Field theories of paramagnetic Mott insulators

Subir Sachdev
Department of Physics, Yale University,
P.O. Box 208120, New Haven CT 06520-8120

Proceedings of the International Conference on
Theoretical Physics, Paris, UNESCO, 22-27 July 2002.

Abstract

This is a summary of a central argument in recent review articles by the author (Physica A **313**, 252 (2002), Annals of Physics **303**, 226 (2003), and Rev. Mod. Phys, July 2003). An effective field theory is derived for the low energy spin singlet excitations in a paramagnetic Mott insulator with collinear spin correlations.

1 INTRODUCTION

In a recent article [1](intended for an audience of experimentalists), the author has reviewed arguments that many aspects of the physics of the cuprate superconductors can be understood by using their proximity to paramagnetic Mott insulators. Further, a distinction was made between Mott insulators with collinear and non-collinear spin correlations, and it was argued that current experimental evidence suggests that we need only consider the collinear class. A phenomenological description of the ground states and excitations of these classes of Mott insulators was provided, along with a discussion of their experimental implications. A more technical discussion (intended for theorists) of such insulators, along with a description of the effective field theories which describe their low energy properties appears in Ref. [2, 3]. Here, we briefly recall the derivation and properties of the effective field theory of Mott insulators with *collinear* spin correlations, which is expressed in terms of a compact U(1) gauge field. The non-collinear class leads naturally

to a Z_2 gauge theory, but we will not consider it here. The reader is referred to these previous reviews [1, 2, 3] for complete citations to the literature.

2 Compact U(1) gauge theory of Mott insulators

We focus on Mott insulators on a d dimensional bipartite lattice of sites j. The SU(2) spin operator $\mathbf{S}_j$ on site j at imaginary time τ can be written as

$$\mathbf{S}_j(\tau) = \eta_j S \mathbf{n}(r_j, \tau); \tag{1}$$

here $\eta_j = \pm 1$ on the two sublattices, r_j is the spatial co-ordinate of site j, $\mathbf{n}$ is a unit length vector in spin space, and S is the (integer or half-odd-integer) angular momentum of each spin. The antiferromagnetic exchange interaction between near neighbor spins implies that $\mathbf{n}(r_j, \tau)$ will be a slowly varying function of its spacetime arguments. A standard analysis of the coherent state path integral of over the SU(2) $\mathbf{S}_j$ spins shows that the low energy quantum fluctuations are described by the following partition function

$$\begin{aligned} \mathcal{Z} &= \int \mathcal{D}\mathbf{n}(r,\tau)\delta(\mathbf{n}^2(r,\tau)-1)\exp\Bigg[-iS\sum_j \eta_j \int d\tau \mathcal{A}_\tau(\mathbf{n}(r_j,\tau)) \\ &\quad - \frac{1}{2gc}\int d^d r d\tau \left((\partial_\tau \mathbf{n})^2 + c^2(\nabla_r \mathbf{n})^2\right)\Bigg], \end{aligned} \tag{2}$$

where c is the spin-wave velocity, and g is a coupling constant which controls the strength of the quantum fluctuations. Excluding the first Berry phase term, this is the action of the O(3) non-linear sigma model in $d+1$ spacetime dimensions. Here we are primarily interested in the consequences of the Berry phases: $\mathcal{A}_\tau(\mathbf{n}(\tau))d\tau$ is defined to be the oriented area of the spherical triangle defined by $\mathbf{n}(\tau)$, $\mathbf{n}(\tau + d\tau)$, and an arbitrary reference point $\mathbf{n}_0$ (which is usually chosen to be the north pole).

The theory (2) can be considered to be the "minimal model" of antiferromagnets. In dimensions $d > 1$ it has at least two phases: at small g there is the conventional magnetically ordered "Néel" phase with $\langle \mathbf{n} \rangle \neq 0$, while at large g there is a "quantum disordered" paramagnetic phase which preserves spin rotation invariance with $\langle \mathbf{n} \rangle = 0$. We are especially interested here in the nature of this paramagnetic state. In this section, we will manipulate

$\mathcal{Z}$ in this large g regime, and derive an alternative formulation which allows easier computation of the integral over the Berry phases.

The key to an analysis of the large g regime is a better understanding of the nature of $\mathcal{A}_\tau$. We will see that $\mathcal{A}_\tau$ behaves in many respects like the time-component of a compact U(1) gauge field, and indeed, this accounts for the suggestive notation. All physical results should be independent of the choice of the reference point $\mathbf{n}_0$, and it is easy to see by drawing triangles on the surface of a sphere that changes in $\mathbf{n}_0$ amount to gauge transformations of $\mathcal{A}_\tau$. If we change $\mathbf{n}_0$ to $\mathbf{n}_0'$, then the resulting $\mathcal{A}_\tau'$ is related to $\mathcal{A}_\tau$ by

$$\mathcal{A}_\tau' = \mathcal{A}_\tau - \partial_\tau \phi(\tau) \tag{3}$$

where $\phi(\tau)$ measures the oriented area of the spherical triangle defined by $\mathbf{n}(\tau)$, $\mathbf{n}_0$, and $\mathbf{n}_0'$. Furthermore, as we will discuss more completely below, the area of any spherical triangle is uncertain modulo 4π, and this accounts for the 'compactness' of the U(1) gauge theory.

We proceed with our analysis of $\mathcal{Z}$. First, we discretize the gradient terms of the O(3) sigma model. We will limit our considerations here to antiferromagnets on d dimensional cubic lattices, but similar considerations apply to other bipartite lattices. We also discretize the imaginary time direction, and (by a slight abuse of notation) use the same index j to refer to the sites of a $d+1$ dimensional cubic lattice in spacetime. On such a lattice we can rewrite (2) as

$$Z = \int \prod_j d\mathbf{n}_j \delta(\mathbf{n}_j^2 - 1) \exp\left(\frac{1}{2g} \sum_{j,\mu} \mathbf{n}_j \cdot \mathbf{n}_{j+\hat{\mu}} - iS \sum_j \eta_j \mathcal{A}_{j\tau} \right), \tag{4}$$

where the sum over μ extends over the $d+1$ spacetime directions. We have also dropped unimportant factors of the lattice spacing and the spin-wave velocity in (4).

As noted above, we are especially interested here in the large g regime where there are strong fluctuations of the $\mathbf{n}_j$. There are strong cancellations from the Berry phases between different spin configurations in this regime, and so the second term in Z has to be treated with great care. We will do this by promoting the field $\mathcal{A}_{j\mu}$ to an independent degree of freedom, while integrating out the $\mathbf{n}_j$. Notice that we have now introduced all $d+1$ components of the compact U(1) gauge field with the index μ, while only the $\mu = \tau$ component appears explicitly in (4). The remaining components

appear naturally as suitable degrees of freedom when we integrate the $\mathbf{n}_j$ out. Formally, the integration over the $\mathbf{n}_j$ can be done by introducing new 'dummy' variables $A_{j\mu}$ and rewriting (4) by introducing factors of unity on each link; this leads to

$$\begin{aligned} Z &= \int \prod_{j\mu} dA_{j\mu} \exp\left(-i2S \sum_j \eta_j A_{j\tau}\right) \int \prod_j d\mathbf{n}_j \delta(\mathbf{n}_j^2 - 1) \delta(\mathcal{A}_{j\mu}/2 - A_{j\mu}) \\ &\qquad \times \exp\left(\frac{1}{2g} \sum_{j,\mu} \mathbf{n}_j \cdot \mathbf{n}_{j+\hat{\mu}}\right) \\ &= \int \prod_{j\mu} dA_{j\mu} \exp\left(-\mathcal{S}_A(A_{j\mu}) - i2S \sum_j \eta_j A_{j\tau}\right). \end{aligned} \tag{5}$$

In the first expression, if the integral over the $A_{j\mu}$ is performed first, we trivially return to (4); however, in the second expression we perform the integral over the $\mathbf{n}_j$ variables first, at the cost of introducing an unknown effective action $\mathcal{S}_A$ for the $A_{j\mu}$. In principle, evaluation of $\mathcal{S}_A$ may be performed order-by-order in a "high temperature" expansion in $1/g$: we match correlators of the $A_{j\mu}$ flux with those of the $\mathcal{A}_{j\mu}$ flux evaluated in the integral over the $\mathbf{n}_j$ with positive weights determined only by the $1/g$ term in (4). Rather than undertaking this laborious calculation, we can guess essential features of the effective action $\mathcal{S}_A$ from some general constraints. First, correlations in the $\mathbf{n}_j$ decay exponentially rapidly for large g (with a correlation length $\sim 1/\ln(g)$), and so $\mathcal{S}_A$ should be local. Second, it should be invariant under the lattice form of the gauge transformation (3)

$$A'_{j\mu} = A_{j\mu} - \Delta_\mu \phi_j/2 \tag{6}$$

associated with the change in the reference point on the unit sphere from $\mathbf{n}_0$ to $\mathbf{n}'_0$, with ϕ_j equal to the area of the spherical triangle formed by $\mathbf{n}_j$, $\mathbf{n}_0$ and $\mathbf{n}'_0$. Finally the area of any triangle on the sphere is uncertain modulo 4π and so the effective action should be invariant under

$$A_{j\mu} \to A_{j\mu} + 2\pi. \tag{7}$$

The simplest local action which is invariant under (6) and (7) is that of *compact U(1) quantum electrodynamics* and so we have

$$Z = \int \prod_{j\mu} dA_{j\mu} \exp\left(\frac{1}{e^2} \sum_\square \cos\left(\Delta_\mu A_{j\nu} - \Delta_\nu A_{j\mu}\right) - i2S \sum_j \eta_j A_{j\tau}\right), \tag{8}$$

for large g; comparison with the large g expansion shows that the coupling $e^2 \sim g^2$. In (8), Δ_μ is the discrete lattice derivative along the μ direction, and the sum over $\Box$ extends over all plaquettes of the $d+1$ dimensional cubic lattice—both notations are standard in the lattice gauge theory literature.

The first term in the action (8) is, of course, the standard 'Maxwell' term of a compact U(1) gauge field. In this language, the Berry phase has the interpretation of a $\int J_\mu A_\mu$ coupling to a fixed matter field with 'current' $J_\mu = 2S\delta_{\mu\tau}$. This corresponds to static matter with charges $\pm 2S$ on the two sublattices. It is this matter field which will crucially control the nature of the ground state.

The remaining analysis of Z depends upon the spatial dimensionality d. In $d = 1$, a dual model of (8) is solvable, and the results are in complete accord with those obtained earlier by Bethe ansatz and bosonization analyses of spin chains. We will consider the $d = 2$ case in the section below. There has been relatively little discussion of the $d = 3$ case (which exhibits both confining and deconfining phases of the gauge theory), and this remains an important avenue for future research.

3 Duality mapping in $d = 2$

As is standard in duality mappings, we first rewrite the partition function in $2 + 1$ spacetime dimensions by replacing the cosine interaction in (8) by a Villain sum over periodic Gaussians:

$$Z = \sum_{\{q_{\bar{j}\mu}\}} \int \prod_{j\mu} dA_{j\mu} \exp\left(-\frac{1}{2e^2} \sum_{\Box} \left(\epsilon_{\mu\nu\lambda}\Delta_\nu A_{j\lambda} - 2\pi q_{\bar{j}\mu}\right)^2 - i2S \sum_j \eta_j A_{j\tau} \right), \tag{9}$$

where $\epsilon_{\mu\nu\lambda}$ is the total antisymmetric tensor in three dimensions, and the $q_{\bar{j}\mu}$ are integers on the links of the *dual* cubic lattice, which pierce the plaquettes of the direct lattice. Throughout this subsection we will use the index $\bar{j}$ to refer to sites of this dual lattice, while j refers to the direct lattice on sites on which the spins are located.

We will now perform a series of exact manipulations on (9) which will lead to a dual *interface* model [4, 5]. This dual model has only positive weights—this fact, of course, makes it much more amenable to a standard statistical analysis. This first step in the duality transformation is to rewrite

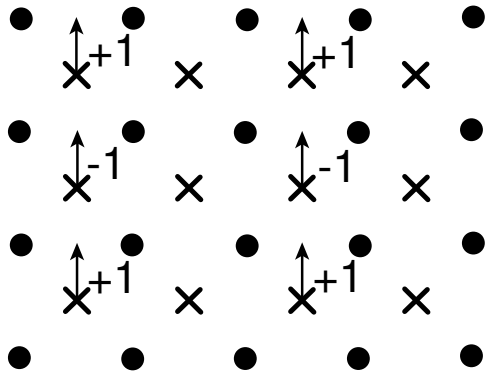

Figure 1: Specification of the non-zero values of the fixed field $a^0_{\bar{j}\mu}$. The circles are the sites of the direct lattice, j, while the crosses are the sites of the dual lattice, $\bar{j}$; the latter are also offset by half a lattice spacing in the direction out of the paper (the $\mu = \tau$ direction). The $a^0_{\bar{j}\mu}$ are all zero for $\mu = \tau, x$, while the only non-zero values of $a^0_{\bar{j}y}$ are shown above. Notice that the a^0 flux obeys (11).

(9) by the Poisson summation formula:

$$\sum_{\{q_{\bar{j}\mu}\}} \exp \left(-\frac{1}{2e^2} \sum_{\Box} \left(\epsilon_{\mu\nu\lambda} \Delta_\nu A_{j\lambda} - 2\pi q_{\bar{j}\mu} \right)^2 \right) = \sum_{\{a_{\bar{j}\mu}\}} \exp \left(-\frac{e^2}{2} \sum_{\bar{j}} a^2_{\bar{j}\mu} - i \sum_{\Box} \epsilon_{\mu\nu\lambda} a_{\bar{j}\mu} \Delta_\nu A_{j\lambda} \right), \quad (10)$$

where $a_{\bar{j}\mu}$ (like $q_{\bar{j}\mu}$) is an integer-valued vector field on the links of the dual lattice (here, and below, we drop overall normalization factors in front of the partition function). Next, we write the Berry phase in a form more amenable to duality transformations. Choose a 'background' $a_{\bar{j}\mu} = a^0_{\bar{j}}$ flux which satisfies

$$\epsilon_{\mu\nu\lambda} \Delta_\nu a^0_{\bar{j}\lambda} = \eta_j \delta_{\mu\tau}, \quad (11)$$

where j is the direct lattice site in the center of the plaquette defined by the curl on the left-hand-side. Any integer-valued solution of (11) is an acceptable choice for $a^0_{\bar{j}\mu}$, and a convenient choice is shown in Fig 1. Using (11) to rewrite the Berry phase in (9), applying (10), and shifting $a_{\bar{j}\mu}$ by the

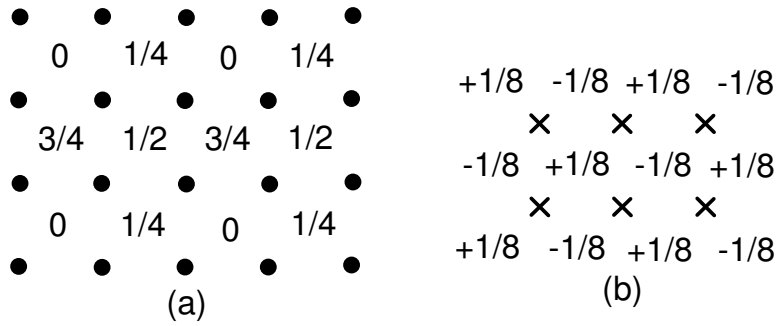

Figure 2: Specification of the non-zero values of the fixed fields (a) $\mathcal{X}_{\bar{\jmath}}$ and (b) $\mathcal{Y}_{j\mu}$ introduced in (14). The notational conventions are as in Fig 1. Only the $\mu = \tau$ components of $\mathcal{Y}_{j\mu}$ are non-zero, and these are shown in (b).

integer $2Sa^0_{\bar{\jmath}\mu}$, we obtain a new exact representation of Z in (9):

$$Z = \sum_{\{a_{\bar{\jmath}\mu}\}} \int \prod_{j\mu} dA_{j\mu} \exp\left(-\frac{e^2}{2} \sum_{\bar{\jmath},\mu} (a_{\bar{\jmath}\mu} - 2Sa^0_{\bar{\jmath}\mu})^2 - i \sum_{\square} \epsilon_{\mu\nu\lambda} a_{\bar{\jmath}\mu} \Delta_\nu A_{j\lambda} \right). \tag{12}$$

The integral over the $A_{j\mu}$ can be performed independently on each link, and its only consequence is the imposition of the constraint $\epsilon_{\mu\nu\lambda}\Delta_\nu a_{\bar{\jmath}\lambda} = 0$. We solve this constraint by writing $a_{\bar{\jmath}\mu}$ as the gradient of a integer-valued 'height' $h_{\bar{\jmath}}$ on the sites of the dual lattice, and so obtain

$$Z = \sum_{\{h_{\bar{\jmath}}\}} \exp\left(-\frac{e^2}{2} \sum_{\bar{\jmath},\mu} (\Delta_\mu h_{\bar{\jmath}} - 2Sa^0_{\bar{\jmath}\mu})^2 \right). \tag{13}$$

This is the promised 2+1 dimensional interface, or height, model in almost its final form.

The physical properties of (13) become clearer by converting the "frustration" $a^0_{\bar{\jmath}\mu}$ in (13) into offsets for the allowed height values. This is done by decomposing $a^0_{\bar{\jmath}\mu}$ into curl and divergence free parts and writing it in terms of new fixed fields, $\mathcal{X}_{\bar{\jmath}}$ and $\mathcal{Y}_{j\mu}$ as follows:

$$a^0_{\bar{\jmath}\mu} = \Delta_\mu \mathcal{X}_{\bar{\jmath}} + \epsilon_{\mu\nu\lambda}\Delta_\nu \mathcal{Y}_{j\lambda}. \tag{14}$$

The values of these new fields are shown in Fig 2. Inserting (14) into (13),

we can now write the height model in its simplest form [4]

$$Z_h = \sum_{\{H_{\bar{\jmath}}\}} \exp\left(-\frac{e^2}{2}\sum_{\bar{\jmath}} \left(\Delta_\mu H_{\bar{\jmath}}\right)^2\right), \tag{15}$$

where

$$H_{\bar{\jmath}} \equiv h_{\bar{\jmath}} - 2S\mathcal{X}_{\bar{\jmath}} \tag{16}$$

is the new height variable we shall work with. Notice that the $\mathcal{Y}_{j\mu}$ have dropped out, while the $\mathcal{X}_{\bar{\jmath}}$ act only as fractional offsets (for S not an even integer) to the integer heights. From (16) we see that for half-odd-integer S the height is restricted to be an integer on one of the four sublattices, an integer plus 1/4 on the second, an integer plus 1/2 on the third, and an integer plus 3/4 on the fourth; the fractional parts of these heights are as shown in Fig 2a; the steps between neighboring heights are always an integer plus 1/4, or an integer plus 3/4. For S an odd integer, the heights are integers on one square sublattice, and half-odd-integers on the second sublattice. Finally for even integer S the offset has no effect and the height is an integer on all sites. We discuss these classes of S values in turn in the following subsections.

3.1 S even integer

In this case the offsets $2S\mathcal{X}_{\bar{\jmath}}$ are all integers, and (15) is just an ordinary three dimensional height model which has been much studied in the literature. Unlike the two-dimensional case, three-dimensional height models generically have no roughening transition, and the interface is always smooth. With all heights integers, the smooth phase breaks no lattice symmetries. So square lattice antiferromagnets with S even integer can have a paramagnetic ground state with a spin gap and no broken symmetries. This is in accord with the exact ground state for a $S = 2$ antiferromagnet on the square lattice found by Affleck *et al.*, the AKLT state [6].

3.2 S half-odd-integer

Now the heights of the interface model can take four possible values, which are integers plus the offsets on the four square sublattices shown in Fig 2a. As in Section 3.1, the interface is always smooth *i.e.* any state of (15) has

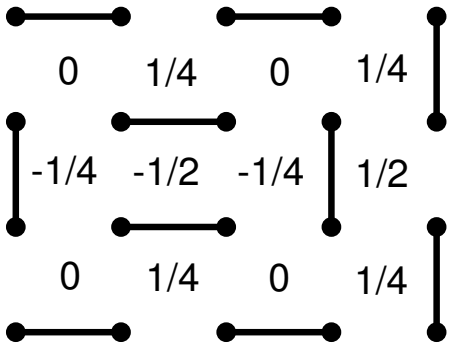

Figure 3: Mapping between the quantum dimer model and the interface model Z in (15). Each dimer on the direct lattice is associated with a step in height of $\pm 3/4$ on the link of the dual lattice that crosses it. All other height steps are $\pm 1/4$. Each dimer represents a singlet valence between the sites, as in Fig 2.

a fixed average interface height $\sum_{\bar{\jmath}} \langle H_{\bar{\jmath}} \rangle$, and *any* well-defined value for this average height breaks the uniform shift symmetry of the height model under which $H_{\bar{\jmath}} \to H_{\bar{\jmath}} \pm 1$. After accounting for the height offsets, we will see below that any smooth interface must also break a lattice symmetry with the development of *bond order*: this allows a number of distinct spin gap ground states of the lattice antiferromagnet.

It is useful, first, to obtain a simple physical interpretation of the interface model in the language of the $S = 1/2$ antiferromagnet [7]. From Fig 2a it is clear that nearest neighbor heights can differ either by $1/4$ or $3/4$ (modulo integers). To minimize the action in (15), we should choose the interface with the largest possible number of steps of $\pm 1/4$. However, the interface is frustrated, and it is not possible to make all steps $\pm 1/4$ and at least a quarter of the steps must be $\pm 3/4$. Indeed, there is a precise one-to-one mapping between interfaces with the minimal number of $\pm 3/4$ steps (we regard interfaces differing by a uniform integer shift in all heights as equivalent) and the dimer coverings of the square lattice: the proof of this claim is illustrated in Fig 3. We identify each dimer with a singlet valence bond between the spins (the ellipses in Fig 2), and so each interface corresponds to a quantum state with each spin locked in the a singlet valence bond with a particular nearest neighbor. Fluctuations of the interface in imaginary time between such configurations correspond to quantum tunneling events between such dimer states, and an effective Hamiltonian for this is provided by the quantum

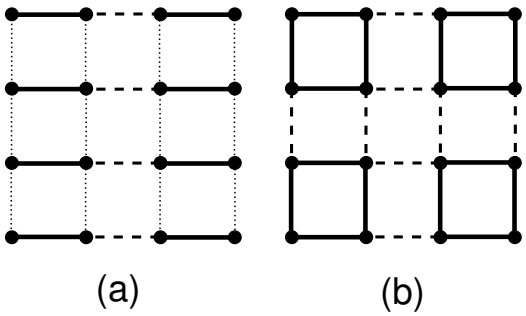

Figure 4: Sketch of the two simplest possible states with bond order for $S = 1/2$ on the square lattice: (a) the columnar spin-Peierls states, and (b) plaquette state. The different values of the $\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle$ on the links are encoded by the different line styles. Both states are 4-fold degenerate; an 8-fold degenerate state, with superposition of the above orders, also appears as a possible ground state of the generalized interface model.

dimer model [8].

The nature of the possible smooth phases of the interface model are easy to determine from the above picture and by standard techniques from statistical theory [4, 7]. Interfaces with average height $\langle H_{\bar{\jmath}} \rangle = 1/8, 3/8, 5/8, 7/8$ (modulo integers) correspond to the four-fold degenerate bond-ordered states in Fig 4a, while those with $\langle H_{\bar{\jmath}} \rangle = 0, 1/4, 1/2, 3/4$ (modulo integers) correspond to the four-fold degenerate plaquette bond-ordered states in Fig 4b. All other values of $\langle H_{\bar{\jmath}} \rangle$ are associated with eight-fold degenerate bond-ordered states with a superposition of the orders in Fig 4a and b.

Support for the class of bond-ordered states described above has appeared in a number of numerical studies of $S = 1/2$ antiferromagnets in $d = 2$ which have succeeded in moving from the small g Néel phase to the large g paramagnet. These include studies on the honeycomb lattice [9], on the planar pyrochlore lattice [10], on square lattice models with ring-exchange and easy-plane spin symmetry [12], and square lattice models with SU(N) symmetry [14].

3.3 S odd integer

This case is similar to that S half-odd-integer, and we will not consider it in detail. The Berry phases again induce bond order in the spin gap state, but

this order need only lead to a two-fold degeneracy.

4 Conclusions

The primary topic discussed in this paper has been the effective field theory of paramagnetic Mott insulators with collinear spin correlations. This field theory is the compact U(1) gauge theory in (8), and applies in all spatial dimensions. We also reviewed duality mappings of (8) which are special to $d = 2$ spatial dimensions, and mapped the theory onto the interface model (15). Finally, we reiterate that paramagnetic Mott insulators with non-collinear spin correlations are described by a Z_2 gauge theory which has not been presented here.

References

[1] S. Sachdev, Rev. Mod. Phys. July 2003, cond-mat/0211005.

[2] S. Sachdev, Physica A **313**, 252 (2002).

[3] S. Sachdev, Annals of Physics **303**, 226 (2003).

[4] N. Read and S. Sachdev, Phys. Rev. Lett. **62**, 1694 (1989); Phys. Rev. B **42**, 4568 (1990).

[5] E. Fradkin and S. A. Kivelson, Mod. Phys. Lett. B **4**, 225 (1990).

[6] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Phys. Rev. Lett. **59**, 799 (1987).

[7] W. Zheng and S. Sachdev, Phys. Rev. B **40**, 2704 (1989).

[8] D. Rokhsar and S. A. Kivelson, Phys. Rev. Lett. **61**, 2376 (1988)

[9] J.-B. Fouet, P. Sindzingre, and C. Lhuillier, Eur. Phys. J. B **20**, 241 (2001); duality mapping on the honyecomb lattice appears in Ref. [4].

[10] J.-B. Fouet, M. Mambrini, P. Sindzingre, and C. Lhuillier, Phys. Rev. B **67**, 054411 (2003); duality mapping for a lattice with the symmetry of the planar pyrochore is in Ref. [11], with a prediction for the bond order observed.

[11] C.-H. Chung, J. B Marston, and S. Sachdev, Phys. Rev. B **64**, 134407 (2001).

[12] A. W. Sandvik, S. Daul, R. R. P. Singh, and D. J. Scalapino, Phys. Rev. Lett., **89**, 247201 (2002); duality mapping on spin models with with easy plane symmetry is in Ref. [13].

[13] K. Park and S. Sachdev, Phys. Rev. B **65**, 220405 (2002).

[14] K. Harada, N. Kawashima, and M. Troyer, Phys. Rev. Lett. **90**, 117203 (2003); the theory (15) applies unchanged to SU(N) antiferromagnets.