ADVANCES IN PHYSICS, 2000, VOL. 49, NO. 1, 93–140

High-order convergent expansions for quantum many particle systems

MARTIN P. GELFAND
Department of Physics, Colorado State University, Fort Collins, CO 80523, USA

and RAJIV R. P. SINGH
Department of Physics, University of California, Davis, CA 95616, USA

[Accepted 11 March 1999]

Abstract

We review recent advances in the calculations of high-order convergent expansions for quantum many-particle systems. Calculations for ground state properties, including correlation functions and static susceptibilities, for spin models as well as for models of many fermions, such as Hubbard and Kondo models, are discussed. A historical perspective to the subject is provided. Recently important technical advances have been made in perturbative calculations of the excitation spectra of quantum many-particle systems, which enable the calculation of these spectra to high orders. The method, along with its applications, are explained. Fairly comprehensive, though simplified, algorithms for generating lists of relevant clusters, their lattice embeddings and subclusters are presented. The perturbative recursion relations and their computer implementation are also discussed in detail. A compilation is made of various series expansion studies that have been carried out for condensed matter problems. The scope and limitations of these methods are explained, and several open problems are noted.

Contents

Advances in Physics ISSN 0001–8732 print/ISSN 1460–6976 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals/tf/00018732.html

1. Introduction

Systems with strong electronic correlations, both idealized models and real materials, have offered a wide variety of challenging and important problems in condensed matter physics. The materials include (but are by no means limited to) the cuprate high-temperature superconductors and their antiferromagnetic, insulating parent compounds; the 'colossal magnetoresistance' manganates and related compounds; heavy-fermion metals and 'Kondo insulators'; and a host of magnetic insulators such as uniformly frustrated antiferromagnets, spin ladders, and spin-Peierls compounds. Some of these materials were only synthesized in the past decade or so, while for others synthetic routes and structures had been determined but their electronic properties had not been thoroughly explored until recently.

The theoretical models which are used to describe the low-energy electronic properties of these materials are by no means new. Hubbard, Kondo, Anderson, double-exchange, and Heisenberg–Ising models have been actively studied for several decades. There has been a resurgence of interest in these models, however, because of calculational techniques—which have been either recently invented or recently rendered more potent due to advances in computer technology—that go beyond traditional many-body methods for calculating ground- and excited-state properties. One example of a venerable technique grown stronger is exact diagonalization of finite clusters, while methods of more recent vintage include dynamical mean-field theory ('$d = \infty$' methods) and the density-matrix renormalization group.

The subject of this paper is *the* most venerable approach to the study of quantum systems which are not exactly diagonalizable by elementary analytic means, namely, non-degenerate Rayleigh–Schrödinger perturbation theory. Our focus will be on modern developments which allow for the efficient generation of high-order expansions, significantly beyond the order of expansion that one can practically carry out 'by hand', for quantum many-body systems defined on infinite lattices. We will refer to the combination of non-degenerate perturbation theory (carried out using modern techniques) and series extrapolation (mostly using techniques brought to physics in the context of high-temperature expansions and classical critical phenomena) as 'the series expansion method'.

For certain problems, particularly for insulators of dimensionality greater than one, we believe that the series expansion method is among the best calculational methods available. These problems include various Heisenberg antiferromagnets (for which series have been constructed about Ising, dimer, and plaquette models, as will be described later), and half-filled Hubbard and Kondo-lattice models. Methods which compete with series expansions most directly are numerical finite-cluster calculations such as exact diagonalization, quantum Monte Carlo, and density matrix renormalization group, which must be supplemented by finite-size scaling to estimate quantities in the thermodynamic limit. One great advantage of the series expansion method is that the associated computational difficulties increase relatively slowly with dimensionality, in contrast with all of the finite-cluster methods. In contrast to quantum Monte Carlo, but like exact diagonalization, there is never a fermion sign problem. Another nice feature is that they are carried out directly for infinite systems: although numerical series extrapolations are required they are distinct from those extrapolations involved in finite-size scaling. This suggests that the inherent limitations of the series expansion method are perhaps different from and in some ways complimentary to those approaches based on non-perturbative calculations for finite systems.

Before proceeding further, we would like to distinguish between more traditional high-temperature expansions and $T = 0$ perturbation expansions for quantum many-body systems. In this paper we will discuss only the latter, even though these methods share some common features such as the basic structure of the connected-graph expansion and various technical details in practical implementation of the calculations. It has also been the experience of the authors that applying both high-temperature and $T = 0$ perturbative methods to the same problem can be very useful. However, the two classes of expansions are sufficiently different in character that they merit separate consideration. For reviews of high-temperature expansions for quantum many-body systems we refer the reader to [1–3].

The outline of the paper is as follows. First, in section 2 we address the question: for what Hamiltonians and for what properties does the series expansion method provide a feasible calculational scheme? Section 3 offers a brief history of the series expansion method. The remainder of the article is independent of this section, but readers may find it useful as an entry point into the literature on methods for series expansions and some of the earlier applications. The three sections that follow, 4, 5 and 6, are devoted to the formalisms underlying the construction of high-order perturbation expansions: first a very general discussion of cluster expansions, then calculations of ground state properties including equal-time correlation functions and static susceptibilities, and finally calculations of excitation spectra and spectral weights. The next two sections, 7 and 8, are concerned with implementation details associated with generating the clusters and evaluating the cluster weights (for both ground-state and excited-state properties). In section 9 various applications of the series expansion method in the condensed-matter literature are described. Some pointers to the field-theory literature are also offered. For the particular case of square-lattice Heisenberg antiferromagnets, a critical comparison of series expansions with 'finite-size' and other numerical approaches to the same problems is offered in section 10. The next section touches on some open technical issues: multiple excitations (both bound states and continua), quenched disorder, and degenerate perturbation theory. A summary and conclusions are presented in section 12.

2. Preliminaries

The series expansion method is suited for *lattice-based* quantum many-body systems. That is, we consider models where the dynamical degrees of freedom, such as spins, are represented by operators on lattice sites and their interactions are specified by a Hamiltonian of the form

$$\mathcal{H} = \sum_i (T_i; H_i) + \sum_{i,j} (T_i, T_j; J_{i,j}), \tag{1}$$

where the T_i are generic operators, the parentheses represent generic functions of those operators, and the 'fields' H_i and 'coupling constants' $J_{i,j}$ parametrize the interactions. All cases treated to date have involved one-body and two-body terms, just as in equation (1), but there is no problem in principle with including multi-body terms in the Hamiltonian, if desired. We will point out in section 4 what modifications in the formalism would be required to include such interactions. It is a useful simplification—but again not an essential one, as will be briefly discussed in section 11—to assume that the Hamiltonian is defined on a periodic lattice and that the interactions are invariant with respect to the translation group of the lattice. This implies that H_i is site-independent and $J_{i,j}$ depends only on $\mathbf{r}_i - \mathbf{r}_j$. A truly essential assumption is that an 'unperturbed Hamiltonian' $\mathcal{H}_0$ exists so that the Hamiltonian of equation (1) is a member of the one-parameter *family* of Hamiltonians

$$\mathcal{H} = \mathcal{H}_0 + \lambda \mathcal{H}_1, \tag{2}$$

where $\mathcal{H}_0$ can be trivially diagonalized in a suitably chosen local basis. By the latter we mean, specifically, that the eigenstates of $\mathcal{H}_0$ can be expressed as a product over states of single sites. We shall also assume that the ground state of $\mathcal{H}_0$ is either non-degenerate or of finite degeneracy. If these above conditions are met, then quite generally it is possible to develop perturbation expansions for the ground state energy and correlation functions of the model defined by equation (2) in powers of λ by means of a connected-cluster expansion, which will be described in section 4. It is also generally possible to develop connected-cluster expansions for the spectra and associated spectral weights of infinitely long-lived *excited states*.

There are certain features of the expansions which depend on the degeneracy (or lack thereof) of the ground state of $\mathcal{H}_0$, so it is worth expanding on the matter.

A non-degenerate unperturbed ground state is associated with an unperturbed Hamiltonian that consists of disconnected sites (i.e. there are no two-body terms). These disconnected sites must have non-degenerate ground states individually, otherwise the ground-state degeneracy of the many-body Hamiltonian would be exponentially large in the number of sites. We denote expansions about this type of $\mathcal{H}_0$ as 'high-temperature type' expansions. (In the literature one can sometimes find such expansions referred to simply as 'high temperature expansions'.) From a practical point of view it is important that the number of states associated with individual sites not be too large. Examples of high-temperature type expansions which have been carried out include sites which comprise a single $S = 1/2$ spin in an external field (large-field expansions), sites which comprise coupled pairs of $S = 1/2$ or $S = 1$ spins (dimer expansions), sites which comprise coupled sets of four $S = 1/2$ spins (plaquette expansions), and sites which comprise a coupled $S = 1/2$ spin and 'orbital' that can hold zero, one, or two fermions. The number of states per site for these examples is 2, 4 or 9, 16, and 8, respectively.

Alternatively, the ground state of $\mathcal{H}_0$ can be degenerate, and the expansion is then denoted 'low-temperature type'. The most frequently studied examples of this class of unperturbed Hamiltonians are Ising models on bipartite lattices, for which the ground state, whether ferromagnetic or antiferromagnetic, is doubly degenerate. Note that it is not possible to do a straightforward expansion for the triangular-lattice antiferromagnet around its Ising limit, as in that case the number of ground states grows exponentially with the number of sites (in a typically-shaped cluster), i.e. there is non-zero entropy per spin at $T = 0$. One can, however, carry out a low-temperature type expansion for the triangular-lattice antiferromagnet by choosing as the unperturbed Hamiltonian a ferromagnetic Ising model in rotated-spin variables, for which the ground states in the original variables correspond to two particular three-sublattice orderings.

Because in every case the unperturbed ground states have a local product structure, the series expansion method appears to be most suitable when the physical system under consideration has finite-range correlations. However, systems with power-law correlations and gapless excitations can also be studied, as limiting points of the series extrapolations, just as classical critical points can be studied by means of high- and low-temperature expansions.

3. A brief history

Here we briefly review the history of methods for constructing high-order convergent perturbation expansions for lattice-based many-body systems. It is not our aim to cite all of the potentially relevant literature, but we do hope to present a clear overview of this particular corner of theoretical physics. Our main worry in writing this section is that, as condensed matter physicists, we may be ignorant of important papers in the field-theory literature, where these calculations go by the name of (Hamiltonian) strong-coupling expansions.

All of the methods fall into two main categories: cumulant expansions and (linked-) cluster expansions. The former has a longer history, is well suited for low-order calculations by hand, and is probably familiar to a fair number of physicists. The techniques to be described in detail in this paper fall in the latter category. Even though the end products of both types of calculations are (or ought to be!) identical, we will argue that in some important respects the linked-cluster methods are preferable.

An elegant presentation of the general cumulant expansion formalism is given by Messiah in his well-known textbook [4]. To get the flavour of this approach, let us examine the structure of the expansion for the energy E of some state which, in the unperturbed limit of the family of Hamiltonians (2), satisfies $\mathcal{H}_0|0\rangle = E_0|0\rangle$ and is not degenerate with any other eigenstate of $\mathcal{H}_0$ (say, $|1\rangle$) unless $\langle 1|\mathcal{H}_1^n|0\rangle = 0$ for all n. Under those conditions, one has $E = \sum_{n\geq 0} \lambda^n E_n$ where

$$E_n = \langle 0| \sum_{k_1,\ldots,k_{n-1}} \mathcal{H}_1 S_{k_1} \mathcal{H}_1 \cdots S_{k_{n-1}} \mathcal{H}_1 |0\rangle, \qquad (3)$$

with $\sum_{i=1}^{n-1} k_i = n-1$, $S_0 = -|0\rangle\langle 0|$, and $S_{k>0} = (E_0 - \mathcal{H}_0)^{-k}(1 + S_0)$. (This expression is not unique but it will suffice for the present purpose, which is merely illustration.) This formula is the basis for perturbation expansions for the energy of the ground state and also states which we will later characterize as ‘single particle’ excited states. Related formulae can be written for the expectation value of any operator in these states.

An expression like (3) naturally leads one to construct ‘space–time’ diagrams. The ‘space’ coordinates are associated with the terms in $\mathcal{H}_1$, which is an infinite sum involving operators at pairs of lattice sites, while the ‘time’ coordinates are associated with the ordering of the various terms in the product for which the expectation value is being evaluated. A calculation then requires constructing all of the diagrams that can contribute at a given order and evaluating them. Fortunately this requires only a finite number of computations even for an infinite system, because only diagrams which are connected (in a particular sense, depending on the detailed form of the expansion) have a non-vanishing contribution.

As hinted above, the first attempts to carry out high-order expansions, which started in the early 1960s, employed cumulant expansion methods. These were in the context of the nearest-neighbour Heisenberg–Ising model

$$\mathcal{H} = \sum_{\langle ij \rangle} S_i^z S_j^z + \lambda (S_i^x S_j^x + S_i^y S_j^y) \tag{4}$$

on bipartite lattices in one, two, and three dimensions [5–9]. These early calculations were carried out for general S and took advantage of detailed properties of the spin operators; for this reason they appear much more technically complicated than modern methods, which require separate calculations for each value of S. It may offer a hint of their complexity that it was not until the work of Parrinello *et al.* [8] in 1973 that correct results were obtained to $O(\lambda^6)$ for the sublattice magnetization and two-point correlation functions. (Note that odd powers in these series vanish, so the terms obtained were only the second non-trivial ones.)

Cumulant expansion methods found other applications in the 1970s, such as the work of Pfeuty and Elliot [10] on the transverse-field $S = 1/2$ Ising model expanding both about the Ising limit

$$\mathcal{H} = \sum_{\langle ij \rangle} S_i^z S_j^z + \lambda \sum_i S_i^x \tag{5}$$

and the 'disordered' limit (where λ is moved from the second term to the first), and the work of Harris [11] on the dimerized $S = 1/2$ Heisenberg chain

$$\mathcal{H} = \sum_i \mathbf{S}_{2i} \cdot \mathbf{S}_{2i+1} + \lambda \sum_i \mathbf{S}_{2i+1} \cdot \mathbf{S}_{2i+2}. \tag{6}$$

None of these calculations went beyond $O(\lambda^4)$, and (unlike some of the later papers on the Heisenberg–Ising model cited above) they were carried out entirely by hand. Harris' work is noteworthy in that it is the first we know of to calculate the form of an excitation spectrum (and not just an excitation gap) to a non-trivial order. In the late 1970s, Hamer *et al.* [12] generated series for the gap between ground and lowest-excited state of one-dimensional quantum rotor models with $O(2)$, $O(3)$, and $O(4)$ symmetries. These models have the form

$$\mathcal{H} = \sum_i \mathbf{L}_i^2 + \lambda \sum_i \mathbf{n}_i \cdot \mathbf{n}_{i+1}, \tag{7}$$

where $\mathbf{n}$ is a unit vector and $\mathbf{L}$ is the corresponding generator of infinitesimal rotations. An interesting distinction between these models and all others that we have mentioned previously is that the state space for any single site is infinite dimensional. That does not cause any problems for the cumulant expansion technique, but as we will see in section 8 it renders untenable the most naive implementation of the cluster expansion technique. The calculations went to order λ^8 for the $O(2)$ model and λ^6 for the other models, and relied on computer codes to generate and evaluate the necessary diagrams.

Cumulant expansion methods were overviewed in the context of lattice gauge theory in the mid 1980s by Hamer and Irving [13].

More recently, a cumulant expansion for energies was applied to the t–J–J_z model at half-filling and with one and two holes by Prelovšek *et al.* [14]. These are expansions about the Ising model. Calculations in which the hopping term alone was the perturbation were carried out to $O(t^{12})$, while double expansions, in both hopping and transverse exchange, were carried out to lower orders.

Most recently, there have been a number of *low-order* perturbation expansions appearing in the literature. We expect that standard cumulant expansion methods were employed in all these hand calculations. Freericks and Monien [15] found that for the Bose–Hubbard model one could essentially reproduce the phase diagram which had been determined using laborious quantum Monte Carlo calculations. Low-order calculations for spin models include dimer and plaquette expansions for the CaV_4O_9 Heisenberg model [16–18] (which will be described in section 9).

Now let us turn to cluster expansions. Relevant formalisms were first described in the early 1980s, by Kadanoff and Kohmoto [19], and Nickel [20] and Marland [21]. The former authors' approach is not very transparent, and seems to have been applied only to a few one-dimensional systems by Kadanoff and associates [22] (and see also section 9.4). The latter approach is very straightforward in principle, has been applied to a wide variety of models, and is the subject of this paper. Since applications of the method are the subject of section 9, and a detailed description of the method forms the next five sections, we will restrict the present discussion to formal developments. But first let us give a very brief description of the cluster expansion method, just enough so that the distinction between it and the cumulant expansion method is clear.

There are two main steps in a cluster expansion. First is the identification of the finite number of relevant connected *real-space* clusters for the Hamiltonian under consideration. Then, for each cluster, one constructs the Rayleigh–Schrödinger perturbation expansion for the extensive quantity (energy, correlation function) under consideration. The results for the various clusters are then combined (via 'subgraph subtraction') so as to yield the quantity per site on the infinite lattice.

The subgraph subtraction step is useful not only because it is essential for obtaining the final result, but it also offers a very strong check on the calculations. As will be seen later, many terms are supposed to exactly cancel and subgraph subtraction makes it obvious if they do not. Failure of any cancellation indicates a flawed calculation: either a relevant cluster has been omitted or there has been an error in one or more of the perturbation expansions for the clusters. Only the highest-order term in the expansion is not checked by subgraph subtraction.

Let us now return to the history of the cluster expansion formalism. An apparent weakness of cluster expansions is that non-extensive properties such as excitation spectra would appear to be beyond their reach. By the mid 1980s two ways to circumvent this difficulty had been proposed and employed. One involves constructing expansions for finite systems with periodic boundaries (and would seem to be feasible only for one-dimensional systems): see Barber and Duxbury [23] and Hornby and Barber [24]. Another, due to Nickel [20], is much closer in spirit to the cluster formalism for extensive properties, but requires consideration of certain *disconnected* clusters as well as the connected ones. For a description of this approach, see He *et al.* [25].

An overview of the cluster expansion formalism in the context of lattice gauge theories was given by Hamer and Irving [26], as a partner to their article on cumulant expansions noted earlier.

In the late 1980s, the cluster expansion formalism for ground-state properties of Marland and Nickel was independently rediscovered by the present authors in collaboration with Huse [27]. We must express some embarrassment over how long it was before we learned of the earlier cluster expansion work, especially since our implementation of the method was not as efficient as that employed by Hamer, Oitmaa, and their collaborators in Australia, who by that time had been applying cluster expansions for several years. Our own collaboration, however, can at least claim credit for some novel applications of the series expansion approach, both to new models and new properties, as will be discussed in section 9.

The latest stage in the formal development of cluster expansion methods came with the recognition by Gelfand [28] that *connected* cluster expansions could be constructed for 'single-particle' excitation spectra and other excited-state properties such as spectral weights. This finally put cluster expansions on par with the cumulant expansion techniques, which from the beginning could treat non-degenerate excited states on the same footing as the ground state.

We had promised at the start of this section some justifications for our favouring cluster expansions over cumulant expansions. Here are two.

First, the clean separation of real-space cluster construction and perturbation theory within the cluster expansion formalism, in contrast with the cumulant expansion formalism, makes it easier to apply optimal techniques to each part of the problem. For example, one can sometimes use 'topological' methods to build the clusters, as will be described in section 7; the analogous procedure would be awkward to implement within the cumulant expansion formalism and we do not know that it has ever been attempted.

Second, the strong self-consistency check provided by subgraph subtraction is a great practical advantage of cluster expansion methods. Although it is possible to unknowingly obtain fully incorrect results using a cluster expansion, the only way to do so is by carrying out a valid calculation for the wrong Hamiltonian or wrong operator. We know this sort of error to be possible from personal experience, but it is not common, and it is no less possible within the cumulant expansion formalism. In contrast, diagrams can be inadvertently omitted in cumulant expansion calculations without producing any obvious warning signs.

4. Cluster expansions: formalism

In this section, we will define the essential features of the cluster expansion formalism. As discussed above, the cluster expansion method is one of several methods for calculating series expansion coefficients of lattice statistical models. We reiterate that we are dealing with exact series expansions so that any method, if implemented correctly, should lead to exactly the same expansion coefficients. The strength of this method is that it is conceptually simple and very convenient for fully automated computer calculations of high-order series coefficients. The central idea of the method is to express the expansion coefficients for properties of a large cluster in terms of expansion coefficients for properties of smaller clusters. Thus, even for an infinite system, the expansion coefficients to a given order are obtained from separate series expansions done for a finite number of finite clusters. The latter calculations are in principle straightforward, since finite clusters have finite-dimensional Hilbert spaces. The role of the infinite lattice is to provide the number and types of different small clusters that can be embedded in it. In other words, in a cluster expansion the calculation of series coefficients is divided into two parts. The first is identifying distinct clusters of various sizes and counting their embeddings in the lattice and the second is calculating properties of finite clusters. The thermodynamic limit is treated at the very outset and only enters the counting of embeddings. Detailed Rayleigh–Schrödinger perturbation theory only needs to be done for one finite cluster at a time. More details on how the finite cluster perturbation expansions are done and how the list of distinct clusters is generated are discussed in the following sections. Here we will set up the general formalism.

Suppose that we are interested in calculating an extensive property P for a model defined on a lattice $\mathcal{L}$. For example, this property P could be the ground-state energy, the uniform susceptibility, or a sum over all second neighbour spin–spin correlations on the lattice. For the sake of concreteness, we assume that we are dealing with a large lattice composed of N sites, with periodic boundary conditions. (The cluster expansion formalism is equally valid for finite systems with free or periodic boundary conditions, but is rarely used in practice.) The key idea in the cluster expansion is to express the quantity P, per lattice site, as a sum over all distinct clusters c

$$\frac{P(\mathcal{L})}{N} = \sum_{c} L(\mathcal{L}, c) \times W(c) . \tag{8}$$

Here, $L(\mathcal{L}, c)$, called the lattice constant of the cluster c, is the number of ways per lattice site that the cluster c can be embedded in the lattice $\mathcal{L}$. The question of what the distinct clusters are, and why the number of their embeddings is proportional to the number of sites of the lattice, is discussed in the following paragraph. The quantity $W(c)$ is called the weight of the cluster. It is expressed as a power series expansion in the appropriate variable and can be obtained from the relation

$$W(c) = P(c) - \sum_{g\subset c}{}' W(g) . \tag{9}$$

Here $P(c)$ is the series expansion for the property defined on the finite cluster c. The prime indicates that the sum runs over proper subsets only. The equations (8) and (9) define the cluster expansion. It is evident that if one has a list of all distinct clusters that can be embedded in the lattice together with their lattice constants and their subclusters and can calculate the property $P(c)$ for the individual clusters in a series expansions, then using relations (8) and (9), one can obtain the series expansion for $P(\mathcal{L})/N$ in the thermodynamic limit. One additional aspect of the cluster expansion method, which makes the cluster expansion practical, is that only smaller clusters will contribute in low orders of the expansion. Hence, only a finite number of clusters will be needed to get the exact coefficients for the thermodynamic system, to any given order.

We now turn to the basic definition of the clusters. The clusters c appearing in equation (8) are finite graphs or pieces of the lattice $\mathcal{L}$, consisting of a certain number of sites and a certain number of bonds connecting pairs of sites. (We are assuming pairwise interactions here but the method is readily generalized to more general interactions.) In addition, graphs which will make identical contribution to desired

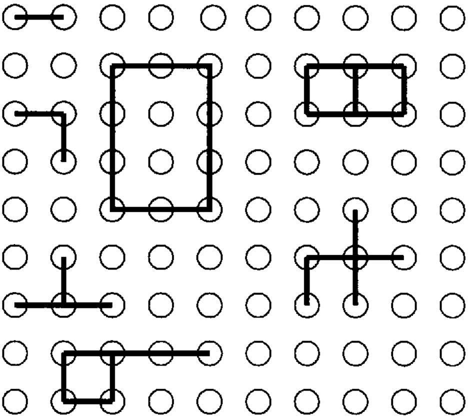


Figure 1. Examples of clusters embedded in the square lattice. The circles represent sites (perhaps individual spins), while the solid lines connecting nearest-neighbour sites correspond to terms in $\mathcal{H}_1$.

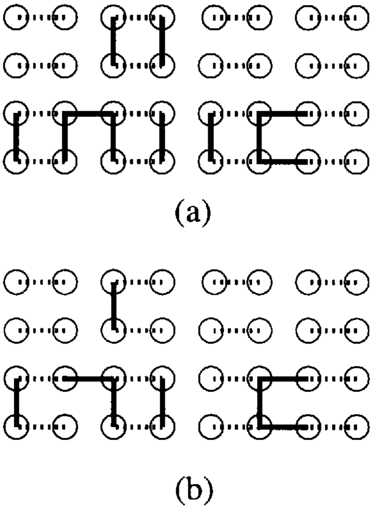


Figure 2. An illustration of the arbitrariness in the set of clusters included in a cluster expansion. Here the circles can be thought of as representing individual spins, and the dotted lines are spin–spin couplings present in $\mathcal{H}_0$. As in the preceding figure, solid lines represent terms in $\mathcal{H}_1$. In (*a*) three clusters are shown that would typically be included in the set of clusters, while in (*b*) three clusters are shown that would typically *not* be included, because pairs of spins would usually be treated as the elementary units in the cluster expansion.

quantities are grouped together as being the same cluster. Their multiplicity is taken into account by the lattice constant. Examples of clusters that can be embedded in the square lattice are shown in figure 1. For most problems, one would consider a list of all possible clusters. However, sometimes it is useful to shorten the list by keeping only some of the clusters and discarding others. For example, in dimer or plaquette expansions, elementary dimers or plaquettes may be treated as elementary units or points. Then a bond connecting two of these points will represent all the interactions between two dimers or plaquettes. Thus the clusters in figure 2 (*a*) will be in our list but those in figure 2(*b*) will not. This arbitrariness in the set of clusters that are included in our list leads to different weights for the clusters. Thus, in reducing the list of clusters one needs to be careful about the lowest order of the expansion in which a given cluster will contribute. These ideas will become clearer when we discuss the cluster expansion from the point of view of multivariable expansions.

Associated with a given cluster will be a cluster Hamiltonian. This will include all interactions between spins within the cluster. In high-temperature type expansions the interactions between the spins in the cluster and those outside the cluster are not included in the cluster Hamiltonian. In low-temperature type expansions, such as Ising expansions, the unperturbed interactions (in this case the Ising part) between spins inside the cluster and those outside are included in the cluster Hamiltonian. For the purpose of the cluster calculation, the outside spins remain frozen in their ground state. Hence, in the cluster Hamiltonian, the interactions with outside spins act as applied local fields on the boundary sites of the cluster.

A cluster is said to be connected, if its sites cannot be decomposed into two non-empty groups, such that the sites of one group have no bonds connecting them to the sites of the other group. The cluster expansions, considered here, involve expansions in connected clusters only. It is straightforward to prove that extensive quantities like the ground-state energy or equal-time ground state correlation functions have a connected cluster expansion. The number of embeddings in the lattice of a given connected cluster is proportional to the number of lattice sites: this follows from the translational invariance of the lattice. Thus the thermodynamic limit and extensivity of the thermodynamic quantities are built into the cluster expansion formalism at the outset.

Next, we wish to understand the lowest order in which a given cluster contributes to a given quantity, so that we can prepare a list of clusters that will ensure that our desired expansions are correct to a predetermined order. For high-temperature type expansions, this order is determined by the number of bonds in the cluster whereas for a low-temperature type expansion, it is determined by the number of sites in the cluster. Note that these determine the lowest order in which the cluster contributes; the total contribution, or weight, of a cluster is typically an infinite series starting at some order. In a high-temperature type expansion the contribution of a cluster is defined as a sum of all those terms in the perturbation theory, where the perturbing Hamiltonian on every bond of the cluster has acted at least once. It thus follows that a cluster with N bonds will contribute in Nth or higher order of the perturbation expansion. In a low-temperature type expansion, the contribution of a given cluster is defined as all terms where the state of every site in the cluster has been altered at least once. Thus the order of the cluster is determined by the number of sites of the cluster. In the case of the much studied Ising expansions for Heisenberg–Ising models, this lowest order is equal to the number of sites in the cluster, because contribution to a trace requires the state of every site to be altered at least twice and in one application of the Hamiltonian the state of exactly two sites is altered. These factors cancel each other leading to the minimum order equalling the number of sites in the cluster. In general, such considerations need to be explored for the model at hand. As an aside, we note that such considerations lead to a ‘rigorous’ minimum order in which a cluster contributes. That is we have a formal proof that the cluster cannot contribute at a lower order. But it need not have a non-zero contribution at this minimum order. In fact, quite often one finds that many clusters contribute only in much higher orders. As an example, for many problems, clusters with ‘free ends’ contribute in much higher than the minimum order. Such graphs are also the most numerous and have the largest lattice constants. Thus an efficient series generation algorithm should include clusters based on their 'true' minimum order, the lowest order in which they actually make a non-zero contribution.

A clear insight into the cluster expansion method is gotten by regarding our perturbation series as a multivariable expansion. Let us assume that associated with each bond of the lattice there is a different interaction variable λ_{ij}. We are interested in expanding some property P for the lattice $\mathcal{L}$ as a power series in the λ_{ij}.

$$P(\mathcal{L}) = p_0 + \sum_{ij} p_{ij}\lambda_{ij} + \sum_{ij}\sum_{kl} p_{ij,kl}\lambda_{ij}\lambda_{kl} + \cdots. \tag{10}$$

We can regroup terms in this expansion by collecting together all terms which depend on one λ_{ij}, those which depend on exactly two λ_{ij}, and so forth, leading to

$$P(\mathcal{L}) = p_0 + \sum_{ij}\sum_{n=1}^{\infty} a_{ij}^{n}\lambda_{ij}^{n} + \sum_{ij,kl}{}'\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} b_{ij,kl}^{m,n}\lambda_{ij}^{m}\lambda_{kl}^{n} + \cdots. \tag{11}$$

Here the prime denotes the fact that λ_{ij} and λ_{kl} are distinct. We can now associate different terms in this multivariable expansion with different clusters. These will be called the weight of the cluster. For the first cluster, consisting of two sites i and j with a bond between them, the weight of the cluster for the property P is

$$W(1) = \sum_{n=1}^{\infty} a_{ij}^{n}\lambda_{ij}^{n}. \tag{12}$$

Similarly, for the second cluster consisting of three sites i, j and k with bonds between sites i and j and between sites j and k, the weight of the cluster is given by

$$W(2) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} b_{ij,jk}^{m,n}\lambda_{ij}^{m}\lambda_{jk}^{n}. \tag{13}$$

It is evident that:

(i) Apart from the constant term p_0, each term in the multivariable expansion is associated with exactly one lattice-embedded cluster. That is,

$$P(\mathcal{L}) = p_0 + \sum_{c^*} W(c^*), \tag{14}$$

where the sum runs over all possible clusters embedded in the lattice. (Let us emphasize that here, unlike the sum in equation (8), *all* clusters rather than *distinct* clusters appear in the sum. We use asterisks on the cluster labels to highlight the distinction.)

(ii) To calculate the weight of a cluster all interactions not in the cluster can be set to zero. This defines the cluster Hamiltonian.

(iii) When the property P is calculated for the finite cluster c^*, apart from the zeroth order term one simply gets a sum of weights of all the subclusters. That is,

$$P(c^*) = p_0^{c^*} + \sum_{s^*\subseteq c^*} W(s^*), \tag{15}$$

where the sum includes the cluster c^*. From this the basic equation for subgraph subtraction in equation (9) follows.

(iv) If the lattice has some symmetries, such as those of translation or rotation then all graphs related to each other by the symmetry will have the same weight when their interactions are equal. Thus one can combine together graphs with the same weight into a counting problem and lead to the basic expression for cluster expansion equation (8).

(v) For properly normalized ground state properties the weights of disconnected clusters will vanish. This follows from the fact that the ground state wavefunction of a disconnected cluster will be a direct product of the ground state wave function of its subclusters. Thus any desired property of a disconnected cluster will be a sum of terms each of which will involve interactions from only one connected group. But that is part of the weight of that particular subcluster and the full disconnected cluster has zero weight left after subgraph subtraction. It is somewhat non-trivial to show that a connected cluster expansion exists for excited state properties as well, and we defer discussion of this matter to section 6.

(vi) For a cluster with n bonds the weight of the cluster can have no less than n powers of the interaction. Thus if we include the weight of all clusters with up to n bonds we will have the correct expansion to order n.

We note that multivariable expansions were invoked only to illustrate the ideas and construct formal proofs. Multivariable expansions are very cumbersome to carry out in practice and do not need to be done. Equations (8) and (9) remain valid when all interactions are set equal to each other, thus reducing the problem to a single variable expansion, or a few-variable expansion when one has just a small number of distinct interactions (for example two interactions corresponding to nearest and second neighbour interactions on a lattice respectively). Equations (8) and (9) can also be derived directly, without considering the multivariable expansion, from the combinatorial principle of inclusion–exclusion. For a discussion of this in context of the closely related high-temperature expansions see [29].

Thus the important simplicity of the cluster expansion method is now apparent. The calculation of series expansions for a thermodynamic system is completely separated into two parts. One is the combinatorial question dealing with the number of embeddings of different clusters in a lattice. This is independent of the degrees of freedom in the Hamiltonian. However, the counting does depend on the nature of the interactions, and will be different with purely nearest versus nearest and second neighbour interactions. The second is the calculation of the weights, which is entirely defined by a series expansion for the desired property of a finite system. Both parts can be fully automated on a computer. The first part is closely related to counting of lattice embeddings in high and low temperature expansions of classical Ising models. It is discussed in more detail in section 7. The second part is computationally rather similar to Lanczos and other exact diagonalization studies, and is discussed in detail in several of the sections that follow.

5. Ground-state properties: formalism

We saw in the last section that the problem of constructing a perturbation expansion for some property P which is additive over disconnected clusters (such as the ground state energy, ground state correlation functions, and $T = 0$ susceptibilities) can be reduced to the calculation of perturbation expansions for P on finite clusters. Let us now see how the latter calculations can be carried out. (This section has been lifted not quite verbatim from [30]; another source which contains the explicit recurrence relations for the multivariable expansions is [31].)

The cluster Hamiltonian has the same form as the full lattice Hamiltonian, which in the simplest case of a one-parameter expansion is just

$$\mathcal{H} = \mathcal{H}_0 + \lambda \mathcal{H}_1, \tag{16}$$

but here all of the $\mathcal{H}$'s refer to restrictions of the full lattice Hamiltonian to the cluster in question so that there are only a finite number of operators present. Writing the ground state energy and wavefunction for the cluster as power series in λ,

$$E = \sum_{i \geq 0} E_i \lambda^i, \tag{17}$$

$$|\Psi\rangle = \sum_{i \geq 0} |\Psi_i\rangle \lambda^i, \tag{18}$$

inserting them into the Schrödinger equation, and collecting power of λ, one readily obtains recurrence relations for the E_i and $|\Psi_i\rangle$ in the basis of the unperturbed eigenstates, that satisfy

$$\mathcal{H}_0 |k\rangle = e_k |k\rangle. \tag{19}$$

Let $|0\rangle$ denote the unperturbed ground state, and label all other unperturbed states with positive integers. Then provided one sets

$$\langle 0|\Psi_i\rangle = \delta_{i,0}, \tag{20}$$

an arbitrary 'normalization convention' that does not affect any physical results but simplifies the intermediate formulae, the recurrence relations (valid for $i \geq 1$) are

$$E_i = \langle 0|\mathcal{H}_1|\Psi_{i-1}\rangle \tag{21}$$

and

$$\langle k|\Psi_i\rangle = (e_o - e_k)^{-1} \left[\langle k|\mathcal{H}_1|\Psi_{i-1}\rangle - \sum_{i'=1}^{i-1} E_{i'} \langle k|\Psi_{i-i'}\rangle \right]. \tag{22}$$

In an implementation of these recurrence relations, one starts out knowing E_0 and $|\Psi_0\rangle$ (e_0 and $|0\rangle$), and then evaluates E_1, $|\Psi_1\rangle$, E_2, $|\Psi_2\rangle$, and so on.

One can then evaluate the ground-state expectation value of an operator $\mathcal{O}$ to order λ^n, provided one has evaluated the wave function to that order, via

$$\langle \mathcal{O} \rangle = \mathcal{N}^{-1} \sum_{i \geq 0} \lambda^i \sum_{i'=0}^{i} \langle \Psi_{i-i'}|\mathcal{O}|\Psi_{i'}\rangle, \tag{23}$$

with the normalization factor $\mathcal{N}$ given by

$$\mathcal{N} = \langle \Psi|\Psi\rangle = \sum_{i \geq 0} \lambda^i \sum_{i'=0}^{i} \langle \Psi_{i-i'}|\Psi_{i'}\rangle. \tag{24}$$

These formulae can be readily generalized to expansions in more than one parameter,

$$\mathcal{H} = \mathcal{H}_0 + \lambda_1 \mathcal{H}_1 + \lambda_2 \mathcal{H}_2 + \cdots. \tag{25}$$

One needs to consider expansions in two variables for the calculation of $T = 0$ susceptibilities (in which case one would want the terms in the ground state energy proportional to λ_2^2) and there are other cases where the Hamiltonians of interest form a space of dimension greater than one. Let us write the expansions for the ground state energy and wavefunction as

$$E = \sum_{i_1 \geq 0, i_2 \geq 0, \ldots} E_{i_1, i_2, \ldots} \lambda_1^{i_1} \lambda_2^{i_2} \cdots, \tag{26}$$

$$|\Psi\rangle = \sum_{i_1 \geq 0, i_2 \geq 0, \ldots} |\Psi_{i_1, i_2, \ldots}\rangle \lambda_1^{i_1} \lambda_2^{i_2} \cdots. \tag{27}$$

For the ground state energy coefficients one finds

$$E_{i_1, i_2, \ldots} = \langle 0|\mathcal{H}_1|\Psi_{i_1 - 1, i_2, \ldots}\rangle + \langle 0|\mathcal{H}_2|\Psi_{i_1, i_2 - 1, \ldots}\rangle + \cdots. \tag{28}$$

For the wavefunction the result is

$$\langle k|\Psi_{i_1, i_2, \ldots}\rangle = (e_0 - e_k)^{-1} \Bigg[\langle k|\mathcal{H}_1|\Psi_{i_1 - 1, i_2, \ldots}\rangle + \langle k|\mathcal{H}_2|\Psi_{i_1, i_2 - 1, \ldots}\rangle + \cdots - \sum_{i_1', i_2', \ldots}{}' E_{i_1', i_2', \ldots} \langle k|\Psi_{i_1 - i_1', i_2 - i_2', \ldots}\rangle \Bigg] \tag{29}$$

for $k > 0$. The primed sum runs over all $i_l' = 0 \ldots i_l$, except for two terms, the one with all $i_l' = 0$ and the one with all $i_l' = i_l$ (so there are $-2 + (i_1 + 1)(i_2 + 1) \cdots$ terms in the sum). In constructing both formulae we have taken $\langle 0|\Psi_{i_1, i_2, \ldots}\rangle = \delta_{i_1, 0}\delta_{i_2, 0} \cdots$, in analogy with equation (20) for the one-variable expansions.

Note that the multivariable recurrence relations can be carried out in several different ways. If one has two expansion parameters, and one is interested in the coefficient E_{n_1, n_2}, the terms could be computed in the order $E_{1,0}$, $|\Psi_{1,0}\rangle$, $E_{2,0}$, $|\Psi_{2,0}\rangle$, and so on up to $|\Psi_{n_1,0}\rangle$; then $E_{0,1}$, $|\Psi_{0,1}\rangle$, $E_{1,1}$, $|\Psi_{1,1}\rangle$, and so on up to $|\Psi_{n_1,1}\rangle$; then $E_{0,2}$ up to $|\Psi_{n_1,2}\rangle$; and continuing until reaching E_{n_1,n_2}. One could just as well increment the second index before the first index. Yet other permutations are possible, but what is essential is that in order to evaluate E_{n_1,n_2} or $|\Psi_{n_1,n_2}\rangle$ *all* of the energy and wave function coefficients with indices i_1, i_2 such that $i_1 \leq n_1$ and $i_2 \leq n_2$ (except of course $(i_1, i_2) = (n_1, n_2)$) must be evaluated first.

6. Excited-state properties: formalism

Until recently it had been thought that the cluster expansion formalism was not suited to the calculation of excited state properties. The reason for this is that excited state properties are not extensive and therefore not additive over disconnected subclusters, so the basis for the cluster expansion formalism seems to be lost.

For example, consider the energy gap between the ground state and a presumed non-degenerate lowest-lying excited state with some given set of quantum numbers (wave number, S^z, etc). On an arbitrary finite cluster it does not make sense to characterize a state by wave number, so it is not at all clear how to proceed. We might hope to determine the gap to the lowest-lying excited state, whatever that happens to be, but even that is apparently beyond reach. If we consider disconnected clusters A and B and label their individual ground state and lowest excited state energies $E_{0,A}$, $E_{1,A}$ (likewise for B), then for the system with both clusters, $A + B$, the ground state energy is $E_{0,A+B} = E_{0,A} + E_{0,B}$ and its lowest excited state energy is

$E_{1,A+B} = E_{0,A+B} + \min(E_{1,A} - E_{0,A}, E_{1,B} - E_{0,B})$. Clearly it will not generally hold that $E_{1,A+B}$ equals $E_{1,A} + E_{1,B}$, and neither is it true in general that $E_{1,A+B} - E_{0,A+B}$ equals $(E_{1,A} - E_{0,A}) + (E_{1,B} - E_{0,B})$. Therefore the weights associated with these properties will not vanish for the combined system $A + B$, and so the formalism which reduces the calculation of ground state expansions to calculations on a finite number of (connected) clusters seems to be inapplicable here.

But cluster expansions for excited state properties are possible, as we will now show in detail. In retrospect, it seems clear that such calculations should be possible by means of cluster expansions, since excited-state properties had been calculated using cumulant expansion methods for many years and the latter require consideration of only a finite number of diagrams.

For any unperturbed Hamiltonian $\mathcal{H}_0$ which is suitable for a cluster expansion, the eigenstates can be expressed as a product of eigenstates of individual sites. (We should emphasize, as was described in section 2, that the sites of $\mathcal{H}_0$ need not correspond to operators at a single point in space—consider for example the dimerized Heisenberg model of equation (6).) Let us denote the eigenstates of $\mathcal{H}_0$ for an individual site, say site s, by $|0^{\{s\}}\rangle$, $|1^{\{s\}}\rangle$ and so forth, in order of increasing energy $\epsilon_0 \leq \epsilon_1 \leq \cdots$. Note that in writing this last sentence we have implicitly assumed that the expansion is of the high-temperature type. For the low-temperature type expansions it does not make sense to refer to the unperturbed energy of a single-site state, since even in $\mathcal{H}_0$ there are intersite couplings. Because the cluster expansion formalism for excited states is simpler to discuss for high-temperature type expansions let us restrict our attention to this class of problems for the moment, and return to the low-temperature type expansions at a convenient point below.

For an expansion of the high-temperature type, ϵ_0 is strictly less than all the other single-site energies, and for an N-site cluster one has (following the notation of the last section) $|0\rangle = |0^{\{1\}}0^{\{2\}} \cdots 0^{\{n\}}\rangle$. Let us now consider the excited states. There are N states with energy $(N - 1)\epsilon_0 + \epsilon_1$, in which a single site is in the excited state 1, and likewise for all the other single-site excited states. Let us refer to such states as 'single-particle' states, and call all other excited states 'multiparticle' states. It should be clear that for the latter the degeneracies will be of order N^2, at minimum.

In the thermodynamic limit, the excitation spectrum of $\mathcal{H}_0$ is characterized by flat (dispersionless) bands. If a small amount of $\mathcal{H}_1$ is added to the Hamiltonian, the bands associated with single-particle states develop dispersion but they remain discrete, in the sense that there is only one state corresponding to each wave vector. The bands associated with the multiparticle states develop not only dispersion but also breadth; at any fixed wave vector there is a spectrum of states with a non-zero range in energy.

The problems we will be able to address directly by a cluster expansion approach involve the single-particle states: calculations of spectra, which will be entirely analogous to calculations of the ground state energy, and properties such as spectral weights, which will be analogous to calculations of ground-state correlations. We will defer discussion of multiparticle excited states to section 11. We will also need to restrict our present considerations to single-particle states which, in the unperturbed limit, are not degenerate with any set of multiparticle states; otherwise the full (and presently unresolved) complications of multiparticle states arise immediately.

The calculations are simplest for single-particle excitations that are not degenerate with any other single-particle excitations which can be connected by powers of

$\mathcal{H}_1$. Consider, for example, the dimer Hamiltonian (6). The single-site spectrum in this case consists of an $S=0$ state of the antiferromagnetically coupled pair of spins (the ground state), and a three-fold degenerate $S=1$ state. The latter leads to three degenerate sets of single-particle excited states. However, the perturbation is spin–rotation invariant, so S^z is a good quantum number not only for $\mathcal{H}_0$ but also for the full $\mathcal{H}$. As a consequence, the single-particle states with $S^z=1$, 0, and -1 are entirely decoupled from one another and one can do a calculation for the $S^z=1$ states alone, as if the other single-particle states did not exist. Now, if it so happens that two or more sets of single-particle excitations are degenerate and not decoupled all is not lost: in all that follows it will just be necessary to append a discrete 'flavour' index labelling the types of excitations to the site indices in all the considerations that follow.

The key insight which makes the cluster expansions possible is that the appropriate quantities to consider are not simply numbers but, rather, are related to effective Hamiltonians for the single-particle excited states. If we use the unperturbed eigenstates as a basis we can represent the states of an N-site cluster by vectors where the first N components are the amplitudes of the single-particle excited states; then the Hamiltonian has the matrix form shown in figure 3 (*a*). By means of various possible similarity transformations

$$\tilde{\mathcal{H}} = \mathcal{S}^{-1}\mathcal{H}\mathcal{S} \tag{30}$$

one can produce a matrix with the structure shown in figure 3 (*b*). (Precisely how that can be achieved will be described below—though of course the zeros in the off-diagonal blocks will only be correct to some finite order in λ, and likewise for $\mathcal{H}^{\mathrm{eff}}$.) Now, suppose we were to carry out such a procedure for two separated clusters, A and B, with N_A and N_B sites. How is $\mathcal{H}^{\mathrm{eff}}$ for the composite system, $\mathcal{H}^{\mathrm{eff}}_{A+B}$, related to those for the individual clusters $\mathcal{H}^{\mathrm{eff}}_A$ and $\mathcal{H}^{\mathrm{eff}}_B$? If we choose appropriate bases for the unperturbed single-particle states and appropriate similarity transforms, then in many cases the answer turns out to be very simple, namely,

$$\mathcal{H}^{\mathrm{eff}}_{A+B} = (\mathcal{H}^{\mathrm{eff}}_A + E_B\mathcal{I}_A) \oplus (\mathcal{H}^{\mathrm{eff}}_B + E_A\mathcal{I}_B), \tag{31}$$

where $\mathcal{I}$'s denote identity operators (with subscripts identifying the space on which they act) and the E's denote ground state energies. Since we know the ground state energies are additive over disconnected clusters this can be rewritten as

$$(\mathcal{H}^{\mathrm{eff}} - E\mathcal{I})_{A+B} = (\mathcal{H}^{\mathrm{eff}} - E\mathcal{I})_A \oplus (\mathcal{H}^{\mathrm{eff}} - E\mathcal{I})_B, \tag{32}$$

with the subscripts now applying to everything within the parentheses: notice that $\mathcal{H}^{\mathrm{eff}} - E\mathcal{I}$ is manifestly additive over disconnected clusters, albeit not arithmetically!

Let us now present appropriate choices for states and transforms. We take as the basis for the unperturbed single-particle states of cluster A the 'site basis', namely

$$\begin{aligned} |1\rangle &= |1^{\{1\}}0^{\{2\}}0^{\{3\}}\cdots 0^{\{N_A\}}\rangle, \\ |2\rangle &= |0^{\{1\}}1^{\{2\}}0^{\{3\}}\cdots 0^{\{N_A\}}\rangle, \\ &\vdots \\ |N_A\rangle &= |0^{\{1\}}0^{\{2\}}0^{\{3\}}\cdots 1^{\{N_A\}}\rangle, \end{aligned} \tag{33}$$

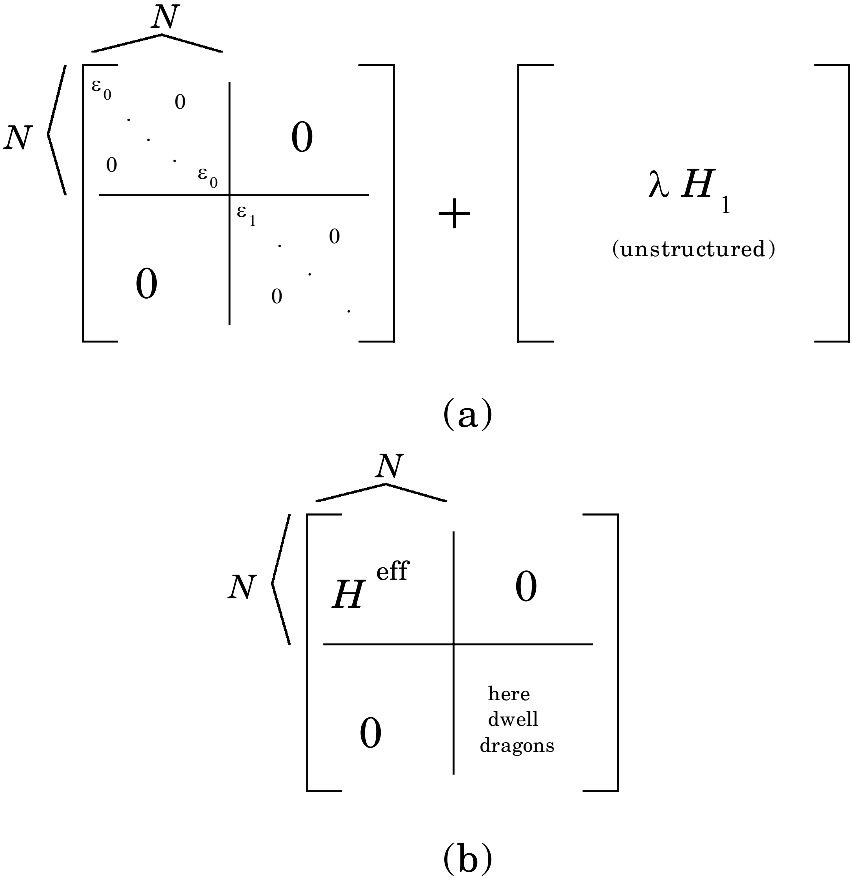


Figure 3. (*a*) The structure of the cluster Hamiltonian in the subspace of unperturbed single-particle excitations and unperturbed states coupled thereto by powers of $\mathcal{H}_1$. (*b*) The structure of the same cluster Hamiltonian, following a similarity transformation that decouples the single-particle excitations from the other states.

and likewise for cluster *B* and composite system $A + B$, where in the latter the sites labelled 1 through N_A are the ones in cluster *A* while sites $N_A + 1$ through $N_A + N_B$ correspond to sites 1 through N_B of cluster *B*.

In order to construct the similarity transformation, we multiply the definition of $\tilde{\mathcal{H}}$ by $\mathcal{S}$ on the left. Since the only part of $\tilde{\mathcal{H}}$ that we are interested in is the $N \times N$ block $\mathcal{H}^{\text{eff}}$, we only need to construct the first N columns of $\mathcal{S}$ which we will consider as a set of state vectors $|\psi^{(1)}\rangle \cdots |\psi^{(N)}\rangle$. Then by expanding every quantity in powers of λ,

$$|\psi^{(l)}\rangle = \sum_{k\geq 0} \lambda^k |\psi_k^{(l)}\rangle, \tag{34}$$

$$\mathcal{H}^{\text{eff}} = \sum_{k\geq 0} \lambda^k \mathcal{H}_k^{\text{eff}}, \tag{35}$$

and collecting like powers of λ in $\mathcal{H}\mathcal{S} = \mathcal{S}\mathcal{H}^{\text{eff}}$ one obtains the set of equations

$$\mathcal{H}_0|\psi_k^{(l)}\rangle + \mathcal{H}_1|\psi_{k-1}^{(l)}\rangle = \sum_{k'=0}^{k}\sum_{l'=1}^{N}|\psi_{k'}^{(l')}\rangle\langle l'|\mathcal{H}^{\mathrm{eff}}|l\rangle. \qquad (36)$$

We will demand that the similarity transformations not mix states within the single-particle manifold, that is, each $\mathcal{S}$ must have an identity matrix in its upper left hand $N \times N$ block, that is

$$\langle l|\psi_k^{(l')}\rangle = \delta_{k,0}\delta_{l,l'}. \qquad (37)$$

This last condition is a generalization of the 'normalization condition' (20) imposed in the context of the ground-state recurrence relations, and it allows us to write down the corresponding recurrence relations (valid for $k \geq 1$) for the effective Hamiltonian

$$\langle l'|\mathcal{H}_k^{\mathrm{eff}}|l\rangle = \langle l'|\mathcal{H}_1|\psi_{k-1}^{(l)}\rangle, \qquad (38)$$

and similarity transform vectors

$$\langle m|\psi_k^{(l)}\rangle = (e_1 - e_m)^{-1}\left[\langle m|\mathcal{H}_1|\psi_{k-1}^{(l)}\rangle - \sum_{k'=1}^{k-1}\sum_{l'=1}^{N}\langle m|\psi_{k'}^{(l')}\rangle\langle l'|\mathcal{H}_{k-k'}^{\mathrm{eff}}|l\rangle\right], \qquad (39)$$

where the states $|m\rangle$ here are eigenstates of $\mathcal{H}_0$ not in the single-particle manifold.

Now consider what happens in the course of constructing the perturbation expansion for $\mathcal{H}^{\mathrm{eff}}$ for the disconnected cluster $A+B$. If it is not possible for the excitation to 'jump' from cluster A to cluster B, that is, $\langle l|(\mathcal{H}_1)^n|l'\rangle = 0$ for any n, any $l = 1, \ldots, N_A$, and any $l' = N_A + 1, \ldots, N_B$, then the off-diagonal $N_A \times N_B$ blocks in $\mathcal{H}^{\mathrm{eff}}$ clearly vanish identically. The only difference between the upper left $N_A \times N_A$ block of $\mathcal{H}_{A+B}^{\mathrm{eff}}$ and $\mathcal{H}_A^{\mathrm{eff}}$ is that the diagonal elements of the two matrices differ by the ground state energy of cluster B. This leads to the result (31) quoted above.

Why should an excitation not jump between disconnected parts of a cluster? The simplest answer would seem to be conservation laws, which must be considered separately in every case. For dimerized Heisenberg models such as (6) the excited states carry spin, which is conserved by the full Hamiltonian. The same is true for the triplet excitations of 'plaquettized' Heisenberg models, where the unperturbed Hamiltonian consists of isolated four-spin clusters. The latter have singlet excitations as well: are the excitations able to move between clusters in that case? The answer appears to be not even in that case, because the two singlet states of the plaquettes have different transformation properties under $\pi/2$ rotation of the plaquettes in real space. The magnon (single spin-flip) excitations of the transverse-field Ising model and the Heisenberg–Ising model also cannot jump, on account of S^z conservation. In the Kondo lattice model, neither charge nor spin excitations can jump, because both charge and spin are conserved in the model.

If the excitations in some model *can* jump from A to B then equation (31) should not hold in general and it is not clear how to proceed: we will take up this point in section 11. In the meantime let us assume that the excitation of interest satisfies suitable conditions such that equation (31) is valid.

At this stage, calculations of excitation spectra are extremely straightforward in principle. One can apply the machinery of cluster expansions, but rather than applying subgraph subtraction to numbers one applies it to the effective Hamiltonians. If cluster C has a subcluster A (with, say, sites 1, 2, and 3 of A corresponding to 3, 2, and 5 of C), then the way one carries out the subtraction for this subcluster would be to subtract $\langle 1|\mathcal{H}_A^{\mathrm{eff}}|3\rangle$ from $\langle 3|\mathcal{H}_C^{\mathrm{eff}}|5\rangle$, and so forth. (For the diagonal elements one must first subtract the cluster ground state energies, which is easily dealt with.)

From the fully subgraph-subtracted $\mathcal{H}^{\mathrm{eff}} - E\mathcal{I}$ for all the clusters with up to a given number of terms of $\mathcal{H}_1$ one can then construct the effective Hamiltonian for excitations on the infinite lattice up to some finite order in λ. The latter is then trivially diagonalized, since its eigenvectors can be expressed as $\exp(i\mathbf{q}\cdot\mathbf{r})$. A convenient way to carry this out is to sum up all of the matrix elements (summing over clusters as well as within each given cluster) associated with a given lattice vector $\mathbf{r}$; if we denote that sum $t_{\mathbf{r}}$ then the spectrum of the excitation is given by

$$\epsilon(\mathbf{q}) = \sum_{\mathbf{r}} t_{\mathbf{r}} \exp(i\mathbf{q}\cdot\mathbf{r}). \tag{40}$$

In fact one can sum up the matrix elements associated with sites at equal distances even earlier in the course of the calculation: subgraph subtraction will work just as well.

It is worth noting that the $\mathcal{H}^{\mathrm{eff}}$ associated with a generic cluster is in general *not* symmetric, either before or following subgraph subtraction. The reason for this is that the similarity transformation constructed following the recurrence relations (38) and (39) is in general *not* a unitary transformation: the vectors $|\psi^{(l)}\rangle$ which make up the first N rows of $\mathcal{S}$ are generally not orthogonal.

While all information concerning the excitation spectrum is encoded in the matrices $\mathcal{H}^{\mathrm{eff}}$ associated with the connected clusters, other excited-state properties must be derived from the vectors $|\psi^{(l)}\rangle$. These vectors are closely related to single-particle excited-state wavefunctions, as one can easily see by considering the equation $\mathcal{H}\mathcal{S} = \mathcal{S}\tilde{\mathcal{H}}$ and the structure of $\tilde{\mathcal{H}}$: the Hamiltonian applied to any of the $|\psi^{(l)}\rangle$ generates a linear combination of the $|\psi^{(l)}\rangle$. The projection onto the single-particle subspace can be achieved by the operator

$$\mathcal{P} = \sum_{l,l'} g^{-1}_{l,l'} |\psi^{(l)}\rangle\langle\psi^{(l')}|, \tag{41}$$

where g is the overlap matrix

$$g_{l,l'} = \langle\psi^{(l)}|\psi^{(l')}\rangle. \tag{42}$$

Note that in equation (41), $g^{-1}_{l,l'}$ refers to elements of the inverse of g and not inverses of elements of g. Since the elements of g are not numbers but rather power series in λ the matrix inversion is not an entirely trivial matter. However, we have in our favour the fact that g is of the form $I + \bar{g}$, with I the identity and $\bar{g}$ of order λ^1. (This is a consequence of the fact that to order λ^0, $|\psi^{(l)}\rangle$ is $|l\rangle$.) Then the identity

$$g^{-1} = (I - \bar{g})(I + \bar{g}^2)(I + \bar{g}^4)(I + \bar{g}^8) + O(\lambda^{16}), \tag{43}$$

reduces the problem of determining g^{-1} to a modest number of additions and multiplications of series.

With the means of constructing $\mathcal{P}$ in hand we can now evaluate spectral weights for the single-particle excitations. There is some single-site operator $\mathcal{O}(\mathbf{r})$ such that $\langle 0|\mathcal{O}|1\rangle\langle 1|\mathcal{O}^{\dagger}|0\rangle = 1$ (that is, it couples the single-site ground state with the excited state of interest) and all other matrix elements of $\mathcal{O}$ vanish. Using the Heisenberg representation, the dynamic structure factor associated with the operator is given by

$$S(\mathbf{q},\omega) = \int \frac{\mathrm{d}t}{2\pi} \exp(-\mathrm{i}\omega t) \sum_{\mathbf{r}} \langle \mathcal{O}(\mathbf{0},0)\mathcal{O}^{\dagger}(\mathbf{r},t)\rangle, \tag{44}$$

where the angular brackets indicate the ground-state expectation value. The frequency-integrated structure factor $S(\mathbf{q})$ is simply expressed in terms of equal-time expectation values and is therefore calculable as a power series expansion in λ using ground-state expansion techniques. But we can now say quite a bit more about the structure factor. At $\lambda = 0$, by construction we have $S(\mathbf{q},\omega) = \delta(\omega - \epsilon)$, where $\epsilon = \epsilon_1 - \epsilon_0$ is the single-site energy gap. For $\lambda \neq 0$, so long as the single-particle excitations remain discrete we anticipate that the structure factor has the form

$$S(\mathbf{q},\omega) = A(\mathbf{q})\,\delta(\omega - \epsilon(\mathbf{q})) + B(\mathbf{q},\omega), \tag{45}$$

where A is the spectral weight and B represents the incoherent background; as $\lambda \to 0$ one has $A \to 1$ and $B \to 0$. The strategy for calculating $A(\mathbf{q})$ is similar to that used for $\epsilon(\mathbf{q})$: one first evaluates a real-space version and then Fourier transforms. The relevant real-space quantity is given by

$$\tilde{A}(\mathbf{r}) = \langle \mathcal{O}(\mathbf{0},0)\mathcal{P}\mathcal{O}^{\dagger}(\mathbf{r},0)\rangle, \tag{46}$$

which is subjected to the cluster-expansion formalism for each value of $\mathbf{r}$; and then $A(\mathbf{q}) = \sum_{\mathbf{r}} \exp(\mathrm{i}\mathbf{q}\cdot\mathbf{r})\tilde{A}(\mathbf{r})$. The ground state expectation values in (46) are straightforward to evaluate on any cluster since the explicit form for the single-particle excited state projection operator $\mathcal{P}$ is known.

Let us now briefly discuss what changes in the strategy detailed above for excited-state calculations are required for low-temperature type expansions. In such cases there is necessarily degeneracy of lowest-lying single-site states, but this degeneracy is broken when the intersite couplings that are part of $\mathcal{H}_0$ are included and a global ground state is selected. Thus the formalism for the excited-state calculations is nearly unchanged; it is just necessary to recognize that the ground state of a single site may depend on which site it is, and likewise for the excited states. Of course one must also take into account the same considerations regarding the choice of clusters (strong versus weak embeddings) that distinguish low- and high-temperature type expansions for ground state properties.

The simplest examples of low-temperature type expansions are provided by expansions about the spin-S ferromagnetic Ising model on lattices in which all sites are equivalent. If one chooses (among the two possibilities) the global ground state to be the one in which $S^z = S$ for all sites, then the single-particle excited states of principal interest are clearly those in which one spin has $S^z = S - 1$. If the unperturbed Hamiltonian is an antiferromagnetic Ising model on a bipartite lattice with all equivalent sites then the single-site ground state is either $S^z = S$ or $-S$, depending on which sublattice the spin sits. There are then two (obviously degenerate) sets of single-particle excitations, those in which a single spin in the 'up' sublattice has $S^z = S - 1$ and those in which a single spin in the 'down' sublattice has $S^z = -S + 1$.

Finally, let us note that for one-dimensional systems a special type of low-temperature type expansion for excited states is possible. In such cases, the unperturbed Hamiltonian supports a finite-energy excitation which is different from changing the state of a single site with respect to the global ground state. Instead one can have a zero-dimensional interface between distinct global ground states; this is generically known as a soliton excitation. Again, the same general formalism applies, but the single-particle (in this case, single-soliton) excited states can not be conveniently expressed in the form (33) since there will be one site, at the interface, which does not have a unique ground state. As a consequence there may be several distinct types of soliton states. These considerations have been discussed in detail, in the context of expansions about Ising antiferromagnetic chains for various values of S, by Singh [32].

7. Constructing clusters: implementation details

In this section we discuss aspects of graph generation and counting. For those readers with experience using Lanczos and other exact diagonalization techniques, the details of weight calculations may appear familiar; thus this section offers the most significant bit of extra technical knowledge required to generate high-order series expansions. Our discussion is oriented towards practical aspects of graph generation, and hence will be kept as informal as possible. We will rely on simple arguments and omit formal proofs. With the help of modern computers, the graphs are constructed essentially by a brute force method, although there are many tricks to make the computation more efficient. From a practical point of view, one may wish to begin with the simplest procedures especially if the bottleneck for computations lies elsewhere (which typically is the case for many quantum cluster expansions). Once the reader is familiar with the basic aspects of graph counting, he can look up more sophisticated algorithms in the literature or invent them for himself. In this section, we first discuss a simple but complete algorithm for generating a list of connected graphs or clusters. Towards the end, we discuss ways of improving the algorithm.

We begin with a few simple definitions. A graph or cluster consists of N nodes and M bonds. A bond joins a given pair of nodes. Our graphs are linear in the sense that there can be at most one bond between two nodes. It would greatly help the reader to keep an example in mind. For this purpose consider the 6-node, 6-bond graph shown in figure 4 (*a*). By an embedding of the graph in a given lattice, we mean an assignment of lattice coordinates to different nodes of the graph, modulo translations. For connected graphs, the number of embeddings is proportional to the number of sites in the lattice (we are working with an infinite lattice). This simply reflects the translational symmetry and hence it makes sense to identify all graphs which are translations of each other, and only discuss counts per lattice site. The aim of the computation is to prepare an exhaustive list of all different types of such graphs with up to $N_{\max}$ nodes or $M_{\max}$ bonds, the node-limited counts applying to low-temperature type expansions, also known as strong embeddings, and the bond-limited counts to the high-temperature type expansions, also known as weak embeddings. Furthermore, one needs to determine the lattice constant and the list of all subgraphs for each graph in the list.

For our simple algorithm, we further consider two graphs to be the same if and only if they are related by a (point-group) symmetry of the lattice. Thus the graphs in figure 4 (*a*) and figure 4 (*b*) are the same but the one in figure 4 (*c*) is different. To get a list of all graphs, we need to represent them in an abstract way. We assume here that we are dealing with high-temperature graphs, and just note that the adaptation to low temperature graphs is straightforward. We will also, for concreteness, restrict our attention to the square lattice, where the sites on which the nodes sit are labelled by two integer coordinates, and interactions are only between nearest-neighbour

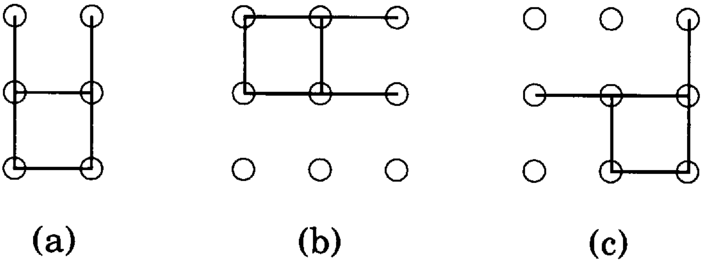


Figure 4. Examples of 6-bond, 6-site graphs embedded in the square lattice.

sites. A bond is associated with the pair of sites that it connects. The list of node coordinates together with the pairs that are connected by bonds completely specifies the graph and its embedding. Note that at this stage how the sites are labelled, namely which site is called first, which second etc. is arbitrary.

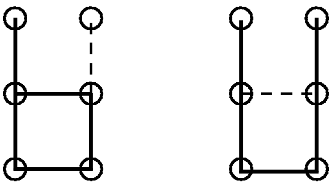
Since graph enumeration involves repeated comparison of one graph with another it is important to have as succinct a representation of the graphs as possible. In a high-temperature type expansion, a graph is completely specified by its bonds. Hence, it would be useful to have coordinates assigned to the bonds. Then these bond coordinates themselves completely specify the graph and one would not need to keep additional information on the sites. In our case, the coordinates of a bond can be defined as the mean of the coordinates of the two sites it connects. Note that this may not always specify the bond uniquely (for example, for second-neighbour interactions on the square lattice two different bonds have the same mean coordinate, and in that case one would need an additional binary index to specify the bond). Thus our graph in figure 4(*a*) is fully specified (if the origin is at the bottom left of the nodes displayed) by saying that there are bonds present at $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$, $(0, 1\frac{1}{2})$, $(1, \frac{1}{2})$, $(1, 1\frac{1}{2})$, $(\frac{1}{2}, 1)$. Once the bonds are drawn at those coordinates, it would be evident what sites are present and what their coordinates are. Note that, in our example, half-integer x-coordinates are associated with horizontal bonds and half-integer y-coordinates with vertical bonds.

So far, our characterization of graphs specifies a graph embedding at a given location and with a particular orientation on the lattice. However, we require a list of distinct graphs or clusters. That is, we need to take only one of the many symmetry-related embeddings as our representative graph. For this we need to construct a simple but definite rule so that, given any graph embedded on the lattice, we can easily determine its corresponding representative graph. To achieve this, we first require that the origin must be one of the nodes of the representative graph. The list of symmetry-related graphs containing the origin is finite. We can choose the representative graph by asserting that among all symmetry-related graphs which include the origin, the representative one is the one with the 'largest' set of bond coordinates, according to a particular lexicographic ordering scheme. In the next paragraph we give an example of such a scheme.

Consider sets of numbers with a given number of elements, for example, the bond coordinates of a graph. Lexicographic ordering works the same way as the ordering of words in a dictionary. To compare two sets of numbers, one first compares the first element of the two sets, followed by the second element, then third element, and so on, as long as the elements agree. The two sets are equal if and only if the entire set of elements is identical. The first time one finds a difference in the elements, the set having the larger element is said to be the larger set. For a given graph, there is an arbitrariness in which bond is called the first, which the second, and so on. This arbitrariness can be removed by ordering the bonds to get the largest set of bond coordinates associated with the graph. Note that for an N-bond graph, the coordinates are represented by $2N$ numbers. We will compare first the x-coordinate of the first bond, then the y-coordinate of the first bond, then the x-coordinate of the second bond, then the y-coordinate of the second bond, and so on. Then for all symmetry-related graphs which include the origin, our representative cluster is the one which leads to the largest set of bond coordinates (after the bond labels are ordered in each individual graph so as to yield the largest set of coordinates for that graph). We can simply refer to this set of bond coordinates as our cluster. We have created a one-to-one correspondence between a cluster and an ordered set of numbers. For our example, the representative cluster embedding has bond coordinates in order: $(2,\frac{1}{2})$, $(1\frac{1}{2},1)$, $(1\frac{1}{2},0)$, $(1,\frac{1}{2})$, $(\frac{1}{2},1)$, $(\frac{1}{2},0)$.

Now, we can obtain the list of required clusters order by order as follows: let the number of bonds in a cluster (or the number of nodes, if one is interested in a node-limited count) denote the order of the cluster. The lowest order cluster is the one with two nodes connected by a bond. All clusters of order $M+1$ can be obtained from some cluster of order M by either adding a bond between existing nodes or adding a bond from an existing node to a new node. For example, the graph of figure 4 (*a*) can be obtained from a 5-bond graph with 5 nodes or a 5-bond graph with 6 nodes, as shown in figure 5. However, if we try to generate all clusters of order $M+1$ by exhaustively adding bonds between nodes or bonds to extra nodes for all clusters of order M, the big problem would be that a given cluster will be generated many times. However, we can find the unique set of numbers that represent the cluster, following the lexicographic ordering discussed earlier, and thus eliminate all duplicates in the cluster generation. Thus we can prepare an exhaustive list of all clusters to the desired order.

The lattice constant of the cluster is simply obtained by performing all the discrete symmetry operations on the cluster and counting the number of embeddings that are not related to each other by translation. For the graph of figure 4 (*a*), the lattice constant is 4. The list of subgraphs can also be obtained by a brute force method. For a cluster with N bonds, one considers all 2^N-2 possibilities, where a given bond is either present or absent (excluding the cases when all bonds are present and all are absent) and looks at the resulting graph of bonds present. One checks to



(a) (b)

Figure 5. The sample graph of figure 4 (*a*) can be obtained from either (*a*) a 5-bond, 5-site graph by adding a site and connecting it by a bond (shown by a thin dashed line) to an existing site, or (*b*) from a 5-bond, 6-site graph, by adding a bond (again shown by a thin dashed line) between existing sites.

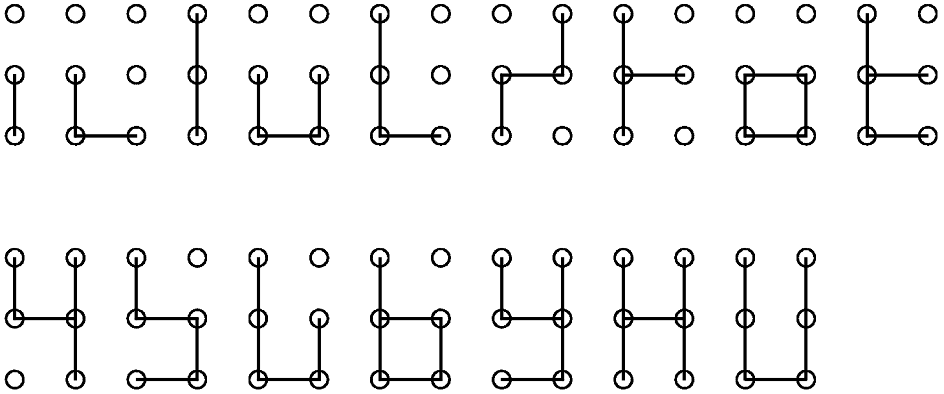


Figure 6. A list of all geometrically distinct subgraphs of the graph in figure 4(*a*). The number of times each subgraph appears in the graph, starting from the upper left hand corner, is 6, 6, 2, 5, 2, 2, 2, 1, 2, 4, 2, 2, 2, 2, 1, and 1, respectively.

see if the present bonds form an acceptable subgraph. This can be done by determining the unique set of bond coordinates for this subgraph and checking with the existing list of smaller clusters to see if it appears on the list. If it does not appear on the list then it must be a disconnected subgraph, and it is discarded. If it appears on the list, then it is a valid subgraph. For the graph of figure 4(*a*), the complete list of its subgraphs is shown in figure 6, with the number of subgraphs of each type stated in the figure caption.

The above scheme is perhaps the simplest conceptually, but it is not particularly efficient. The drawback of the scheme arises from the fact that the number of distinct clusters grows very rapidly with order. Many graphs which are not related to each other by a lattice symmetry operation can still have the same contribution to the physical quantities desired. In an efficient scheme they should be identified as the same cluster, thus only adding to the lattice constant of the cluster and not to the number of distinct clusters, but this is not being done in the method described above. However, the advantage of this scheme, apart from its simplicity, is that every symmetry-unrelated embedding is treated separately. Thus all information required for calculating spatially dependent correlations is present in this list. With modest computational facilities, this scheme is typically practical for 10th order calculations. For problems where the size of the basis set grows rapidly with number of sites or elementary units, such as the Kondo or Anderson lattice, or double-exchange models, or plaquette expansions for Heisenberg models, graph generation is likely to be the easiest part of the calculation, and the above scheme is recommended for its simplicity and generality. On the other hand, for transverse-Ising or Heisenberg–Ising models, where the weight calculations are fast and simple, the graph generation becomes a computational bottleneck, and one needs a more efficient scheme for carrying it.

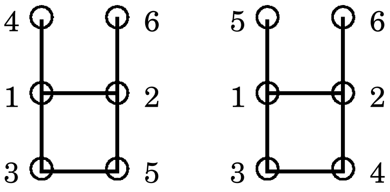
For some properties, such as ground state energy or uniform susceptibility, the weight of a graph only depends on its topology or connectedness, which we will define below. The key to developing a more efficient scheme is to combine together all graphs which have the same topology. Let us order the N nodes of our graphs in some arbitrary manner. We can represent our graph by an $N \times N$ matrix A, with

$A_{ij} =$ if i and j are connected by a bond and zero otherwise. The connectedness or topology of the graph is fully specified by the adjacency matrix A. It is evident on inspection that if two graphs have the same adjacency matrix their associated finite-dimensional quantum mechanical Hamiltonians will be identical and thus they would have identical ground-state energies and uniform susceptibilities. However, in general two such graphs will have different embeddings on the lattice and thus their contributions to any spatially-dependent quantity, such as the spin–spin correlation function for spins at a given spatial separation, will be different. An example of two graphs with the same topology, which were treated as distinct before are the graphs in figures 4 (*a*) and (*c*).

In order to work with the topological graphs or clusters, we need to define a unique adjacency matrix for each graph. Since our labelling of nodes is arbitrary, the adjacency matrix specifies the graph only up to a permutation of those labels. To get a unique representation for the graph, we can lexicographically order the adjacency matrices and, among all permutations of the node labels, choose the one which gives the largest adjacency matrix. This leads to the labelling of nodes for our sample graph as shown in figure 7 (*a*). The adjacency matrix becomes

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This way of picking an adjacency matrix is very inefficient, as the $N!$ permutations required for a cluster of N-nodes rapidly becomes very large. A way around it is to first order the nodes in order of decreasing valence (defined as the number of bonds that meet at that node) and then only consider permutation of nodes with the same valence choosing the largest adjacency matrix among this restricted set. With this scheme, the labelling of nodes for our graph becomes as in figure 7 (*b*) and the adjacency matrix becomes



(a) (b)

Figure 7. Two different labellings of nodes for the sample graph of figure 4 (*a*).

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

One can easily convince oneself that this defines a unique adjacency matrix for the graph. Note that there may be more than one permutation of the node labels which leads to the same adjacency matrix. This would reflect symmetries of the graph, which could in principle be exploited in the weight calculations.

One can now begin by generating all distinct embeddings and prepare a list of topologies by collecting together all graphs with the same adjacency matrix. However, this scheme is again not very efficient. It is much better to obtain directly a list of topological graphs. All topological graphs of order $N+1$ can be obtained from those of order N by the operations of adding a bond between two existing nodes, or adding an extra node to the system and connecting it by a bond to an existing node. After removing all duplicates in the process, one has a complete list of all topological graphs.

To find the lattice constant of a given topological graph is non-trivial, as it is not simply related to the symmetries of the lattice. One needs to exhaustively try out all possible embeddings of the graph on the lattice (barring those related by translations) to determine this number. This can be achieved as follows: one places the first node at the origin of the lattice, and with that point fixed, exhaustively enumerates all possible ways of embedding the rest of the nodes in the lattice, ensuring that no two nodes fall on the same site and that those connected by a bond are indeed nearest neighbours on the lattice. Embeddings that are translations of a previously obtained one are discarded. An excellent recursive algorithm for calculating lattice constants is discussed by Martin [33]. Treated as a topological graph, the lattice constant for our sample graph is 16.

At this stage, spatial correlations can be generated from the list of topological graphs as follows: first calculate all pair-wise correlations for the topological graph. However, the distance associated with any pair will depend on the specific embedding. In order to correctly assign distances to pair-correlations, one needs to go through the list of all embeddings of the graph (as is needed in obtaining the lattice constant) and identify the corresponding spatial distances for each pair. It is then possible to calculate the spatial correlations and wave vector dependent properties.

An even more efficient method for generating a complete list of graphs is to first reduce the graphs to a skeleton graph or topology by ignoring all nodes with valence two. The resulting skeletons contain the basic shape of the graph. The bonds are now appropriately thought of as bridges between the nodes, which can have various lengths. Furthermore, there can be more than one bridge between two nodes. One can now begin by generating a list of skeleton graphs up to certain order and then by assigning various lengths to bridges recover the full list of graphs. From a mathematical point of view, the set of all skeleton graphs can be divided into two categories, (i) star-graphs (also known as multiply connected graphs) and (ii) tree-graphs. The simplest tree-graph is a Cayley tree, which has no closed loops, and its basic elements are linear segments which are joined together in some tree-like pattern. For the most general tree-graph, the basic elements are star-graphs, which are connected together in a tree-like pattern. Then, it is theoretically possible to obtain a listing and count of all connected clusters starting from star-graphs alone by simple algebraic relations. This is significant as the number of star-graphs is much smaller than the number of all graphs, and excellent generation algorithms and high-order listings of star-graphs exist in the literature. However, in practice, getting counts for tree-graphs from those of star-graphs is a non-trivial task and in most cases one proceeds by a brute force enumeration of all possible graphs.

It is evident that there is no unique way to generate a list of all graphs. A variety of methods have been used in the past. We hope our discussion here helps the reader in understanding the various steps involved in the calculation, and in creating their own schemes for generating graphs. For more details in certain cases we refer the reader to the literature [2].

8. Weight evaluations: implementation details

In this section we will describe the nuts and bolts of weight calculations. We aim to make the description sufficiently detailed that an interested reader can write code which is both functional and reasonably efficient. After giving an essentially complete discussion, we will share with the reader two 'tricks of the trade' which make it possible to carry out weight calculations even more efficiently. We are indebted to the Zheng Weihong, J. Oitmaa and C. J. Hamer, at the University of New South Wales, for sharing with us the second of those ideas.

We begin with an overview of the simplest weight calculation algorithm. Let us restrict our attention, for convenience, to a single-variable expansion. Let us also consider our calculation to be complete once the expansions for the ground state energy and eigenvector (see equations (21) and (22)), or effective Hamiltonian and similarity transform (equations (38) and (39)), have been obtained, since the further calculation of expectation values or spectral weights is relatively straightforward.

Step 0. *Construction of 'basic matrix elements'*. In order to calculate the matrix elements of $\mathcal{H}_1$ for any given cluster it is necessary to know the matrix elements associated with each term of the perturbing Hamiltonian (which will typically be associated with two sites). These 'basic matrix elements' need to be calculated only once, before the individual cluster calculations are carried out. One wants to know these matrix elements in the basis in which the single-site (unperturbed) Hamiltonian is diagonal. For some kinds of expansions, such expansions in transverse exchange or transverse field about the $S = 1/2$ Ising model, the matrix elements are trivial to write down by hand. In other cases, such as dimer or plaquette expansions, it is useful to do as much as possible using the computer, in order to avoid insidious errors. (Errors in the basic matrix elements may not manifest themselves in the weight calculations until relatively high order in perturbation theory, and can be very difficult to track down.) In the latter cases, we have found it convenient to do things in the most straightforward way, going from the site-eigenstate basis to the spin-S^z basis, using this for the calculation of the matrix elements, and then converting back to the site-eigenstate basis. Note that when two dimers are coupled by a bilinear term $\mathbf{S_a} \cdot \mathbf{S_b}$, the matrix elements depend on *which* spin of the first dimer $\mathbf{S_a}$ refers to, and likewise for $\mathbf{S_b}$ of the second dimer. It is in general necessary to construct several basic matrices, and to know how to associate each term in the perturbation with the appropriate basic matrix.

Step 1. *Identifying the relevant states for a given cluster.* Typically the perturbation does not destroy all of the symmetries of the unperturbed Hamiltonian, so that there is some obvious subset of the complete set of states of a given cluster for which $\langle k|\Psi_i\rangle$ (or $\langle m|\psi_k^{(l)}\rangle$) can be non-zero. In the context of spin models, total S^z is often a good quantum number; for example, in many cases of expansions about dimer Hamiltonians for ground-state properties, only states with vanishing total S^z need to be taken into account.

In order to make this identification of relevant states useful, it is necessary to construct a list of these states; if there are N_r such states, one needs to have a function which, given an integer between 1 and N_r, returns the site eigenstates associated with that state of the cluster. We have found it most effective to store the site-eigenstate representation of the states as a single integer, in the obvious way: if there are M site eigenstates, denoted 0 through $M-1$, and the state has the form (following the discussion in section 6) $|m_1^{(1)} m_2^{(2)} m_3^{(3)} \cdots\rangle$ then the numerical representation is $m_1 + Mm_2 + M^2 m_3 + \cdots$. For N-site clusters this representation will run into practical difficulties once M^N becomes larger than the largest integer that can be represented efficiently on one's computer (nowadays, typically 2^{31}); however this has not been a problem in our applications. One advantage of storing the site-eigenstate representation as a single integer rather than a list of N integers is clearly the reduced memory requirement. One disadvantage is that one needs to convert to the latter representation many times in the course of the calculations. However, this penalty in computer time is more than compensated for by savings in computational effort that will be discussed in Step 5.

Step 2. *Constructing the unperturbed energies.* This is probably the most straightforward step in the calculation. In principle one could calculate the e_k on demand and further reduce the amount of memory required for the calculation, but we have not tried to implement this.

Step 3. *Executing the recurrence formulae.* It is also not difficult to code the recurrence formulae, either for the ground-state or excited-state calculations. One useful point to recall is that one can determine the $(n+1)$st order energies from the nth order wavefunctions with very little effort. Since memory will typically be the limiting factor in the calculations and the wavefunctions account for the bulk of the code's memory usage, it is always worth taking advantage of the extra order in the energies that one gets almost for free.

In the recurrence formulae, the bulk of the calculations are hidden in the calculations of the matrix elements such as $\langle k|\mathcal{H}_1|\Psi_{i-1}\rangle$. These we consider a separate step.

Step 4. *Evaluating matrix elements of* $\mathcal{H}_1$. Such calculations are, essentially, a long series of look-ups in the table of basic matrix elements. Rather than trying to calculate a single matrix element such as $\langle k|\mathcal{H}_1|\Psi_{i-1}\rangle$, it is easier and more efficient to calculate a vector such as $\mathcal{H}_1|\Psi_{i-1}\rangle$ all at once: starting with each unperturbed eigenstate $|p\rangle$ such that $\langle p|\Psi_{i-1}\rangle \neq 0$, apply each term in $\mathcal{H}_1$ to $|p\rangle$. That last step is where the basic matrix elements come in. The resulting non-zero elements, multiplied by the relevant entry in $|\Psi_{i-1}\rangle$, should then be accumulated in a new vector, provided the non-zero element does not correspond to a state of the low-energy manifold (see equations (20) and (37)).

Note that in the evaluation of the matrix elements, as described above, we have implicitly assumed that the unperturbed eigenstates are represented as lists of single-site eigenstates. In fact, one needs to know what these matrix elements are when the states are represented by integers from 1 to N_r, as described in Step 1. In Step 1 we have already described (at least in principle) how to go from that integer to the list of single-site eigenstates. Going the other way is sufficiently important to do properly that we will treat it as another separate step.

Step 5. *Finding the label for an unperturbed eigenstate given the single-site eigenstate representation.* In principle, one could carry out this step by searching through all N_r relevant states until one finds a match. However, we would not advise it! A far more efficient procedure is possible, provided that one uses an integer representation (as described in Step 1) to encode the single-site eigenstate representation. When carrying out Step 1, it is natural to generate the relevant states in such a way that the integer representations are an increasing function of the state labels. Thus one has an ordered list of the relevant states. Searching such an ordered list can be done efficiently using a binary search: this step is then reduced from an $O(N_r)$ process to an $O(\log_2 N_r)$ process.

This works very well for the ground-state calculations. In such cases, the unperturbed ground state will inevitably carry the lowest integer representation and will be the ‘first’ on the list of N_r states. However, there is a bit of a complication for the excited-state calculations. Then, it is convenient for the purposes of Step 3 that the unperturbed low-energy manifold, comprising say L states, to carry the labels 1 through L. However, those states will not have the L lowest integer representations among the relevant states. One can still code Step 3 simply and carry out Step 5 efficiently, but it is necessary to construct two additional arrays of integers with N_r elements. In Step 1, one first constructs an integer-representation-ordered list of relevant states. One then rearranges the order of this list so that the low-energy manifold comes first; but when doing this rearranging one generates an index array, and from the latter one generates an inverse index array. The index array shows how to go from the ‘current’ list of states, in which the low-energy manifold comes first, to the integer-representation-ordered list. That is, suppose that the ‘first’ state in the low-energy manifold was the thirteenth state in the integer-representation-ordered list. Then the first element of the index array is 13, and the thirteenth element of the inverse index array is 1.

With the integer-representation-ordered list, the index array, and the inverse index array all at hand, the procedures described above are modified as follows. In order to carry out Step 5, one does a binary search through the integer-representation-ordered list, then uses the inverse index array to determine the state label. In order to carry out Step 4, starting with a state label one uses the index array to find where the integer representation of the state sits in the ordered list, and then uses the integer representation as before.

Let us now turn to the promised tricks of the trade.

First, use a sparse matrix representation for the basic matrices. This trick is very easy to implement, and if there are several states per site it can produce significant savings in computational time. The point is that by using a sparse matrix representation one avoids all look-ups of basic matrix elements that are zero. Since the large majority of basic matrix elements tend to be zeros when the individual sites have many states (and the perturbation does not break too many symmetries), one can get a substantial speed-up in Step 4.

Second, identify the 'truly' relevant states rather than just the 'potentially' relevant states as indicated in the description of Step 2 above. That is, one can determine, for a given graph and order n to which some ground-state property is desired, exactly the unperturbed eigenstates $|k\rangle$ which one needs to evaluate the $\langle k|\Psi_i\rangle$ (and analogously for excited state calculations). Consider, for example, the ground state energy to order n. A straightforward implementation of the recurrence relations (21) and (22) would require the evaluation of $\langle k|\Psi_{n-1}\rangle$ for all states $|k\rangle$ for which it is non-zero. However, all one *really* needs to know it for are for $|k\rangle$ such that $\langle 0|\mathcal{H}_1|k\rangle$ is non-zero. Let us look at this from another angle. A calculation of the ground state energy to nth order, fully written out, involves terms of the form

$$\frac{1}{(\text{energy denominator})}\langle 0|\mathcal{H}_1|k_1\rangle\langle k_1|\mathcal{H}_1|k_2\rangle\cdots\langle k_{n-1}|\mathcal{H}_1|0\rangle. \tag{47}$$

Each application of $\mathcal{H}_1$ potentially takes one farther away from the unperturbed ground state $|0\rangle$, but after n such steps one has to be back at $|0\rangle$. Hence the intermediate states which require the most applications of $\mathcal{H}_1$ to reach from the ground state must be in the middle term; these are states $|k\rangle$ which have $\langle k|\mathcal{H}_1^q|0\rangle \neq 0$, where $q = n/2$ for even n and $(n-1)/2$ for odd n.

Similar arguments allow one to determine the truly relevant states for other properties such as susceptibilities, excited state energies, and ground state expectation values (but see below). By only storing the $\langle k|\Psi_i\rangle$ for the truly relevant states, enormous savings in memory requirements and computation time are possible. Let us briefly outline how the program outline given above needs to be changed. Steps 0 and 1 go as before, and then one immediately goes on to determine which of the potentially relevant states are truly relevant. This involves applying $\mathcal{H}_1$ to the unperturbed ground state some number of times and keeping track of which states are generated in the process. One might imagine that this step would take almost as much time as calculating the ground state energy using the full set of potentially relevant states, but in fact it goes much faster because no arithmetic is being done. Once the truly relevant states are determined, one then goes to Step 2 and continues.

A potential problem with the improved algorithm is that it has fewer self-consistency checks than the simpler, slower algorithm. When one is keeping all of the potentially relevant states in the calculation, there is a strong check in Step 5, which will fail if the state for which one is trying to find the corresponding label is not a potentially relevant state. That is, if one seems to be finding that $\langle k|\mathcal{H}_1|\Psi_i\rangle$ is non-zero for some $|k\rangle$ which is not potentially relevant, then there must be an error somewhere in the code. However, when only the truly relevant states are being dealt with, such 'failures' in Step 5 are entirely routine. Consequently, we have found it useful to first carry out calculations using the potentially relevant states up to some order (typically fourth) by which point one expects any coding errors to manifest themselves. After the simple code is thoroughly debugged the calculation using the truly relevant states can be run, and one can check the consistency of the latter's results with those of the former.

Another weakness of the improved algorithm, when ground-state expectation values are of interest, is that it is only guaranteed to be useful when the quantities being evaluated are *diagonal* in the unperturbed basis. For such quantities the truly relevant states are the same as for the ground state energy. For quantities which are off-diagonal in the unperturbed basis, additional states are truly relevant—possibly so large a fraction of the potentially relevant states that the computational overhead associated with identifying them could exceed the time saved in carrying out the rest of the calculation.

9. Applications

In this section we present a catalogue of calculations utilizing high-order convergent Rayleigh–Schrödinger perturbation expansions. In some cases, we will mention work based on low-order expansions (which are typically carried out by hand via cumulant expansion rather than by the methods described in this review) which is of interest for one reason or another, but we have made no effort to include all such calculations. High-order strong-coupling expansions have found many applications in the context of field theory, but our coverage of that literature is not intended to be comprehensive. We leave a thorough review of that work to those with better qualifications in that field. In keeping with our condensed-matter bias, we will try to express the degrees of freedom in the various models, as much as possible, in terms of quantum spins.

We have two principal purposes in presenting this overview of the literature. One is to give the reader a relatively easy way to find out what problems in condensed matter physics have been addressed using high-order perturbation theory. The other is to give the prospective user of these methods a better idea of what sorts of problems might be addressed.

All the works we will describe here have several standard components. There is a family of models of interest, which is decomposed into an unperturbed Hamiltonian $\mathcal{H}_0$ and a perturbation $\lambda\mathcal{H}_1$ (or perhaps several perturbations). Power series in λ are computed for various quantities of interest, which are extrapolated and interpreted.

In order to organize this compendium, we will use the *unperturbed Hamiltonians* which have underlied the perturbation expansions. Under each $\mathcal{H}_0$, we will then mention the models of interest and the quantities that were calculated, giving references to the literature.

9.1. *Ising models*

Ising models are characterized by the value of S of the quantum spins that are the degrees of freedom in the model, and exchange couplings of the form $S_i^z S_j^z$. Typically the couplings are between nearest-neighbour sites on a bipartite lattice, but more general couplings are possible; the crucial feature, for the purpose of perturbation expansions, is that there should be just two ground states related by $S_i^z \leftrightarrow -S_i^z$.

Most of the calculations in the literature are concerned with $S = 1/2$ models, so when S is not explicitly mentioned that value should be assumed.

Ising models are the obvious starting points for calculations of Néel-ordered states in antiferromagnets. They have also proven useful in calculations for models with algebraic Néel order (such as the $S = 1/2$ Heisenberg chain), and even in some cases where Néel order is absent (where demonstrating the limit of stability of the Néel-ordered phase is the main problem).

9.1.1. *Expansions for Heisenberg models*

As noted in section 3, some of the earliest applications of the cumulant expansion techniques [5–9] were to the Heisenberg–Ising models,

$$\sum_{\langle ij\rangle} S_i^z S_j^z + \lambda \sum_{\langle ij\rangle} (S_i^x S_j^x + S_i^y S_j^y), \tag{48}$$

which interpolate between the Ising model at $\lambda = 0$ and the Heisenberg model at $\lambda = 1$.

There have been more studies of Heisenberg models by series techniques than of any other class of models. Section 10 describes how the series calculations have fit into the whole framework of numerical studies of the square-lattice, $S = 1/2$ Heisenberg model. We will not repeat the review of that literature here. One related work not listed in that section is Oitmaa and Zheng [31], in which expansions were carried out not only in the transverse exchange couplings but also in the strengths of staggered longitudinal and uniform transverse fields.

Ising expansions can also be applied to the square lattice with further than nearest-neighbour exchange interactions. Including second-neighbours, Oitmaa and Zheng [34] have constructed expansions in the transverse exchange for order parameters, susceptibilities, and energy gaps, in order to determine the domains of stability of magnetically ordered phases.

$S = 1$ square-lattice Heisenberg model ground state properties (ground-state energy, sublattice magnetization, transverse susceptibility, and spin stiffness) have been estimated by expansions around the Ising model by Singh [35] and by Hamer *et al.* [36].

Oitmaa *et al.* [37] have considered the nearest-neighbour Heisenberg–Ising antiferromagnet on the honeycomb lattice, generating expansions for the ground-state energy, triplet gap, magnetization, and staggered longitudinal susceptibility. In this same paper they also studied the 'XY-Ising' honeycomb antiferromagnet (taking as the perturbation $\sum_{\langle ij\rangle} S_i^x S_j^x$ rather than both the xx and yy terms) as well as the triangular lattice Heisenberg–Ising ferromagnet. These authors have also examined the XY-Ising square lattice ferromagnet [38], carrying out series expansions for the ground state energy and energy gap, magnetization, longitudinal susceptibility, and the two transverse susceptibilities. Later, they constructed a similar collection of perturbation expansions for the $S = 1/2$ and $S = 1$ Heisenberg–Ising antiferromagnets and the $S = 1/2$ XY-Ising model on various three-dimensional lattices, and made detailed comparisons with the results of spin-wave theory (which they extended to order $1/S^3$) [39].

$S = 1/2$ Heisenberg–Ising chain correlation functions were studied by Singh *et al.* [40]. The properties of interest were both the longitudinal and transverse correlation functions, which were found to exhibit different critical properties as the Heisenberg model was approached.

Spatially anisotropic $d = 2$ Ising models were investigated by Ising expansions (as well as other expansions) by Affleck *et al.* [41]. The issue being addressed was whether Néel order survived in the Heisenberg model for strongly anisotropic couplings. Quantities for which series were generated included the magnetization and static correlation functions (from which the correlation length anisotropy was derived).

The $S = 1$ chain was examined using expansions about the Ising model by Singh and Gelfand [42]. In fact, a wider class of models than the $S = 1$ Heisenberg–Ising chain was considered, in that 'dimerization' (or 'bond alternation') was another parameter. To be precise, the models under consideration had the form

$$\sum_i (S^z_{2i} S^z_{2i+1} + \delta S^z_{2i+1} S^z_{2i+2}) + \lambda \sum_i (\text{transverse components}). \qquad (49)$$

This model has a rich phase structure in the δ, λ plane, which was explored in part by expansion in λ of the staggered magnetization and correlation functions. (It is worth noting that the phase diagram presented in that paper is not entirely correct, in that the phase boundary between the Haldane phase and the XY-ordered phase does not actually go into the multicritical point: see Kitazawa *et al.* [43]. That particular phase boundary was not accessible to any of the series expansions, and so was drawn entirely speculatively.)

Elementary excitation spectra for $S = 1/2$, 1 and 3/2 Heisenberg–Ising chains have been calculated by Singh [32]. There are two noteworthy features of this work. One is that topological excitations were studied. The second was the use of an extra term in the perturbing Hamiltonian for $S = 1$, in addition to the transverse exchange terms, in order to make the Heisenberg model accessible to the expansion about the Ising model. The Heisenberg–Ising model for $S = 1$ exhibits a phase transition before the Heisenberg model is reached; the latter lies in the 'Haldane phase', in which the spin-chain exhibits short-range magnetic order. In order to estimate the magnon spectrum for the $S = 1$ Heisenberg chain it is necessary to somehow bypass that phase transition. The method used in that paper was to add a staggered field $\sum_i (-)^i S^z_i$ to the Hamiltonian with coefficient $(1 - \lambda)$, that is, it was added to both $\mathcal{H}_0$ and $\mathcal{H}_1$, in such a way that the Heisenberg model did not have the term at all. However, for $\lambda < 1$ the degeneracy of the Néel states was broken, and the transition between the Ising phase and the Haldane phase was avoided.

Recently, there has been much interest in Heisenberg models on lattices with larger-than-minimal unit cells. This interest has been spurred by the recognition that the magnetic properties of various compounds ought to be well represented by such models. Such lattices include 'spin ladders' (comprising two or more chains), 'bilayers' and various 'CAVO' lattices (after the compound CaV_4O_9 which it is supposed to represent; the reason for the plural is that there have been several proposals regarding the relative strength of the various exchange interactions). Most of the systems which have motivated the calculations are actually magnetically *disordered* down to anomalously low temperatures, so expansions about the Ising model serve mainly to locate the domains of stability of the Néel-ordered phases in the models. Ising expansions for ground state and excited state properties for the spins ladders were carried out by Oitmaa *et al.* [44], and for one class of CAVO models by Zheng *et al.* [45]. Another property that has been calculated for spin ladders based on Ising expansions is the magnetization in a finite external field [46]. Zheng [47] has carried out a comprehensive set of expansions about the Ising model for antiferromagnetically coupled square-lattice bilayers. Bilayers are supposed to describe the magnetic interactions in certain high-T_c parent compounds (such as

$YBa_2Cu_3O_6$) well, so in that case the properties of the magnetically ordered phase are in fact of main interest. Similar Ising expansions for $S = 1$ bilayers were later presented by Gelfand *et al.* [48].

One area that has barely been touched by series expansion methods is *inhomogeneous* systems. The only real difference between inhomogeneous and homogeneous systems, insofar as the cluster expansion formalism is concerned, lies in the identification of the clusters and their lattice constants. Gelfand and Glöggler [49] have presented results (which were actually generated by R. R. P. Singh) for semi-infinite Heisenberg–Ising chains based on Ising expansions. The quantities of interest in this case were local susceptibilities, that is, the derivative of the local transverse magnetizations with respect to the strength of a uniform transverse field. The series were not particularly well behaved; whether that is a general feature of perturbation expansions in inhomogeneous systems will not become clear until more such calculations are performed.

9.1.2. *Transverse-field Ising model*

The transverse-field Ising model is given by the family of Hamiltonians

$$J\sum_{\langle ij\rangle} S_i^z S_j^z - h\sum_i S_i^x \tag{50}$$

in which $\langle ij\rangle$ typically runs over nearest-neighbour pairs on a lattice. Note that for a bipartite lattice, the sign of J makes no difference, and in fact the ferromagnetic case, for which the Ising model always has a trivial ground state, has traditionally been of interest. Either the Ising or the transverse-field part of the model can serve as an unperturbed Hamiltonian, and in fact most studies of this model attack it from both ends. The interest in this model comes mostly from theoretical considerations, such as demonstrating universality of the order–disorder transition for various lattices in a given dimension, and demonstrating that the critical points lie in the same universality class as the finite-temperature Ising transition in one dimension greater.

The earliest perturbative treatment for this class of models seems to have been carried out by Pfeuty and Elliot [10], who constructed series for the magnetization to third order in the transverse field on the square and simple cubic lattices.

Marland [21], applying 'modern' cluster expansion techniques for the first time, constructed series for the ground state energy, magnetization, (longitudinal) susceptibility, and the next two higher derivatives of the energy with respect to longitudinal magnetic field on the linear chain, square lattice, and triangular lattice. The order of the series depended on the property and lattice, but was never less than $(h^2)^6$. (It should be clear by symmetry that all the properties mentioned above are even in h.) Oitmaa *et al.* [50] extended the energy, magnetization, and susceptibility series and also calculated the energy gap, on the square, triangular, and honeycomb lattices. Zheng *et al.* [51] carried out the analogous calculations for simple cubic, body-centred cubic, and face-centred cubic lattices, with most of the series coefficients determined to order $(h^2)^{12}$.

9.2. *Non-collinear Ising models*

It is easy to envisage (and there exist many examples in nature) of magnetic order in which the net magnetizations on various sublattices are not collinear with each other. Since Ising expansions have proven extremely reliable in estimating the effects of quantum fluctuations on Heisenberg models which exhibit Néel order, as discussed in section 10, one might imagine that expansions about Hamiltonians which trivially exhibit a desired non-collinear spin ordering might be be similarly useful.

This type of calculation has been carried out by Singh and Huse [52] for the nearest-neighbour triangular and kagomé lattice antiferromagnets. Unperturbed Hamiltonians were constructed which exhibited two degenerate ground states with three-sublattice order. Expansions in the 'transverse' exchange components (everything in the Heisenberg model that was not in $\mathcal{H}_0$) were constructed for ground state properties (notably the sublattice magnetization), the analysis of which suggested that the kagomé antiferromagnet is not magnetically ordered while the triangular antiferromagnet is significantly less ordered than implied by spin-wave theory. The former conclusion is consistent with other studies, but the latter conclusion is not consistent with the best finite-size calculations [53, 54].

9.3. *Large-field models*

A Hamiltonian which includes only a uniform magnetic field, $\sum_i S_i^z$, can serve as the unperturbed Hamiltonian for a 'high-temperature type' expansion. This approach has been used in calculations for the transverse-field Ising model, approaching the critical point from the disordered side.

In addition to papers mentioned in the subsection on Ising expansions [10, 51], calculations are also presented in papers by Hamer and Irving [26], Hamer and Guttman [55], and He *et al.* [25]. In each case the ground state energy, energy gap, and transverse susceptibility were calculated for square and triangular lattice models; in later papers the series were generated to higher order. The last paper is noteworthy for including a very nice description of the cluster expansion method.

Kadanoff and Kohmoto [19] used the ground-state energy of the one dimensional transverse-field model as a test of their particular cluster expansion formalism.

9.4. *Dimer and plaquette models*

A pair of spins with the same value of S, when coupled antiferromagnetically, has a non-degenerate ground state. Thus Hamiltonians consisting of isolated 'dimers' can serve as bases for high-temperature type expansions. Likewise, if four $S = 1/2$ spins are coupled antiferromagnetically in the form of a square, with arbitrary couplings across the diagonals, the ground state is non-degenerate except for special ratios of the 'side' to 'diagonal' couplings; these 'plaquettes' can also serve as the bases for expansions. The perturbing Hamiltonian in either case couples the originally disconnected spins. Unlike any of the expansions discussed in preceding subsections, the dimer and plaquette expansions can have rotationally invariant (in spin space) interactions for all values of the perturbation.

The first dimer expansion was carried out by Harris [11] for the alternating Heisenberg chain. As mentioned in section 3, this work was noteworthy in that an expansion for the entire elementary excitation spectrum was constructed. Some of the first dimer expansions done to high order, by cluster expansion methods, were reported by Singh *et al.* [27]. The calculations were of ground state correlations and the antiferromagnetic susceptibility, for three different lattice geometries—the chain, and two distinct dimer coverings of the square lattice (in which the remaining nearest-neighbour bonds defined the perturbation). Shortly thereafter, dimer expansions of ground state correlations and $T = 0$ susceptibilities were applied to the square-lattice Heisenberg model with further neighbour couplings [56, 57]. Dimer expansions for nearest-neighbour square-lattice Heisenberg models with spatially anisotropic couplings (again, for ground state correlations and susceptibilities) were presented by Affleck *et al.* [41].

Several years prior to the calculations by Singh *et al.* [27], Kohmoto *et al.* [58] carried out detailed analysis of a one-dimensional quantum Hamiltonian version of the Ashkin–Teller model, using high-order perturbation expansions constructed following the formalism of Kadanoff and Kohmoto [19]. As noted in their paper, the quantum Hamiltonian is equivalent to a dimerized *XXZ* chain, that is, the interaction is not rotationally invariant but instead has the more general form $S_i^x S_j^x + S_i^y S_j^y + \gamma S_i^z S_j^z$. For any value of γ the two-site problem remains exactly solvable and for $\gamma > -1$ the ground state is non-degenerate and a dimer expansion is possible. The calculation by Kohmoto *et al.* was in fact an expansion about the dimer limit. This work was extended more recently by Yamanaka *et al.* to examine the case of ferromagnetic interdimer coupling in the *XXZ* chain [59], and Heisenberg interdimer couplings (both ferro- and antiferromagnetic) between *XXZ* dimers in a chain [60]. All of these calculations were for the ground state energy and various derivatives of the ground state energy with respect to parameters in the Hamiltonian.

Dimer expansions have been constructed for Heisenberg chains with second neighbour interactions (all of which are included in $\mathcal{H}_1$, in addition to half of the nearest neighbour couplings), as discussed briefly by Singh and Gelfand [61]. Another variation on dimer expansion for Heisenberg models is possible when $S \geq 1$. In that case, the Hamiltonian can contain terms beyond the 'bilinear' coupling $\mathbf{S}_i \cdot \mathbf{S}_j$, because higher powers (such as $(\mathbf{S}_i \cdot \mathbf{S}_j)^2$, the 'biquadratic' interaction) are independent operators, and they can be put in $\mathcal{H}_0$ as well as $\mathcal{H}_1$. Expansions for ground state correlations and susceptibilities involving such Hamiltonians for $S = 1$ chains have been carried out by Singh and Gelfand [42].

Dimer expansions have been applied to models of coupled Heisenberg chains and planes. In these cases, $\mathcal{H}_0$ consists of the interchain or interplane coupling, with all of the intra-chain–plane interactions put into $\mathcal{H}_1$. For coupled planes, ground state correlations and susceptibilities were first treated by Hida [62]; these calculations were extended to higher order and expanded to include the excitation spectrum by Gelfand [63] and Zheng [47]. Matsushita *et al.* [64] have considered a more general class of bilayer models, in which the interplane couplings were allowed to be different (even of different sign) in the two planes. These calculations were restricted to the excitation spectrum. Another more general class of Heisenberg bilayers was considered by Gelfand *et al.* [48]. They carried out dimer expansions for the energy gap and antiferromagnetic susceptibility for models with $S = 1$ up to $S = 4$, in order to see how the critical inter/intra-layer coupling ratio varied with S.

For two-chain ladders, dimer expansions to eighth order for the triplet excitation spectrum have been constructed by Oitmaa *et al.* [44].

Recently, Barnes *et al.* [65] have constructed dimer expansions for the triplet spectrum (and also spectral weights and bound state energies!) for the alternating Heisenberg chain. Some of the low-order results were obtained using perturbation theory (standard cumulant expansion formalism). However, many of the series coefficients were actually determined by means of 'multi-precision' (in effect, exact) finite-size calculations. Given values of, say, the ground state energy per site, for various values of the perturbation parameter, it is clear that one can in principle extract the coefficients of the polynomial which gives the analytic connection between perturbation parameter and energy. In practice the reconstruction of the coefficients is a numerically unstable with finite-precision arithmetic, but if the energies are known exactly and the coefficients are known to be rational fractions with relatively small denominators one can carry out the procedure. A finite system with periodic boundary conditions has the same coefficients as the infinite chain up to some order determined by the system size, so Barnes *et al.* were limited by the size of system they could diagonalize using multi-precision arithmetic. We should note that this approach should work for any one-dimensional system, but does not generalize (in any obvious way) to higher-dimensional systems. It also strikes us as being computationally less efficient than perturbation theory. However, these considerations in no way detract from the many nice results Barnes *et al.* have obtained.

Plaquette expansions were first motivated by the suggestion of Ueda *et al.* [18] that the spin Hamiltonian for CaV_4O_9 might best be thought of in terms of plaquettes. Those authors also carried out, analytically, a second-order plaquette expansion for the ground state energy and triplet gap. Higher-order plaquette expansions in the context of CAVO Heisenberg models were carried out by Gelfand *et al.* [66] and Zheng *et al.* [45]. Note that dimer expansions can also be applied to the CAVO Hamiltonian. In addition to the works noted above [18, 45], a dimer expansion for the ground state energy was constructed by Meshkov and Foerster [67].

The attempt to qualitatively determine the appropriate spin Hamiltonian for CaV_4O_9 did not end with the work of Ueda *et al.* An *ab initio* calculation by Pickett [68] leads to a rather different class of Hamiltonians—which remarkably enough are best described in terms of interacting plaquettes, but different plaquettes than those proposed by Ueda *et al.* A neutron scattering study [69] seems more consistent with Pickett's model just on the basis of lowest-order perturbation theory, and the experimental spectrum has been analysed in light of third-order plaquette expansions by Fukumoto and Oguchi [70]. (We should note that some features of the spectral weight data seem difficult to explain by means of the Hamiltonian that works well for the spectrum.)

9.5. *Strong-coupling symmetric Kondo lattices*

We now turn our attention to models which cannot be described solely in terms of localized spins. It is obvious that many of the most important models in condensed matter physics are models of *fermion* rather than *spin* dynamics. Unfortunately, at least from the standpoint of convergent perturbation expansions, the familiar exactly solvable models such as free fermions on a lattice (i.e. 'tight-binding models') are gapless in the thermodynamic limit for generic parameter values.

However, there are some significant exceptions. Consider, for example, the symmetric Kondo lattice model. Each lattice site is associated with a $S = 1/2$ spin and a single orbital. The spin is exchange-coupled to the spin density associated with the orbital, that is, one term in the Hamiltonian has the form

$$J \sum_i \mathbf{S}_i \cdot (c^{\dagger}_{i\alpha} \sigma_{\alpha\beta} c_{i\beta}), \qquad (51)$$

where $c^{\dagger}_{i\alpha}$ creates an electron with spin α on site i and σ represents the Pauli matrices. The factor J is the Kondo coupling strength. We will apply the Einstein summation convention to the spin indices. Another term in the Hamiltonian is the electronic kinetic energy. In the narrow-band limit this is well represented by nearest-neighbour 'hopping'

$$-t\sum_{\langle ij\rangle} c_i^{\dagger}c_j + c_j^{\dagger}c_i. \qquad (52)$$

Finally, there is a chemical potential for the electrons, but in the symmetric model this is taken to be zero.

Associated with each site there are 8 possible states (two states for the fixed spins and four states for the orbital), and for $J > 0$ a single site has a unique ground state. This state has exactly one electron in the orbital, which is singlet-paired with the localized spin. Thus one can construct convergent perturbation expansions in the ratio t/J. These calculations are somewhat more challenging to carry out than the corresponding calculations for spin systems because one needs to take into account fermion anticommutation in order to correctly determine the sign of kinetic energy matrix elements.

Expansions for ground state correlations and susceptibilities have carried out by Shi *et al.* [71] for linear, square, and simple-cubic lattices. For the latter two lattices, these were used to establish phase boundaries between magnetically disordered and ordered phases. Note that these symmetric Kondo lattice models are always insulators, that is, there is always a gap to states which have more or less than one electron per orbital.

9.6. *Hubbard and t–J models at half-filling*

The one-band Hubbard model and the *t–J* model have been subject to an extraordinary degree of theoretical analysis over the past decade, thanks to the discovery of high-temperature superconductivity in the cuprate perovskites and the theoretical proposals (starting with Anderson) that such models embody the relevant physics. At half-filling, where the *t–J* model reduces to the Heisenberg model and so does the Hubbard model, at least in the limit of large on-site repulsion U, it is possible to apply convergent perturbation expansions and obtain interesting results. The 'trick' which makes that possible is to break spin–rotational symmetry in intermediate stages of the calculation.

For the *t–J* model at half-filling, one can take the Ising model for $\mathcal{H}_0$, and then put the transverse exchange as well as the kinetic energy into the perturbation. Of course, precisely at half-filling this is nothing more than an Ising expansion for the Heisenberg model. However, states with one electron short of half-filling can be treated as a manifold of discrete excitations, and so can multi-hole bound states. This program has been carried out using cumulant (rather than cluster) expansion methods by Prelovšek *et al.* [14].

For the Hubbard model at half-filling a somewhat more radical strategy is necessary. In the absence of kinetic energy the ground state is 2^N-fold degenerate. The degeneracy must be broken before any convergent expansion can be carried out. Shi and Singh [72, 73] approached this problem by putting the antiferromagnetic Ising model (in addition to the on-site interaction and the chemical potential term) in $\mathcal{H}_0$, and including the opposite term in $\mathcal{H}_1$. When the coefficient of the perturbing term reaches 1, one is left just with the Hubbard model. They were then able to calculate many characteristics of the antiferromagnetic half-filled Hubbard model as a function of the U/t ratio, such as the sublattice magnetization, transverse susceptibility, and spin stiffness. Both large U/t (Heisenberg model limit) and small U/t (spin-density-wave limit) were accessible using this approach.

9.7. *Boson–Hubbard models at integer filling*

Another model of current interest that involves 'mobile' degrees of freedom is the boson Hubbard model. At integer fillings there is a unique ground state when the coefficient of the kinetic energy term vanishes, with a gap (of order U) to other states which do not have equal occupancy at every site. Thus one can perturb in the kinetic energy. A low-order calculation, most notably for the gap to excited states with particle numbers that differ by one from half-filling (which vanishes at the conductor–insulator transition) along these lines has been carried out by Freericks and Monien [15]

9.8. *Other models from condensed matter physics*

There are a few other models for which convergent perturbation expansions have been constructed which (i) were motivated by problems in condensed matter physics but (ii) are not readily described in terms of quantum spins, bosons, or fermions. Typically the $T = 0$ properties of these models are known (or expected) to be closely related to the finite-temperature properties of some model in one spatial dimension higher. One example is the relationship between the d-dimensional $T = 0$ transverse-field Ising model and the $(d + 1)$-dimensional finite-temperature Ising model.

A quantum Hamiltonian version of the 3-state Potts model has been examined by high-order perturbation expansions in two dimensions by Hamer *et al.* [74, 75]; and, later, in three dimensions (on various lattices) [76]. In fact, this model was subjected to both high-temperature type and low-temperature type expansions. Quantities studied include the ground state energy, energy gap, magnetization, and susceptibility. In three dimensions, these models undergo first-order order–disorder transitions, so it was essential to have both types of series in order to obtain physically useful results.

Earlier work, not using cluster expansion methods, included calculations of properties including gaps for $d = 1$ quantum versions of the ANNNI (axial next-nearest neighbour Ising) model [23] and chiral clock model [22].

9.9. *Models from field theory*

Here we will rather quickly describe a variety of strong-coupling calculations from the field-theory literature; let us emphasize again that the following compilation is by no means complete, but may serve as a useful entry-point to the literature.

One class of models which have been extensively studied by means of high-order convergent expansions is the quantum rotor models. The unperturbed Hamiltonian associated with the $O(n)$ rotor model is of the form

$$\sum_i \mathbf{L}_i^2, \tag{53}$$

where $\mathbf{L}$ is the n-dimensional angular momentum operator. In the unperturbed ground state, each site has angular momentum zero. Coupling between the angular momenta takes place via a Heisenberg-type coupling, $\mathbf{n}_i \cdot \mathbf{n}_j$, where the $\mathbf{n}$ are angular orientation operators.

Expansions for $n = 2$, 3, and 4 quantum rotor models were carried out for one-dimensional systems by Hamer *et al.* [12]. A higher order expansion for the energy gap for the $n = 2$ model was carried out by Hornby and Barber [24], while Hamer and Guttman [55] considered $n = 2$ and 3 in both one and two dimensions.

Other models which have been studied by strong-coupling expansions include a range of pure gauge theories. Let us just note the gauge groups: Z_N [77, 78], $U(1)$ [77, 79], $SU(2)$ [79], $SU(3)$ [80]. At least one gauge theory including fermions, the massive Schwinger model [81], has also been the subject of strong-coupling expansions.

10. An illustrative example: the square lattice Heisenberg antiferromagnet

To illustrate the power of the perturbation expansions discussed here and to compare them with other numerical methods, we discuss the case of spin-half square-lattice Heisenberg antiferromagnets. Consider first the nearest-neighbour model defined by the Hamiltonian

$$\mathcal{H} = \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j. \tag{54}$$

Here $\mathbf{S}_i$ represents a spin-half operator at site i and the sum runs over all nearest neighbour pairs on the lattice. This model has been studied extensively by numerical methods over the past decade or so, especially because of its relevance to the cuprate high-T_c materials. The use of Ising expansions for this model go back to the work of Davis in 1960 [5]. However, it became a controlled numerical method only after Huse pointed out how the singular behaviour could be accommodated in the series extrapolations [82]. Furthermore, the development of cluster expansion methods has made it possible to calculate long series by fully automated computer programs.

Ising expansions have been developed for this model for a number of different quantities including:

(i) basic ground state properties, such as ground state energy and sublattice magnetization [83, 84];
(ii) static response functions, such as uniform susceptibility and spin-stiffness constant [36, 83–85];
(iii) equal time spin–spin correlation functions and the elementary excitation spectrum in the single mode approximation [86];
(iv) complete spin-wave excitation spectrum [87];
(v) wave vector dependent spectral weights associated with the spin-wave spectrum [87];
(vi) frequency moments for the Raman response functions [88].

Although many of these quantities were first obtained by spin-wave theory, its reliability was not *a priori* obvious for $S = 1/2$ antiferromagnets in two dimensions. To our knowledge, all these quantities, except for ground-state energy and sublattice magnetization, were first calculated in a controlled numerical way by the Ising expansion. As discussed first by Huse, the gapless spin waves lead to singular corrections to the thermodynamic quantities as the Heisenberg model is approached from the Ising model. However, the power laws associated with the singularities are entirely determined by non-interacting spin-wave theory. To be more explicit, let our expansion parameter be $\lambda = J_{xy}/J_I$, with the Heisenberg model corresponding to $\lambda = 1$. Then quantities like the sublattice magnetization will have singularities of the form,

Table 1. Numerical results for the $S = 1/2$ square-lattice Heisenberg model. E_0 is the ground state energy, M the sublattice magnetization, $\chi_\perp$ the uniform perpendicular susceptibility, ρ_s the spin-stiffness constant and c the spin-wave velocity. QMC stands for quantum Monte Carlo, GFMC for Green's function Monte Carlo and SSE for stochastic series expansions.

Quantity	Ising expansions	Cluster QMC	GFMC	SSE
E_0	– 0.6693(1)		– 0.66934(3)	– 0.669437(5)
M	0.307(1)	0.3084(2)	0.3075(35)	0.3070(3)
$\chi_\perp$	0.0659(10)		0.0669(7)	0.0625(9)
ρ_s	0.182(5)	0.185(2)		0.175(2)
c	1.655(12)	1.68(1)	1.55(4)	1.673(7)

$$M = M_0 + A(1 - \lambda^2)^{1/2}.$$

This implies that the Nth partial sum of the series will converge as a power of $1/N$. For the sublattice magnetization the partial sums behave as [83]

$$S_N = M_0 + 1/N^{1/2}.$$

Thus the magnetization for the Heisenberg model M_0 is readily determined from a plot of S_N as a function of $N^{-1/2}$. Alternatively, properties such as M_0 can also be determined by changing to a new variable such that the leading singularity is removed, and carrying out Padé extrapolations [85].

The numerical values generated by the analysis of Ising expansions have since been confirmed by other methods, and there are no serious discrepancies. Recent quantum Monte Carlo simulations using cluster algorithms [89], the stochastic series expansion (SSE) method [90], and Green's function Monte Carlo methods [91, 92] have produced results of comparable accuracy to the best series expansion results [36, 84]. Table 1 gives a listing of some of the basic quantities calculated by different numerical methods along with the uncertainties quoted by these respective authors. The small differences imply that there are systematic errors in various methods but it is not possible to say which results are the most accurate.

Perhaps the clearest advantage of the series expansion method is in dealing with systems with frustration, which lead to minus signs in quantum Monte Carlo simulations. The other advantage may be that one can get reasonably accurate results with modest computer resources (or perhaps even without using a computer—note that 4th order Ising expansions for the ground state energy and magnetization can be done by hand and when properly analysed give an accuracy of a few percent). This would clearly not be possible for the Monte Carlo methods, although this is perhaps not a serious consideration with the availability of significant computing power on one's desktop in recent years.

Let us now consider the square-lattice Heisenberg model with nearest and second neighbour antiferromagnetic interactions, also known as the J_1–J_2 model. This model has the usual antiferromagnetic order at small J_2. For large J_2, a collinear 4-sublattice order, in which neighbouring spins align ferromagnetically along one of the axes of the square-lattice and antiferromagnetically along the other axis is known to be promoted by quantum fluctuations—an order-by-disorder phenomenon [93]. In our view (though others may disagree), by far the most reliable numerical results

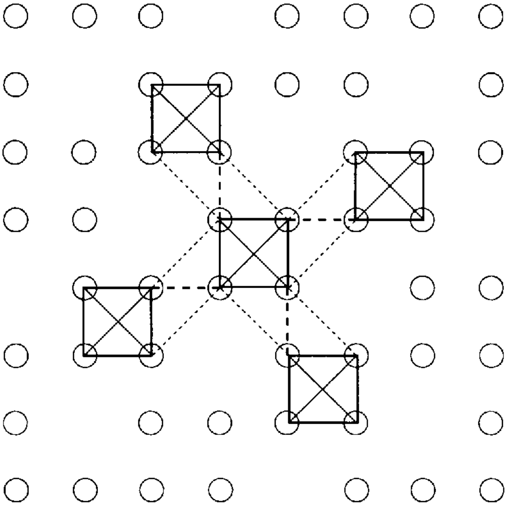


Figure 8. The one-fifth depleted square lattice relevant to the material CaV_4O_9.

for intermediate ratios of the couplings have come from series expansions. Gelfand *et al.* [56] and later Oitmaa and Zheng [34], using Ising and dimer expansions, have shown that there is an intermediate phase where there is a finite energy gap and no magnetic order. This phase exists approximately in the parameter range $0.4 < J_2/J_1 < 0.6$. The nature of this phase is still not fully understood, although it most likely has broken translational symmetry [94]. For these frustrated models, the only other numerical method which has been used extensively is exact diagonalization [95].

Another two dimensional Heisenberg model which has received considerable attention recently is the $\frac{1}{5}$th depleted square-lattice relevant to the material CaV_4O_9 [18]. The lattice is shown in figure 8. From the geometry, one can identify two distinct nearest neighbour couplings and two distinct second neighbour couplings. An appropriate Heisenberg Hamiltonian can be written as:

$$\begin{aligned}\mathcal{H} = {} & J_1 \sum_{(i,j)} \mathbf{S}_i \cdot \mathbf{S}_j + J_1' \sum_{(i,k)} \mathbf{S}_i \cdot \mathbf{S}_k \\ & + J_2 \sum_{(i,l)} \mathbf{S}_i \cdot \mathbf{S}_l + J_2' \sum_{(i,m)} \mathbf{S}_i \cdot \mathbf{S}_m, \end{aligned} \qquad (55)$$

where the sums run over nearest-neighbour bonds within plaquettes (J_1), nearest-neighbour bonds between plaquettes (J_1'), second-neighbour bonds within plaquettes (J_2), and second-neighbour bonds between plaquettes (J_2'). Observe that this Hamiltonian has the property that every spin is equivalent, just as every vanadium atom in CaV_4O_9 is equivalent.

For the nearest neighbour model, the most extensive calculation has been done by quantum Monte Carlo simulations [96]. Results for the phase diagram and energy gap obtained from series expansions [45, 66] are in close agreement with those of the Monte Carlo simulations. However, for frustrated models, series expansions remains the method of choice, although the density matrix renormalization group method has also been used successfully for this model [97].

11. Open problems

In this section we will discuss three methodological open problems for cluster expansion calculations: how to treat problems with quenched disorder, how to treat 'multi-particle' excited states, and how to carry out expansions when the ground state of the unperturbed Hamiltonian is highly degenerate.

11.1. *Quenched disorder*

It is clear from the multivariable expansions discussed in section 4 that the cluster expansion method works even when each term in $\mathcal{H}_1$ has a different coefficient. Thus it is possible to carry out an expansion for any realization of a disordered system, when the disorder is associated with the perturbation. If the disorder is discrete, such as bimodal, then the finite number of bond configurations of a given cluster (that is, the set of possible cluster Hamiltonians) can all be enumerated. Thus one can carry out disorder-averaged calculations for static ground state properties, by giving each possible cluster Hamiltonian, say all 2^m of them for bimodal disorder and a cluster with m terms of $\mathcal{H}_1$, the appropriate weight. Note that these calculations require significantly more effort than the analogous calculations for uniform systems, since each graph of the uniform system corresponds to many graphs of the disordered system. Calculations of this type have been carried out for the random transverse Ising chain [98]. But further difficulty arises in the analysis of such series. Consider, for example, what is known about the random transverse Ising chain [99] by nearly rigorous real-space renormalization group arguments. They have very strong Griffith's singularities; their properties have very broad distributions; and they may be gapless even when their correlation lengths are finite. It is not obvious how series analysis methods can accommodate these physical features.

(It is also not obvious how to develop cluster expansions when there is continuous disorder in $\mathcal{H}_0$. This would be required for expansions around gapless localized limits, such as a Bose glass [100]. We suspect that the cluster expansion formalism, as developed here, is just not applicable in this case.)

Looking at the excitation spectra poses further problems. Disorder-averaging of the effective Hamiltonian discussed in section 6 is tantamount to an effective medium approximation and does not reflect the true excitation spectra. One way around some of these difficulties is to consider a finite but large system with a given realization of disorder, and calculate the series expansions for all states of the system and construct the density of states and other relevant quantities. Such calculations have not been attempted so far.

11.2. *Multi-particle excited states*

Questions associated with two- (or more) particle excitations pose interesting challenges for cluster expansion methods. A useful example to consider is the Heisenberg ferromagnet [101]. In that case, the single particle excitations can be obtained exactly. For the two-particle excitations, one needs to solve an integral equation to find the spectrum and lifetimes. For the general case of interest to us, with quantum fluctuations in the ground state, it seems that we should in principle be able to write down the integral equation for the two-particle excitations, order by order in perturbation theory. In a recent paper Damle and Sachdev [102] have calculated the S matrix for the spin-ladder model in the strong coupling expansion to second order. It would be useful to develop the cluster 'machinery' so that a wide variety of such calculations could be done by (largely) automated computer programs.

However, even in the case of multiparticle bound states we have not been able to overcome some fundamental conceptual difficulties which arise for certain models. For the Ising model on the square lattice (or any bipartite lattice in $d > 1$) there are two-particle bound states consisting of adjacent flipped spins in an ordered background. When transverse exchange is included perturbatively one anticipates that these excitations develop dispersion but retain their character as discrete excitations. For the ferromagnet, there appears to be no difficulty encountered in constructing a perturbation expansion for these excitations, because by conservation of total S^z the bound state cannot move from one cluster to another disconnected cluster under the action of the perturbation. However, for the antiferromagnet, the action of each term of the transverse exchange on the Néel state is to create a two-particle bound state. Consequently, the formalism for cluster expansions appears to be inapplicable in this case, since the excitations can jump from one component to another of a disconnected graph. What is not clear to us is whether this implies that a series expansion in integer powers of the transverse exchange is not possible, or that the cluster expansion machinery cannot be used, or that we just need to think more clearly to derive a cluster expansion method for this problem.

11.3. *Massive ground-state degeneracy in the unperturbed Hamiltonian*

There are many natural candidates for perturbation theory where the ground state of the unperturbed Hamiltonian is highly degenerate. Examples of such problems are the large-U (or, better, small-t) limit of the Hubbard model, or the strong 'rung' coupling limit of the 3-chain Heisenberg model. The standard way of treating such problems is degenerate perturbation theory, which produces an effective Hamiltonian within the unperturbed ground-state manifold that is then used for further calculations. Thus the Hubbard model reduces to the Heisenberg or t–J model; and the 3-chain ladder reduces to the single Heisenberg chain problem. This obviously works well for very strong coupling. Extending this scheme beyond leading order is quite cumbersome [103], and has rarely been used to generate numerical results. One way around this difficulty is to develop strong coupling expansions at finite temperatures. The methodology for carrying out such high-order expansions on the computer has recently been developed [104] and should prove useful for the odd-chain Heisenberg ladder.

12. Conclusions

In this review we have presented details of the $T = 0$ series expansion method for quantum statistical models on a lattice. The method consists of a straightforward high-order perturbation theory in the coupling constants, followed by use of series extrapolation methods borrowed from the study of classical critical phenomena. This article focuses entirely on calculations of series coefficients. We have tried to present sufficiently detailed descriptions of the algorithms so that it would possible for someone new to the field to set up working computer programs within a reasonable period of time.

The basic structure of the method discussed here is very simple and applies to many different problems. The perturbative problem on the infinite lattice is broken up into many small finite cluster problems. Generating exact series coefficients for ground and excited state properties of the thermodynamic systems is reduced to an enumeration of clusters on the lattice and perturbation theory on those finite clusters. We discuss how these clusters can be generated, and the perturbation theory can be carried out, by means of fully automated computer programs.

The class of problems where these methods apply is limited in some sense, because one needs a simple starting Hamiltonian to expand around. On the other hand, the list of problems where these methods can provide a quick, and in many cases the quantitatively most accurate, answers is growing everyday. For example, many new materials have been recently discovered which exhibit spin-gap phenomena. That is, thermodynamic quantities such as specific heat and susceptibility exhibit activated behaviour at low temperatures. Many of these materials can be thought of as weakly coupled spin-clusters, where the spins within each cluster interact much more strongly than any pair of spins in different clusters. The series expansion method is ideal for studying such systems. So far, calculations have been mostly limited to clusters of two and four $S = 1/2$ spins; but, with the growth in computer power, much larger clusters or higher spin clusters are also feasible. As the cluster sizes grow, even uniform systems are likely to show improved convergence in low orders.

Applications of these methods beyond the problems involving only spin degrees of freedom have been limited so far. There have been calculations of Hubbard, Kondo lattice, Bose–Hubbard etc. models, but all around the insulating phases. Extension of these methods to metallic phases and to quenched random systems are some of the most important challenges remaining in the formal development of series expansion methods. We hope our review attracts more attention to these problems, and encourages more use of these most traditional calculational tools.

References

[1] Baker Jr, G. A., Rushbrooke, G. S, and Wood, P. W., 1974, *Phase Transitions and Critical Phenomena*, Vol. 3, edited by C. Domb and M. S. Green (London: Academic Press).

[2] Domb, C., 1974, *Phase Transitions and Critical Phenomena*, Vol. 3, edited by C. Domb and M. S. Green (London: Academic Press).

[3] Elstner, N., 1997, *Int. J. mod. Phys.* B, **11**, 1753.

[4] Messiah, A., 1961, *Quantum Mechanics*, Vol. II (Amsterdam: North-Holland), pp. 712–720.

[5] Davis, H. L., 1960, *Phys. Rev.*, **129**, 789.

[6] Boon, M. H., 1961, *Nouvo Cim.*, **21**, 885.

[7] Arai, T, and Goodman, B., 1967, *Phys. Rev.*, **155**, 514.

[8] Parrinello, M., Scirè, M, and Arai, T., 1973, *Lett. Nouvo Cim.*, **6**, 138.

[9] Parrinello, M, and Arai, T., 1974, *Phys. Rev.* B, **10**, 265.

[10] Pfeuty, P, and Elliott, R. J., 1971, *J. Phys.* C, **4**, 2370.

[11] Harris, A. B., 1973, *Phys. Rev.* B, **7**, 3166.

[12] Hamer, C. J., Kogut, J. B, and Susskind, L., 1979, *Phys. Rev.* D, **19**, 3091.

[13] Hamer, C. J, and Irving, A. C., 1984, *Nucl. Phys.* B, **230**, 336.

[14] Prelovšek, P., Sega, I, and Bonvca, J., 1990, *Phys. Rev.* B, **42**, 10706.

[15] Freericks, J. K, and Monien, H., 1996, *Phys. Rev.* B, **53**, 2691.

[16] Katoh, N, and Imada, M., 1995, *J. Phys. Soc. Jpn.*, **64**, 4105.

[17] Sano, K, and Takano, K., 1996, *J. Phys. Soc. Jpn.*, **65**, 46.

[18] Ueda, K., Kontani, H., Sigrist, M, and Lee, P. A., 1996, *Phys. Rev. Lett.*, **76**, 1932.

[19] Kadanoff, L. P, and Kohmoto, M., 1981, *J. Phys.* A, **14,** 1291.
[20] Nickel, B. G., 1980, unpublished.
[21] Marland, L. G., 1981, *J. Phys.* A, **14,** 2047.
[22] Howes, S., Kadanoff, L. P, and Den Nijs, M., 1983, *Nucl. Phys.* B, **215,** 169.
[23] Barber, M. N, and Duxbury, P. M., 1982, *J. stat. Phys.*, **29,** 427.
[24] Hornby, P. G, and Barber, M. N., 1985, *J. Phys.* A, **18,** 827.
[25] He, H.-X., Hamer, C. J, and Oitmaa, J., 1990, *J. Phys.* A, **23,** 1775.
[26] Hamer, C. J, and Irving, A. C., 1984, *J. Phys.* A, **17,** 1649.
[27] Singh, R. R. P., Gelfand, M. P, and Huse, D. A., 1988, *Phys. Rev. Lett.*, **61,** 2484.
[28] Gelfand, M. P., 1996, *Solid State Commun.*, **98,** 11.
[29] Mckenzie, S., 1980, *Phase Transitions, Cargese*, edited by M. Levy, J. C. LeGuillou and J. Zinn-Justin (New York: Plenum).
[30] Gelfand, M. P., Singh, R. R. P, and Huse, D. A., 1990, *J. Stat. Phys.*, **59,** 1093.
[31] Oitmaa, J, and Zheng, W., 1992, *J. Phys.* A, **25,** 5849.
[32] Singh, R. R. P., 1996, *Phys. Rev.* B, **53,** 11582.
[33] Martin, J. L., 1974, *Phase Transitions and Critical Phenomena*, Vol. 3, edited by C. Domb and M. S. Green (London: Academic Press).
[34] Oitmaa, J, and Zheng, W., 1996, *Phys. Rev.* B, **54,** 3022.
[35] Singh, R. R. P., 1990, *Phys. Rev.* B, **41,** 4873.
[36] Hamer, C. J., Zheng, W., and Oitmaa, J., 1994, *Phys. Rev.* B, **50,** 6877.
[37] Oitmaa, J., Hamer, C. J, and Zheng, W., 1992, *Phys. Rev.* B, **45,** 9834.
[38] Hamer, C. J., Oitmaa, J, and Zheng, W., 1991, *Phys. Rev.* B, **43,** 10789.
[39] Oitmaa, J., Hamer, C. J, and Zheng, W., 1994, *Phys. Rev.* B, **50,** 3877.
[40] Singh, R. R. P., Fisher, M. E, and Shankar, R., 1989, *Phys. Rev.* B, **39,** 2562.
[41] Affleck, I., Gelfand, M. P, and Singh, R. R. P., 1994, *J. Phys.* A, **27,** 7313; 1995, *ibid.*, **28,** 1787 (erratum).
[42] Singh, R. R. P, and Gelfand, M. P., 1988, *Phys. Rev. Lett.*, **61,** 2133.
[43] Kitazawa, A., Nomura, K, and Okamoto, K., 1996, *Phys. Rev. Lett.*, **76,** 4038.
[44] Oitmaa, J., Singh, R. R. P, and Zheng, W., 1996, *Phys. Rev.* B, **54,** 1009.
[45] Zheng, W., Gelfand, M. P., Singh, R. R. P., Oitmaa, J, and Hamer, C. J., 1997, *Phys. Rev.* B, **55,** 11377.
[46] Zheng, W., Singh, R. R. P, and Oitmaa, J., 1997, *Phys. Rev.* B, **55,** 8052.
[47] Zheng, W., 1997, *Phys. Rev.* B, **55,** 12267.
[48] Gelfand, M. P., Zheng, W., Hamer, C. J, and Oitmaa, J., 1998, *Phys. Rev.* B, **57,** 392.
[49] Gelfand, M. P, and Glöggler, E. F., 1997, *Phys. Rev.* B, **55,** 11372.
[50] Oitmaa, J., Hamer, C. J, and Zheng, W., 1991, *J. Phys.* A, **24,** 2863.
[51] Zheng, W., Oitmaa, J, and Hamer, C. J., 1994, *J. Phys.* A, **27,** 5425.
[52] Singh, R. R. P, and Huse, D. A., 1992, *Phys. Rev. Lett.*, **68,** 1766.
[53] Bernu, B., Lecheminant, P., Lhuillier, C, and Pierre, L., 1994, *Phys. Rev.* B, **50,** 10048.
[54] Bernu, B., Lecheminant, P., Lhuillier, C, and Pierre, L., 1995, *Phys. Rev.* B, **52,** 9162.
[55] Hamer, C. J, and Guttmann, A. J., 1989, *J. Phys.* A, **22,** 3653.
[56] Gelfand, M. P., Singh, R. R. P, and Huse, D. A., 1989, *Phys. Rev.* B, **40,** 10801.
[57] Gelfand, M. P., 1990, *Phys. Rev.* B, **42,** 8206.
[58] Kohmoto, M., den Nijs, M, and Kadanoff, L. P., 1981 *Phys. Rev.* B, **24,** 5229.
[59] Yamanaka, M., Hatsugai, Y, and Kohmoto, M., 1994, *Phys. Rev.* B, **50,** 559.
[60] Yamanaka, M., Hatsugai, Y, and Kohmoto, M., 1993, *Phys. Rev.* B, **48,** 9555.
[61] Singh, R. R. P, and Gelfand, M. P., 1989, *Phys. Rev. Lett.*, **62,** 840.
[62] Hida, K., 1992, *J. Phys. Soc. Jpn.*, **61,** 1013.
[63] Gelfand, M. P., 1996, *Phys. Rev.* B, **53,** 11309.
[64] Matsushita, Y., Gelfand, M. P, and Ishii, C., 1997, *J. Phys. Soc. Jpn.*, **66,** 3648.
[65] Barnes, T., Riera, J, and Tennant, D. A., preprint (cond-mat/9801224).
[66] Gelfand, M. P., Zheng, W. H., Singh, R. R. P., Oitmaa, J, and Hamer, C. J., 1996, *Phys. Rev. Lett.*, **77,** 2794.
[67] Meshkov, S. V, and Foerster, D., 1996, *J. Phys.: Condens. Matter*, **8,** 7917.
[68] Pickett, W. E., 1997, *Phys. Rev. Lett.*, **79,** 1746.

[69] Kodama, K., Harashina, H., Sasaki, H., Kobayashi, Y., Kasai, M., Taniguchi, S., Yasui, Y., Sato, M., Kakurai, K., Mori, T, and Nishi, M., 1997, *J. Phys. Soc. Jpn.*, **66,** 793.
[70] Fukumoto, Y, and Oguchi, A., 1998, *J. Phys. Soc. Jpn.*, **67,** 2205.
[71] Shi, Z.-P., Singh, R. R. P., Gelfand, M. P, and Wang, Z., 1995, *Phys. Rev.* B, **51,** 15630.
[72] Shi, Z. P, and Singh, R. R. P., 1995, *Europhys. Lett.*, **31,** 219.
[73] Shi, Z. P, and Singh, R. R. P., 1995, *Phys. Rev.* B, **52,** 9620.
[74] Hamer, C. J., Aydin, M., Oitmaa, J, and He, H. X., 1990, *J. Phys.* A, **23,** 4025.
[75] Hamer, C. J., Oitmaa, J, and Zheng, W., 1992, *J. Phys.* A, **25,** 1821.
[76] Zheng, W., Hamer, C. J, and Oitmaa, J., 1994, *J. Phys.* A, **27,** 5789.
[77] Irving, A. C, and Hamer, C. J., 1984, *Nucl. Phys.* B, **230,** 361.
[78] Hamer, C. J., Oitmaa, J, and Zheng, W., 1994, *Phys. Rev.* D, **49,** 535.
[79] Hamer, C. J., Oitmaa, J, and Zheng, W., 1992, *Phys. Rev.* D, **45,** 4652.
[80] Hamer, C. J., 1989, *Phys. Lett.* B, **224,** 339.
[81] Hamer, C. J., Zheng, W. H, and Oitmaa, J., 1997, *Phys. Rev.* D, **56,** 55.
[82] Huse, D. A., 1988, *Phys. Rev.* B, **37,** 2380.
[83] Singh, R. R. P., 1989, *Phys. Rev.* B, **39,** 9760.
[84] Zheng, W., Oitmaa, J, and Hamer, C. J., 1991, *Phys. Rev.* B, **43,** 8321.
[85] Singh, R. R. P, and Huse, D. A., 1989, *Phys. Rev.* B, **40,** 7247.
[86] Singh, R. R. P., 1993, *Phys. Rev.* B, **47,** 12337.
[87] Singh, R. R. P, and Gelfand, M. P., 1995, *Phys. Rev.* B, **52,** R15695.
[88] Singh, R. R. P., Fleury, P. A., Lyons, K. B, and Sulewski, P. E., 1989, *Phys. Rev. Lett.*, **62,** 2736.
[89] Beard, B. B, and Wiese, U.-J., 1996, *Phys. Rev. Lett.*, **77,** 5130.
[90] Sandvik, A. W., 1997, *Phys. Rev.* B, **56,** 11678.
[91] Runge, K., 1992, *Phys. Rev.* B, **45,** 7229.
[92] Runge, K., 1992, *Phys. Rev.* B, **45,** 12292.
[93] Chandra, P., Coleman, P, and Larkin, A. I., 1990, *Phys. Rev. Lett.*, **64,** 88.
[94] Zhitomirsky, M. E, and Ueda, K., 1996, *Phys. Rev.* B, **54,** 9007.
[95] Dagotto, E., 1994, *Rev. Mod. Phys.*, **66,** 763.
[96] Troyer, M., Imada, M, and Ueda, K., 1997, *J. Phys. Soc. Jpn.*, **66,** 2957.
[97] White, S. R., 1996, *Phys. Rev. Lett.*, **77,** 3633.
[98] Singh, R. R. P., 1995, unpublished.
[99] Fisher, D. S., 1995, *Phys. Rev.* B, **51,** 6411.
[100] Fisher, M. P. A., Weichman, P. B., Grinstein, G, and Fisher, D. S., 1989, *Phys. Rev.* B, **40,** 546.
[101] Mattis, D. C., 1981, *The Theory of Magnetism I* (Berlin: Springer-Verlag).
[102] Damle, K, and Sachdev, S., 1998, *Phys. Rev.* B, **57,** 8307.
[103] Macdonald, A. H., Girvin, S. M, and Yoshioka, D., 1988, *Phys. Rev.* B, **37,** 9753.
[104] Elstner, N, and Singh, R. R. P., 1998, *Phys. Rev.* B, **57,** 7740.