The one-dimensional Hubbard model with first-, second- and third-nearest-neighbour hopping in the strong-coupling limit

1994 J. Phys.: Condens. Matter 6 481

(http://iopscience.iop.org/0953-8984/6/2/019)

J. Phys.: Condens. Matter 6 (1994) 481–493. Printed in the UK

The one-dimensional Hubbard model with first-, second- and third-nearest-neighbour hopping in the strong-coupling limit

M W Long, C W M Castleton and C A Hayward
School of Physics, Birmingham University, Edgbaston, Birmingham B15 2TT, UK

Received 21 June 1993, in final form 8 November 1993

Abstract. Using an extension of the Jordan–Wigner transformation, we solve the one-dimensional Hubbard model at $U = \infty$ in the limit of dominant nearest-neighbour hopping combined with infinitesimal hopping over slightly longer ranges. We find several possible phases at zero temperature, including ferromagnetism, paramagnetism (with and without a spin gap), and even long-range antiferromagnetism with a particular limiting procedure. Our solutions are always spin–charge separated, and we give evidence that the charge degrees of freedom are best described by bosonic statistics.

1. Introduction

The Hubbard model is probably the simplest description that can be used to discuss metals with 'Mott insulating' atoms. In the most extreme case, where one type of charge fluctuation is banished from the problem altogether, the model reduces to a *single-parameter* model; the t model:

$$H = -t \sum_{\langle ii' \rangle \sigma} (1 - c^{\dagger}_{i\bar{\sigma}} c_{i\bar{\sigma}}) c^{\dagger}_{i\sigma} c_{i'\sigma} (1 - c^{\dagger}_{i'\bar{\sigma}} c_{i'\bar{\sigma}}) \tag{1.1}$$

where $c^{\dagger}_{i\sigma}$ ($c_{i\sigma}$) creates (annihilates) an electron of spin σ (complementary spin $\bar{\sigma}$) on an atom i. The model hops electrons between nearest-neighbour atoms, denoted by $\langle ii' \rangle$, and the factors $(1 - c^{\dagger}_{i\bar{\sigma}} c_{i\bar{\sigma}})$ ensure that sites can never become doubly occupied, hence enforcing the constraint that one charge state for each atom is eliminated: that with two electrons on the atom.

Applied to the square lattice, this model may well provide an explanation for high-temperature superconductivity [1], although it is far from clear whether or not an essential ingredient has been omitted [2]. Although this model has been heavily studied over the past few years, the results remain in the realm of folklore and conjecture [3], with 'hard' facts restricted to unphysical limits and one dimension [4].

The $U = \infty$ limit of the square lattice Hubbard model has been a source of interest, and various ideas have emerged. In contradiction to the predictions of standard mean-field treatments, the low-density limit is paramagnetic [5]. The ground state involves highly correlated electrons, and the physics has not been resolved. Close to one electron per site lies a possibility of a ferromagnetic phase, although the evidence for this is really only Nagaoka's theorem [6], which is a solution of the model in a non-thermodynamic limit. We will interpret our solutions in these terms, and we find rigorous examples of these phases in one dimension.

0953-8984/94/020481+13$07.50

The behaviour of electrons under the action of the Hubbard model may be very different from the usual behaviour in normal metals. Normally, electrons 'run around' and effectively 'ignore' each other. Although the electrons are repelled from each other the consequences are usually minor. All that usually occurs is that the probability of the electrons coming together is slightly reduced, but the particles are still free to go everywhere. For the $U = \infty$ Hubbard model, however, the electrons are not allowed to doubly occupy atoms, and so there is a new restriction on the places where the electrons may go. Two possibilities are plausible. Firstly, there may be sufficient 'gaps' between electrons for the particles to move around effectively 'ignoring' each other. Secondly, the motion of the particles may become highly correlated, with the motion becoming *collective* and particles moving by 'pushing' each other out of the way. For one-dimensional systems there is no way around other particles, and so we are forced into the collective motion picture. This has led to one of the most exciting facts to have emerged in the analysis, that the one-dimensional solution is *spin–charge separated*, which means that the quasi-particles have quantum numbers distinct from those found on the constituent electrons. The real issue is then whether or not these results extend to the experimentally relevant two dimensional case.

It has been argued quite forcefully that one dimension is special. The fact that particles cannot orbit each other is thought to lead to both the exact solutions and the spin–charge separation, and that the two-dimensional system will show neither characteristic. We do not believe this, and believe on the contrary that the two-dimensional systems are *more* susceptible to the many-particle interactions that cause the problems. In this article we will solve some one-dimensional models that involve particles orbiting each other. Our solutions can be interpreted in similar terms to those that have been used to discuss the two-dimensional square model, and the role of the many-body interactions is seen to be larger in the two-dimensional system.

Although in two dimensions there are gaps through which electrons can pass in order to avoid other electrons, the physics is not that simple: electrons are fermions. One can view the quantum mechanical problem of electronic motion as that of a superposition of all possible ways for particular configurations of particles to be encountered in the motion, combined with a statistical phase in the superposition. If particles are exchanged in motion along distinct routes, then the fermionic phase difference between these different paths can phase cancel, leading to reduced motion. In practice, this is a very strong force reducing the fermionic motion. With a careful choice of spin wavefunction, it is possible for this statistical phase cancellation to be reduced, and for the charge motion to become *controlled* by the spin wavefunction. The picture is that the electrons move predominantly when the local spin configuration favours phase coherence and do not move when phase cancellation is probable. We will see that this idea is central to an interpretation of our solutions, and that indeed the spin wavefunction almost totally controls the charge motion in our paramagnetic phases. Recently one of the present authors has investigated this idea for the two-dimensional Hubbard model in a separate collaboration, and the evidence in favour of the uniform phase motion is excellent [7].

Unfortunately, it has not proved possible to solve the Hubbard model in anything other than an extreme limit, although this limit is quite a physically attractive one. We take the $U = \infty$ one-dimensional Hubbard model with nearest-neighbour hopping and solve the model with infinitesimal hopping over short but finite distances. In practice, this constitutes using degenerate perturbation theory to lift the spin degeneracy found in the $U = \infty$ limit. The one-dimensional Hubbard model has been solved by the Bethe *ansatz* [4], and the role of weak Heisenberg interactions between the spins in the separated spin system is well

understood. The critical difference between these one-dimensional systems and the two-dimensional square lattice of physical interest is the existence of square loops around which particles can permute. These loops lead to very strong interactions between the spins, which could easily dominate the weak Heisenberg interactions for sizeable doping. In our analysis, it is precisely this type of square-loop permutation we are modelling, yielding insight into this fundamental difference between the one- and two-dimensional models.

The reason that we have been able to solve the present limit is that we have devised a representation for the state space that naturally describes spin–charge-separated solutions, and since *all* of our solutions fall into this class they are tractable. Our representation is a horribly non-linear transformation away from the original spin-full fermion starting point, and this transformation constitutes the underlying mathematics. In concept the transformation is akin to the Jordan–Wigner transformation [8], which solves the one-dimensional $x-y$ model, and we have generalized the transformation to take account of the spin degree of freedom present on the electrons.

In section 2 we present the transformation, in section 3 we apply the transformation to our Hubbard model, solving it in our chosen limit in section 4. In section 5 we interpret the solution, and in section 6 we conclude.

2. The representation and the transformation

The basic reason for spin–charge separation is that the charges move around on a time-scale for which the spin degrees of freedom are frozen. To leading order the charges 'concertina' against each other, and to move a charge about requires shuffling *all* the charges between the two relevant end points, exchanging them in such a way that the spin configuration along the chain is conserved. In order to describe this sort of state it is useful to rerepresent the states by

$$c^\dagger_{i_1\sigma_1}c^\dagger_{i_2\sigma_2}\ldots c^\dagger_{i_n\sigma_n}\,|\,0\rangle \equiv f^\dagger_{i_1}f^\dagger_{i_2}\ldots f^\dagger_{i_n}\,|\,\sigma_1\sigma_2\ldots\sigma_n\rangle \tag{2.1}$$

where the $f^\dagger_i$ create particles with fermionic statistics on site i and the spins σ_α are presumed to be ordered along the chain. It is crucial to realize that we must deal with free boundary conditions in our derivation, since our ordering must start somewhere. We have elected to label the first atom in the chain by 1. The key to this representation is to find the transformation that maps the original operators, the $c^\dagger_{i\sigma}$, onto the new operators, the $f^\dagger_i$ for the charges and $\hat{S}_\alpha$, which act on the spins. Although it is quite tricky to write down this transformation on all states, it has proved possible to write it down for each subset of states with fixed electron number. The operators which move charges around are 'simply'

$$c^\dagger_{i\sigma}c_{i+n\sigma} = f^\dagger_i f_{i+n}\sum_{\chi_1=0}^{1}\sum_{\chi_2=0}^{1}\cdots\sum_{\chi_{i+n}=0}^{1}\prod_{m=1}^{i+n}\left[f^\dagger_m f_m\right]^{\chi_m}\left[1-f^\dagger_m f_m\right]^{1-\chi_m}\frac{1}{2}\left[1+\sigma\hat{S}^z_{\alpha_l}\right]$$
$$\times\prod_{\alpha=\alpha_l}^{\alpha_{l+n}-1}\left[\frac{1}{2}+2\hat{S}_\alpha\cdot\hat{S}_{\alpha+1}\right] \tag{2.2}$$

where the χ_l measure whether or not an electron is on a particular site i, and in terms of which the $\alpha_l=\sum_{m=1}^{l}\chi_m$ count how many electrons come before a particular site, making a useful spin label. The first two f operators move the charge. The summations over the

χ variables break the states down into all possible charge configurations up to final point in the charge transfer. The operator involving the z component of spin ensures that the electron moved has the correct spin, and the string of spin operators shuffle the spins along conserving their order along the chain. Each term $\left[\frac{1}{2} + 2\hat{S}_\alpha \cdot \hat{S}_{\alpha+1}\right]$ provides an elementary permutation of the two spins involved, and so the product is just a simple cyclic permutation of the relevant spin variables. Due to the central role that these cyclic permutations hold in our analysis, we will introduce a notation for such a permutation, $\hat{R}_{\alpha,\alpha+n}$, where

$$\hat{R}_{\alpha,\alpha+n} = \prod_{\beta=\alpha}^{\alpha+n-1} \left[\frac{1}{2} + 2\hat{S}_\beta \cdot \hat{S}_{\beta+1}\right] \tag{2.3}$$

and we are using the letter R to represent the idea of 'ring exchange' which is how this concept has often been labelled in the literature.

We will not be concerned with applications of this formalism to finite systems, but only with the thermodynamic limit. We assume, without proof, that the final result does *not* depend on boundary conditions, and then the particular starting point for the α label is irrelevant, and only the difference between any two is relevant. Careful analysis of the formula will show that the only dependence on the χ_j, for $j < i$, is in the combination $\sum_{j=1}^{i-1} \chi_j$, which is simply the total number of electrons before the site i. In the thermodynamic limit, we would expect this number of electrons to diverge and to become irrelevant. If this is the case, then there is no direct dependence on χ_j for $j < i$, and the summations can be performed yielding unity. In the thermodynamic limit, we need only break up the charge configurations *between* the two end-points of the charge transfer.

We provide examples of the first few such operators in the next section.

3. The one-dimensional Hubbard model with infinitesimal short-range hopping

The reason that we can solve the current model is that we can map it onto a pair of models that have previously been independently solved: the charge motion is controlled by a non-interacting spinless-fermion model and the spin dynamics are mapped onto the Heisenberg model with short-range magnetic interactions. The fermion problem is elementary, although the Heisenberg model requires the Bethe *ansatz* for a complete solution [9]. Solutions for other values of short-range interaction parameters have also been given in the literature [10].

In the absence of next-nearest-neighbour hopping, the spin degeneracy remains and the charge degrees of freedom map onto a spinless fermion model:

$$H_0 = -t_0 \sum_i \left[f_i^\dagger f_{i+1} + \text{cc}\right] \tag{3.1}$$

which comes directly from equation (2.2), for our representation. This model is diagonal in reciprocal space, and since it is a non-interacting theory all the correlation functions can be found from the single-particle correlations:

$$n_n = \langle f_i^\dagger f_{i+n}\rangle = \frac{\sin(\pi n n_0)}{\pi n} \tag{3.2}$$

in terms of the average electron number n_0 per site.

Let us commence with second-nearest-neighbour hopping. Once again, equation (2.2) provides us with the interaction for our basis:

$$H_1 = -t_1 \sum_i \left[f_i^\dagger f_{i+2} \left[1 - f_{i+1}^\dagger f_{i+1} + f_{i+1}^\dagger f_{i+1} \hat{R}_{\alpha_i,\alpha_i+1} \right] + \text{cc} \right] \tag{3.3}$$

where we are using our new spin operator, $\hat{R}_{\alpha,\alpha+1}$ which cyclically permutes spin indices. Obviously, $\hat{R}_{\alpha,\alpha+1} = \frac{1}{2} + 2\hat{S}_\alpha \cdot \hat{S}_{\alpha+1}$, as has already been explained.

If this additional interaction is infinitesimal, then the spin degeneracy will be lifted. Using degenerate perturbation theory requires finding the effective spin–spin interaction when the charge degrees of freedom are in their ground state. In practice this is elementary and the interaction is simply

$$H_1^{\text{S}} = 4t_1(n_1^2 - n_0 n_2) \sum_\alpha \left[\hat{S}_\alpha . \hat{S}_{\alpha+1} - \frac{1}{4} \right] \tag{3.4}$$

i.e. the Heisenberg model with nearest-neighbour coupling. Evaluating these correlation functions demonstrates that the interaction is always antiferromagnetic, a result which will be physically explained in the next section.

Second-nearest-neighbour hopping involves the introduction of *triangles* into the geometry, and this is only natural for a geometrically frustrated system. The square lattice is an unfrustrated bipartite geometry, and the simplest infinitesimal bipartite inclusion is third-nearest-neighbour hopping. In our representation this interaction is simply

$$H_2 = -t_2 \sum_i \Big[f_i^\dagger f_{i+3} \Big[(1 - f_{i+1}^\dagger f_{i+1})(1 - f_{i+2}^\dagger f_{i+2}) + f_{i+1}^\dagger f_{i+1}(1 - f_{i+2}^\dagger f_{i+2}) \hat{R}_{\alpha_i,\alpha_i+1} + (1 - f_{i+1}^\dagger f_{i+1}) f_{i+2}^\dagger f_{i+2} \hat{R}_{\alpha_i,\alpha_i+1} + f_{i+1}^\dagger f_{i+1} f_{i+2}^\dagger f_{i+2} \hat{R}_{\alpha_i,\alpha_i+2} \Big] + \text{cc} \Big] \tag{3.5}$$

where we have extended the definition of cyclic permutations to that for three objects, $\hat{R}_{\alpha,\alpha+2}$. The effective spin interaction for this case is of the form

$$H_2^{\text{S}} = J_1 \sum_\alpha \left[\hat{S}_\alpha \cdot \hat{S}_{\alpha+1} - \frac{1}{4} \right] + J_2 \sum_\alpha \left[\hat{S}_\alpha \cdot \hat{S}_{\alpha+2} - \frac{1}{4} \right] \tag{3.6}$$

where

$$J_1 = 4t_2(n_3 n_0^2 + n_1^3 + n_1 n_2^2 - n_3 n_1^2 - 2n_0 n_1 n_2 + 2(n_1 n_2 - n_0 n_3))$$

is the nearest-neighbour effective Heisenberg exchange and

$$J_2 = -2t_2(n_3 n_0^2 + n_1^3 + n_1 n_2^2 - n_3 n_1^2 - 2n_0 n_1 n_2)$$

is the next-nearest-neighbour effective Heisenberg exchange. We now find some next-nearest-neighbour spin interactions. Unlike the previous case, there is now a competition between the various types of behaviour as the chemical potential is moved, and we find that these coefficients can change sign as depicted in figure 1.

We can extend these ideas to longer-range hopping, generating longer-range ring exchange. Unlike the current short-range hops, ring exchange of four or more particles is *not* equivalent to a pure quadratic exchange interaction, and a sizeable biquadratic interaction is induced. The physics of these more exotic Hamiltonians has not been heavily studied in the literature and remains quite mysterious to the authors.

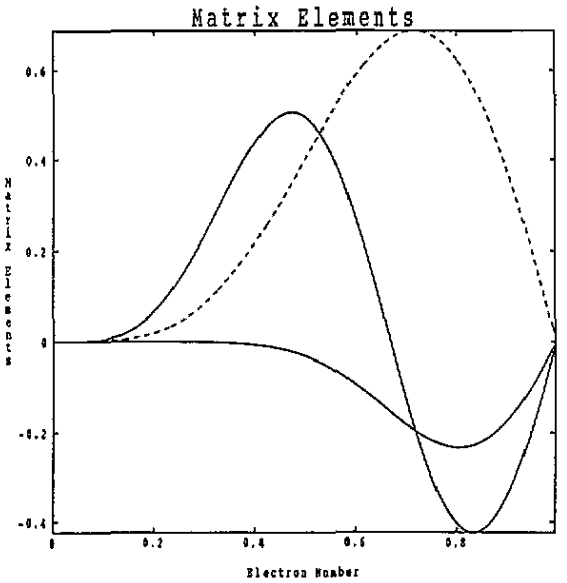

Figure 1. The Heisenberg exchange matrix elements J_1 and J_2 for nearest-neighbour and next-nearest neighbour exchange, respectively, in units of the infinitesimal inclusion t_n to the hopping. The broken curve is for next-nearest neighbour hopping, t_1, and is purely nearest-neighbour exchange, whereas the two full lines correspond to next-next-nearest neighbour hopping, t_2, yielding a weak negative J_2 and a strong but wildly varying J_1.

4. The solution

In the previous section we found the effective interactions between the spins for our Hubbard model with infinitesimal short-range hopping. A fair amount is known about the particular spin models encountered, and in this section we detail a fraction of this knowledge and then comment on the relevance to the original Hubbard model.

For next-nearest-neighbour hopping, the character of the solution depends only on the sign of the next-nearest-neighbour hopping-matrix elements. This is because of a symmetry inherent to the Hubbard model: if we reverse the relative phase of the states in the two natural sublattices, the physics remains unaltered but the hops between the two sublattices reverse their sign. The phase of the nearest-neighbour hops is therefore irrelevant, but the phase of the next-nearest-neighbour hops is physically important. If this phase is negative, then we find a purely ferromagnetic ground state with fermionic excitations. If, on the other hand, the phase is positive, then we find a paramagnetic spin state involving the spin-half Heisenberg ground state sequentially along the chain. This spin wavefunction then feeds back an interaction between the charges, attracting them to each other. The effective charge Hamiltonian is

$$H_1^{\mathrm{C}} = -t_1 \sum_i \left[f_i^\dagger f_{i+2} \left[1 - f_{i+1}^\dagger f_{i+1} 2 \ln 2 \right] + \mathrm{cc} \right] \tag{4.1}$$

where the additional interaction originates from the fact that two fermions in a singlet exchange with the opposite sign to two fermions in a triplet, and the probability of finding two neighbouring particles in a singlet is $\ln 2$ in the ground state of the Heisenberg model. This result indicates that the motion of the particles becomes much more bosonic than

fermionic, since the same model with $\ln 2 = 0.693\,15$ replaced by unity would provide the precise fermionic description for hard-core bosons moving under the action of infinitesimal next-nearest-neighbour hopping. The effective charge motion is much more like that of bosons than that of fermions, and this method of comparing the interactions in a definite representation appears to provide a useful technique for comparing statistics of excitations in one dimension.

In the introduction we introduced the idea that the spin wavefunction could control the motion of the charges. It is important to realize that this idea underlies the current solution. The spin wavefunction along the chain is the Heisenberg ground state and so is *invariant* with respect to the actual positions of the particles. Although the electrons have the freedom to hop past each other, they do not do so independently: when an electron hops past another one, the spin state in the ground state remains the same. In practice, this means that there is a reduced probability for the electron to make the hop, and in fact it only hops into a state with the *reordered* spins having the Heisenberg ground state. Alternatively, we can say that as the electrons hop past each other the spin wavefunction for the previous ordering changes in such a way as to produce the same spin wavefunction for the new ordering. This argument is *equivalent* to phase cancellation in the the superposition interpretation: if two routes lead to the same state with opposite phase, then there is a cancellation and the state is included with a smaller amplitude. We can interpret the physics either in terms of electrons not hopping or in terms of electrons hopping and cancelling; it is a matter of taste. The phases are such that electrons exchanging in triplets phase cancel, while electrons exchanging in singlets are phase coherent. Electrons therefore hop across each other mainly when they are in local singlet configurations, and do not usually hop over when parallel: This is the physical reason why the Heisenberg ground state is stabilized. At this level the argument is not very strong, because all possibilities can also be thought to take place, and the lack of motion as triplets can alternatively be understood in terms of phase cancellation between all the possible ways of arriving in the new state, a phenomenon which also occurs in the non-interacting free-electron gas. In this picture the correlations are associated with the different phase that the electrons pick up, when comparing hops past vacant and occupied atoms: the phase when vacant is the same as for a spinless fermion, but the phase when occupied is on average opposite, because there is an additional phase from the antisymmetry of the spin singlet. When we move on to next-next-nearest-neighbour exchange, we will see that there is a more direct effect of the local electronic correlations.

Unlike second-nearest-neighbour hopping, third-nearest-neighbour hopping provides a host of phase transitions and fairly violent changes of behaviour. The phase diagram can immediately be deduced from figure 1. Once again, if the relative phase of the states on the two natural sublattices is reversed, we see that both hops, namely nearest- and next-next-nearest-neighbour, reverse the signs of their matrix elements, and the physics does not therefore depend on the absolute phases of these matrix elements, but only on their relative phase. In the physically more reasonable case, when the two matrix elements have the same sign, then at densities below the critical density, $n_c = 0.667\,504\,5$, we find a quantum paramagnet that ranges from the Heisenberg ground state at low charge densities to an ordered antiferromagnet in the limit where the density approaches the critical density from below. Above the critical density we find saturated ferromagnetism, a variant of 'Nagaoka ferromagnetism' [6]. For the case where the two matrix elements are of opposite sign, we find the reversed result of ferromagnetism for low densities together with a quantum paramagnet for high densities. The effective Heisenberg model for the paramagnet is the much studied 'railroad trestle' geometry [11]. In the limit tending to the critical density from above, we find the pathological case of two disconnected rails, while in the limit tending to

one electron per site we find the exactly solvable Majumdar–Ghosh model [10], for which the system exhibits a spontaneously broken spatial symmetry. We will now discuss each of these phases in a little more detail.

Our ferromagnetic phase corresponds directly to the phase plausibly predicted by Nagaoka in his infamous paper dealing with one hole in the $U = \infty$ limit [6]. Since then this ferromagnetic phase has been heavily criticized for the two-dimensional square lattice, of which a variety of analytical and numerical calculations find little trace [12], although a few calculations seem to support its existence [13]. For our one-dimensional model we have unambiguously proved the existence of the phase, but in a subsequent article we shall show that if the *range* of hopping is allowed to diverge, a necessary requirement to try to move towards two dimensions, then the stable region of ferromagnetism shrinks down to nothing.

The quantum antiferromagnet below the critical density is the more interesting phase, because of the obvious connection with perovskite superconductivity. The spin wavefunction is the ground state of the Heisenberg model with nearest-neighbour and next-nearest-neighbour interactions, the so-called 'railroad trestle' geometry. As can be seen from figure 1, the second-nearest neighbour exchange is always ferromagnetic, and so will promote both ferromagnetic *and* antiferromagnetic Néel correlations. The first-nearest-neighbour interactions are antiferromagnetic below the critical density, but become progressively less important as we approach n_c from below. At low densities the nearest-neighbour hopping dominates and we find the Heisenberg ground state, but as we increase the electron density the Néel order in the spin wavefunction increases until at n_c we find pure Néel order, since the next-nearest-neighbour interaction requires both sublattices to be saturated and infinitesimal antiferromagnetic nearest-neighbour exchange then forces the sum of these two spins to form a total-spin singlet. Worries about lack of order in one-dimensional systems with short-range interactions are eliminated by observing that at the critical density itself there is additional *degeneracy*, associated with freely rotating the two sublattices with respect to each other, and this allows the order to survive, with the fluctuations leading to additional ground states, and not loss of order. Although the spin system becomes ordered, this does not mean that there is any experimentally observable magnetism present, as is explained in the next section: although the spins achieve order, the order is not directly tied to the fundamental atoms, but rather to the electron order which is not directly measured in most experiments.

For next-next-nearest-neighbour hopping there is now a more direct effect of the spin wavefunction on the local electronic motion. Instead of just vacant and occupied atoms, as before, there are now hops past zero, one and two electrons. The system can therefore gain its energy from different types of local correlation: ferromagnets prefer hops past pairs of electrons, so the holes will be kept apart, while paramagnets prefer hops past single electrons, so pairing of holes becomes beneficial. The electrons take account of both the local charge and spin configurations in deciding whether to move.

The more exotic paramagnet originating from the frustrated choice of signs requires some interpretation. Close to one electron per site, the spin wavefunction has a broken symmetry, but this will *not* be easily seen experimentally over distances greater than the hole–hole separation, because most probes couple to fixed atoms and if the number of holes fluctuates between any two atoms the correlations also fluctuate. There should be some consequences to the broken symmetry, but they will be quite subtle to assess. The critical concentration where the nearest-neighbour interaction vanishes no longer marks the phase transition between ferromagnetism and paramagnetism, although it is not difficult to convince oneself that the phase transition occurs to a *classical* spiral of infinitesimal pitch.

This spiral-ferromagnetism transition occurs at a second critical point, $n_c' = 0.581\,689\,9$, a little way below n_c. This result can be deduced immediately from the observation that the ferromagnetic to spiral phase transition occurs when $J_2 = -J_1/4 > 0$ for the Heisenberg model, and n_c' is the concentration at which this particular ratio occurs. We will briefly discuss the status of this ferromagnetic to spin-spiral phase transition for the Heisenberg model, which is not a rigorous result as far as we know, in the appendix. We do not believe that this frustrated limit is worthy of more serious study at present.

5. Interpretation

One of the most confusing aspects to the current solutions is the spin–charge separation. In order to help with some insight, we will devote some space to explaining some of the conceptual difficulties and some of the physical reasons that the model behaves in the way that it does.

Let us commence with the ferromagnetic phases, which one might expect to be trivial. In fact, even for this case the system is spin–charge separated and we have to consider the behaviour carefully. There are two types of excitation: charge-only excitations, which involve adding or subtracting an electron parallel to the existing ferromagnetism, and spin-only excitations, which involve a flipped spin travelling very slowly along the spin chain. This picture is very similar to that for *itinerant* ferromagnets, where we have spin-polarized band structure combined with *collective* spin waves: the charged excitations are the single-particle excitations, and the spin excitations correspond to the spin waves. For itinerant ferromagnets one can consider electronic states with electrons traveling with the 'wrong' spin, namely the higher-lying spin-polarized single-particle bands, but for our model these excitations are completely irrelevant, having been carried off to infinity as $U \to \infty$. All that remains are the collective spin waves. It is the ability to describe collective effects that is the great strength of the analysis.

The reason that the ferromagnet can be understood fairly readily in terms of single-particle and two-particle physics is because the main subtlety associated with labelling via the atoms or electrons is irrelevant. For a ferromagnet all the electrons are parallel, and so if we label with atoms or by electron number along the chain from one end, we find no spin dependence. For the quantum paramagnets this result is completely obliterated, since as we have seen, the spin wavefunction can be described only with an electron label and not with an atom label: the spin configuration along the chain remains invariant while the spin of an atom fluctuates as different electrons fluctuate on and off that atom.

Probably the most important phase is the quantum paramagnet found below the critical density with the unfrustrated choice of relative phases. This phase is the closest analogy to the phase expected to be superconducting in the perovskite superconductors. The charged excitations once again map onto those of a saturated spin-polarized band, but there is no longer a direct analogy to a spin-polarized theory. This is a very important consideration in understanding the physics of the problem, so we will devote space to explaining this. To set up a single-particle spin-polarized band in the limit of infinite U, all that one needs to do is to define a spin quantization direction for each atom and then force an electron arriving on that atom to arrive parallel to the chosen direction. The physical content of spin-polarized Hartree–Fock theory is then to choose these orientations so as to optimize the kinetic energy. The non-interacting character of such a band ensures that only one orientation can be found on any one atom, since if both were possible, both would happen and the Hubbard repulsion would be prohibitively expensive. What our description allows is for electrons of either

spin to hop onto a particular atom, but only to move to that atom when there is no other electron present, thereby avoiding the Hubbard repulsion completely. The electrons 'flow' along quite freely, but they carry their spin with them and do not change it according to which atom they reside on. The key to the theory is to label the spins according to the electron that they are found on, and not according to the atom on which they reside. Although the electrons carry their spins around, the charge motion is actually *collective* and the dominant motion comes from 'slopping' a region of charge along, without altering the relative positions of the electrons in that region appreciably; in practice, one can think of the charge as a 'baton' in a relay race, and each electron passes on the baton to its neighbour as the charge flows around in the system, but retains its spin.

In the absence of longer-range hopping, the spin order along the chain never changes, and the inclusion of infinitesimal longer-range hopping provides a mechanism for the spins to move on a corresponding extremely long time-scale. The process by which the spins move is that of short-range ring exchange as an electron hops past others. Unlike previous calculations, where the type of spin state predicted is fixed, as the concentration of charge carriers varies, there is competition between different types of exchange which can stabilize different types of magnetic correlation. It is this competition that controls the spin physics. For our unfrustrated next-next-nearest-neighbour hopping there are two types of spin interaction induced: ring exchange of two and three particles. When an electron hops the longer distance, there are three possibilities: if there are no particles hopped over then the spin configuration along the chain remains invariant. If a single electron is hopped over on one of the two intermediate atoms then there is a two-particle ring exchange, and if both intermediate sites are occupied then there is a three-particle exchange. Obviously the probabilities of these three possibilities vary strongly as a function of electron concentration, and this is what causes the changes of behaviour.

The type of spin interaction induced is controlled by the relative phase of the two methods of transferring the charge in the system. In the absence of longer-range hopping, the electrons can move into any spatial configuration, conserving the spin configuration along the chain. Pairs of configurations are therefore present for which a single electron appears to have been moved across others. The way the motion actually occurs, is via each electron playing the role of its nearest neighbour, i.e. *collective* motion of sliding all the electrons along exchanging places as they go. The inclusion of direct hopping allows a second route between two such charge configurations, with an electron actually passing over the others. We are in the limit where the collective motion is dominant, and so the spin configuration is controlled by the overlap between the two different resulting spin configurations and the relative phase in the charge wavefunction, and there are no compensations in the charge motion to make use of the new routes available. At first sight one might think that the Fermi phases in the free-electron gas wavefunction might present a problem, but this is not so, because for one dimension, spinless fermions map onto hard-core bosons and the hard-core boson wavefunction is positive definite. This allows us to deduce that the relative phase for collective motion is always unity, and the difference in phase is controlled by the Fermi minus sign for the real physical exchange of the electrons and the change in the spin wavefunction. Armed with these ideas, let us now look at our actual problem.

When an electron hops past another, there is an exchange of fermions and a resulting Fermi minus sign. If the two electrons had parallel spins, then this minus sign would lead to destructive interference and would lose energy, whereas two antiparallel electrons are in an anti-symmetric singlet that compensates the Fermi minus sign yielding constructive interference and a gain in energy. This effect strongly promotes low-spin correlations, and at low densities, where there is a very low probability of electrons hopping across pairs of

other electrons, this interaction dominates and forces the Heisenberg ground state on the spin system. When an electron hops across a pair of other electrons, then we find two Fermi minus signs, which cancel, promoting spin states that are 'symmetric' under ring exchange. In practice, ferromagnetism is the only spin state that is symmetric under all such ring exchanges, and near one electron per atom, where the probability of hopping past pairs of electrons dominates, this interaction is dominant and we find that ferromagnetism is stable. Near the critical density, where the competition is fierce, the situation is more subtle and worthy of further investigation. It is quite instructive to consider the critical point itself, where the phenomena are quite exotic: In the ferromagnetic phase the electrons hop over both single and paired electrons in just such a way that the lost energy from the first possibility *exactly* balances the gain in energy from the second. In the Néel antiferromagnet, however, the electrons *never* hop over each other at all, and only gain energy from hops over vacant sites. Although these two phases are precisely degenerate, the manner in which they achieve their low energy is quite different, with the ferromagnet yielding non-interacting motion while the Néel antiferromagnet has such strong correlations that the electrons only move if there is nothing in the way.

If we were to raise the strength of the infinitesimal interaction to finite values, it is instructive to ask which of the two predicted phases would become relatively stable. The ferromagnet has optimized all of its available degrees of freedom and remains in an identical state until the longer-range hops are a sizeable fraction of the nearest-neighbour hops, gaining nothing from the longer-range hops. For the antiferromagnetic Néel state, however, it is clear that the charge-motion is controlled by

$$H_{\text{Néel}} = -t_2 \sum_i \left[f_i^\dagger f_{i+3} (1 - f_{i+1}^\dagger f_{i+1})(1 - f_{i+2}^\dagger f_{i+2}) + \text{cc} \right] \tag{5.1}$$

involving interactions between the charges which will induce correlations, which will in turn save energy and stabilize the state with respect to the ferromagnet. Of course these charge correlations would feed back modifications to the spin interactions which would then change, feeding back further modifications to the charge state, ad nauseam. These charge correlations are collective effects much akin to pairing correlations in BCS theory and can be used to suggest why perovskite superconductors work [14].

6. Conclusions

We have mapped the one-dimensional Hubbard model with dominant nearest-neighbour hopping and infinitesimal next-nearest- and next-next-nearest-neighbour hopping onto a pair of simpler models for the separated charge and spin degrees of freedom. The charges move around under the action of the spinless fermion model, and the spins move around under the action of the Heisenberg model with both nearest- and next-nearest-neighbour spin interactions.

For the physically interesting case of unfrustrated next-next-nearest-neighbour hopping, the spin physics changes dramatically as a function of electron concentration. At low concentration there is a highly quantum spin paramagnet, while at high concentrations there is a non-interacting saturated ferromagnet. This agrees with the simple picture of Kanamori paramagnetism at low electron concentrations [5], changing to Nagaoka ferromagnetism at high electron concentration [6]. Although this agreement exists, and further our phase diagram is to all intents and purposes rigorous, we do *not* believe that the result is relevant to

the two-dimensional square lattice. We have been able to show, for extensions of our model to longer-range hopping, that the ferromagnetic phase has a shrinking range of stability, which vanishes when the range of the hopping diverges. Although our extended model is still not the two-dimensional square lattice, we are more convinced by the lack of a ferromagnetic region for the extended model than by the current stable ferromagnetic region.

The charge motion for these two phases is very different, with the ferromagnet being a non-interacting problem, while the paramagnet is a strongly interacting problem with particles moving only when the local configuration of their neighbours is favourable. Of course, the predominant motion is that of a free-fermion gas, but the longer-range hopping that lifts the spin degeneracy feeds back sophisticated many-particle interactions between the fermions when considered for the fixed spin wavefunction ground state. The source of energy is the avoidance of the phase cancellation inherent in a fermionic description: the fermionic minus signs, originating in exchange, can be compensated by the symmetry properties of the spin wavefunction, leading to enhanced motion. Since the enhancements occur only when particles pass others, it is not surprising that the electrons are attracted to each other in order to make use of this additional motion.

The modifications expected to the charge motion, fed back from the spin wavefunction, involve strong attractions between the charges, making them behave in a more hard-core boson manner than spinless fermion manner. The use of the spin wavefunction to avoid fermionic phase cancellation allows the particles to freely explore all of space, in a similar manner to the motion of hard-core bosons, and not to be restricted to avoiding lines in space as often happens with fermions.

When ferromagnetism ceases to be stable as a function of hole doping, the state that replaces it is fundamentally antiferromagnetic. This result comes purely from the charge motion and does *not* arise from any Heisenberg interactions which we have omitted from the model. This basic result, that the Heisenberg interactions yield local spin correlations very similar to those induced by permutations around square loops, should be born in mind when thinking about perovskite superconductors. It is quite possible that the Heisenberg interactions are irrelevant to the system in the important superconducting regime, where the motion around loops could easily take over the role of dominant interaction.

The use of the Jordan–Wigner transformation, to convert a bosonic representation to a fermionic representation and vice versa, may well prove to be a powerful technique for comparing the motion of charges in highly correlated systems. For the current models we see that the motion in our paramagnetic phases follows quite closely the behaviour of a hard-core Bose gas acting under equivalent longer-range hopping interactions, in complete agreement with the previous numerical work in two dimensions [7].

Appendix. The J_1–J_2 Heisenberg model

Although it is not central to our argument, since we map our Hubbard models onto the J_1–J_2 Heisenberg model, we ought to say what is currently known about the phase diagram of this model. As well as the obvious phase transition that occurs when $J_1 = 0$ and $J_2 < 0$ between a long-range ordered antiferromagnet (in the limit) and a ferromagnet, there are *three* other necessary phase transitions. Firstly, the ferromagnetism must destabilize again, and we believe that this occurs when $4J_2 = -J_1 > 0$. Secondly, it is known that when $2J_2 = J_1$ there is a gap in the spectrum, and so there must be two phase transitions, into and out of this gapped phase.

Ferromagnetism is *classical*, in the sense that there are no quantum fluctuations. Any second-order phase transition must therefore also be classical, and it is simple to verify that

the classical phase boundary between ferromagnetism and a spiral is when $4J_2 = -J_1 > 0$. This argument is 'true' but not rigorous, and further there could be a previous first-order transition to a singlet phase dominated by quantum mechanics. If one performs some exact diagonalization studies, then one is immediately led to a remarkable result that establishes the validity of this phase boundary to any reasonable person. For any finite loop with periodic boundary conditions there is an *exact* degeneracy between a saturated ferromagnetic ground state and a total-spin singlet spin-spiral ground state, when $4J_2 = -J_1 > 0$. Since small changes in parameters stabilize the expected ground state, it is difficult to see how this point could not correspond to the exact phase transition.

The other two phase transitions are much more subtle, since they occur between total-spin singlet states that cannot be found analytically. We believe that one transition may occur when $J_1 = 0$ and $J_2 > 0$, although we have no direct evidence for this conjecture, and the other transition must occur between $J_2 = 0$ and $2J_2 = J_1 > 0$. We believe that this second transition occurs exactly when $4J_2 = J_1$, but the evidence for this will be published at a later date.

References

[1] Anderson P W 1987 *Science* **235** 1196
[2] Phonons, plasmons, 'spin fluctuations' and many more have been suggested.
[3] Long M W 1991 *Int. J. Mod. Phys.* B **5** 865
[4] Lieb E H and Wu F Y 1968 *Phys. Rev. Lett.* **20** 1445
van Donge P and Vollhardt D 1989 *Phys. Rev.* B **40** 7252
[5] Kanamori J 1963 *Prog. Theor. Phys.* **30** 275
[6] Nagaoka Y 1966 *Phys. Rev.* **147** 393
[7] Long M W and Zotos X 1993 *Phys. Rev.* B **48** 317
[8] Jordan P and Wigner E 1928 *Z. Phys.* **47** 631
[9] Bethe H A 1931 *Z. Phys.* **71** 205
[10] Majumdar C K and Ghosh D K 1969 *J. Math. Phys.* **10** 1338
[11] Long M W and Siak S 1991 *J. Phys.: Condens. Matter* **3** 4901
[12] Putikka W O, Luchini M U and Ogata M 1992 *Phys. Rev. Lett.* **69** 2288
Barbieri A, Riera J A and Young A P 1990 *Phys. Rev.* B **41** 11 697
Trugman S A 1990 *Phys. Rev.* B **42** 6612
[13] von der Linden W and Edwards D M 1991 *J. Phys.: Condens. Matter* **3** 4917
[14] Long M W 1991 *J. Phys.: Condens. Matter* **3** 6375
Long M W 1991 *J. Phys.: Condens. Matter* **3** 6387