Electron correlations in copper-oxide layers and chains

1989 J. Phys.: Condens. Matter 1 1621

(http://iopscience.iop.org/0953-8984/1/9/008)

J. Phys.: Condens. Matter **1** (1989) 1621–1640. Printed in the UK

Electron correlations in copper-oxide layers and chains

J H Jefferson
Royal Signals and Radar Establishment, St Andrews Road, Malvern, Worcestershire WR14 3PS, UK

Received 28 March 1988, in final form 15 June 1988

Abstract. An extended, but orbitally non-degenerate, Hubbard Hamiltonian which models the CuO_2 planes in the high-T_c superconductors is considered. Effective Hamiltonians are derived for different parameter ranges leading to the identification of various phases and hole-pairing interaction terms which may be relevant to superconductivity mechanisms which are purely electronic in origin. The relationship with other models of electron correlations in the superconducting oxides is discussed and finally an extension of the model to include orbital degeneracy and crystal-field splitting is considered.

1. Introduction

Since the discovery of high-temperature superconductivity in doped La_2CuO_4 (Bednorz and Muller 1986) and $YBa_2Cu_3O_{9-\delta}$ (Chu *et al* 1987) there has been renewed speculation about a mechanism for superconductivity which is primarily electronic in origin and can give rise to higher transition temperatures than is possible with the usual electron–phonon interaction. Anderson (1987) has conjectured that the antiferromagnetic exchange interaction between electrons on neighbouring Cu^{2+} sites will favour singlet pairing and that these pairs will behave like a quantum spin liquid in the CuO_2 planes of La_2CuO_4. The insulating (singlet) ground state may be pictured as mobile singlet pairs with double-site occupancy precluded due to the large on-site Coulomb repulsion and is thus a superposition of different pairings of electrons, which has been called the 'resonating valence bond' (RVB) state (Anderson 1973), resembling the exact solution for the ground state of the linear antiferromagnetic chain (Bethe 1931). Doping La_2CuO_4 with Ba or Sr removes electrons from the Cu^{2+} sites allowing the singlet pairs to propagate as superconducting pairs. A spin-$\frac{1}{2}$ Hubbard Hamiltonian underlies the RVB state and a number of papers have appeared which develop the theory (Anderson 1987, Baskaran *et al* 1987, Zou and Anderson 1988) or propose similar mechanisms based on the Hubbard Hamiltonian and antiferromagnetic pairing (Cyrot 1986, 1987a, b). Stabilisation of the pairing due to lattice deformations has been considered by Kivelson *et al* (1987) and Hirsch (1987a) and the effects of nearest-neighbour Coulomb repulsions by Ruckenstein *et al* (1987).

A different mechanism has been proposed by Emery (1987), who argues that doping La_2CuO_4 will result in excess holes on the oxygen rather than the copper atoms and that these holes form superconducting pairs. Hirsch (1987b) uses Emery's model and again argues that this will lead to propagating hole-pairs in a strong-coupling limit. There have

been a number of photoemission and XPS experiments (Sarma *et al* 1987, Bianconi *et al* 1987a, b, Fuggle *et al* 1988) which support the view that holes reside on the oxygen (or, equivalently, that there is an absence of Cu^{3+}), though these results are somewhat controversial and we must await further experimental evidence. From the theoretical point of view, the resolution of this dichotomy is imperative.

In this paper we shall consider the consequences of an extended Hubbard model in which it will be assumed (in common with all the theoretical papers mentioned above) that the Coulomb repulsion between electrons on the same Cu site is somewhat larger than the hopping interaction with nearest-neighbour O sites. This assumption enables us to reduce the space in which the initial Hamiltonian operates to obtain effective Hamiltonians for the ground manifold (including the low-lying excitations), which are, of course, the only states relevant to superconductivity. The motivation for this work has been the following.

(i) To show that various d electron correlation models (referred to in the above-mentioned papers) emanate from the same basic (and rather simplified) model.

(ii) To derive rigorous expressions, using degenerate or quasi-degenerate perturbation theory, for effective Hamiltonians for a number of cases depending on the relative positions of the Cu (3d) and O (2p) energy levels.

(iii) To use these effective Hamiltonians as a basis for further work—both numerical (simulations) and analytical (e.g. mean-field theory)—and to explore the relevance of the model to high-T_c superconductivity.

(iv) To provide a framework for extensions of the model to include refinements such as crystal-field effects, lattice distortions etc.

It will be shown that there are, in addition to those previously considered, other interesting situations, such as the mixed valence regime, which may bear some relevance to the physics of superconductivity in these materials. We also identify various interaction terms, such as spin-dependent scattering, of the same order as the pairing terms, which may be important in inhibiting the tendency towards electronically driven superconductivity.

In § 2 the basic model is defined and the perturbation method outlined. This is followed by a treatment of the case of one d hole per Cu site which reduces to an antiferromagnetic spin-Hamiltonian for the Cu sites only, through superexchange via the O atoms. Although this well known result has been obtained many times in various ways (see, e.g. Anderson 1959) we include it here to demonstrate the perturbation method used throughout the paper and to point out that there is an extra contribution to the exchange interaction, usually ignored, which may be important for the superconducting oxides. We go on to show how this perturbation method may be used more generally to eliminate the O sites, provided the 2p energy levels are sufficiently low, yielding an effective Hamiltonain for the Cu holes only. It is pointed out that this gives the same result as would be obtained from the usual single-band Hubbard model *only in a certain limit* and it is by no means clear at the present time that this limit is appropriate. The various terms in the effective Hamiltonian are discussed, including hole–hole pairing in both the upper and lower Hubbard bands. It should be emphasised that when we speak of holes as residing either on the Cu sites or the O sites we are really referring to quasi-particles which, in the present framework, only have meaning in the sense of perturbation theory. The hopping-matrix element, t, is quite large ($\geqslant 1$ eV) and there will be significant hybridisation of the 'bare' 3d and 2p orbitals as evidenced by band

theory, i.e. mean-field theory (Mattheiss 1987, Mattheiss and Hamann 1987, Yu *et al* 1987).

However, with large intra-atomic Coulomb repulsions we can, in some circumstances, eliminate either the O or the Cu sites by the perturbation method (or a canonical transformation), yielding an effective Hamiltonian for quasi-holes located either on the Cu or the O sites (at least for a subset of the eigenstates of the original Hamiltonian, usually chosen to be the ground manifold). However, as will be demonstrated, this is not always possible and we are then in a truly mixed valence situation with the quasi-holes having significant amplitudes as a result of their being on either Cu or O sites. This mixed valence regime is discussed in § 6. In § 5 we consider the situation when the quasi-holes reside on the O sites, as conjectured by Emery (1987) and Hirsch (1987b), and in § 6 we discuss the limitations of the basic model and indicate how it may be extended to more realistic situations which will ultimately be necessary for reliable quantitative calculations.

2. The basic Hamiltonian and the perturbation method

It is now widely accepted that superconductivity in the oxide metals $La_{2-x}B_xCuO_4$ (B = Ca, Sr or Ba), $YBa_2Cu_3O_{9-\delta}$ and related compounds is largely confined to CuO_2 planes in the perovskite structure, which are widely separated along the c axis compared with the Cu–O bond lengths in the layers. Pure La_2CuO_4 is a semiconductor in which the Cu exists as Cu^{2+}, having donated its outer 4s and one of its 3d electrons to a neighbouring O site, thereby completing its 2p shell. The remaining electrons needed to fill all the 2p shells in the CuO_2 planes (and the O atoms in the La–O planes) are provided by the lanthanum which exists as La^{3+}. It is only energy levels derived from the CuO_2 planes (with possible contributions from the 2p orbitals of the O atoms in neighbouring La–O planes) that lie near the Fermi energy and are thus important for conduction. That La_2CuO_4 is *not* a conductor is believed to be due to the large intra-ionic repulsion of the d electrons on the Cu sites which ensures that each Cu ion has just one d hole. Doping with Ca, Ba or Sr, which are acceptors, injects holes into the CuO_2 layers. These holes can propagate and the material becomes a metal (albeit a poor one) at high temperatures and a superconductor at low temperatures. In stoichiometric $YBa_2Cu_3O_{9-\delta}$ the chemistry is such that extra holes already exist in the CuO_2 layers without further doping. A simple Hamiltonian which models the physics of this situation is the extended Hubbard Hamiltonian

$$H = H_0 + V \tag{1}$$

where

$$H_0 = \sum_i (\varepsilon_d n_i + U n_{i\uparrow} n_{i\downarrow}) + \varepsilon_p \sum_j n_j$$

and

$$V = t \sum_{\langle ij\rangle\sigma} (d^\dagger_{i\sigma} p_{j\sigma} + \text{HC}).$$

Here ε_d is the energy of a 3d hole on a Cu site, ε_p the energy of a 2p hole on an O site, t is the hopping matrix element between nearest-neighbour sites and U the intra-atomic Coulomb repulsion in the d shell (i.e. the energy of two d holes is $2\varepsilon_d + U$). n_i is the

number operator for d electrons at site i, i.e. $n_i = n_{i\uparrow} + n_{i\downarrow}$ where $n_{i\sigma} \equiv d^\dagger_{i\sigma} d_{i\sigma}$ and $d^\dagger_{i\sigma}$ is a creation operator for a hole with spin σ at site i. Similarly for the O orbitals, $n_j = n_{j\uparrow} + n_{j\downarrow}$ where $n_{j\sigma} \equiv p^\dagger_{j\sigma} p_{j\sigma}$ and $p^\dagger_{j\sigma}$ creates a spin-σ hole at site j. These localised orbitals are chosen to be orthonormal and the p and d operators thus satisfy the usual Fermi anticommutation rules. The sum $\langle ij \rangle$ in the last term is over nearest-neighbour Cu–O sites.

This model Hamiltonian is, of course, already grossly simplified. The fivefold orbital degeneracy of the d states and threefold degeneracy of the p states has been omitted as have many of the electron–electron interactions, both on the same site and between sites. The former interactions give rise to atomic multiplet structure, whereas the latter are responsible for crystal fields and additional hybridisation (hopping) terms. These interactions may be regarded as accounted for to some extent (in a mean-field or Hartree–Fock sense) by the renormalisation of the energy parameters in (1) (ε_d, ε_p, t and U). The neglect of orbital degeneracy will be discussed further in § 7. The simplified Hamiltonian (1) has the merit of being the simplest model Hamiltonian which contains both the dominant intra-ionic interaction (U) and electron transfer terms (V).

The 'unperturbed' Hamiltonian, H_0, is diagonal in the localised basis, i.e. its eigenvectors are completely determined by the occupation numbers of the Cu and O orbitals at the various sites and may be written $|n'_{1\uparrow}\, n'_{1\downarrow}\, n'_{2\uparrow}\, n'_{2\downarrow} \ldots\rangle$ where each $n'_{k\sigma} = 0$ or 1 for k ranging over all Cu and O sites. For all cases of interest there will be a set of lowest-lying states that are well separated in energy from the remaining states. This set of states will be referred to as the ground manifold of H_0 and the subspace of the total Fock space in which they lie, the model subspace. The ground manifold will be different in different situations (depending on the parameters ε_d, ε_p, U and t in equation (1)) and will be described in detail for each case in the following sections.

If we imagine the hopping interaction (V) being gradually switched on, then the ground manifold of H_0 will develop into exact eigenstates of H with accompanying energy shifts and lifting of degeneracy. It may be shown (see e.g., Lindgren and Morrison 1986) that these true eigenstates of H may be obtained from an effective Hamiltonian, H_{eff}, operating in the model subspace:

$$H_{\text{eff}} = PH\Omega \tag{2}$$

where P is the projection operator for all base states in the model subspace, i.e.

$$P = \sum_\alpha P_\alpha$$

where $P_\alpha = |\alpha\rangle\langle\alpha|$ with $\{|\alpha\rangle\}$ the set of all eigenstates of H_0 in the ground manifold. Ω is the 'wave operator' which may be expressed as a power series in the perturbation V, i.e.

$$\Omega = \Omega_0 + \Omega_1 + \Omega_2 + \Omega_3 + \ldots \tag{3}$$

where

$$\Omega_0 = P$$

$$\Omega_1 = \sum_\alpha R_\alpha V P_\alpha$$

$$\Omega_2 = \sum_\alpha R_\alpha V R_\alpha V P_\alpha - \sum_\alpha R_\alpha \Big(\sum_{\alpha'} R_{\alpha'} V P_{\alpha'} \Big) V P_\alpha$$

$$\Omega_3 = \sum_\alpha R_\alpha V R_\alpha V R_\alpha V P_\alpha - \sum_\alpha R_\alpha V R_\alpha \Big(\sum_{\alpha'} R_{\alpha'} V P_{\alpha'} \Big) V P_\alpha$$
$$- \sum_\alpha R_\alpha \Big(\sum_{\alpha'} R_{\alpha'} V P_{\alpha'} \Big) V R_\alpha V P_\alpha$$
$$- \sum_\alpha R_\alpha \Big(\sum_{\alpha'} R_{\alpha'} V R_{\alpha'} V P_{\alpha'} \Big) V P_\alpha$$
$$+ \sum_\alpha R_\alpha \Big[\sum_{\alpha'} R_{\alpha'} \Big(\sum_{\alpha''} R_{\alpha''} V P_{\alpha''} \Big) V P_{\alpha'} \Big] V P_\alpha$$

and

$$R_\alpha = (1 - P)/(E_\alpha - H_0)$$

with E_α satisfying $H_0|\alpha\rangle = E_\alpha|\alpha\rangle$. $(1 - P)$ is, of course, the projection operator for the high-energy (intermediate) states.

This is just a generalisation of ordinary Rayleigh–Schrödinger perturbation theory to the quasi-degenerate problem. It is to be preferred over other types of perturbation expansion, such as the Brillouin–Wigner method, since the effective Hamiltonian (2), to any order, does not depend on the energy of a particular eigenstate of H, nor does it contain spurious terms which are proportional to arbitrary powers of the volume (Thouless 1972). When H_0 is an independent-electron Hamiltonian, the expansion (3) may be simplified by expressing the series in terms of Feynman diagrams and it may be shown that only unlinked diagrams, including the so-called backwards or folded diagrams, are to be retained (Sandars 1969, Lindgren 1974, Brandow 1976). However, the H_0 of equation (1) is *not* an independent electron Hamiltonian because of the intra-atomic correlation term

$$U \sum_i n_{i\uparrow} n_{i\downarrow}$$

and the usual linked-diagram theorem cannot be invoked without some modification. We shall, therefore, deal with the Rayleigh–Schrödinger expansion directly. Fortunately this task is less formidable than it appears since many of the terms vanish.

The expansion (2) and (3) is formally exact, when taken to all orders, and will yield exact eigenvalues and eigenvectors of H, i.e. if $|\Phi_n\rangle$ is an eigenvector of H_{eff} satisfying

$$H_{\text{eff}}|\Phi_n\rangle = E_n|\Phi_n\rangle$$

then

$$H|\Psi_n\rangle = E_n|\Psi_n\rangle$$

where

$$|\Psi_n\rangle = \Omega|\Phi_n\rangle.$$

Furthermore, H_{eff} will yield the complete low-lying excitation spectrum of H (which comprises the only states relevant to superconductivity). In practice the power series expansion for H_{eff} may be truncated in low order since it is rapidly convergent. This implies that $t \ll \Delta E$, where ΔE is the *smallest* energy difference between a state $|\alpha\rangle$ in the model subspace and an intermediate state which has a non-zero overlap with $V|\alpha\rangle$. For the cases of interest, the model subspace can always be defined so that this condition is fulfilled.

It is emphasised that obtaining explicit expressions for H_{eff}, which is the main purpose of this paper, does not solve the problem (as it does in the non-degenerate case) but merely re-expresses it in a form in which high-energy states have been eliminated in the sense of perturbation theory. However, apart from obtaining effective Hamiltonians which operate in a smaller space (with obvious advantages for numerical work), one hopes that the new forms of the effective Hamiltonians will give greater insight into low-energy solutions by providing alternative descriptions of the true ground manifold, as has been the case for magnetic problems (see, e.g., Stevens 1976).

We conclude this section by pointing out that the basic Hamiltonian (1) may itself be regarded as an effective Hamiltonian since it is incapable of yielding very high energy states, such as those which would result from exciting core electrons. We can in principle, of course, account for very high energy states by the perturbation method just outlined. One effect of this would be to further renormalise the parameters t, U, ε_p and ε_d from what they would be if H included such very high energy states explicitly. It will be assumed that this has been done and in this sense our basic Hamiltonian already describes quasi-holes.

3. One hole per Cu site—the antiferromagnetic spin Hamiltonian

As a simple application of the perturbation method described in the previous section, we consider briefly the case where there is just one hole per Cu site. A similar treatment for electrons, leading to the usual antiferromagnetic spin-Hamiltonian but with an enhanced exchange constant, has been presented elsewhere (Jefferson 1988a).

In stoichiometric La_2CuO_4 the CuO_2 planes have an average of one hole per CuO_2 cell. Furthermore, it will be assumed that the O p orbitals are somewhat deeper than the Cu d orbitals, i.e. for hole energies, $\varepsilon_p > \varepsilon_d$ with $\varepsilon_p - \varepsilon_d \gg t$. Hence, from (1), the ground manifold of $H_0 \equiv H(t = 0)$ for N cells will consist of 2^N states each with energy $N\varepsilon_d$, corresponding to one hole per Cu site with either spin up or spin down. Since odd powers of the perturbation V (equation (1)) can never connect states in the degenerate manifold the effective Hamiltonian to fourth order is, from (2) and (3),

$$H_{\text{eff}} = P(E_0 + VR_0V + VR_0VR_0VR_0V - VR_0^2VPVR_0V)P \tag{4}$$

where $R_0 = (1 - P)/(E_0 - H_0)$ is the same for all (degenerate) states in the model subspace, enabling the summations over α to be performed. Equation (4) is in fact the degenerate Rayleigh–Schrödinger result first obtained by Bloch (1958). The second-order terms merely renormalise the energy of the ground manifold and

$$P(E_0 + VR_0V)P = \tilde{E}_0P$$

where

$$\tilde{E}_0 = N\left(\varepsilon_d - \frac{\nu t^2}{\varepsilon}\right) \qquad \text{and} \qquad \varepsilon = \varepsilon_p - \varepsilon_d$$

with ν the number of nearest-neighbour O sites (2 for a linear chain, 4 for CuO_2). This constant term may be dropped since it is the same for all states in the ground manifold. There is some mutual cancellation between the fourth-order terms in (4) rather like that which takes place in the linked-diagram theorem (when H_0 is an independent-electron

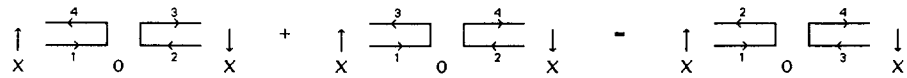


Figure 1. Cancellation of singlet processes in $PVR_0VR_0VR_0VP$ (LHS) and $PVR_0^2VPVP_0VP$ (RHS). The numbers refer to ordering of the hops. ×, copper, ○, oxygen.

Hamiltonian). This cancellation is shown in figure 1 and occurs only for singlet processes which do not flip spins. The residual processes can be written in terms of spin operators in just the same way as for electrons (Jefferson 1988a) and

$$H_{\text{eff}} = J_1 \sum_{\langle ii' \rangle} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} - \tfrac{1}{4}) + J_2 \sum_{\langle ii' \rangle} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} + \tfrac{1}{4}) \tag{5}$$

where

$$J_1 = 2t^4/\varepsilon^2 U \qquad \text{and} \qquad J_2 = 2t^4/\varepsilon^3.$$

The first term in (5) arises from intermediate states in which a hole has been transferred from a Cu ion to a neighbouring O ion and was first considered by Anderson (1959). The interaction between neighbouring Cu spins is attractive when they are antiparallel. On the other hand, the second term in (5) arises from intermediate states in which two holes are on the same O atom and is a repulsive interaction between neighbouring Cu spins when they are parallel. Since there is exactly one hole per Cu site the constant terms may be dropped and

$$H_{\text{eff}} = J \sum_{\langle ii' \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} \tag{6}$$

where $J = J_1 + J_2$. Thus the distinction between attractive singlet and repulsive triplet interactions is unimportant and the overall effect is an enhancement of the antiferromagnetic exchange constant. However, as we show later, the distinction *is* important when there is more than one hole per Cu site and there is charge transport.

The second term in (5), not usually considered explicitly, is negligible when $\varepsilon \gg U$. At the present time it is not clear that this is the situation in the superconducting oxides. Indeed, if the O 2p levels are sufficiently high in energy for doping to create holes on the O rather than the Cu sites, as has been speculated, then intermediate states with holes on O sites give the *largest* contribution to the exchange interaction.

4. Elimination of the oxygen

In this section we consider what happens in the metallic case when extra holes are introduced into the insulating CuO_2 planes described in the previous section. It is shown that provided the O p levels are sufficiently deep, the holes will be confined largely to Cu sites and the O sites may be completely eliminated from the problem by the basic perturbation method described in § 2. Although the resulting effective Hamiltonian only involves Cu sites, it is shown that it cannot be obtained from a single-band Hubbard model (i.e. a Hubbard model for Cu sites only), except in a certain limit which is unlikely to be a reasonable approximation in practice.

Doping La_2CuO_4 with Sr or Ba, which are acceptors, has the effect of introducing holes into the CuO_2 planes, the mean hole occupation per cell becomes greater than unity and the material becomes a conductor. In the $YBa_2Cu_3O_{9-\delta}$ compounds the

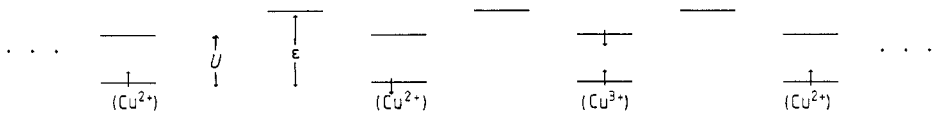


Figure 2. Energy diagram of a typical base state in the ground manifold of a linear chain for which the O hole orbitals are unoccupied.

mean hole occupation of the CuO_2 planes exceeds unity without further doping. If $\varepsilon_d + U < \varepsilon_p$ then, according to (1), the ground manifold of H_0 will consist of states in which the O sites are unoccupied and the Cu sites are either singly or doubly occupied with holes. If x is the degree of doping in $La_{2-x}B_xCuO_4$ (B = Ca, Sr or Ba) then $N(1-x)$ sites will be singly occupied and Nx doubly occupied. There are thus $2^{N(1-x)}\,{}^{N}C_{Nx}$ states in the ground manifold each with energy $N[\varepsilon_d + x(\varepsilon_d + U)]$ (states with holes on the O sites are, of course, higher in energy). These base states may be represented by energy diagrams in which each Cu site has two energy levels at ε_d (occupied by the 'first' hole) and $\varepsilon_d + U$ (occupied by the 'second' hole). Such diagrams, first used by Anderson (1961) for the magnetic impurity problem, correctly reproduce the two-hole energy, $2\varepsilon_d + U$, including intra-ionic coulomb repulsion. A typical energy diagram for a linear chain is shown in figure 2 for a state in the ground manifold. Note that O hole levels are all unoccupied (since $\varepsilon > U$) with the Cu levels either singly or doubly occupied.

If the hopping interaction (V) is switched on, the degenerate manifold of states will develop into eigenstates of H which, according to the results of § 2, may be described by an effective Hamiltonian. Since the ground manifold of H_0 is exactly degenerate, the effective Hamiltonian to fourth-order is given by (4) with $E_0 = N[\varepsilon_d + x(\varepsilon_d + U)]$. After some manipulation, the effective Hamiltonian may be written in the form (dropping a constant, renormalised energy term):

$$\begin{aligned} H_{\text{eff}} = {} & -t_0 \sum_{\langle ii'\rangle\sigma} d^\dagger_{i\sigma} d_{i'\sigma} \delta_{12} \\ & + \sum_{\langle ii'\rangle} \{[J_1(\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} - \tfrac{1}{4}) + J_2(\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} + \tfrac{1}{4})]\delta_{11} + \Delta_1 \delta_{12} + \Delta_2 \delta_{22}\} \\ & + \sum_{\langle iki'\rangle\sigma} [(t_1 d^\dagger_{i\sigma} n_{k\sigma} d_{i'\sigma} + t_2 d^\dagger_{i\sigma} d^\dagger_{k,-\sigma} d_{k\sigma} d_{i',-\sigma} \\ & + t_3 d^\dagger_{i\sigma} n_{k,-\sigma} d_{i'\sigma}) \delta_{112} + t_4 d^\dagger_{i\sigma} d_{i'\sigma} \delta_{122}] \end{aligned} \tag{7}$$

where

$$t_0 = t^2/(\varepsilon - U) + O(t^4) \qquad J_1 = 2t^4/\varepsilon^2 U \qquad J_2 = 2t^4/\varepsilon^3$$

$$\Delta_1 = [t^4/\varepsilon(\varepsilon - U)][1/\varepsilon + 1/(\varepsilon - U)] \qquad \Delta_2 = 2t^4/(\varepsilon - U)^3$$

$$t_1 = [t^4/\varepsilon(\varepsilon - U)](1/U + 1/\varepsilon) \qquad t_2 = t^4/\varepsilon(\varepsilon - U)U$$

$$t_3 = t^4/\varepsilon^2(\varepsilon - U) \qquad t_4 = t^4/(\varepsilon - U)^3 = \Delta_2/2.$$

The δ-factors ensure that neighbouring Cu sites have a specific hole occupation. For example, δ_{112} is unity when sites i and k are singly occupied and site i' doubly

occupied, and is zero otherwise. It can be written formally in terms of number operators, i.e.

$$\delta_{112} \equiv \delta_{n_i:1}\delta_{n_k:1}\delta_{n_{i'}:2}$$

where

$$\delta_{n:1} \equiv n(2-n) \equiv n - 2n_\uparrow n_\downarrow$$

and

$$\delta_{n:2} \equiv \tfrac{1}{2}n(n-1) \equiv n_\uparrow n_\downarrow .$$

Similar equivalences follow straightforwardly for the other δ-operators in (7). It should be noted that H_{eff} is indeed an operator which lies in the low-energy subspace in which the O sites are unoccupied (and therefore disappear from the problem) and the Cu sites are never unoccupied. The ratio of two-hole to one-hole sites (the degree of doping, x, in $La_{2-x}B_xCuO_4$) remains fixed, the hopping terms merely causing rearrangement. It is to be noted also that H_{eff} is Hermitian despite the fact that this is not generally the case with degenerate Rayleigh–Schrödinger perturbation theory (see, e.g., Lindgren 1974). Its Hermiticity is readily proved, term by term, by writing the δ-operators in a more symmetric form. For example,

$$\sum_{\langle ii'\rangle} d_{i\sigma}^\dagger d_{i'\sigma}\delta_{12} = \sum_{\langle ii'\rangle} d_{i\sigma}^\dagger d_{i'\sigma}\delta_{n_i:1}\delta_{n_{i'}:2} = \sum_{\langle ii'\rangle} \delta_{n_i:2} d_{i\sigma}^\dagger d_{i'\sigma}\delta_{n_{i'}:2}$$

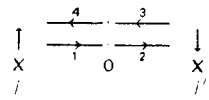
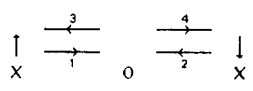
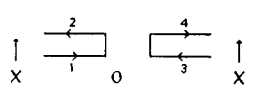
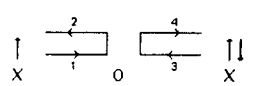
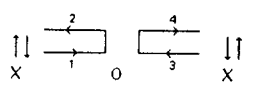
which is obviously Hermitian.

Let us consider the various terms in (7) and, briefly, the processes in the perturbation expansion which gave rise to them. The first term (the only second-order term apart from energy renormalisations which have been dropped) represents hopping from a doubly occupied Cu site to a neighbouring singly occupied site with effective hopping, $t^2/(\varepsilon - U)$. In fourth order, there are a number of important differences when compared with an equivalent second-order perturbation treatment applied to the simpler Cu-only Hubbard model (see below). In particular, extra terms appear which *cannot* be obtained unless intermediate states with holes on the O site are treated explicitly. These are the terms with coefficients J_2, Δ_1, Δ_2, t_3 and t_4 in (7). The antiferromagnetic exchange terms, J_1 and J_2, were considered in § 3. For that insulating case the J_2 term merely enhanced the exchange and the constant factors $\pm\frac{1}{4}$ were dropped. This cannot be done when the holes are able to move because of the occupation factors, δ_{11}. Neighbouring holes which are antiferromagnetically aligned are bound with energy $J_1 + J_2/2$ whereas ferromagnetically aligned pairs have a repulsive energy $J_2/2$. This result still applies when doping produces holes on the O site (§ 5) or in the mixed valence phase (§ 6) since it only involves singly occupied sites.

Some care is needed in deriving the fourth-order terms since there is much cancellation, analogous to that which takes place in the linked-cluster theorem for independent fermions. An example of such a cancellation for singlet-exchange processes was shown in figure 1. Tables 1 and 2 show the surviving fourth-order processes after all such cancellations have been accounted for. As a general rule, all processes in $PVR_0^2VPVR_0VP$ are cancelled by processes in $PVR_0VR_0VR_0VP$, *except* when they are prevented from doing so by the exclusion principle.

The diagonal terms $\Delta_1\delta_{12}$ and $\Delta_2\delta_{22}$ represent repulsions between holes and hole pairs on adjacent sites. The latter interaction may be regarded as a repulsion between

Table 1. Allowed diagonal and spin-flip processes (fourth-order) which are not cancelled.

Term in H_{eff} (equation (10))	Process	Comments
$J_1(S_i \cdot S_{i'} - \frac{1}{4})\delta_{11}$		$\in PVR_0VR_0VR_0VP$†
$J_2(S_i \cdot S_{i'} + \frac{1}{4})\delta_{11}$		$\in PVR_0VR_0VR_0VP$‡
		$\in PVR_0^2VPVR_0VP$§
$\Delta_1\delta_{12}$		$\in \text{PVR}_0^2VPVR\ VP$‖
$\Delta_2\delta_{22}$		

† Spin flip and non-spin flip.
‡ Spin flip only.
§ Holes must have the same spin (triplet repulsion).
‖ Both holes transferred must have the same spin.

'extra' holes. The remaining three-site hopping terms involve transferring a hole to a next-nearest-neighbour Cu site with and without spin flips. Note that the last term (see also table 2) may be regarded as the simultaneous movement of two hole pairs on adjacent sites, which favours the pairing of 'extra' holes, in contrast to the Δ_2 term. Such competition will be important if the O levels are not too deep, i.e. if $\varepsilon - U \ll U$. Retaining only the largest terms in (7) gives

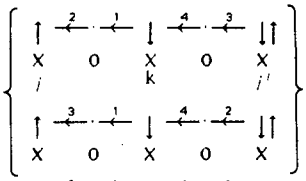
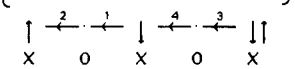
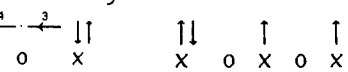
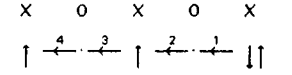
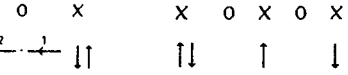
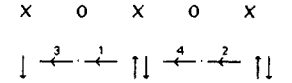
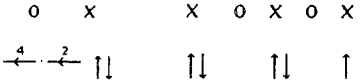
$$H_{\text{eff}} = -\frac{t^2}{(\varepsilon - U)} \sum_{\langle ii'\rangle\sigma} d_{i\sigma}^{\dagger} d_{i'\sigma} \delta_{12} + \frac{2t^4}{(\varepsilon - U)^3} \sum_{\langle ii'\rangle} \delta_{22} + \frac{t^4}{(\varepsilon - U)^3} \sum_{\langle iki'\rangle\sigma} d_{i\sigma}^{\dagger} d_{i'\sigma} \delta_{122}. \tag{8}$$

This is, formally, the infinite correlation limit ($\varepsilon, U = \infty$ with $\varepsilon - U$ finite) for which the Cu sites can never be unoccupied, even as intermediate states. The eigenstates of this Hamiltonian have been obtained exactly for the case of one 'extra' hole and numerically for the case of two extra holes (Jefferson 1988b). These solutions show that there is no binding between the extra holes and this result applies even when $t > |\varepsilon - U|$ for which the perturbation theory of this section breaks down (see § 6).

Finally, consider the single-band Hubbard model for the Cu site only, which has been used as the basic Hamiltonian for CuO_2 layers by a number of authors. Applying the basic perturbation method to the Hubbard Hamiltonian

$$\bar{H} = \bar{\varepsilon}_{\text{d}} \sum_{i\sigma} n_{i\sigma} + \bar{U} \sum_{i} n_{i\uparrow} n_{i\downarrow} - \bar{t} \sum_{\langle ii'\rangle\sigma} d_{i\sigma}^{\dagger} d_{i'\sigma} \tag{9}$$

Table 2. Allowed fourth-order, next-nearest-neighbour hopping processes which are not cancelled.

Term in H_{eff}	Initial state	Final state	Comments
$t_1 d_{i\sigma}^{\dagger} n_{k\sigma} d_{i'\sigma} \delta_{112}$			$\in PVR_0VR_0VR_0VP$†
$t_2 d_{i\sigma}^{\dagger} d_{k-\sigma}^{\dagger} d_{k\sigma} d_{i'\sigma} \delta_{112}$			$\in PVR_0VR_0VR_0VP$‡
$t_3 d_{i\sigma}^{\dagger} n_{k-\sigma} d_{i'\sigma} \delta_{112}$			$\in PVR_0^2VPVR_0VP$§
$t_4 d_{i\sigma}^{\dagger} d_{i'\sigma} \delta_{122}$			$\in PVR_0VR_0VR_0VP$‖

† Non-spin flip.
‡ Spin flip.
§ Non-spin flip. Sites i and k must have the same spin in the initial state.
‖ Hole pair–pair hopping.

yields the following second-order effective Hamiltonian operating in the ground manifold (see also Hirsch 1985):

$$\bar{H}_{\text{eff}} = -\bar{t} \sum_{\langle ii' \rangle \sigma} d_{i\sigma}^{\dagger} d_{i'\sigma} \delta_{12} + \frac{2\bar{t}^2}{\bar{U}} \sum_{\langle ii' \rangle} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} - \tfrac{1}{4}) \delta_{11}$$
$$+ \frac{\bar{t}^2}{\bar{U}} \sum_{\langle iki' \rangle \sigma} (d_{i\sigma}^{\dagger} n_{k\sigma} d_{i'\sigma} + d_{i\sigma}^{\dagger} d_{k,-\sigma}^{\dagger} d_{k\sigma} d_{i',-\sigma}) \delta_{112}. \tag{10}$$

Thus, comparing with (7), we see that a number of important terms are missing and the form (10) (and hence the Hubbard Hamiltonian which gave rise to it) are inadequate for describing the low-lying excitations. Note that for $\varepsilon \gg U$, (10) and (7) are approximately equivalent, with $\bar{t} = t^2/\varepsilon$ and $\bar{U} = U$, reflecting the fact that in this limit the basic Hamiltonian (1) and the Hubbard Hamiltonian (9) are equivalent on energy scales $\ll \varepsilon$. This limit is, however, unlikely to be relevant to the cases of interest for which ε and U are of comparable magnitudes.

In summary, the main result of this section is that for situations in which the holes are confined mainly to the Cu sites with $t < \varepsilon - U$, the oxygen may be eliminated from the problem and the Hamiltonians (1) and (7) have the same low-energy spectrum to fourth order.

5. Hole transport on the oxygen

As pointed out by Emery (1987), if the O 2p orbitals are sufficiently high in energy then doping La_2CuO_4 will inject holes on to the O sites. This will occur if $\varepsilon < U$.

Furthermore, if $U - \varepsilon < t$ the Cu-charge degrees of freedom (hole transport) may be eliminated from the problem by perturbation theory, just as the oxygen was in the previous section. From (1) we see that the degenerate ground manifold of H_0 has energy $E_0 = N(\varepsilon_d + x\varepsilon_p)$ with the nearest excited states being ε or $U - \varepsilon$ higher in energy. The ground manifold consists of all states in which each Cu site has just one (immobile) hole with the remaining Nx holes distributed amongst the O sites. The degeneracy is very large (since the Cu sites and singly occupied O sites can have either spin), though we do not need to know it explicitly in order to derive the effective Hamiltonian which is, to second order, (cf (4))

$$H_{\text{eff}} = P(E_0 + VR_0V)P. \tag{11}$$

Substituting V from equation (1) and dropping a constant energy term, H_{eff} may be written in the form

$$H_{\text{eff}} = -\nu\tau_1 \sum_j \delta_{n_j:0} + 2\tau \sum_{\langle ij\rangle} (\boldsymbol{S}_i \cdot \boldsymbol{s}_j - \tfrac{1}{4})\delta_{n_j:1} - \nu\tau_2 \sum_j \delta_{n_j:2}$$
$$+ \sum_{\langle ij'\rangle\sigma} (\tau_1 p^\dagger_{j\sigma} n_{i\sigma} p_{j'\sigma} - \tau_2 p^\dagger_{j\sigma} n_{i,-\sigma} p_{j'\sigma} + \tau p^\dagger_{j\sigma} d^\dagger_{i,-\sigma} d_{i\sigma} p_{j,-\sigma}) \tag{12}$$

where

$$\tau_1 = t^2/\varepsilon \qquad \tau_2 = t^2/(U - \varepsilon) \qquad \tau = \tau_1 + \tau_2$$

and ν is the number of nearest-neighbour Cu sites, which is two for both CuO chains and CuO_2 planes. It is emphasised that H_{eff} operates in the ground manifold of H_0 for which all the Cu sites are always singly occupied. It thus describes the interaction between a system of fixed spins (the Cu sites) and a system of mobile holes (on the O sites) which are scattered and induce spin flips. Examples of the processes in equation (12) are shown in figure 3. The energy renormalisations in (12) (first and third terms) are not very important since for low doping they give an approximately constant

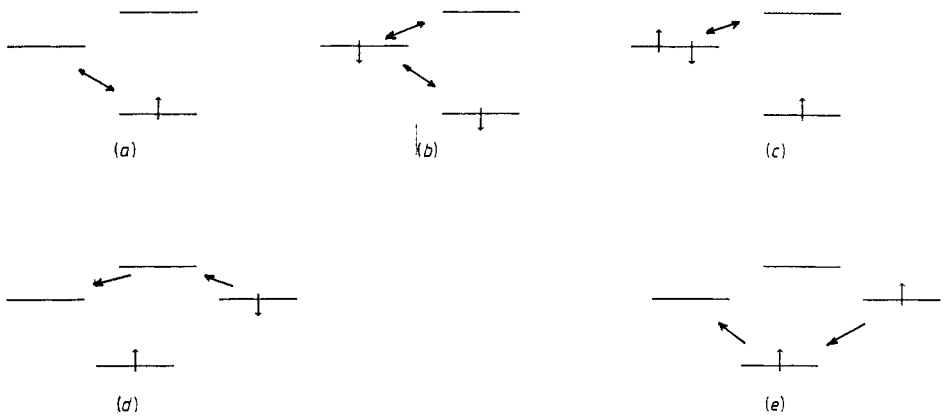


Figure 3. Second-order processes: (*a*) energy renormalisation; (*b*) Cu–O antiferromagnetic exchange; (*c*) energy renormalisation; (*d*) O hole hopping (depends on the spin of the intervening Cu site which may flip in the process); (*e*) further O hole hopping (hole transferred from Cu site first).

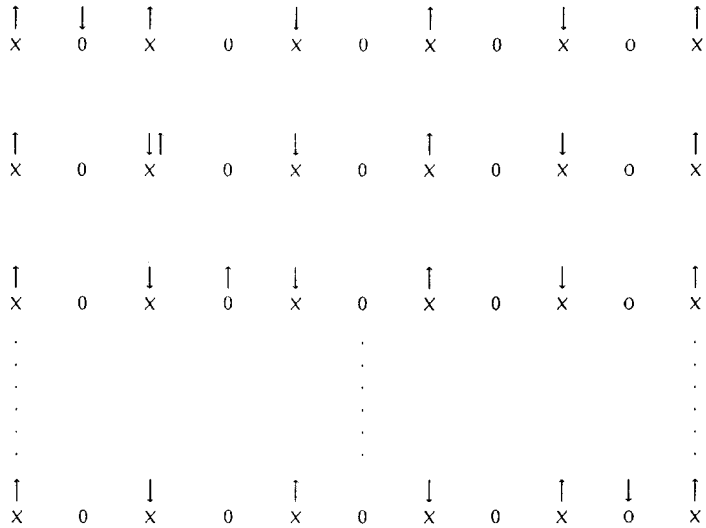


Figure 4. Hole propagation which preserves ferromagnetic alignment of nearest neighbours.

contribution given that the number of doubly occupied O sites is very small†. It is to be noted that some of the remaining terms of the same order of magnitude survive in the infinite correlation limit ($\varepsilon \to \infty$, $U - \varepsilon$ finite). It has been suggested that the antiferromagnetic Cu–O exchange (second term) should induce a *ferromagnetic* interaction between neighbouring Cu spins (Aharony *et al* 1988). It is not immediately apparent that these ferromagnetic correlations will survive under under the influence of the hopping interaction (last term) which can induce spin flips. However, the exact result presented in § 6 shows that, at least for a linear chain, an O hole *will* carry with it a predominantly ferromagnetic alignment of its nearest-neighbour Cu sites as it propagates. The mechanism of this process is shown in figure 4. The hole propagates from left to right by a series of hops in which it is always sandwiched between two Cu sites of opposite spin whenever it occupies an O site. Note also that the pattern of Cu spins to the right of the O hole (antiferromagnetically ordered in figure 4) shifts one place to the left as the hole propagates.

As one might expect, these second-order processes dominate the hole transport giving rise to a low-lying band of width $\sim t^2/(U - \varepsilon)$ (see § 6). However, certain fourth-order processes are still important, particularly when the number of O holes is small (low doping), since they largely determine the behaviour of the majority of the Cu spins, which have neighbouring O sites that are unoccupied. As in §§ 3 and 4 they give rise to antiferromagnetic exchange terms in the effective Hamiltonian which may be written (cf (5) and (7)):

$$H_{\mathrm{ex}} = \sum_{\langle iji' \rangle} [J_1(\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} - \tfrac{1}{4}) + J_2(\boldsymbol{S}_i \cdot \boldsymbol{S}_{i'} + \tfrac{1}{4})]\delta_{101} \tag{13}$$

† Double occupation will also be inhibited by the Coulomb repulsion between holes on the same O site, which was neglected in (1). (This is offset to some extent by direct O–O hopping, which favours double occupation. However, estimates of the relative magnitudes give $t_{\mathrm{O-O}} \ll U_{\mathrm{p}}$ and there is little doubt that the amplitudes for double occupation are small.) In addition, it is shown in § 6 that the holes repel, reducing further the probability of double occupation.

where $J_1 = 2t^4/\varepsilon^2 U$ and $J_2 = 2t^4/\varepsilon^3$, as before, and $\delta_{101} \equiv \delta_{n_i:1}\delta_{n_j:0}\delta_{n_{i'}:1}$ ensures that the Cu sites are singly occupied with the intervening O site unoccupied. Thus we still expect antiferromagnetic pairing of Cu holes which are able to move in unison as the O hole propagates (see figure 4). In this sense one can still envision an RVB state of pre-existing pairs as conjectured by Anderson (1987) for the case when the holes are confined to the Cu sites.

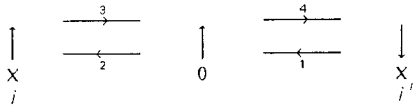


Figure 5. Singlet process with intervening O site occupied. Note that Cu spin flips are not allowed.

Hirsch (1987b) has pointed out that in two dimensions a hole will break antiferromagnetic bonds as it propagates and argues that this will favour the pairing of holes on O sites, since such a pair can propagate without leaving broken bonds behind. It should be noted, however, that this effective attraction will have to overcome effective repulsions between nearest-neighbour O holes, $\sim t^4/(U-\varepsilon)^3$, in addition to the (second-order) kinetic-energy gains in keeping the holes far apart. These effects would seem to mitigate against hole pairing and this is further supported by numerical solutions for the two-hole case (Jefferson 1988b and § 6) which show that, for a linear chain in the infinite correlation limit ($\varepsilon, U \to \infty$, $U - \varepsilon$ finite), the holes do indeed repel. It is not essential, of course, to have bound pairs of O holes in order to circumvent the problem of broken bonds since the Cu spins can rearrange themselves under the influence of spin flips contained in (13) when a single hole propagates.

In addition to (13) there are numerous fourth-order processes which involve O sites that are occupied. For example, the singlet process shown at the top of figure 3 is still allowed when the intervening O site is singly occupied but spin flips must be excluded. This process, shown in figure 5, can only take place in the direction $i' \to i$ giving a contribution to the effective Hamiltonian is thus (cf (12)):

$$\frac{J_1}{2}(S_l^{(z)}S_l^{(z)} - \tfrac{1}{4})\delta_{111}.$$

This, and the remaining fourth-order processes, are not expected to be very important since they will be swamped by the second-order processes in (12). Hence we will not consider them further here, though their derivation would be relatively straightforward (but tedious!) were this necessary.

Neglecting double occupancy, the final effective Hamiltonian for hole transport on the oxygen becomes:

$$H_{\text{eff}} = \sum_{\langle jij'\rangle\sigma} (\tau_1 p_{j\sigma}^\dagger n_{i\sigma} p_{j'\sigma} - \tau_2 p_{j\sigma}^\dagger n_{i,-\sigma} p_{j'\sigma} + \tau p_{j\sigma}^\dagger d_{i,-\sigma}^\dagger d_{i\sigma} p_{j'\sigma})$$
$$+ 2\tau \sum_{\langle ij\rangle} (\boldsymbol{S}_i \cdot \boldsymbol{s}_j - \tfrac{1}{4})\delta_{n_j:1} + H_{\text{ex}} \qquad (14)$$

where H_{ex} is given by (13).

6. The mixed valence case

When the hopping interaction, t, is comparable to (or greater than) the energy to transfer a hole between a doubly occupied Cu site and an O site, then the perturbation expansions of §§ 4 and 5 break down and we cannot regard the (extra) holes as being confined to either Cu or O sites. This is a truly mixed valence situation and the low-lying eigenstates of H (equation (1)) will be coherent superpositions of basis states for which the amplitudes of hole occupation on the Cu and O sites are comparable. However, although $t \geqslant |\varepsilon - U|$ (see figure 2), the energy required to remove the hole from a singly occupied Cu ion (i.e. to create Cu^+ from Cu^{2+}) will still be large, i.e. $t \ll \varepsilon, U$. The ground manifold of H_0 will not be exactly degenerate (except when $\varepsilon = U$) but will consist of all states in which all Cu sites have one hole with the remaining Nx holes distributed among Cu and O sites. The energy of a state with $N\delta$ 'extra' holes on the Cu sites and thus $N(x-\delta)$ holes on the O sites is, from (1), $E_0(\delta) = N[\varepsilon_d + \delta(\varepsilon_d + U) + (x-\delta)\varepsilon_p]$ with $0 \leqslant \delta \leqslant x$. There will, of course, be many states with energy $E(\delta)$ (δ fixed) due to spin degeneracy and the various ways in which the extra holes may be rearranged. The energy levels $E(\delta)$, for all δ, form a uniformally spaced set with separation $|\varepsilon - U|$ and the set of all corresponding states is what constitutes our ground manifold, or model subspace, for this intermediate valence situation. The total number of states in this set (dimension of the model subspace) is extremely large but all states must be included in order that the perturbation expansion for H_{eff} be convergent (see below). (If, for example, we took only the degenerate states with the lowest energy the perturbation expansion would contain terms with coefficients $\sim t^{n+1}/|\varepsilon - U|^n$ and hence be divergent since, by assumption $t \geqslant |\varepsilon - U|$.) It is not necessary to know the size of the model subspace in order to obtain an effective Hamiltonian which may be derived using the more general 'quasi-degenerate' perturbation theory of § 2 (cf (2) and (3)). The effective Hamiltonian to fourth order is

$$H_{\text{eff}} = t\sum_{\langle ij\rangle\sigma}(d_{i\sigma}^{\dagger}p_{j\sigma}\delta_{n_i:1} + \text{HC}) + (\varepsilon - U)\sum_j n_j + \frac{2t^2}{\varepsilon}\sum_{\langle ij\rangle}(\boldsymbol{S}_i\cdot\boldsymbol{s}_j - \tfrac{1}{4})\delta_{n_i:1}\delta_{n_j:1} + \frac{t^2}{\varepsilon}\sum_{\langle jij'\rangle\sigma}(p_{j\sigma}^{\dagger}n_{i\sigma}p_{j'\sigma} + p_{j\sigma}^{\dagger}d_{i,-\sigma}^{\dagger}d_{i\sigma}p_{j',-\sigma})\delta_{n_i:1} + H_{\text{ex}} \tag{15}$$

where H_{ex} is given by (13). Again we have retained only the terms up to second-order which involve occupied O sites in the model subspace. In a sense, equation (15) is more general than the cases considered in the previous sections since it also applies to those cases where the holes are confined to either the Cu or O sites. In fact the perturbation method may be applied again to (15) for the cases $t \ll |\varepsilon - U|$ to recover the previous results ((7) and (12)).

If we neglect the second- and fourth-order terms in (15) (the infinite correlation limit), then the problem of a single (extra) hole may be solved exactly and the two-hole case numerically in one dimension (Jefferson 1988b). These results show that there are three bands of single-hole energies:

$$E_0(k) = \Delta \qquad \text{(dispersionless)}$$

and

$$E_{\pm}(k) = \tfrac{1}{2}\Delta \pm \tfrac{1}{2}[\Delta^2 + 16t^2(1 + \tfrac{1}{2}\cos 2\pi k/N)]^{1/2} \tag{16}$$

where $\Delta = \varepsilon - U$. Thus, in the mixed valence case with $|\Delta| \ll t$ the band width of the ground manifold is $\sim t$, whereas, for $t \ll |\Delta|$, equation (16) may be expanded to give for the lowest band

$$E_{-}(k) \simeq -(4t^2/|\Delta|)(1 + \tfrac{1}{2}\cos 2\pi k/N) + \tfrac{1}{2}(\Delta - |\Delta|)$$

which are the same as tight-binding results for non-interacting fermions. For the case $\Delta \gg t$ (hole on Cu), it may be shown that (15) is approximately equivalent to a Hubbard model for Cu only in the large-U limit (Cu sites never unoccupied). This problem has been solved exactly (Klein and Seitz 1979) and is indeed equivalent to a non-interacting gas of spinless fermions. Thus E_{-} (above) is a limiting case of this situation.

The eigenstates with energy $E_{\pm}(k)$ are

$$|\Psi_{\pm}(k)\rangle = A\left(|\mathrm{I}k\rangle + \frac{t[1 + \exp(\mathrm{i}2\pi k/N)]}{E_{\pm} - \Delta}|\mathrm{II}k\rangle + \frac{t\sqrt{2}}{E_{\pm} - \Delta}|\mathrm{III}k\rangle\right) \tag{17}$$

where A is a normalising constant. The state $|\mathrm{I}k\rangle$ is a running wave with the extra hole on the Cu sites, i.e.

$$|\mathrm{I}k\rangle = N^{-1/2}\sum_{n}\exp(\mathrm{i}2\pi kn/N)|1n\rangle$$

where $|1n\rangle$ is a base state with the extra hole on Cu site n. Similarly, states $|\mathrm{II}k\rangle$ and $|\mathrm{III}k\rangle$ are running waves with the extra hole on the oxygen:

$$|\mathrm{II}k\rangle = N^{-1/2}\sum_{j}\exp(\mathrm{i}2\pi kj/N)|2j\rangle$$

where $|2j\rangle$ is a base state with the extra hole on oxygen site j, having its spin antiparallel to neighbouring Cu spins (figure 4, top);

$$|\mathrm{III}k\rangle = N^{-1/2}\sum_{j}\exp(\mathrm{i}2\pi kj/N)[|3j\rangle + |4j-1\rangle]$$

where $|3j\rangle$ is the base state with the extra hole on site j having its spin parallel to the Cu spin to its right and antiparallel to the Cu spin to its left, and vice versa for state $|4j\rangle$.

We see from (16) and (17) that, in the mixed valence regime ($|\Delta| \ll t$), the amplitudes for cases where there is the extra hole on the Cu and O sites are comparable, as expected. Conversely, when $t \ll |\Delta|$ the hole will be either on the Cu sites (Δ positive) or on the O sites (Δ negative), i.e.

$$|\Psi_{-}\rangle \rightarrow |\mathrm{I}k\rangle$$

and

$$|\Psi_{+}\rangle \rightarrow \text{constant}\,(|\mathrm{II}k\rangle + a(k)|\mathrm{III}k\rangle)$$

where

$$a(k) = \{2\sqrt{2}[1 + \exp(\mathrm{i}2\pi k/N)]\}^{-1}.$$

Hence, when the hole is confined to the O sites it is predominantly in the state $|\mathrm{II}k\rangle$ (with probability ~ 0.97 for small k) and the antiferromagnetic alignment with

neighbouring Cu spins is preserved as the hole propagates, as stated in the previous section.

Finally let us point out that although (16) reduces to the tight-binding results for $t \ll |\Delta|$, implying the non-interacting fermion result referred to earlier, the extra holes surely do interact when corrections due to finite t/Δ are accounted for. This is supported by the perturbation theory of § 4 which gave terms in the effective Hamiltonian corresponding to hole–hole repulsion and hole–pair hopping and these terms survive in the infinite correlation limit (see (7) and (8)). These results imply that the extra holes probably repel though we cannot be sure from the form of the effective Hamiltonian alone and in any case the behaviour may be different in the mixed valence regime for which (8) is no longer valid. However, detailed numerical calculations of the two-hole states in the infinite-correlation limit confirm that the holes do indeed repel over the full range of interest (i.e. holes on O, Cu or the mixed valence regime; see Jefferson 1988b). Although the foregoing exact and numerical results are only strictly valid in one dimension, they do give some insight into the more difficult two-dimensional problem. There is little doubt that the spin polarisation caused by the propagating O hole and its broad band width ($\sim t^2/|\varepsilon - U|$ or $\sim t$ in the mixed valence case) will occur in two dimensions.

7. Summary and outlook

In this paper we have analysed a simplified Hamiltonian which models electron correlations in copper-oxide chains and planes. It has been shown that the Hilbert space of the initial Hamiltonian can always be reduced to produce effective Hamiltonians which adequately describe the low-lying excitations. The choice of base states on which the effective Hamiltonian operates (i.e. the model subspace) depends to some extent on the relative positions of the Cu and O energy levels (ε_p and ε_d). When $\varepsilon_p \gg \varepsilon_d + U$ the holes are confined to Cu sites which are never unoccupied in the model subspace and this yields, to fourth order, the effective Hamiltonian given by (7), describing the true low-energy states of the system. It reduces to the usual antiferromagnetic spin Hamiltonian (equation (6)) when there is just one hole per Cu site (corresponding to pure La_2CuO_4). In the general case (i.e. with doping) there are a number of extra terms in the effective Hamiltonian when compared with a similar treatment for a Hubbard Hamiltonian for the Cu sites only (cf (7) and (10)). These include triplet repulsions between holes in the lower Hubbard band and further repulsions between holes in the upper Hubbard band (i.e. the holes due to doping). The former interactions show that the antiferromagnetic singlet-binding energy is less than the spin-flip energy (they are equal for the Hubbard model), whereas the latter interactions imply that pairing between 'extra' holes is unlikely. This is, in fact, shown to be the case for the infinite-correlation limit (equation (8)) for which numerical solutions show that two extra holes always repel.

When the O levels are sufficiently shallow for doping to put holes on the O (i.e. $\varepsilon < U$) then the effective Hamiltonian (equation (14)) contains *second-order* antiferromagnetic Cu–O spin interactions in addition to the fourth-order exchange interactions between Cu spins (which again give the usual antiferromagnetic spin Hamiltonian for the undoped case (6)). These second-order exchange terms imply a *ferromagnetic* alignment between adjacent Cu spins when an O spin intervenes. This is indeed shown to be the case for the infinite-correlation limit (Cu sites always

occupied) despite the second-order hopping terms (including spin flips) and the exact solution presented in § 6 shows that as the O hole propagates it drags with it the ferromagnetic alignment of its neighbouring Cu spins (see figure 6).

Although the basic Hamiltonian (1) probably contains the main physics of copper-oxide chains and planes for the parameter ranges considered, i.e. $t < \varepsilon < \infty$ with $U \gg t$, any detailed calculations for comparison with experiment will need to take into account further electron–electron interactions and orbital degeneracy. The most important interactions which have been omitted are the remaining Coulomb and hopping interactions on the O and nearest-neighbour Cu and O. Thus we add to (1) the terms:

$$H_1 = V_{dp} \sum_{\langle ij \rangle} n_i n_j + U_p \sum_j n_{j\uparrow} n_{j\downarrow} + V_p \sum_{\langle jj' \rangle} n_j n_{j'} + t_p \sum_{\langle jj' \rangle \sigma} p^\dagger_{j\sigma} p_{j'\sigma} \tag{18}$$

where V_{dp} is the Coulomb repulsion between holes on adjacent Cu–O sites, U_p is the Coulomb repulsion (Hubbard U) between holes on the same O site, V_p is the Coulomb repulsion between holes on nearest-neighbour O sites and t_p is the direct hopping between O sites. It is straightforward to apply the basic perturbation method to $H + H_1$, again using the kinetic-energy terms (the t-terms) as the perturbation. The main effect will be to renormalise the terms in the effective Hamiltonians by changing energy denominators.

A rigorous treatment of degeneracy is more problematic though a number of recent papers have begun to address this question. An effective spin-$\frac{1}{2}$ treatment of the Cu sites is obviously incorrect when $Cu^{3+}(d^8)$ states are important since, according to Hund's rule, the two Cu holes will align to give a spin triplet lowest. This point has been discussed by Aoki and Kamimura (1987) and Kamimura (1987) who have shown how the Cu only Hubbard Hamiltonian may be generalised to include intra-ionic exchange on the Cu and how this effects the super-exchange between Cu ions. It would be straightforward, though tedious, to extend this analysis to include the O explicitly. The lifting of orbital degeneracy has been considered recently by Chakraverty *et al* (1987), Zhang and Rice (1988), Mila *et al* (1988) and Ashkenazi and Kuper (1988). The lowest-lying orbitals on the Cu ions are the $d_{x^2-y^2}$ orbitals which hybridise with the p_x and p_y orbitals on the O to form σ-bonds, with π-bonds formed from the remaining p_x, p_y and p_z orbitals lying somewhat higher in energy. Although retention of only those orbitals which form the lowest-lying σ-bonds will remove the orbital degeneracy, it is not clear that the higher-lying levels are unimportant and they should be included in the initial Hamiltonian before being eliminated, where possible, by the perturbation method described in this paper.

Finally, we mention some recent work by Gagliano *et al* (1988), Balseiro (1988) and Hirsch (1988) who consider the basic Hamiltonian (equation (1) with one or more terms from H_1, equation (18)) in the regime $\varepsilon < t$, which has not yet been discussed in this paper. It may be regarded as a mixed valence regime in which there is strong hybridisation between $Cu^{2+}(d^9) + O^{2-}(p^6)$ and $Cu^{+}(d^{10}) + O^{-}(p^5)$ and has been discussed earlier by Varma *et al* (1987) who referred to the process as a charge-transfer excitation, forming (largely unscreened) Frenkel excitons. Although $\varepsilon < t$, we may still apply the basic perturbation method to obtain an antiferromagnetic exchange Hamiltonian between Cu spins provided V_{dp} is sufficiently large, giving an insulating ground manifold for the undoped case. (Higher-lying states involving $Cu^{3+}(d^8)$ can also be eliminated, of course.) Upon doping the CuO_2 planes, there can be pairing between the 'extra' holes (contrary to the result of § 6), again provided V_{dp} is sufficiently large. This can be appreciated from a simple electrostatics argument which shows that

for the model Hamiltonian (1) including the V_{dp} term in (18), with $t = 0$, the hole-paired state is lower in energy provided that $V_{dp} > \varepsilon$. Such pairing correlations might be expected to persist when t is switched on and this was shown to be the case for finite planes by numerical solutions. Similar hole-pair states are possible when other interactions in (18) are switched on. Because of the large uncertainty in the magnitude of the energy parameters, further experimental and theoretical evidence is needed before we can decide whether or not this pairing mechanism is feasible for the oxide superconductors.

Acknowledgments

I would like to thank Professor E R Pike (Kings College and RSRE) and Dr S Sarkar (RSRE) for stimulating discussions on various aspects of this work. Helpful discussions with Dr L Feiner (Philips, Eindhoven) and Dr F Mila (ETH, Zurich) are also gratefully acknowledged.

Note added in proof. Since this work was first completed there has been mounting evidence (particularly from photoemission and EELS experiments) that the itinerant holes are primarily of O (2p) character, for which the results of § 5 are appropriate.

References

Aharoney A, Birgeneau R J, Coniglio A, Kastner M A and Stanley H E 1988 *Phys. Rev. Lett.* **60** 1330
Aoki H and Kamimura H 1987 *Solid State Commun.* **63** 665
Anderson P W 1959 *Phys. Rev.* **115** 2
—— 1961 *Phys. Rev.* **124** 41
—— 1973 *Mater. Res. Bull.* **8** 153
—— 1987 *Science* **235** 1196
Ashkenazi J and Kuper C G 1988 *Workshop on Superconductivity (Les Houches)* (oral presentation)
Balseiro C A 1988 *Workshop on Superconductivity (Les Houches)* (oral presentation)
Baskaran G, Zou Z and Anderson P W 1987 *Solid State Commun.* **63** 973
Bednorz J G and Muller K A 1986 *Z. Phys.* B **64** 189
Bethe H 1931 *Z. Phys.* **71** 205
Bianconi A, Congiu-Castellano A, De Santis M, Rudolf P, Lagarde P, Flank A M and Marcelli A 1987 *Solid State Commun.* **62** 821
—— 1987b *Solid State Commun.* **63** 1015
Bloch C 1958 *Nucl. Phys.* **6** 329
Brandow B H 1976 *Lectures in Theoretical Physics* (New York: Gordon and Breach)
Chakraverty B K, Feinberg D, Zheng H and Avignon M 1987 *Solid State Commun.* **64** 1147
Chu C W, Hor P H, Meng R L, Gas L, Huang Z J and Wang Y Q 1987 *Phys. Rev. Lett.* **58** 405
Cyrot M 1986 *Solid State Commun.* **60** 253
—— 1987a *Solid State Commun.* **62** 821
—— 1987b *Solid State Commun.* **63** 1015
Emery V J 1987 *Phys. Rev. Lett.* **58** 2794
Fuggle J C, Weijs P J W, Schoorl R, Sawatzky G A, Fink J, Nucker N, Durham P J and Temmerman W M 1988 *Phys. Rev.* B **37** 123
Gagliano E R, Rojo A G, Balseiro C A and Alascio B 1988 *Solid State Commun.* **64** 901
Hirsch J E 1985 *Phys. Rev. Lett.* **54** 1317
—— 1987a *Phys. Rev.* B **35** 8726
—— 1987b *Phys. Rev. Lett.* **59** 228
—— 1988 *Workshop on Superconductivity (Les Houches)* (oral presentation)
Jefferson J H 1988a *J. Phys. C: Solid State Phys.* **21** L193
—— 1988b *J. Phys. C: Solid State Phys.* **21** L425

Kamimura H 1987 *Proc. Adriatico Res. Conf. on High Temperature Superconductors* (Singapore: World Scientific)
Kivelson S, Sethna J and Rokshar D 1987 *Phys. Rev.* B **35** 8865
Klein D J and Seitz W A 1979 *Phys. Rev.* B **10** 321
Lindgren I 1974 *J. Phys. B: At. Met. Phys.* **7** 2441
Lindgren I and Morrison J 1986 *Atomic Many-Body Theory* (Berlin: Springer)
Mattheiss L F 1987 *Phys. Rev. Lett.* **58** 1028
Mattheiss L F and Hamann D R 1987 *Solid State Commun.* **63** 395
Mila F, Zhang F C and Rice T M 1988 *Workshop on Superconductivity (Les Houches)* (oral presentation)
Ruckenstein A E, Hirschfield P J and Appel J 1987 *Phys. Rev.* B **36** 857
Sandars P G H 1969 *Adv. Chem. Phys.* **14** 365
Sarma D D *et al* 1987 *Phys. Rev.* B **36** 2371
Stevens K W H 1976 *Phys. Rep.* **24** 1
Thoules D J 1972 *The Quantum Mechanics of Many-Body Systems* (New York: Academic)
Varma C M, Schmitt-Rink S and Abrahams E 1987 *Solid State Commun.* **62** 681
Yu J, Freeman A J and Xu J H 1987 *Phys. Rev. Lett.* **58** 1035
Zhang F C and Rice T M 1988 *Phys. Rev.* B **37** 3759
Zou Z and Anderson P W 1988 *Phys. Rev.* B **37** 627