully observe that the $s = 1/2$ data shown in Fig. 8.6(b) is indeed consistent with the results that have been presented in Secs. 7.4 and 8.4. In Sec. 7.4, $T\chi$ was calculated for the $\{V_{30}\}$ magnetic molecule (i.e., the $s = 1/2$ icosidodecahedron) for $k_B T/J \gtrsim 1/2$, or equivalently, $k_B T/J_c \gtrsim 2/3$. At that temperature, $\langle \text{sgn}(W_i)\rangle'$ is *extremely* small ($< 10^{-2}$), so it was not possible to extend to lower temperatures without prohibitively long computations. (This important connection between $\langle \text{sgn}(W_i)\rangle'$ and computation time will be clarified shortly.) The values of $\langle \text{sgn}(W_i)\rangle'$ can also be (at least approximately) inferred from the histograms that were shown in Figs. 8.3 and 8.4. In both of these figures, examine the data corresponding to $k_B T/J_c = 1.4$. At this temperature, the useful data—remaining after cancellation—is somewhat larger than the $W_i < 0$ data, implying[21] that $\langle \text{sgn}(W_i)\rangle'$ is somewhat larger than 1/3, which is consistent with Fig. 8.6(b).

Finally, it is important to be clear as to the important connection between the value of $\langle \text{sgn}(W_i)\rangle'$ and the computation time that is necessary in order to obtain useful results. When we obtain a value of $\langle \text{sgn}(W_i)\rangle' = 10^{-2}$, that means that 99% of the data that was recorded during the QMC calculation "disappeared" due to cancellation between the positive and negative values of W_i. For a calculation that *would* take one second,[22] were it not for the NSP, 100 seconds—i.e., a factor of $1/\langle \text{sgn}(W_i)\rangle'$ longer—would instead be necessary in order to obtain the same statistical errors. For a single temperature, this means investing less than 2 minutes, which is certainly not unreasonable. However, it is clear in Fig. 8.6 that by the time $\langle \text{sgn}(W_i)\rangle'$ has become this small, a small decrease in temperature causes $\langle \text{sgn}(W_i)\rangle'$ to become exponentially smaller, leading to exponentially longer computation times.

As an example of this increase in computation time, we again consider the AFM icosido-

[21]If the data for $W_i < 0$ exactly coincided with the "difference" that remained after cancellation, then one would find $\langle \text{sgn}(W_i)\rangle' = 1/3$.

[22]One second is a typical computation time for the calculation of a single temperature, $k_B T \approx 1$, when the NSP is not present.
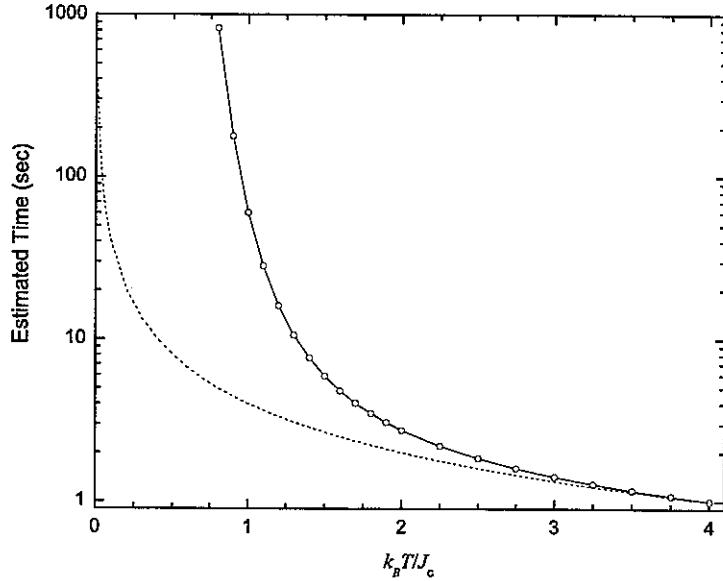
Figure 8.7   The estimated time that would be necessary in order to perform
calculations with fixed (small) statistical error-bars for the AFM
icosidodecahedron, of arbitrary spins s, as a function of k_BT/J_c.
The solid curve with symbols was obtained using Eq. (8.16), as
described in the text, and takes into account the temperature
dependence of $\langle \mathrm{sgn}(W) \rangle'$. The dashed line shows the computa-
tion time that would be necessary for the same system if the
NSP were *not* present (e.g., the FM icosidodecahedron), and
the computation time were proportional to $(k_BT/J_c)^{-1}$.

decahedron. For $k_BT/J_c = 4.0$, $\langle \mathrm{sgn}(W_i) \rangle' \approx 1$, and very small statistical errors were obtained
with approximately one second of computation. This computation time is roughly independent
of s for any fixed value of k_BT/J_c; and as k_BT/J_c is varied, $\langle \mathrm{sgn}(W_i) \rangle'$ depends only weakly
on s. Therefore, in Fig. 8.7 we are able to plot estimated computation times[23] for an icosido-
decahedron of *arbitrary* spins s. The solid curve was obtained by starting with the observation
that one second of computation time was involved for $k_BT/J_c = 4.0$, for which $\langle \mathrm{sgn}(W_i) \rangle' \approx 1$.
Then, the computation time for lower temperatures was estimated to be,

$$\text{computation time} \approx \frac{1 \text{ sec}}{4 \times \frac{k_BT}{J_c} \times \langle \mathrm{sgn}(W_i) \rangle'}, \tag{8.16}$$

accounting for both the $J_c/(k_BT)$ scaling of the computation time, described at the end of

---

[23]Since $\langle \mathrm{sgn}(W_i) \rangle'$ does have some weak dependence on s, Fig. 8.7 was produced by averaging $\langle \mathrm{sgn}(W_i) \rangle'$ for
all values of s in the range $1/2 \le s \le 9/2$.

Chapter 4, and the factor of $1/\langle\text{sgn}(W_i)\rangle'$, described above. The dashed curve assumes instead that $\langle\text{sgn}(W_i)\rangle' = 1$ (i.e., that there is no NSP) and describes the $J_c/(k_BT)$ scaling. This dashed data can be thought of as representing the same (icosidodecahedron) but with all FM interactions instead of all AFM interactions.

Note in Fig. 8.7 that the dashed curve (without the NSP) does not exceed one minute of computation time per temperature until $k_BT/J_c \ll 1$. This is to be contrasted with the solid curve, which takes the NSP into account, and exceeds one minute for all $k_BT/J_c \lesssim 1$. As the temperature is lowered below $k_BT/J_c = 1$, this computation time then grows *so* rapidly that days of computation would be necessary in order to obtain reasonable data for $k_BT/J_c = 0.5$; and for yet lower temperatures, the estimated computation time would soon grow to years. Not only would such calculations be impractical from the standpoint of time, but would also be potentially problematic in terms of numerical precision. As the cancellation becomes nearly perfect, many significant figures would be necessary in order to retain the important data, and could exceed the numerical precision of the computer being used. The combination of these factors hence make it very important to have some initial estimation of the feasibility of QMC calculations, as provided in this chapter.

Bibliography

[1] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[2] R. Moessner and A. P. Ramirez, Physics Today **59(2)**, 24 (2006).

[3] G. Misguich, and C. Lhuillier, cond-mat/0310405.

[4] J. R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (University Science Books, 1982).

[5] M. Takasu, S. Miyashita, and M. Suzuki, Prog. Theor. Phys. **75**, 1254 (1986).

[6] J. E. Hirsch, R. L. Sugar, D. J. Scalapino, and R. Blankenbecler, Phys. Rev. B **26**, 5033 (1982).

[7] J. E. Hirsch, Phys. Rev. Lett. **51**, 1900 (1989).

[8] J. Shumway, in *Computer Simulations Studies in Condensed Matter Physics XVII*, edited by D. P. Landau, S. P. Lewis, and H. B. Schtter (Springer Verlag, 2006), pp. 181–195.

[9] N. Hatano and M. Suzuki, Phys. Lett. A **163**, 246 (1992).

[10] T. Nakamura, Phys. Rev. B **57**, R3197 (1998).

[11] S. Chandrasekharan and U.-J. Wiese, Phys. Rev. Lett. **83**, 3116 (1999).

[12] P. Henelius and A. W. Sandvik, Phys. Rev. B **62**, 1102 (2000).

[13] S. Bergkvist, P. Henelius, and A. Rosengren, Phys. Rev. E **68**, 016122 (2003).

CHAPTER 9. Conclusions and outlook

9.1 Capabilities and limitations

As we have described in Sec. 2.4, and shown schematically in Fig. 2.1, the theoretical methods that have been previously employed in studying magnetic molecules each have a serious limitation, making calculations very challenging—and sometimes impossible—for certain magnetic molecule systems. These limitations provided the primary motivation for the present work, which has focused on adding an additional method (that has not been previously employed in the study of these types of systems) to the "toolbox" that is used for studying magnetic molecules. To this end, we have implemented a quantum Monte Carlo (QMC) algorithm—described in Chapter 4, and based upon the principles from Chapters 2 and 3—to perform calculations for complex systems of interacting quantum spins. This method is applicable to arbitrarily large quantum Hilbert spaces D (defined in Sec. 2.2), enabling, for example, calculations for which $D \gg 10^{100}$ in Chapter 6, while simultaneously avoiding any systematic approximations.

In our current implementation of the QMC method, we have included the capability to study all values of intrinsic spins in the range $s \leq 9/2$, which has proven quite adequate, both for the study of general systems (in Chapters 5 and 6) and for the analysis of specific magnetic molecules (in Chapter 7). Although we have not studied values of $s > 9/2$, there is no fundamental limitation precluding such calculations. The implementation would simply require the calculation of more matrix elements (see Sec. 4.7.2), and the Monte Carlo updating would subsequently involve more directed loop diagrams (see Sec. 4.7.3). Therefore, calculations for larger values of s will be a topic for future study, if motivated by either the synthesis of new systems for which $s > 9/2$, or open theoretical questions for which $s \leq 9/2$ is not sufficient.

This method is also applicable to arbitrary combinations of different ions—such as the $\{Cr_{10}Cu_2\}$ and $\{Cr_{12}Ni_3\}$ magnetic molecules that are discussed in Secs. 7.1 and 7.2—by sampling the quantity that is derived in Appendix A. Furthermore, calculations are possible (at least to some extent) for arbitrary geometries. However, when the "negative sign problem" (NSP) is present, we are restricted to some minimum temperature, which is determined by both the geometry and the interactions of a given system. In some situations this minimum temperature can be relatively low (e.g., < 2 K for the $\{Ni_{12}\}$ magnetic molecule of Sec. 7.3), such that QMC calculations are able to provide a great deal of useful information, in spite of the NSP. For other systems though, the minimum temperature can be rather large (e.g., > 100 K for the $\{V_{30}\}$ magnetic molecule of Sec. 7.4), in which case QMC calculations are of limited usefulness. When the NSP is not present, calculations for larger systems and/or lower temperatures can proceed by simply investing more computation time and/or more computer processors, as we have described in Sec. 4.7.4. However, when the NSP occurs, the investment of additional computational resources quickly becomes impractical, as we have described at the end of Sec. 8.6. This limitation is very important, and has hence been discussed in detail in Chapter 8.

Throughout Chapters 5–8, results have been presented for various systems, all of which have been described by the *isotropic* Heisenberg Hamiltonian,[1]

$$\mathcal{H} = \sum_{\langle j,k \rangle} J_{j,k} \vec{s}_j \cdot \vec{s}_k + \mu_B \vec{H} \cdot \sum_{j=1}^{N} g_j \vec{s}_j. \tag{9.1}$$

However, we have also implemented calculations which include certain (very specific) types of anisotropic interactions, because (as we describe below) their inclusion required only minimal modifications to the method that was described in Chapter 4. Upon including these anisotropies, the most general Hamiltonian for which we are currently able to perform QMC calculations is

$$\mathcal{H} = \sum_{\langle j,k \rangle} J_{j,k} \left[s_j^z s_k^z + \Delta_{j,k} \left(s_j^x s_k^x + s_j^y s_k^y \right) \right] + \sum_{j=1}^{N} \left[\mu_B H g_j s_j^z + D_j \left(s_j^z \right)^2 \right], \tag{9.2}$$

---

[1]This is identical to Eq. (2.11), with the exception that we now allow for g_j to vary from site to site.

which includes exchange anisotropy (if $\Delta_{j,k} \neq 1$) and single-ion anisotropy (if $D_j \neq 0$). Note though that the direction of these anisotropies is currently restricted to the z-axis, defined by the direction of the external magnetic field. Given this restriction, these anisotropies do not provide a realistic description of magnetic molecules, and hence results have only been included for Eq. (9.1).

The reason that these anisotropies are restricted to the z-axis lies in the details of Chapter 4. A value of $\Delta_{j,k} \neq 1$ would change the numerical prefactor in Eq. (4.9b), and hence also Eqs. 4.43 and 4.44. Similarly, a nonzero D_j would change the values of the *existing* diagonal matrix elements in Eq. (4.42). These changes would then yield different numerical values of W_i in Eq. (4.21), but our QMC calculations would be otherwise unaffected. Consider, however, a more general anisotropy, such as single-ion anisotropy with a direction[2] that varies from site to site, $\sum_j D_j \left(\vec{e}_j \cdot \vec{\underline{s}}_j \right)^2$. This term would produce *new* off-diagonal operators and matrix elements, which would be different (than the ones described in Sec. 4.7.2) in that they would not conserve the quantum number M_S (defined in Sec. 2.2). For this reason, the construction of directed loops (described in Secs. 4.6.2 and 4.6.3) would need to be greatly modified. The implementation of such modifications should be feasible, and has been mentioned in Refs. [1] and [2]. The inclusion of these interactions will thus be a topic of future work, and should allow us to study of other types of magnetic molecules, for which the isotropic Heisenberg model is known to be insufficient.

9.2   General discussion

In conclusion, we have implemented the QMC method described in Chapter 4; and it has proven very useful for the study of many interesting systems. In particular, the magnetization of Heisenberg rings was studied in detail, as described in Chapter 5. This study revealed the striking result that for large values of intrinsic spin ($s \geq 3/2$), the low-temperature magnetization is *independent* of s, with the exception of a particular (power-law) scaling relation, as shown in Fig. 5.7. Also described in Chapter 5, we have determined the energy gap between

---

[2]This direction is specified in the following by the unit vector $\vec{e}_j$, which can depend on the index j.

the ground state and first excited state for these systems, from which we have inferred the functional dependence of this quantity on the rings size N for several values of s. These data are shown in Fig. 5.8, providing results that are immediately applicable to experiments involving magnetic molecules, and are much more accurate than the assumptions that are typically employed. For example, the results of Chapter 5 were applied to the $\{Fe_{12}\}$ magnetic molecule, as described in Sec. 5.4, leading to an improved estimated of the exchange energy for this system.

In Chapter 6, the zero-field susceptibility was calculated for a wide variety of structures, for the purpose of determining the temperature range for which a classical spin model (and hence classical Monte Carlo calculations) will provide an accurate approximation to a system of *quantum* spins. Our findings are shown in Figs. 6.3 and 6.5, wherein we show that a classical model provides accurate results when the thermal energy $(k_B T)$ is several times the exchange energy[3] J_s. Specifically, for the values of s for which classical approximations are typically used $(3/2 \leq s \leq 5/2)$, classical and quantum results are in good agreement for $k_B T / J_s \gtrsim 5$, if the system is described by antiferromagnetic interactions, whereas this temperature is considerably higher $(k_B T / J_s \gtrsim 10)$ for ferromagnetic interactions.

In Chapter 7, we have reported the results of our analysis (which relied upon QMC calculations) for four different species of magnetic molecules. Two of these systems, $\{Cr_{10}Cu_2\}$ and $\{Cr_{12}Ni_3\}$, were previously analyzed using approximate theoretical methods, which we have shown to have produced poor results. By instead performing QMC calculations—which do *not* introduce such approximations—we have reanalyzed these systems, obtaining good agreement with the existing experimental data, and providing predictions for future experiments. The third species of magnetic molecules, $\{Ni_{12}\}$, was newly synthesized, so had not been previously analyzed using an approximate method. The process of analyzing this system differed from that of $\{Cr_{10}Cu_2\}$ and $\{Cr_{12}Ni_3\}$ in a number of other respects as well: (1) For this system, there are four different types of bonds, each of which (as determined by our analysis) turn out to give rise to different strengths of interactions. Hence, there was a corresponding four-dimensional

---

[3]Here we assume that all values of $J_{j,k}$ in Eq. 9.1 have the same value, J_s.

parameter space to be searched in order to determine the Hamiltonian that best describes these magnetic molecules. (2) The NSP is present for this system,[4] so QMC calculations were limited to $T \gtrsim 2$ K. This did not hinder the calculation of (and comparison with) weak-field susceptibility data, but prohibited the use of QMC to calculate the low-temperature magnetization. (3) For this system, $D < 10^6$, so calculations could also be performed by diagonalizing the Hamiltonian, although this required several hours of computation for any given point in parameter space. Due to the combination of these three factors, $\{Ni_{12}\}$ provides an ideal example of a system for which QMC and diagonalization effectively *complement* one another. We first performed QMC calculations for thousands of different points in parameter space, and compared these results with the measured susceptibility data in order to determine the ideal Hamiltonian. Then, the resulting Hamiltonian was diagonalized in order to obtain the associated energy spectrum, and the predicted low-temperature magnetization. The fourth magnetic molecule reported in Chapter 7 was $\{V_{30}\}$. For this system, the NSP prohibited QMC calculations[5] for $T \lesssim 120$ K, which allowed us to determine the average strength of the interactions, but did not allow a more detailed analysis.

As we have already mentioned, the negative sign problem provides a serious limitation to QMC calculations for certain geometries of magnetic molecules. It is therefore important to understand whether or not the NSP will be present for a given system, a description of which has been provided in Sec. 8.2. Furthermore, if the NSP is present, it is then necessary to know whether or not the desired calculations will still be feasible. To that end, some typical examples have been provided in Sec. 8.6. Finally, it should be stressed once again that these calculations are applicable only to *equilibrium* properties, e.g., the magnetization and magnetic susceptibility, of magnetic molecules. However, it has been shown[3] that a similar QMC strategy can be used to estimate Green's functions, from which dynamical correlation functions can be derived. The ability to perform such calculations would be quite valuable, and will be explored as another possibility for future research. In conclusion, the quantum Monte Carlo method that has been implemented and used for this thesis can provide a great deal

---

[4]A description of why the negative sign problem is present for this system has been provided in Sec. 8.2.
[5]See Chapter 8 for further details.

of information for certain systems (although sometimes limited by the NSP) and for certain quantities (although not yet applicable to dynamic calculations). It has therefore proven to be an effective *tool*, although other tools are certainly more appropriate for certain tasks.

Bibliography

[1] O. F. Syljuåsen, Phys. Rev. E **67**, 046701 (2003).

[2] F. Alet, S. Wessel, and M. Troyer, Phys. Rev. E **71**, 036706 (2005).

[3] A. Dorneich, and M. Troyer, Phys. Rev. E **64**, 066701 (2001).

APPENDIX A. Derivation of the magnetization for dissimilar ions

In this appendix we derive a formula for the magnetization which is applicable to situations in which dissimilar ions—and hence multiple values of g—are present in the *same* magnetic molecule. This formula (and its proof) could not be found in the literature, but is absolutely vital in order for the method that was described in chapter 4 to be applicable to magnetic molecules. For example, this situation occurs in the $\{Cr_{10}Cu_2\}$ and $\{Cr_{12}Ni_3\}$ molecules (which are described in Secs. 7.1 and 7.2, respectively), hence providing the motivation for this derivation. In order to be useful, we specifically desired the magnetization in a form that is readily amenable to importance importance sampling—i.e., in the form of Eq. (4.17)—such that the methods described in chapters 3 and 4 could be immediately applied. This desired result is eventually obtained in Eq. (A.15c), as we describe in the following. Repeating the same process a second time, an analogous result for dM/dH is obtained, which is provided in Eq. (A.16)

In chapter 2, the general form of the magnetization M was given in terms of the partition function Z as

$$M = \frac{1}{\beta Z} \frac{\partial Z}{\partial H}. \tag{A.1}$$

The form of the partition function that we have used for our QMC calculations was given in Eq. (4.21), which we use here as well, but in (at least initially) a slightly different notation. Note that this new notation has been adopted solely because it is more convenient for the present purpose. In particular, we now write the partition function as

$$Z = \sum_{i'} W_{i',n'}, \tag{A.2}$$

with each term given by

$$
\begin{aligned}
W_{i',n'} &= \frac{\beta^{n'}(L'-n')!}{L'!} \langle\psi_1|\mathcal{H}_{b_1}|\psi_2\rangle\langle\psi_2|\mathcal{H}_{b_2}|\psi_3\rangle\cdots\langle\psi_{L'}|\mathcal{H}_{b_{L'}}|\psi_1\rangle \qquad\text{(A.3a)} \\
&= \frac{\beta^{n'}(L'-n')!}{L'!}\pi_{i',n'}, \qquad\text{(A.3b)}
\end{aligned}
$$

where $\pi_{i',n'}$ represents the product of n' matrix elements in the term $W_{i',n'}$. Eq. (A.3) is identical to Eq. (4.21), with the following (notational) exceptions: Here, we truncate the expansion at $L' \equiv L + 1$ total operators, with $n' \equiv n + 1$ bond operators, and $n' \equiv n + 1$ corresponding matrix elements.[1] We also have now labeled each term $W_{i',n'}$ with the additional index n', indicating the number of matrix elements that it contains. Finally, in Eq. (A.3) we have not distinguished between the diagonal and off-diagonal portions of the bond operators $\mathcal{H}_{b_k}$. Again, this is merely a more convenient notation,[2] since, for a given pair of states, ψ_k and ψ_{k+1}, the matrix element $\langle\psi_k|\mathcal{H}_{b_k}|\psi_{k+1}\rangle$ will be *either* diagonal *or* off-diagonal, but obviously not both.

Using the present notation, and combining the previous equations, the magnetization is

$$
M = \frac{1}{\beta Z}\sum_{i'}\frac{\partial W_{i',n'}}{\partial H}. \qquad\text{(A.4)}
$$

Differentiating $W_{i',n'}$ with respect to H yields many zeros (from both the identity operators, and the off-diagonal matrix elements), and n'_d non-zero terms, where we now define n'_d to be the number of *diagonal* matrix elements in the term $W_{i',n'}$. Differentiating these matrix elements gives

$$
\frac{\partial\langle\psi_k|\mathcal{H}_{b_k}|\psi_{k+1}\rangle}{\partial H} = \mu_B\frac{g_{b_k}(1)m_{b_k}(1)}{\xi_{b_k}(1)} + \frac{g_{b_k}(2)m_{b_k}(2)}{\xi_{b_k}(2)} \qquad\text{(A.5)}
$$

for each term. There are n'_d such terms, so

$$
\frac{\partial W_{i',n'}}{\partial H} = \mu_B\frac{\beta^{n'}(L'-n')!}{L'!}\sum_k\left[\left(\frac{g_{b_k}(1)m_{b_k}(1)}{\xi_{b_k}(1)} + \frac{g_{b_k}(2)m_{b_k}(2)}{\xi_{b_k}(2)}\right)\pi_{i,n}\right], \qquad\text{(A.6)}
$$

---

[1]The $L' - n'$ identity operators have matrix elements that are equal to unity, so we henceforth use the term "matrix elements" to refer *only* to the remaining bond operator matrix elements, not the identity operator matrix elements.

[2]When one derives Eq. (4.21), as we showed in chapter 4, it is desirable to distinguished between the diagonal and off-diagonal operators simply for the sake of convenience during the diagonal update (described in Sec. 4.6.1). For our present purpose, it is instead more convenient to leave the full bond operators $\mathcal{H}_{b_k}$.

where the summation over k includes n'_d values of k. Now, in Eq. (A.6), $\pi_{i,n}$ represents the product of the n matrix elements that remain after the kth matrix element has been "removed" as we have shown in Eq. (A.5). Inserting Eq. (A.6) into Eq. (A.4), the magnetization becomes

$$M = \frac{\mu_B}{\beta Z} \sum_{i'} \frac{\beta^{n'}(L'-n')!}{L'!} \sum_k \left[\left(\frac{g_{b_k}(1) m_{b_k}(1)}{\xi_{b_k}(1)} + \frac{g_{b_k}(2) m_{b_k}(2)}{\xi_{b_k}(2)} \right) \pi_{i,n} \right]. \qquad (A.7)$$

Note that the product $\pi_{i,n}$ also appears in a *different* term $W_{i,n}$ (as opposed to $W_{i',n'}$) with only n matrix elements (instead of n' matrix elements) and truncated at a total of L operators (instead of L' operators). This other term is

$$W_{i,n} = \frac{\beta^n (L-n)!}{L!} \pi_{i,n} \qquad (A.8)$$

from Eq. (4.21), and is a part of the partition function written

$$Z = \sum_i W_{i,n}. \qquad (A.9)$$

By comparing Eq. (A.7) with Eqs. (A.8) and (A.9), we now wish to show how the magnetization can be written in terms of *these* terms, $W_{i,n}$. Some of the (apparent) differences between Eqs. (A.7) and (A.8) are trivial to rectify. Namely, $\frac{\beta^{n'}}{\beta} \to \beta^n$ and $(L'-n')! \to (L-n)!$. The *challenge* is then that we must be very careful to properly treat the two sums; i.e., the sum over all terms i' in Eq. (A.7), and the sum over all i in Eq. (A.9). This is addressed in the following:

(a) Note first that for any *fixed* k, the sum over i' includes *all* N_b possible values that the bond b_k can take in the matrix element $\langle \psi_k | \mathcal{H}_{b_k} | \psi_k \rangle$. Writing this out explicitly, we have

$$M = \frac{\mu_B}{Z} \sum_{i'_a} \frac{\beta^n (L-n)!}{L!} \frac{1}{L'} \sum_k \left[\sum_{b_k=1}^{N_b} \left(\frac{g_{b_k}(1) m_{b_k}(1)}{\xi_{b_k}(1)} + \frac{g_{b_k}(2) m_{b_k}(2)}{\xi_{b_k}(2)} \right) \pi_{i,n} \right], \qquad (A.10)$$

where the sum $\sum_{i'_a}$ does *not* include the N_b different values of b_k that were included in the sum $\sum_{i'}$. Note however that

$$\sum_{b_k=1}^{N_b} \left(\frac{g_{b_k}(1) m_{b_k}(1)}{\xi_{b_k}(1)} + \frac{g_{b_k}(2) m_{b_k}(2)}{\xi_{b_k}(2)} \right) = \sum_{j=1}^{N} g_j m_{j,k}, \qquad (A.11)$$

where g_j is the g-value for the jth spin site, and $m_{j,k}$ is the quantum number corresponding to the operator $\underline{s}_j^z$, acting on the state $|\psi_k\rangle$, of the term W_i. The magnetization therefore becomes

$$M = \frac{\mu_B}{Z} \sum_{i'_a} \frac{\beta^n (L-n)!}{L!} \frac{1}{L'} \sum_k \left(\sum_{j=1}^{N} g_j m_{j,k} \pi_{i,n} \right). \qquad (A.12)$$

(b) Next, we must take into account the various values of k (being summed over in $\sum_{i'_a} \sum_k$) that can result in the creation of a term $W_{i,n}$ with the quantum numbers $m_{j,k}$. To do so, note that there are L' values of k in the sum $\sum_{i'_a}$ that produce the *same* value of $\pi_{i,n}$ (upon differentiating $\langle \psi_k | \mathcal{H}_{b_k} | \psi_{k+1} \rangle$), but multiplied by different values of $g_j m_{j,k}$, each corresponding to a matrix element in the product $\pi_{i,n}$. Writing this portion of the sum $\sum_{i'_a} \sum_k$ out explicitly, we have

$$M = \frac{\mu_B}{Z} \sum_{i'_b} \frac{\beta^n (L-n)!}{L!} \frac{1}{L'} \sum_{k=1}^{L'} \left(\sum_{j=1}^{N} g_j m_{j,k} \pi_{i,n} \right), \qquad (A.13)$$

where the sum over all of the L' possible values of k has been included explicitly in $\sum_{k=1}^{L'}$, and has been removed from $\sum_{i'_b}$.

(c) There is, however, one slight problem with Eq. (A.13). The sum over k extends over all L' states that exist in the terms $W_{i',n'}$, but there are only L states in the resulting terms $W_{i,n}$. To account for this, we note that the replacement

$$\sum_{k=1}^{L'} \rightarrow \frac{L'}{L} \sum_{k=1}^{L} \qquad (A.14)$$

will leave the total sum unchanged, as the slight reduction in the sum is compensated by the ratio $\frac{L'}{L}$. Now writing the magnetization in its final form, with the summation $\sum_i$

extending over all terms W_i, we have

$$M = \frac{\mu_B}{Z} \sum_i \frac{\beta^n (L-n)!}{L!} \frac{1}{L} \sum_{k=1}^{L} \left(\sum_{j=1}^{N} g_j m_{j,k} \pi_{i,n} \right) \tag{A.15a}$$

$$= \mu_B \frac{\sum_i \left[\left(\frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right) W_i \right]}{\sum_i W_i} \tag{A.15b}$$

$$= \mu_B \left\langle \frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right\rangle, \tag{A.15c}$$

which also appears in Eq. (7.1a).

Following all of the same steps as we showed above, the derivation of the analogous formula for dM/dH is now straightforward. The only difference is that one should begin by considering terms that include $L'' \equiv L + 2$ total states, and $n'' \equiv n + 2$ bond operators. Then, since the calculation of dM/dH involves differentiating the partition function *twice* with respect to H, the familiar terms W_i are eventually obtained. The resulting formula for dM/dH is

$$\chi = \beta \mu_B^2 \left[\left\langle \left(\frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right)^2 \right\rangle - \left\langle \frac{1}{L} \sum_{k=1}^{L} \sum_{j=1}^{N} g_j m_{j,k} \right\rangle^2 \right], \tag{A.16}$$

which also appears in Eq. (7.1b).

APPENDIX B.   Triangular Geometrical and Magnetic Motifs Uniquely Linked on a Spherical Capsule Surface

A paper published in Angewandte Chemie International Edition[1]

A. Müller,[2] A. M. Todea,[2] J. van Slageren,[3] M. Dressel,[3] H. Bögge,[2] M. Schmidtmann,[2] M. Luban,[4] L. Engelhardt,[4] and M. Rusu[5]

Polygons can be placed on spherical surfaces such that periodical structures of a cyclic nature result, while these can be considered as discrete models for two-dimensional (extended) structures. If we wish to construct a chemical structure on a spherical capsule surface in the same way, we have to remember that 1) pentagons are the basic units for sphere constructions, as is well known, for example, from virus structures, 2) they exist, for example, in the form of $\{(Mo^{VI})Mo^{VI}_5\}$ type units, and that 3) they occur in Keplerates of the type $\{(Mo^{VI})Mo^{VI}_5\}_{12}\{Linker\}_{30}$ [1-5] (linker can be of the mononuclear M (M=metal center) or dinuclear type M_2 ; for the definition of Keplerates, see ref. [5b]). However, until now it was not possible to synthesize a spherical capsule surface directly by the addition of linkers to the pentagonal units that are available in a dynamic library.[1-5] It is significant that in the Keplerates the linkers describe generic Archimedean solids: in the case of dinuclear linkers M_2 a distorted truncated icosahedron, $\{M_2\}_{30}$, and in the case of mononuclear linkers the unique icosidodecahedron (Figure B.1)[6] $\{M_{30}\}$, which has-geometrically speaking-linked M_3 triangles. Surprisingly the related consequences for chemistry have not been discussed until now. In the $\{M_{30}\}$ situation, there is a network of corner-shared trian-

[2]*Faculty of Chemistry, University of Bielefeld, 33501, Bielefeld, Germany*

[3]*Department of Physics, University of Stuttgart, 70550, Stuttgart, Germany*

[4]*Department of Physics and Astronomy & Ames Laboratory, Iowa State University, Ames, Iowa 50011, USA*

[5]*Faculty of Chemistry, Babes-Bolyai University, 3400, Cluj-Napoca, Romania*

gles on the sphere surface, this can result unique magnetic properties as in the case of the "classical" Keplerate $\{(Mo^{VI})Mo^{VI}_5\}_{12}Fe^{III}_{30}$.[7, 8] This is the first laboratory example of a "zero-dimensional" system that at low temperatures embodies characteristics of geometrical frustration/magnetic ordering[8b] which otherwise have only been observed in selected one-, two-, and three-dimensional lattice spin systems.[9] Herein we report on the spherical cluster **1a** where the twelve $\{(Mo^{VI})Mo^{VI}_5\}$ type units fix 30 d^1 V^{VI} linkers/centers with spin $S = 1/2$ in the form of an icosidodecahedron, and thus 1) demonstrating for the first time that the spherical capsule/Keplerate can be directly constructed from the mononuclear linkers and the appropriate molybdate library,[5c] 2) providing the chance to obtain new information regarding the unique molecular magnetism of the $\{M_{30}\}$ type network of linkers/triangles, and 3) clarifying the quantum effects of the spin $S = 1/2$ vanadyl linkers especially in connection with the two-dimensional $S = 1/2$ Kagomé lattice which contains linked triangles and exhibits unique magnetic properties.[9a]

After adding vanadyl sulfate to an acidified molybdate solution, in the presence of K^+ ions, compound **1** precipitates after some time in high yield. (A simpler expression for the cluster anion **1a** without referring to structural differences is given as well.)

$$Na_8K_{14}(VO)_2[\{(Mo^{VI})Mo^{VI}_5O_{21}(H_2O)_3\}_{10}\{(Mo^{VI})Mo^{VI}_5O_{21}(H_2O)_3(SO_4)\}_2$$
$$\{V^{IV}O(H_2O)\}_{20}\{V^{IV}O\}_{10}(\{KSO_4\}_5)_2] \approx 150H_2O \tag{1}$$

$$[K_{10} \subset \{(Mo^{VI})Mo^{VI}_5O_{21}(H_2O)_3(SO_4)\}_{12}\}\{(V^{IV}O)_{30}(H_2O)_{20}\}]^{26-} \tag{1a}$$
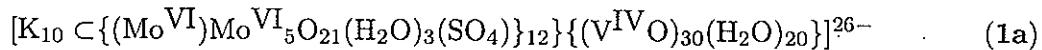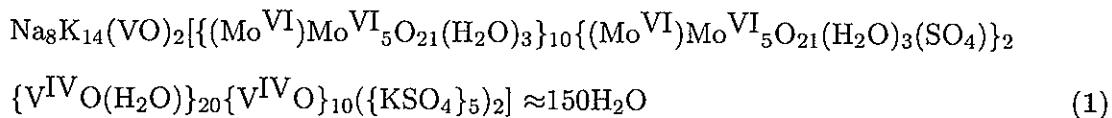
Compound 1, which crystallizes in the monoclinic space group $C2/c$, was characterized by elemental analysis, thermogravimetry (to determine the crystal water content), redox titrations (to determine the number of V^{IV} centers), spectroscopic methods (IR, Raman, UV/Vis), single-crystal X-ray structure analysis (including bond valence sum (BVS) calculations),[10] and susceptibility measurements (including related quantum Monte Carlo calculations).
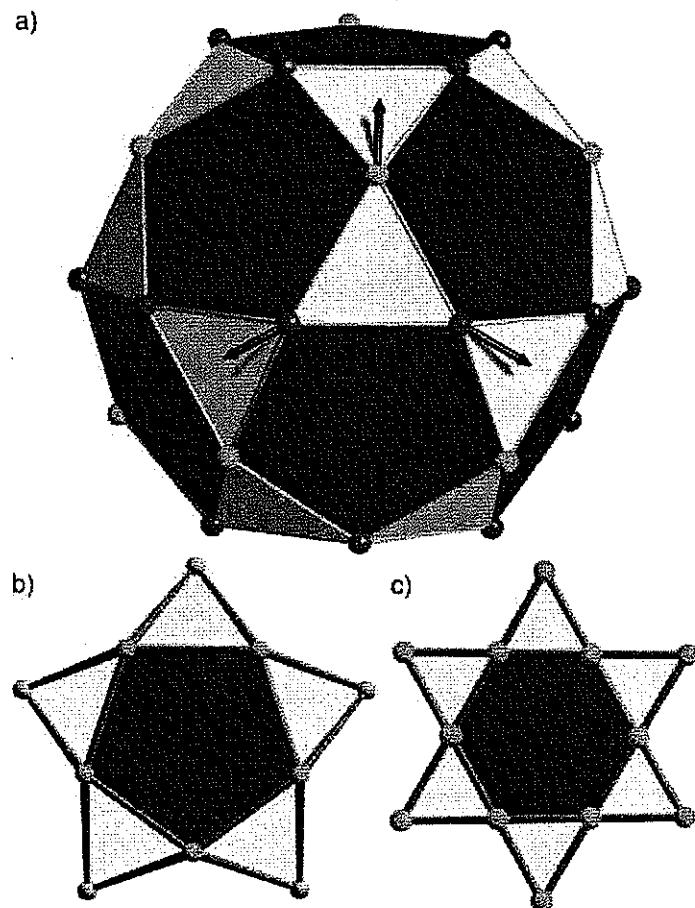
Figure B.1   a) The M centers (small colored spheres) of the $\{(Mo^{VI})Mo^{VI}{}_5\}_{12}M_{30}$ type Keplerates (e.g., M=V^{IV}, Fe^{III}) describe the icosidodecahedron shown, which is unique among the icosahedral Archimedean solids as all edges are equivalent and all dihedral angles equal. Referring to the special situation of M=Fe^{III}, there are three groups ("sublattices") of 10 spins (colors: red, blue, green), with all spins of a sublattice pointing in the same direction, while nearest-neighbor spin vectors (three are highlighted) differ in angular orientation by 120°. Also shown: b) A fragment highlighting five linked triangles around a pentagon. c) A fragment of a planar Kagomé lattice with six linked triangles around a hexagon.
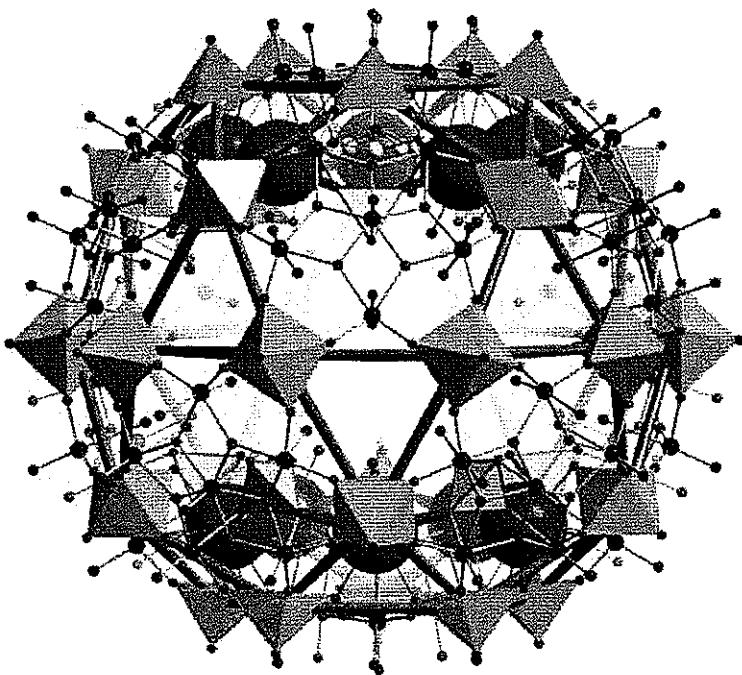
Figure B.2   Combined polyhedral and ball-and-stick representation of the
structure of 1a showing the triangles and pentagons of the icosi-
dodecahedron (green sticks), and additionally the basic {VO₅}
and {VO₆} units as green polyhedra; as the interaction between
the "lattice K⁺" ions and the 20 pores is not homogeneous, this
interaction was not considered here (blue Mo, red O, purple K,
yellow tetrahedra: {SO₄} groups; yellow spheres: disordered S
atoms).

The cluster anion **1a** of **1** is of the expected (Pentagon)$_{12}$- (Linker)$_{30}$ type and is a slightly
compressed sphere, while the heptacoordinate Mo^{VI} centers of the 12 pentagonal units corre-
spondingly describe a slightly distorted icosahedron and the 30 V^{IV} centers–acting as linkers
for the pentagonal {(Mo^{VI})Mo^{VI}$_5$} type units–describe a (slightly distorted) icosidodecahe-
dron (Figure B.2; the V^{IV}-V^{IV} distances in the distorted Archimedean solid vary from 6.3
to 6.6 Å). The distortion is in agreement with the fact that 20 V^{IV} centers in the equatorial
region have octahedral coordination and the two sets of five V^{IV} centers in the polar area have
squarepyramidal coordination; the distances from the 10 equatorial V^{IV} units to the center of
the cluster are a little shorter (10.3 Å) than the related distances of the other 20 V^{IV} units
(10.6 Å). Ten of the twelve [SO$_4$]$^{2-}$ ligands are coordinated by three oxygen atoms to three
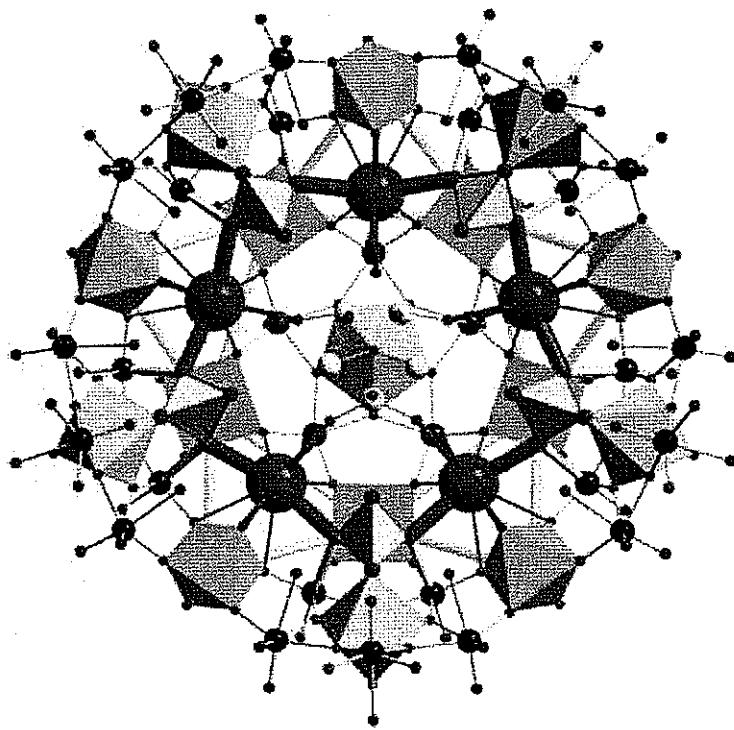
Figure B.3   Combined polyhedral and ball-and-stick representation of a
fragment of **1a** (view in direction of the C_5 axis) showing one of
the two $\{KSO_4\}_5$ rings and the coordination of $\{SO_4\}$ groups
to Mo centers as well as the disorder of one of the two sulfate
groups (color code as in Figure 2, additional light blue pentag-
onal bipyramids: $\{MoO_7\}$).

adjacent Mo^{VI} centers of the $\{(Mo^{VI})Mo^{VI}_5\}$ groups such that two $\{KSO_4\}_5$ rings parallel

to the equator result, with the K^+ ions (formally) bridging the $[SO_4]^{2-}$ ions (Figure B.3). The

other two sulfate groups are disordered and act as ligands to the two polar $\{(Mo^{VI})Mo^{VI}_5\}$

groups. The structure of **1a** comprising the twenty triangular and twelve pentagonal faces of

the icosidodecahedron built up by 30 V^{IV} centers shows an interesting relation to the much

less symmetrical cluster anion **2a** which has a non-complete spherical $\{V_3\}$ type net. In **2a**, a

strongly distorted icosidodecahedron is described by 10 Mo^{VI} and 20 V^{VI} centers, while the

equatorial $\{V_{20}\}$ belt–formed by 10 linked $\{V_3\}$ triangles–is identical to the related equatorial

segment of **1a**.

$$[\{Mo^{VI}O_3(H_2O)\}_{10}\{(Mo^{VI})Mo^{VI}{}_5O_{21}(H_2O)_3\}_{10}-$$

$$\{V^{IV}O(H_2O)\}_{20}(\{Mo^{VI}O_2(H_2O)_2\}_{5/2})_2(\{NaSO_4\}_5)_2]^{20-} \qquad (2a)$$
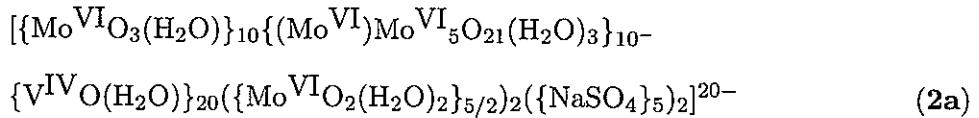
The presence of K^+ and $[SO_4]^{2-}$ ions in the reaction medium seems to be of fundamental importance for the structure formation,. as the potassium cations of the two $\{KSO_4\}_5$ rings of **1a** attract the two negatively charged polar $\{(Mo^{VI})Mo^{VI}{}_5\}$ units thus causing the slight compression of the sphere. This distortion leads to an inclination of the adjacent $\{VO_5\}$ polyhedra and thus prevents an octahedral coordination of the 10 polar V^{IV} centers. The sixth (H_2O) ligand required for octahedral coordination would be too close to the $[SO_4]^{2-}$ ligands of the $\{KSO_4\}_5$ rings.

The investigations nicely show that $\{(Mo^{VI})Mo^{VI}{}_5\}$ type units are potentially available in a dynamic polymolybdate library; remarkably, they can be "used" in the present case as virtual units in the presence of potential linkers. In aqueous solution at low pH values, the pentagonal structural unit occurs in the $[Mo^{VI}{}_{36}O_{112}(H_2O)_{16}]^{8-}$ ion,[11] which is the only(!) abundant species under those conditions. Correspondingly, **1a** is formed from that solution in the presence of VO^{2+} linkers by a "split-and-link" process with the $\{(Mo^{VI})Mo^{VI}{}_5\}$ unit being formed from the $\{Mo_{36}\}$ species after the addition of the linkers. The option to extend this to mixed-metal species such as $\{Mo^V{}_8V^{IV}{}_{22}\}$,[12] $\{Mo^V{}_4V^{IV}{}_{26}\}$, or $\{Fe^{III}{}_{22}V^{IV}{}_8\}$[13] will be reported elsewhere.

Turning to the magnetic properties, two circumstances are of pivotal importance for the possible occurrence of geometrical frustration in the type of system considered herein: First, the 30 mononuclear magnetic Keplerate linkers occupy the vertices of an icosidodecahedron, which may be pictured as 20 linked (corner sharing) triangles arranged around 12 pentagons and corresponds to an equidistant distribution of the spins on the surface of a spherical cluster; second, each magnetic center ("spin vector") interacts with its four nearestneighbors by isotropic antiferromagnetic exchange as a consequence of the special geometry of the unique $\{M_{30}\}$ type quasi-regular solid (Figure B.1). Analogous to what occurs for the Kagomé spin system (planar lattice of triangles framed around hexagons; Figure B.1c),[9] the geometric frus-

tration of the individual Keplerate can be achieved by the cooperative interactions among the full set of spin vectors. In the special case of the above mentioned $\{(Mo^{VI})Mo^{VI}_5\}_{12}Fe^{III}_{30}$, we may refer to it as a "classical" Keplerate (because of the relatively high spin, $S = 5/2$, of individual Fe^{III} centers) and the spin frustration/magnetic ordering may therefore be visualized in geometrical terms[8c] (Figure B.1). The 30 spin vectors are composed of three groups ("sublattices") of 10 spins each; all spins of a given sublattice point in the same direction, and any pair of nearest-neighbor spin vectors differ in angular orientation by 120° (Figure B.1).[8c] In **1a** we have replaced the Fe^{III} centers by VO^{2+} ions which have the much smaller spin of $S = 1/2$, that is, "quantum spins". In addition, in **1a** the 3d electrons are not "localized" at the vertices of the icosidodecahedron as is approximately the case in the $\{Fe^{III}_{30}\}$ Keplerate.[14] However, the spin frustration of these quantum spins can not be visualized in geometric terms. More generally the magnetism of the "quantum" Keplerate **1a** is expected to be significantly different from that of its classical counterpart $\{Fe_{30}\}$, and more properties are expected to emerge.

Our experimental susceptibility data versus T, recorded for an applied field of $H = 0.1$ T, and corrected for the d^1 centers of two VO^{2+} ions which are magnetically/structurally independent from the cluster skeleton **1a**, are shown in Figure B.4.[15] These results show the strong antiferromagnetic coupling in **1a**, in contrast to the $\{(Mo^{VI})Mo^{VI}_5\}_{12}Fe^{III}_{30}$ case.[16] The behavior of $T\chi$ at low T is qualitatively what could be expected for a spin system having a ground state with $S = 0$ and with very strong exchange coupling. This situation can be explained by a strong delocalization of the 3d electrons which arises because the 3d V levels are comparable in energy with the LUMOs of the molybdate fragment system.[16] This is a completely different situation than the classical $\{Fe_{30}\}$ type Keplerate where the exchange interaction is very weak, and therefore the room temperature value of $T\chi$ corresponds to 30 uncoupled $S = 5/2$ ions. The quantum Monte Carlo (QMC) method provides accurate values of susceptibility for the Heisenberg model of the $\{V_{30}\}$ system for $T > 0.5J/k_B$, and as seen in Figure B.4, a very good fit to $T\chi$ is achieved for $T > 120$ K for the choices $J/k_B = 245$ K and $g = 1.95$. Unfortunately, reliable results cannot be obtained for lower temperatures using the.
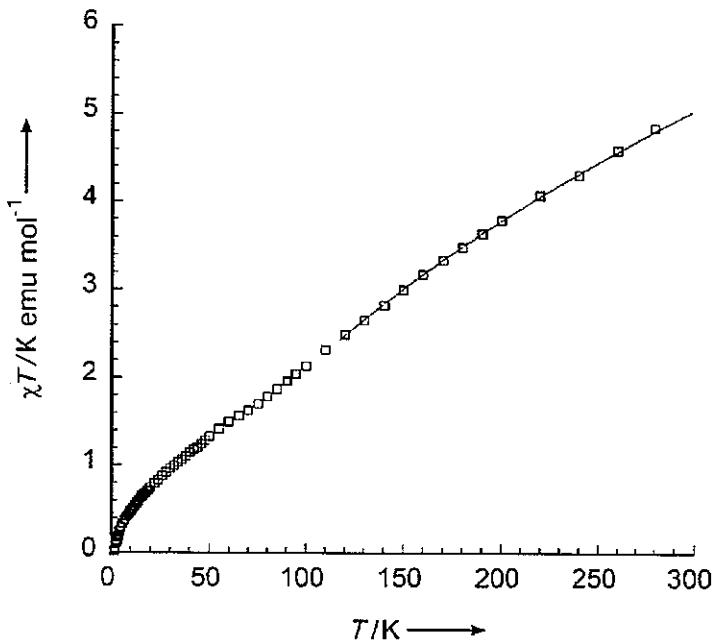
Figure B.4   Magnetic susceptibility of $\{V_{30}\}$ versus temperature: Exper-
imental data corrected for the two d^1/VO^{2+} centers ($\square$);[15]
quantum Monte Carlo results (solid curve).

QMC method owing to the "sign problem" that occurs for spin systems with antiferromagnetic
exchange based on lattice geometries where the classical counterpart exhibits spin frustration.

To summarize:We have demonstrated that it is possible to "use" pentagonal units as "build-
ing units" which play, geometrically speaking, the same role as the pentagonal units in other
sphere-based constructions, such as spherical viruses, fullerenes, and geodesic domes; in our
context they are used as a glue for trapping magnetic centers, such that triangles are linked
to form an icosidodecahedron, that is, a part of a Keplerate. As the "quantum" Keplerate **1a**
obtained is a new example of a frustrated magnetic system which shares a topological feature
with the classical Keplerate $\{(Mo^{VI})Mo^{VI}_5\}_{12}Fe^{III}_{30}$ and the Kagomé-lattice antiferromag-
net, its exploration is expected to provide a deeper understanding of basic aspects of magnetic
frustration and the role played by the size of the intrinsic spin of the interacting magnetic ions.
This study should also shed light on the parallel problem, and the focus of an intense effort, to
characterize and understand the $S = 1/2$ Kagomé lattice, which is considered to have unique

magnetic/electronic properties originating in the small value of S.[9] The behavior of the magnetization of the quantum Keplerate **1a** in high magnetic fields will be of key importance in studies aimed at elucidating the characteristics of its magnetic frustration.[17]

A solution of $VOSO_4 \cdot 5H_2O$ (2.53 g, 10 mmol) in H_2O (35 mL) was added to a stirred solution of $Na_2MoO_4 \cdot 2H_2O$ (2.42 g, 10 mmol) in H_2SO_4 (0.5m; 8 mL) in an conical flask. The resulting dark purple mixture was stirred at room temperature for 30 min (flask closed with a rubber stopper) and then treated with KCl (0.65 g, 8.72 mmol). After additional stirring for 30 min the solution was stored in the flask which was closed with a rubber stopper. After 5 days, the purple-black rhombic crystals of **1** were collected by filtration, washed with cold water, and finally dried in air. Yield: 1 g; elemental analysis: calcd (%) for $Na_8K_{24}Mo_{72}V_{32}S_{12}O_{538}H_{412}$: Na 0.96, K 4.92, V 8.55, S 2.02; found: Na 1.0, K 5.1, V 8.5, S 2.1. IR (KBr pellet): $\tilde{\nu} = 1622$ (m), ($\delta(H_2O)$) 1198 (w), 1130 (w), 1055 (w) ($\nu_{as}(SO_4)$ triplet), 964 (s) ($\nu(V=O)/\nu(Mo=O)$), 791 (vs), 631 (w), 575 (s), 449 (w) cm^{-1}; FT-Raman (solid; $\lambda_e = 1064$ nm): $\tilde{\nu} = 941$ (w, $\nu(V=O)/\nu(Mo=O)$), 872 (s, $A_{1g}O_{br}$ breathing) cm^{-1}; UV/Vis (in H_2O): $\lambda = 510$ (vs), 689 (w), 845 (w) nm.

Bibliography

[1] A. Müller, E. Beckmann, H. Bögge, M. Schmidtmann, A. Dress, Angew. Chem. 2002, 114, 1210 1215; Angew. Chem. Int. Ed. 2002, 41, 1162-1167.

[2] a) In this context we should refer first to the fact that molybdenum oxide building units-abundant as virtual ingredients in a dynamic library in aqueous solution-display a unique type of flexibility and that the units can adjust their size and shape by a type of "split-and-link process" depending on the relevant boundary conditions.[1] (Pentagonal $\{(Mo^{VI})Mo^{VI}_5\}$ building blocks, for instance, occur in spherical capsules/Keplerates of the $\{Pentagon\}_{12}\{Linker\}_{30}$ type.[1-5]) The flexibility is nicely shown by the reaction of $[PMo_{12}O_{40}]^3$ with Fe^{III} leading to a fragmentation of the Keggin ion while the fragments and the Fe^{III} ions form the (present type of) spherical system $\{(Mo^{VI})Mo^{VI}_5\}_{12}Fe^{III}_{30}$ which has the pentagonal units and encapsulated non-decomposed Keggin ions; see b)

A. Müller, S. K. Das, H. Bögge, M. Schmidtmann, A. Botar, A. Patrut, Chem. Commun. 2001, 657-658.

[3] A. Müller, S. Roy, Coord. Chem. Rev. 2003, 245, 153-166.

[4] A. Müller, S. Roy in The Chemistry of Nanomaterials: Synthesis, Properties and Applications (Eds.: C. N. R. Rao, A. Müller, A. K. Cheetham), Wiley-VCH, Weinheim, 2004, pp. 452-475.

[5] a) A. Müller, P. Kögerler, C. Kuhlmann, Chem. Commun. 1999, 1347-1358; b) A. Müller, P. Kögerler, A. W. M. Dress, Coord. Chem. Rev. 2001, 222, 193-218; c) The related direct synthetic conditions, simply based on addition of the finally appearing linker to the relative medium, are reported herein for the first time. Earlier Keplerate syntheses were more complex.

[6] The (quasi-regular) icosidodecahedron is a hybrid of the icosahedron and the dodecahedron and thus contains 20 trigonal and 12 pentagonal faces which intersect at each corner. Unlike the other related Archimedean solids, such as the $\{M_{60}\}$-type truncated icosahedron, it comprises not only equivalent vertices but also equivalent edges and equal dihedral angles (Figure B.1); see a) M. O'Keeffe, B. G. Hyde, Crystal Structures, I. Patterns and Symmetry, Mineralogical Society of America, Washington, DC, 1996; b) H. S. M. Coxeter, Introduction to Geometry,Wiley, New York, 1989; c) H. S. M. Coxeter, Regular Polytopes, 3rd ed., Dover, New York, 1973, in which we read (p. 18):"*A quasiregular polyhedron (like the icosidodecahedron) is defined as having regular faces, while its vertex figures, though not regular, are cyclic and equiangular (i.e., inscriptible in circles and alternate-sided). It follows from this definition that the edges are all equal, say of length 2L, that the dihedral angles are all equal, and that the faces are of two kinds, each face of one kind being entirely surrounded by faces of the other kind.*"); d) A. Holden, Shapes, Space and Symmetry, Dover, New York, 1991.

[7] a) A. Müller, S. Sarkar, S. Q. N. Shah, H. Bögge, M. Schmidtmann, Sh. Sarkar, P. Kögerler, B. Hauptfleisch, A. X. Trautwein, V. Schönemann, Angew. Chem. 1999, 111, 3435-3439;

Angew. Chem. Int. Ed. 1999, 38, 3238-3241; b) D. Gatteschi, R. Sessoli, A. Müller, P. Kögerler in Polyoxometalate Chemistry: From Topology via Self-Assembly to Applications (Eds.: M. T. Pope, A. Müller), Kluwer, Dordrecht, 2001, pp. 319-328; c) D. Gatteschi, Europhys. News 2003, 34(6), 214–216 (Magnetism Special Issue); D. Gatteschi in Organic conductors, superconductors and magnets: from synthesis to molecular electronics (Eds.: L. Ouahab, E. Yagubskii), Kluwer, Dordrecht, 2004, pp. 179-196; d) see also A. Müller, M. Koop, H. Bögge, M. Schmidtmann, F. Peters, P. Kögerler, Chem. Commun. 1999, 1885–1886. In the related compound discrete hydrated VO^{2+} groups are also present (evidenced by ESR); these can, as usual, be easily integrated into the crystal-water lattice, which leads to a corresponding lower alkali cation concentration.

[8] a) M. Axenovich, M. Luban, Phys. Rev. B 2001, 63, 100407; b) C. Schröder, H. Nojiri, J. Schnack, P. Hage, M. Luban, P. Kögerler, Phys. Rev. Lett. 2005, 94, 017205; c) A. Müller, M. Luban, C. Schröder, R. Modler, P. Kögerler, M. Axenovich, J. Schnack, P. Canfield, S. Bud'ko, N. Harrison, ChemPhysChem 2001, 2, 517–521.

[9] When frustration originates purely from the geometry or topology of the lattice it is termed geometric frustration. The corresponding compounds are called geometrically frustrated antiferromagnetic materials (GFAF) and can have exotic magnetic ground states; see: a) J. Greedan, J. Mater. Chem. 2001, 11, 3753; b) Magnetic Systems with Competing Interactions (Ed.: H. Diep), World Scientific, Singapore, 1994; c) Quantum Magnetism (Eds.: U. Schollwöck, J. Richter, D. J. J. Farnell, R. F. Bishop), Springer, Berlin, 2004; d) Y. Narumi, K. Katsumata, Z. Honda, J.-C. Domenge, P. Sindzingre, C. Lhuillier, Y. Shimaoka, T. C. Kobayashi, K. Kindo, Europhys. Lett. 2004, 65, 705–711.

[10] Crystal data for 1: $H_{412}K_{24}Mo_{72}Na_8O_{538}S_{12}V_{32}$, $M_r = 19068.10$ gmol^{-1}, monoclinic, space group $C2/c$, $a = 47.155(2)$, $b = 42.5364(18)$, $c = 26.5344(12)$Å, $\beta = 90.2510(10)°$, $V = 53222(4)$Å^3, $Z = 4$, $\rho = 2.380$ gcm^{-3}, $\mu = 2.530$ mm^{-1}, $F(000) = 36848$, crystal size= $0.30 \times 0.20 \times 0.02$ mm^3. Crystals of 1 were removed from the mother liquor and immediately cooled to 183(2) K on a Bruker AXS SMART diffractometer (three circle goniome-

ter with 1 K CCD detector, $Mo_{\kappa\alpha}$ radiation, graphite monochromator; hemisphere data collection in ω at 0.3° scan width in three runs with 606, 435, and 230 frames ($\phi = 0, 88$, and 180°) at a detector distance of 5 cm). A total of 158364 reflections ($1.50 < \Theta < 26.99°$) were collected of which 57718 reflections were unique (R(int)=0.0612). An empirical absorption correction using equivalent reflections was performed with the program SADABS. The structure was solved with the program SHELXS-97 and refined using SHELXL-97 to R= 0.0567 for 40563 reflections with $I > 2\sigma(I)$, R=0.0916 for all reflections; max/min residual electron density 3.456 and -2.026 e$Å^{-3}$. (SHELXS/L, SADABS from G. M. Sheldrick, University of Göttingen **1997**; structure graphics with DIAMOND2.1 from K. Brandenburg, *Crystal Impact GbR*, **2001**.) As usual the cations not bound to the capsule as well as the crystal water molecules could not be found completely due to disorder. As this is also the case for the $[VO(H_2O)_5]^{2+}$ centers, the H_2O ligand sphere was not considered in the formula. Further details on the crystal structure investigations may be obtained from the Fachinformationszentrum Karlsruhe, 76344 Eggenstein-Leopoldshafen, Germany (fax: (+49) 7247-808-666; e-mail: crysdata@fiz-karlsruhe.de), on quoting the depository number CSD-415055.

[11] See for example, a) P. Kögerler, A. Müller in Polyoxometalate Molecular Science (Eds.: J. J. Borrás-Almenar, E. Coronado, A. Müller, M. T. Pope), Kluwer, Dordrecht, 2003, pp. 297–323; b) A. Müller, F. Peters, M. T. Pope, D. Gatteschi, Chem. Rev. 1998, 98, 239-271.

[12] A. Müller, B. Botar, H. Bögge, P. Kögerler, A. Berkle, Chem. Commun. 2002, 2944-2945.

[13] A.Müller, S. K. Das, H. Bögge, M. Schmidtmann, M. Dressel, J. van Slageren, unpublished results.

[14] In spite of these expected differences, there are some important features common to the two Keplerates, such as similar characteristics of the ground state and the low-lying excitation states, which arise because of their otherwise common geometrical structure and the existence of antiferromagnetic exchange between nearest-neighbor spins. As one ex-

ample, the total spin quantum number of the ground state eigenvector for the "quantum" Keplerate **1** is expected to be $S = 0$, as it is for the "classical" Keplerate {Fe_{30}}, and its direct manifestation would be in the vanishing of $T\chi$ (proportional to $\langle S^2 \rangle$), in the low-temperature regime; χ is the zero-field susceptibility.

[15] The raw $M/H (\equiv \chi)$ data display a rapid rise on decreasing the temperature below 50 K that is very accurately simulated by a term proportional to $1/T$ (Curie behavior). This observation suggests the presence of additional non-interacting paramagnetic centers in the sample. In fact, fitting the low temperature data leads to a value of approximately 2 $S = 1/2$ centers. These centers are considered to be the non-interacting VO^{2+} (d^1) ions. In large cluster systems like **1**, they can take the same place as the other diamagnetic cations, which are usually disordered in the large voids between the clusters and can therefore not be discovered by single-crystal structure analysis (see ref. [10]). On subtracting the contribution of these discrete paramagnetic centers, we obtain a corrected magnetization M' as well as the corrected susceptibility data, $T\chi \equiv T'/H$, shown in Figure B.4. The Curie behavior of the paramagnetic centers shows that they do not interact with each other and with the {V_{30}} cluster. The possibility that the finite $T\chi$ value at low temperatures is due to a non-zero spin ground state can be precluded. Regarding the presented $T\chi$ curve: As always there is uncertainty as to the appropriate choices for diamagnetic and temperature-independent paramagnetism (TIP) corrections. Additionally, because of the large voids between the clusters the VO^{2+} groups need not be present stoichiometrically in the compound; correspondingly, there is a very small error limit in the given/used number of two VO^{2+} groups which influences the correction of the raw magnetic data (($VO)_{1.8}$ could, for example, correspond to $K_{14.4}$).

[16] The exchange coupling difference between the classical and quantum Keplerates discussed herein is analogous to that of the cluster pair with 6 V^{IV} and 6 Fe^{III} centers embedded in the {Mo_{57}} type skeleton; see D. Gatteschi, R. Sessoli, W. Plass, A. Müller, E. Krickemeyer, J. Meyer, D. Sölter, P. Adler, Inorg. Chem. 1996, 35, 1926-1934.

[17] Note added in proof, April 27, 2005: Very recent measurements by H. Nojiri (Tohoku University) of M versus H at 0.5 K up to 27 Tesla, as well as ESR measurements at 190 GHz for several temperatures, show features which are fully consistent with our physical interpretation of a strong intracluster exchange constant and approximately two VO^{2+} ions per formula unit that are magnetically independent of the $\{V_{30}\}$ cluster. Full details will be published elsewhere.

ACKNOWLEDGEMENTS