Charles L. Cleveland and Rodrigo Medina A.

## Related Articles

## Additional information on Am. J. Phys.

# Obtaining a Heisenberg Hamiltonian from the Hubbard model

Charles L. Cleveland and Rodrigo Medina A.*
*School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332*
(Received 17 April 1974; revised 9 May 1975)

A pedagogically useful method of obtaining a Heisenberg spin–spin interaction from a familiar many-electron Hamiltonian by means of ordinary degenerate perturbation theory is presented.

## I. INTRODUCTION

In recent years much attention has been given to the Hubbard model,[1] a simplified picture which contains in a qualitative way some of the more important aspects of electron behavior in an atomic lattice. We shall obtain from it a spin–spin interaction of the Heisenberg form (e.g., $S \cdot S$) by the elementary and pedagogically valuable means of Rayleigh–Schrödinger perturbation theory.

In Sec. II the Hubbard model is briefly discussed, the Hubbard Hamiltonian is introduced, and an attempt is made to show how perturbation theory might be a reasonable approach to the model.

In Sec. III we apply degenerate perturbation theory to yield the desired spin–spin interaction.

Section IV consists of concluding remarks.

## II. THE HUBBARD MODEL AND HAMILTONIAN

The Hubbard model can be thought of as a lattice of atoms supporting only a single atomic state, which can hold up to two electrons with opposite spins.

The Hamiltonian contains one part that can be thought of as the contribution to the energy due to the electrons' motion as they hop from site to site (the "hopping" part), and another part which represents the Coulomb repulsion between two electrons on the same site (the "Coulomb repulsion" part).

Regarding the system as being close to the atomic limit, which we shall do throughout this paper, we can reasonably speak of an electron being "on" a particular site and have some justification for neglecting the Coulomb repulsion between two electrons on different sites. We further presume that each atom contributes one electron to the system, so that there are just as many electrons as sites.

Explicitly, the Hamiltonian is

$$H = \sum_{ij}' \sum_{\sigma} T_{ij} c_{i\sigma}{}^{+} c_{j\sigma} + \tfrac{1}{2} I \sum_{i\sigma} n_{i\sigma} n_{i-\sigma}. \qquad (1)$$

Indices $i$ and $j$ represent lattice sites, while $\sigma$ is a spin index. The operator $c_{i\sigma}{}^{+}$ $(c_{i\sigma})$ creates (destroys) an electron of spin $\sigma/2$ on site $i$. The operator $n_{i\sigma}$ is defined by

$$n_{i\sigma} = c_{i\sigma}{}^{+} c_{i\sigma}$$

and counts the number of electrons of spin $\sigma/2$ on site $i$. The prime on the first sum of Eq. (1) excludes the terms such that $i = j$, and causes Eq. (1) and the Hamiltonian of Hubbard's 1963 paper[1] to differ by a constant. Since $H$ is presumed to be Hermitian, the numbers $T_{ij}$ must satisfy the relation

$$T_{ij} = T_{ji}{}^{*}.$$

The first sum is the part of $H$ we have called the hopping part. The action of one of its terms is to move an electron of spin $\sigma/2$ from site $j$ to site $i$. The hopping part represents the kinetic energy of the electrons together with the energy due to the periodic potential. The larger the contribution of this sum to the energy, the more freely the electrons move about the lattice.

The second sum is the Coulomb repulsion part. The positive number $I$ is the electrostatic energy of two electrons on the same site, while the remainder of the Coulomb repulsion part counts the number of sites with a pair of electrons of opposite spin. Clearly, the interaction between electrons on different sites has been neglected.

Since the system in its ground state will attempt to minimize its energy, the hopping part will prefer the situation where the electrons move freely about the lattice, not being localized to particular lattice sites (band behavior). This situation, however, has superpositions of electrons on any given site, which makes the Coulomb repulsion part larger. The Coulomb repulsion part prefers the highly localized situation where there is one and only one electron on any site. Thus the Hubbard Hamiltonian contains two competing parts: one which encourages nonlocalized and the other localized electronic behavior. Since we are close to the atomic limit, we can expect localized behavior to dominate; that is, we can expect the contribution of the Coulomb repulsion part to be much larger than the contribution of the hopping part. This sets the stage for a perturbative approach to the model, taking the hopping part as the perturbation.

## III. OBTAINING THE SPIN–SPIN INTERACTION

In the framework of degenerate Rayleigh–Schrödinger perturbation theory, we will write the Hamiltonian as

$$H = H_0 + H_1, \qquad (2)$$

where the unperturbed Hamiltonian, $H_0$, is given by

$$H_0 = \tfrac{1}{2} I \sum_{i\sigma} n_{i\sigma} n_{i-\sigma}, \qquad (3)$$

and the perturbation is

$$H_1 = \sum_{ij}' \sum_{\sigma} T_{ij} c_{i\sigma}{}^{+} c_{j\sigma}. \qquad (4)$$

We consider now the eigenvalues and eigenstates of $H_0$, which play an important role in the perturbation theory. Since $\sum_{i\sigma} n_{i\sigma} n_{i-\sigma}/2$ acting on a state merely counts the number, $n$, of sites with a pair of electrons, the eigenvalues of $H_0$ depend only on $n$ and we may write

$$H_0 \left| nk \right\rangle_0 = E_n^{(0)} \left| nk \right\rangle_0 = nI \left| nk \right\rangle_0 .$$

The state $\left| nk \right\rangle_0$ is not uniquely determined by $n$, but may be uniquely specified by listing, in detail, how many electrons of each spin are on each site. $k$ may be thought of as providing this listing.

As usual, we must apply the perturbation theory for one value of $n$ at a time. For the purposes of this paper we will only do it for $n = 0$, that is, for the ground states of $H_0$, for which there is only one electron on each site. Besides simplicity, reasons for not repeating the calculations for different values of $n$ are given in Sec. IV.

The multiplicity of these ground states is large; since on any site the spin may be flipped without affecting the energy, the degeneracy is $2^N$, where $N$ is the total number of sites. The first order of perturbation theory gives the eigenvalue equation[2]

$$E_{0k}^{(1)} \delta_{lk} = {}_0\langle 0l | H_1 | 0k \rangle_0 . \qquad (5)$$

In general, this equation gives the first-order energy corrections and removes the degeneracy of the unperturbed states. But, from (4), we can see that each term of $H_1$ acting on a state with one electron on each site must produce a state with precisely one pair on some site. That is,

$${}_0\langle nl | H_1 | 0k \rangle_0 \neq 0 \quad \text{only if } n = 1. \qquad (6)$$

So the matrix denoted by Eq. (5) is entirely filled with zeroes, all the first-order corrections $E_{0k}^{(1)}$ are zero, and the degeneracy is completely unremoved. We then must pass to the second-order equation (which has to be considered, as the first-order equation was, as an eigenvalue equation). In this case it reads as follows[2]:

$$E_{0k}^{(2)} \delta_{jk} = {}_0\left\langle 0j \left| H_1 \sum_{pl}' \frac{|pl\rangle_0 \, {}_0\langle pl|}{E_0^{(0)} - E_p^{(0)}} H_1 \right| 0k \right\rangle_0 .$$

The prime omits from the sum all $p$ such that $E_p^{(0)} = E_0^{(0)}$. Now from Eq. (6) we note that only those terms in the sum which have $p = 1$ are not zero in the last equation. So, without loss of generality, we replace the denominator, $E_0^{(0)} - E_p^{(0)}$, by $-I$, and see that we may drop the prime on the sum since, by Eq. (6), the prime excludes only terms that are zero anyway. We obtain

$$E_{0k}^{(2)} \delta_{jk} = - {}_0\left\langle 0j \left| H_1 \left( \sum_{pl} |pl\rangle_0 \, {}_0\langle pl| \right) H_1 \right| 0k \right\rangle_0 I^{-1},$$

but the sum in parentheses is just the identity operator since the $\{|pl\rangle_0\}$ form a complete set of states. So

$$E_{0k}^{(2)} \delta_{jk} = - {}_0\langle 0j | H_1^2 | 0k \rangle_0 / I$$

or, since $E_{nk}^{(0)} = E_{nk}^{(1)} = 0$, we have to second order

$$E_{0k} \delta_{jk} = - {}_0\langle 0j | H_1^2 | 0k \rangle_0 / I.$$

That is, to second order the eigenvalues $E_{0k}$ of $H$ are just the eigenvalues of $-H_1^2/I$ restricted to the subspace corresponding to the lowest eigenvalue of $H_0$. We write then for our effective second-order Hamiltonian

$$H_{eff} = P_0(- H_1^2/I)P_0 , \qquad (7)$$

where $P_0$ is the projection operator that restricts the action of $H_{eff}$ to the proper subspace of states with $n = 0$.

We now obtain an expression for $H_{eff}$ in terms of spin operators. First we use the expression given by Eq. (4) for $H_1$ to obtain

$$H_{eff} = P_0\left( -\sum_{ij}' \sum_{kl}' \sum_{\sigma\eta} T_{ij} T_{kl} c_{i\sigma}^+ c_{j\sigma} c_{k\eta}^+ c_{l\eta} / I \right) P_0 .$$

Then we note that many terms in the sum are zero. Because of the projection operator on the right, any term acts on a state with one electron on each site, and because of the projection operator on the left, that term contributes only if it produces another state with one electron on each site. Since $i \neq j$ and $k \neq l$, only those terms for which $j = k$ and $i = l$ contribute. We now have

$$H_{eff} = P_0\left( -\sum_{lk}' \sum_{\sigma\eta} T_{lk} T_{kl} c_{l\sigma}^+ c_{k\sigma} c_{k\eta}^+ c_{l\eta} / I \right) P_0 .$$

Rearranging via the usual anticommutation rules and using the fact that $T_{ij} = T_{ji}^*$ yield

$$H_{eff} = P_0\left( -\sum_{lk}' \sum_{\sigma\eta} |T_{kl}|^2 c_{l\sigma}^+ c_{l\eta}(\delta_{\sigma\eta} - c_{k\eta}^+ c_{k\sigma})/I \right) P_0 .$$

Explicitly carrying out the sum over $\eta$, we obtain

$$H_{eff} = P_0\left( \sum_{lk}' |T_{lk}|^2 \right.$$
$$\left. \times \sum_\sigma (n_{l\sigma} n_{k\sigma} + c_{l\sigma}^+ c_{l-\sigma} c_{k-\sigma}^+ c_{k\sigma} - n_{l\sigma})/I \right) P_0 . \qquad (8)$$

Spin operators for the electronic spin on a site can be simply defined in terms of the fermion operators as follows:

$$S_i^z = \tfrac{1}{2}\sum_\sigma n_{i\sigma}, \quad z \text{ component} \qquad (9)$$

and

$$S_{i\sigma} = c_{i\sigma}^+ c_{i-\sigma} , \qquad (10)$$

raising ($\sigma = +1$) and lowering ($\sigma = -1$). From (9) we have quickly that

$$P_0 \sum_\sigma n_{l\sigma} n_{k\sigma} P_0 = P_0(2 S_l^z S_k^z + \tfrac{1}{2}) P_0$$

C. L. Cleveland and R. Medina A.    45

and, since

$$P_0 \sum_\sigma n_{l\sigma} P_0 = P_0 ,$$

Eq. (8) transforms to

$$H_{\text{eff}} = P_0 \left[ \sum_{kl}{}' \, |T_{kl}|^2 \left( \sum_\sigma S_{l\sigma} S_{k-\sigma} + 2 S_l{}^z S_k{}^z - \tfrac{1}{2} \right) I^{-1} \right] P_0 .$$

But since

$$\mathbf{S}_l \cdot \mathbf{S}_k = S_l{}^z S_k{}^z + \tfrac{1}{2} \sum_\sigma S_{l\sigma} S_{k-\sigma} ,$$

we finally obtain

$$H_{\text{eff}} = P_0 \left( \sum_{kl}{}' \, |T_{kl}|^2 (2\mathbf{S}_l \cdot \mathbf{S}_k - \tfrac{1}{2})/I \right) P_0 , \qquad (11)$$

that is, a Heisenberg-type Hamiltonian with strictly positive coefficient. Our system is approximately a Heisenberg antiferromagnet.

## IV. CONCLUDING REMARKS

Something like what we have done here has been done before (see, for example, Ref. 3, in which a nearest-neighbor interaction of the Heisenberg type is obtained from other than the Hubbard model per se), but to our knowledge not in a manner so straightforward or so assimilable to the student.

This desire that the paper be pedagogically valuable is one of the reasons we have performed our calculation for the ground state only. The arguments have a special simplicity for this case which we hope will minimize confusion on the part of the reader, who may realize at this point that the loss of simplicity for higher levels might imply that for these higher levels one begins to obtain contributions of a character different from the Heisenberg interaction. This is indeed the case. Consider, the example, the level $E_1{}^{(0)} = I$, the states of which have one site containing a pair of electrons and one site empty of electrons. The nature of the perturbation theory alone tells us that we need only concern ourselves with transitions between these states (although, of course, at some intermediate point in the transition the state of the system may not belong to this level), but whereas the pair might be on one site initially, after the transition (since the Hub-

bard Hamiltonian "moves" electrons) the pair might be on another site. In other words, transitions between states belonging to the level $E_1{}^{(0)} = I$ can permit a net motion of electrons.

This inspires one to think that for nonground states one may begin to see conduction phenomena. These phenomena are not part of the topic of this paper and are easily discarded by considering the ground states only. As for the gross magnetic behavior of the system, it seems unlikely that it will be altered much by the existence of one site with a pair of electrons, or a million such sites, out of Avogadro's number of sites. So there is nothing to prevent one from applying this Heisenberg-type interaction to higher levels, where it will be approximately valid. Spin waves and other phenomena associated with the Heisenberg model can now be interpreted from the point of view of the conceptually simple Hubbard model.

The reader familiar with the Heisenberg interaction will also have realized that our Hamiltonian $H_{\text{eff}}$ is formally the Hamiltonian for a Heisenberg antiferromagnet, since the coefficients of the $\mathbf{S} \cdot \mathbf{S}$ term are positive. This goes a long way toward justifying the assertion (often supported with much more qualitative arguments in the classroom) that the ground state of the Hubbard model is antiferromagnetic, since the Hubbard model itself presumes that one is not far from the atomic limit and therefore not far from the range of applicability of the perturbation theory.

This same reader, it is hoped, will be gratified at seeing a physical process which gives rise to at least part of the "exchange integral," that is, the coefficient of $\mathbf{S} \cdot \mathbf{S}$, a number usually arrived at empirically.

[1] J. Hubbard, Proc. R. Soc. A **276**, 238 (1963).

[2] Good treatments of this theory may be found in many textbooks on quantum theory, such as the following, for example: L. D. Landau and E. M. Lifshitz, Quantum Mechanics, Non-Relativistic Theory (Addison-Wesley, Reading, MA, 1965), 2nd ed., pp. 129–135; E. Merzbacher, Quantum Mechanics (Wiley, New York, 1970), 2nd ed., Chap. 17.

[3] P. W. Anderson, Solid State Phys. **14**, 166 (1963).