Insulating phases of the $d = \infty$ Hubbard model

Insulating phases of the $d = \infty$ Hubbard model

David E Logan†‡, Michael P Eastwood† and Michael A Tusch†

† Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK
‡ Institut Laue–Langevin, Avenue des Martyrs, BP 156X, 38042 Grenoble Cédex 9, France

Abstract. A theory is developed for the $T = 0$ Mott–Hubbard insulating phases of the d^∞ Hubbard model at $\frac{1}{2}$-filling, including both the antiferromagnetic (AF) and paramagnetic (P) insulators. Local moments are introduced explicitly from the outset, enabling ready identification of the dominant low-energy scales for insulating spin-flip excitations. Dynamical coupling of single-particle processes to the spin-flip excitations leads to a renormalized self-consistent description of the single-particle propagators that is shown to be asymptotically exact in strong coupling, for both the AF and P phases. For the AF case, the resultant theory is applicable over the entire U-range, and is discussed in some detail. For the P phase, we consider in particular the destruction of the Mott insulator, the resultant critical behaviour of which is found to stem inherently from proper inclusion of the spin-flip excitations.

1. Introduction

Since its inception more than thirty years ago [1], the Hubbard model has become the canonical model of interacting fermions on a lattice. Although possibly the simplest model for describing competition between electron itinerancy and localization, with attendant implications for a host of physical phenomena from magnetism to metal–insulator transitions, its simplicity is superficial and an exact solution exists only for $d = 1$ dimension [2].

Recently, Metzner and Vollhardt [3] have pointed to the importance of the opposite extreme, $d = \infty$. In suppressing spatial fluctuations, the many-body problem here simplifies considerably, reducing to a dynamical single-site mean-field problem. Motivated in part by the expectation that an understanding of the $d = \infty$ limit will serve as a starting point for systematic investigation of finite dimensions, and by the knowledge that some important vestiges of finite-d behaviour remain inherent in the $d = \infty$ limit, intense study of the $\frac{1}{2}$-filled $d = \infty$ Hubbard model on bipartite lattices has since ensued; for recent detailed reviews, see references [4–6].

The true ground state of the model is an antiferromagnet (AF) for all interaction strengths $U > 0$. One aim of the present work [7] is to develop a theory for the $d = \infty$ AF which, in contrast to previous theories for the AF phase [8–10], is reliable over the entire U-range, and in particular becomes exact in the $U \to \infty$ strong-coupling limit both at $\frac{1}{2}$-filling where the Hubbard model maps onto the AF Heisenberg model, and in the one-hole sector where it reduces to the t–J model [11]

The majority of previous work [4–6] on the $d = \infty$ Hubbard model has focused on the paramagnetic (P) phase that results, even for $T = 0$, simply by neglecting the magnetic ordering (or suppressing it via frustration [5]). One highlight of this work has been the

emergence of a detailed description of the Mott metal–insulator transition, although here too the picture is not complete: for example, a firm understanding of the mechanism by which the $T = 0$ Mott insulating solution is destroyed, and even whether it is continuous or first order, remains elusive [5, 12]. A second aim of this paper is to focus on the insulating state of the P phase, and to develop a theory for it on a footing essentially identical to that for the AF, which likewise becomes exact in strong coupling and which permits an analysis of the destruction of the Mott insulator.

In seeking to develop a 'unified' description of the AF and P insulating phases, we adopt a rather different approach to that taken in previous work [4–6] by introducing explicitly, and from the outset, the notion of site local moments. To this end we consider first a conventional $T = 0$ mean-field approach to the problem in the form of unrestricted Hartree–Fock (UHF), together with a random-phase approximation (RPA) for transverse spin excitations of the mean-field state. Despite the limitations of such an approach *per se*, its importance resides in enabling identification of key low-energy scales for insulating spin-flip excitations. Since spatial fluctuations are suppressed for $d = \infty$, the low-energy spin-flip excitations are found to be Ising-like and (for each phase) characterized by a single scale, ω_s. This has a simple physical interpretation. For the AF, $\omega_s = \omega_p(U)$ is essentially just the energy cost of flipping a spin in the Néel ordered background; the ubiquity of antiferromagnetism for all $U > 0$ leads naturally to $\omega_p > 0$ for all U, with $\omega_p \sim 1/U$ as $U \to \infty$ as one expects in the Heisenberg limit. For the P phase by contrast, where magnetic ordering is absent, the fact that a given spin is equally likely to be surrounded by ↑- or ↓-spins and thus (for $d = \infty$) has as many ↑- as ↓-spin neighbours, ensures that the corresponding spin-flip energy cost $\omega_s = 0$ for all U in the insulating state.

Identification of the low-energy spin-flip scales, while crucial to the present work, is preliminary: to transcend the limitations of the conventional 'static' mean-field approach, single-particle processes must subsequently be coupled dynamically to the transverse spin-flip excitations. It is this which, in leading as we shall describe to a self-consistent description of the single-particle Green functions, enables the aims outlined in the preceding paragraphs to be achieved.

The Hubbard Hamiltonian, in standard notation, is

$$H = -t \sum_{(ij)\sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \qquad t = t_*/\sqrt{2Z} \tag{1.1}$$

with the (ij) sum over nearest-neighbour sites on a bipartite lattice of coordination number Z: a Bethe lattice (on which in practice we shall largely focus), or a d-dimensional hypercube. To ensure a non-trivial limit as $d \to \infty$ [3], the hopping is scaled as $t = t_*/\sqrt{2Z}$. The paper is organized as follows. UHF + RPA, and the spin-flip scales referred to above, are discussed in section 2. Emphasis is also given here to simple physical arguments which, in highlighting both the deficiencies and virtues of UHF + RPA, indicate what is required to go beyond it; particular attention being given in this regard to UHF for the P phase, in view of its close relation to the early work of Hubbard [13] and the Falicov–Kimball model [14].

Dynamical coupling of single-particle processes to the transverse spin excitations is considered in section 3, leading (section 3.1) to a renormalized self-consistent approximation for the ($T = 0$) single-particle Green functions upon which we subsequently concentrate. In section 3.2 the strong-coupling behaviour is examined analytically, and shown to be asymptotically exact for both the P and AF phases. Results are given in section 3.3, focusing in particular on single-particle spectra for the AF phase from strong to weak coupling, and on a discussion of the localization characteristics of the single-particle excitations—the latter

being quite subtle, and pointing to the delicacy of the limit $U \to \infty$ for the AF phase. For the P insulator, single-particle spectra are discussed briefly in section 3.3, before considering the destruction of the Mott insulating solution in section 4. The single-particle gap is found to close continuously, with an exponent $\nu = 1$, at a critical $U_c = 3.41t_*$. The origins of this behaviour are found to stem from inclusion of the $\omega_s = 0$ spin-flip scale in the interaction self-energies, pointing to the importance of such throughout the entire insulating regime, and not solely in obtaining the exact strong-coupling limit. The results of section 4 are in good agreement with recent numerical work [12], as discussed in section 5.

2. Conventional mean-field approach

We focus on the zero-temperature single-particle Green functions, defined by

$$G_{ii;\sigma} = -i \langle T\{c_{i\sigma}(t)c_{i\sigma}^{\dagger}\}\rangle \equiv G_{ii;\sigma}^{+}(t) + G_{ii;\sigma}^{-}(t) \tag{2.1}$$

(for the site-diagonal element); and separated for later purposes into retarded $(+, t > 0)$ and advanced $(-, t \leqslant 0)$ components. The essential feature of d^{∞} is that the corresponding interaction self-energy is site diagonal [3, 15], $\tilde{\Sigma}_{ij;\sigma}(\tilde{\omega}) = \delta_{ij}\tilde{\Sigma}_{i\sigma}(\tilde{\omega})$; here, and throughout, $\tilde{\omega}$ denotes frequency relative to the Fermi level, namely $\tilde{\omega} = \omega - U/2$. $G_{ii;\sigma}(\tilde{\omega})$ may be written as

$$G_{ii;\sigma}(\tilde{\omega}) = [\tilde{\omega} - \tilde{\Sigma}_{i\sigma}(\tilde{\omega}) - S_{i\sigma}(\tilde{\omega})]^{-1} \tag{2.2a}$$

where $S_{i\sigma}$ is the 'medium' self-energy—which alone survives in the non-interacting limit—expressing hopping of σ-spin electrons to neighbouring sites. Simple application of Feenberg's renormalized perturbation theory [16, 17] shows that, for $d = \infty$ but regardless of lattice type, $S_{i\sigma}$ is a functional solely of the $\{G_{jj;\sigma}\}$. The functional dependence is particularly simple for the Bethe lattice (BL) on which we concentrate, namely

$$S_{i\sigma}(\tilde{\omega}) = \sum_j t_{ij}^2 G_{jj;\sigma}(\tilde{\omega}) \tag{2.2b}$$

with $t_{ij} = t_*/\sqrt{2Z}$ the nearest-neighbour hopping element. Note that this is quite general; no assumption has been made about magnetic ordering or otherwise.

We consider now a conventional mean-field approach to the single-particle Green functions.

2.1. UHF

For both the AF and P phases, a Hartree–Fock approximation—by which we emphasize is here meant spin-*unrestricted* Hartree–Fock (UHF)—is the simplest non-trivial mean-field approximation, in which the notion of site local moments (μ_i), regarded as the first effect of electron interactions, enters from the outset. In the AF case, the local moments are naturally ordered in an A/B two-sublattice Néel state, with $\mu_i = \pm|\mu|$ for site i in the A/B sublattice respectively [18]. For the P phase by contrast, the local moments are randomly oriented: a site is equally likely to be A type as B type [19]. In either case the essential—and limiting—feature of UHF is that it is a static approximation, with solely elastic scattering of electrons and ω-independent interaction self-energies approximated by

$$\tilde{\Sigma}_{A\sigma}^{0} = -\frac{\sigma}{2}U|\mu| = -\tilde{\Sigma}_{B\sigma}^{0}. \tag{2.3}$$

For the AF phase, the UHF Green functions ($G^0_{ii;\sigma} \equiv G^0_{\alpha\sigma}$ with $\alpha = $ A or B) follow from equations (2.2), (2.3) for the BL as

$$G^0_{A\sigma}(\tilde{\omega}) = \left[\tilde{\omega} + \frac{\sigma}{2}U|\mu| - \frac{1}{2}t_*^2 G^0_{B\sigma}(\tilde{\omega})\right]^{-1}$$

$$G^0_{B\sigma}(\tilde{\omega}) = \left[\tilde{\omega} - \frac{\sigma}{2}U|\mu| - \frac{1}{2}t_*^2 G^0_{A\sigma}(\tilde{\omega})\right]^{-1} \qquad \text{AF} \qquad (2.4)$$

where the 'medium' self-energy part reflects the two-sublattice structure of the Néel state. Equations (2.4) are a closed set, with the UHF local moment $|\mu| = |\mu_0|$ found self-consistently via the usual gap equation (see e.g. [20]), which may be written formally as

$$|\mu_0| = \int_{-\infty}^{0} d\tilde{\omega} \; (D_{A\uparrow}(\tilde{\omega}) - D_{A\downarrow}(\tilde{\omega})) \qquad (2.5)$$

in terms of the corresponding spectral densities. And the total Green function is given by

$$G^0(\tilde{\omega}) = \frac{1}{2}[G^0_{A\sigma}(\tilde{\omega}) + G^0_{B\sigma}(\tilde{\omega})] \qquad (2.6)$$

such that $D^0(\tilde{\omega}) = -\pi^{-1}\,\text{sgn}(\tilde{\omega})\,\text{Im}\,G^0(\tilde{\omega})$ gives the total single-particle spectrum.

For the P phase by contrast,

$$G^0_{A\sigma}(\tilde{\omega}) = \left[\tilde{\omega} + \frac{\sigma}{2}U|\mu| - \frac{1}{2}t_*^2 G^0(\tilde{\omega})\right]^{-1}$$

$$G^0_{B\sigma}(\tilde{\omega}) = \left[\tilde{\omega} - \frac{\sigma}{2}U|\mu| - \frac{1}{2}t_*^2 G^0(\tilde{\omega})\right]^{-1} \qquad \text{P.} \qquad (2.7)$$

The sole difference from equation (2.4) occurs in the medium self-energy (see equation (2.2b)), since the nearest neighbours to any site are equally likely to be A- or B-type sites. Equations (2.6) are a closed set for $G^0(\tilde{\omega})$ and the $G^0_{\alpha\sigma}(\tilde{\omega})$ in the P phase; the UHF local moment is again found from equation (2.5).

For either phase there are two basic symmetries, namely

$$D^0_{A\sigma}(\tilde{\omega}) = D^0_{B-\sigma}(\tilde{\omega}) \qquad (2.8a)$$

and

$$D^0_{A\sigma}(\tilde{\omega}) = D^0_{A-\sigma}(-\tilde{\omega}) \qquad (2.8b)$$

reflecting the $\uparrow/\downarrow$-spin symmetry ($G^0_{A\sigma}(\tilde{\omega}) = G^0_{B-\sigma}(\tilde{\omega})$) and particle–hole symmetry ($G^0_{A\sigma}(\tilde{\omega}) = -G^0_{A-\sigma}(-\tilde{\omega})$) respectively; and note therefore from equation (2.6) that $G^0(\tilde{\omega})$ is naturally independent of the spin, σ.

We add further that UHF yields the correct atomic limit (where $|\mu_0| = 1$) for either phase, as is clear from equations (2.4), (2.6), (2.7) with $t_* = 0$.

2.1.1. Antiferromagnet. UHF for the AF has been widely studied since the early work of Penn [18]. Here we mention only that for any $d > 1$ the exact ground state of the $\frac{1}{2}$-filled Hubbard model on a bipartite lattice is an AF insulator for all $U > 0$, and this is qualitatively well captured at UHF level: for all $U > 0$, the mean-field ground state is a two-sublattice Néel AF, with a gap in the single-particle spectrum $D^0(\tilde{\omega})$ given by $\Delta(U) = U|\mu|$; figure 1 shows $D^0(\tilde{\omega})$ at $U/t_* = 4$ for the d^∞ BL.

The deficiencies of UHF are however most clearly seen in strong coupling, $U \to \infty$, where for the AF the single-particle spectrum reduces to $D^0(\tilde{\omega}) = \frac{1}{2}[\delta(\tilde{\omega} + U/2) + \delta(\tilde{\omega} - U/2)]$—as for the atomic limit, $t_* = 0$. The physical origin of this is simple:
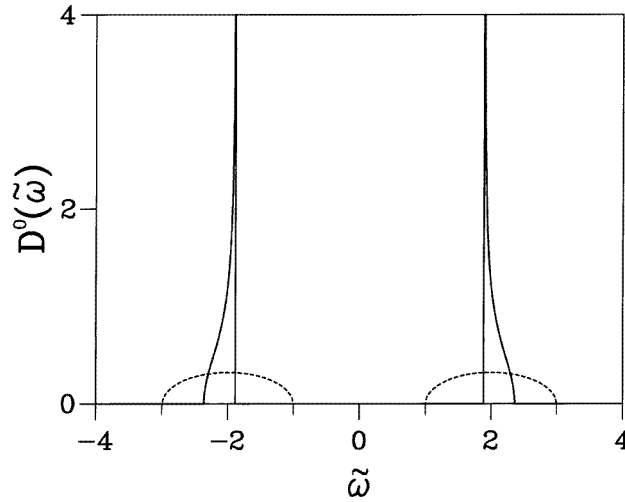
Figure 1. The UHF single-particle spectrum, $D^0(\tilde{\omega})$, versus $\tilde{\omega} = \omega - U/2$ (in units of t_*) for the d^∞ Bethe lattice. At $U/t_* = 4$, for the AF phase (full line) and the P phase (dashed).

consider for example the upper Hubbard band in strong coupling, and imagine adding an ↑-spin electron to a site (B type) already occupied by a ↓-spin. Since UHF is an *independent* (albeit interacting) electron approximation, only the added ↑-spin can potentially hop to nearest-neighbour (NN) sites. But it cannot do so in the strong-coupling limit, since for the AF all NNs to the ↓-spin B site are ↑-spins (A type). The added ↑-spin thus effectively 'sees' the ↓-spin site as an isolated site, hence the emergence of atomic limit behaviour as the strong-coupling limit at UHF level. But while physically transparent, this behaviour is wrong. In strong coupling, and for the one-hole sector appropriate to the lower Hubbard band (or one-doublon sector for the upper Hubbard band), the Hubbard model maps onto the t–J model [11]

$$\hat{H}_{tJ} = -t \sum_{(i,j)\sigma} \bar{c}^\dagger_{i\sigma} \bar{c}_{j\sigma} + \frac{1}{2} J_\infty \sum_{(i,j)} \left(\mathbf{S}_i \cdot \mathbf{S}_j - \frac{1}{4} n_i n_j \right) \qquad (2.9)$$

where the hole moves in a restricted subspace of no doubly occupied sites ($\bar{c}^\dagger_{i\sigma} = c^\dagger_{i\sigma}(1 - n_{i-\sigma})$); and in the fluctuating spin background provided by the Heisenberg part of $\hat{H}_{tJ}$, with NN exchange coupling $J_\infty = 4t^2/U$. Although it is exact in the atomic limit, UHF by itself can evidently say essentially nothing about the strong-coupling limit.

2.1.2. Paramagnet. UHF for the $T = 0$ paramagnetic phase warrants separate discussion, in part because of its very close relation to two other well known approaches. The first is that due to Hubbard [13], with 'spin-disorder scattering' only. This is often called the Hubbard III (HIII) approximation, and we here refer to it thus (noting that 'resonant broadening' contributions are additionally included in reference [13]). HIII is equivalent to UHF, but with a saturated local moment. Thus, with $|\mu| = 1$, the resultant cubic equation for $G^0(\tilde{\omega})$ on the d^∞ BL, obtained from equations (2.6), (2.7) above, coincides precisely with the HIII approximation for any U; see e.g. equation (34) of reference [21]. Although Hubbard's original formalism is very different, its physical content is that of a static approximation to an alloy analogy description [22]; a close relationship to UHF is thus to be expected.

The second connection is to the Falicov–Kimball (FK) model [14], a simplified version of the Hubbard model in which electrons of only one spin type are mobile, and which for d^∞ is exactly soluble [23–25]. For the paramagnetic phase of the FK model the single-particle Green function reduces precisely to that of HIII for any U [22], i.e. to the above-mentioned cubic for $G^0(\tilde{\omega})$ on the d^∞ BL; see e.g. equations (7.3) and (4.4) of reference [25].

For the P phase, figure 1 shows the UHF $D^0(\tilde{\omega})$ for $U/t_* = 4$ on the d^∞ BL, contrasted with its AF counterpart. Ordering or otherwise of the preformed local moments clearly has a significant effect on the spectra. In the AF ordered case, for example, the interior edges of the Hubbard bands have characteristic square-root divergences, while for the P phase all band edges vanish with square-root behaviour. More significantly, while the two-sublattice structure of the Néel ordered state ensures a band gap $\Delta = U|\mu|$ for all $U > 0$, the single-particle UHF gap vanishes in the P phase at a critical $U_c \simeq 1.9t_*$ given by $U_c|\mu(U_c)| = \sqrt{2}t_*$—or correspondingly $U_c = \sqrt{2}t_*$ for HIII/FK—signalling an insulator–metal transition. UHF/HIII fails of course in the metallic phase, there being no well defined Fermi surface or quasiparticles [22, 26]. This is inevitable for any inherently static approximation with a frequency-independent self-energy $\tilde{\Sigma}_{\alpha\sigma}$, since the essence of Fermi liquid behaviour is the inelasticity of electron scattering near the Fermi level, $\tilde{\omega} = 0$ [27].

However, even in the P insulating phase of interest here, UHF/HIII is deficient. As for the AF this is seen most clearly in strong coupling, $U \to \infty$, where although the centres of the Hubbard bands are separated by U, each has a non-vanishing width. In the strong-coupling P phase, and for the one-hole (doublon) sector corresponding to the lower (upper) Hubbard band, the Hubbard model maps onto the t–J model (equation (2.9)) in a *random* spin background; and the exact full bandwidth of either band is given for the d^∞ BL by [28, 29]

$$W_\infty = 2\sqrt{2}t_* \qquad U \to \infty. \tag{2.10a}$$

Note that this is also the single-particle bandwidth in the other extreme of the non-interacting limit, reflecting physically that in strong coupling the hole/doublon behaves essentially as a free particle [30].

In contrast, the strong-coupling bandwidth at UHF/HIII level is

$$W_\infty^0 = 2t_* \qquad \text{UHF/HIII}. \tag{2.10b}$$

UHF or HIII does not therefore give the exact strong-coupling limit for the Hubbard model, contrary to what has been suggested recently [31]; but, as is clear from the above discussion, gives instead the strong-coupling limit of the FK model [25]. The physical origin of equation (2.10b) is however both simple and revealing. Consider again the upper Hubbard band in strong coupling, and imagine adding an ↑-spin electron to a B-type ↓-spin site. Within a static approximation such as UHF/HIII, only the added ↑-spin can hop; and it can do so in the first instance to any of ↓-spin NNs (B-type sites) only—the effective coordination number for which is $Z_{\text{eff}} = \frac{1}{2}Z$. Since Z_{eff} for the propagating ↑-spin electron is reduced by a factor of 2 below the full coordination number, and since the bandwidth of the d^∞ BL is proportional to $\sqrt{Z_{\text{eff}}}$, the strong-coupling UHF/HIII width is thus diminished by $\sqrt{2}$ from the corresponding non-interacting value which, as in equation (2.10a) above, is also the exact strong-coupling limit; equation (2.10b) thus results.

The distinction between equations (2.10a) and (2.10b) is however qualitative, and not solely a matter of degree, reflecting the need to take seriously—even in strong coupling—the *correlated* dynamics of the electrons. Whenever, say, an ↑-spin electron is added to a site occupied by a ↓-spin electron, the added ↑-spin can indeed propagate in the P phase,

scattering elastically off successive neighbouring ↓-spins; and as sketched above this is well captured at UHF/HIII level. But, having added the ↑-spin to a ↓-spin site, the latter can itself clearly hop off the site—to a neighbouring ↑-spin site—leaving behind it a spin-flip on the original site. The energy cost for the spin flip is zero, since we are considering the P phase where, for d^∞, a given spin is equally likely to be surrounded by ↑- or ↓- spins and has as many ↑- as ↓-spin neighbours (whence there is no 'exchange penalty' for a spin flip). Thus, whether the added ↑-spin or the ↓-spin already present hops off the site, the initially created doublon propagates as a free particle [30]; equation (2.10*a*) thus results. To describe correctly the electron dynamics, *both* types of process above—and therefore the interference between them—must be included. A static approximation such as UHF/HIII cannot handle this, since such dynamics reside in the frequency dependence of the full interaction self-energy $\tilde{\Sigma}_{i\sigma}(\tilde{\omega})$, as considered in section 3.
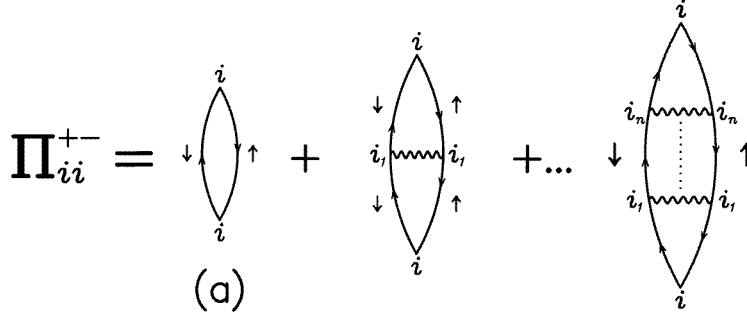

Figure 2. The particle–hole ladder sum in the transverse spin channel, for RPA Π_{ii}^{+-}. Bare (UHF) propagators are denoted by solid lines, on-site interactions by wiggles. For $d = \infty$, all intermediate sites $i_1, \ldots, i_n$ are equal to i.

2.2. RPA

In contrast to single-particle spectra—probing states one hole or particle away from $\frac{1}{2}$-filling—RPA probes fluctuations about the mean-field state, and thus excitations of the $\frac{1}{2}$-filled state itself. For the insulating phases, with a gap to charge excitations, *transverse* spin excitations are of lowest energy. These are reflected in the transverse spin-polarization propagators $\Pi_{ij}^{+-}(t) = i\langle T\{S_i^+(t)S_j^-\}\rangle$ and $\Pi_{ij}^{-+}(t)$, given within RPA by

$$\mathbf{\Pi}^{+-}(\omega) = {}^0\mathbf{\Pi}^{+-}(\omega)[\mathbf{1} - U\,{}^0\mathbf{\Pi}^{+-}(\omega)]^{-1} \tag{2.11a}$$

where $[\mathbf{\Pi}^{+-}(\omega)]_{ij} = \Pi_{ij}^{+-}(\omega)$, $[\mathbf{1}]_{ij} = \delta_{ij}$ and ${}^0\Pi_{ij}^{+-}$ is the pure UHF transverse spin-polarization bubble. Equation (2.11*a*) leads directly to a familiar diagrammatic 'bubble sum'. Alternatively, since the interaction is solely on-site, this may be recast as a 'ladder sum' of repeated particle–hole interactions in the transverse spin channel, as shown in figure 2; bare UHF propagators are denoted by solid lines, and the on-site interactions (conserving spin at each vertex end) by wiggly lines. From the basic symmetries (equation (2.8)), it follows that $\Pi_{ii}^{-+}(\omega) = \Pi_{ii}^{+-}(-\omega)$ for $i = $ A or B; and $\Pi_{BB}^{+-}(\omega) = \Pi_{AA}^{-+}(\omega)$.

For finite d, intermediate sites in the ladder sum for Π_{ii}^{+-} (figure 2) are in general different from i. But since ${}^0\Pi_{ij} \sim O(d^{-m})$ for sites i and j mth-nearest neighbours, *all* intermediate sites in Π_{ii} are equal to i for d^∞; i.e. $i = i_1 = i_2 = \cdots = i$ whence Π_{ii}^{+-} (or

Π_{ii}^{-+}) is purely algebraic; that is

$$\Pi_{ii}^{+-}(\omega) = {}^{0}\Pi_{ii}^{+-}(\omega)/[1 - U \, {}^{0}\Pi_{ii}^{+-}(\omega)] \qquad d^{\infty}. \qquad (2.11b)$$

The spectral density of transverse spin excitations is naturally reflected in $\mathrm{Im}\,\Pi_{ii}^{+-}(\omega)$, as now considered for the AF and P phases.
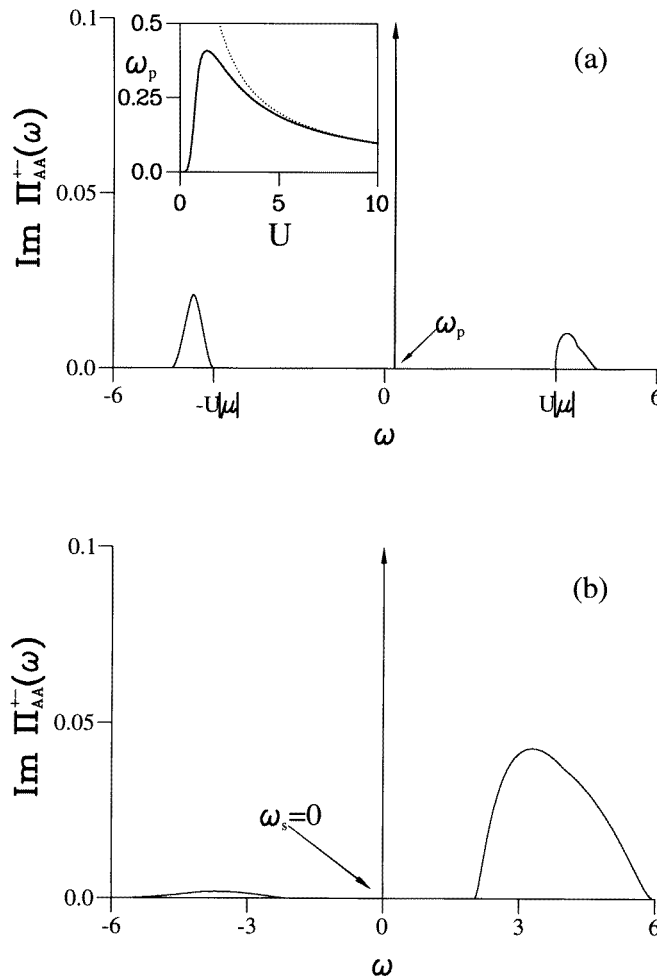

Figure 3. $\mathrm{Im}\,\Pi_{AA}^{+-}(\omega)$ versus ω/t_* at $U/t_* = 4$ for d^{∞} BL. (a) For the AF phase; the inset shows the U/t_* dependence of the AF spin-flip pole ω_p, with the dotted line denoting the $U \to \infty$ asymptote $\omega_p^{\infty} = t_*^2/U$. (b) For the P phase, where the spin-flip pole $\omega_s = 0$ for all U in the insulating state.

2.2.1. The AF phase. For the AF phase, figure 3(a) shows $\mathrm{Im}\,\Pi_{AA}^{+-}(\omega)$ at $U/t_* = 4$ for the d^{∞} Bethe lattice. Two distinct features are apparent: a low-frequency spin-flip pole (discussed below), and a high-energy Stoner-like band. The latter consists simply of weakly renormalized Hartree–Fock excitations across the gap in the mean-field single-particle spectrum. Spectral density for the Stoner bands does not therefore begin until

precisely $|\omega| = U|\mu|$ (see figure 1), and their maximum density occurs for $|\omega| \simeq U$. This is as found also for finite d; see e.g. reference [32].

The central feature in figure 3(a) is a low-ω pole at ω_p, located via equation (2.11b) from $U \, {}^0\Pi_{ii}^{+-}(\omega_p) = 1$, and occurring for all $U > 0$ (figure 3(a), inset). This is the sole remnant, for d^∞, of the spin-wave-like component of the transverse spin spectrum studied recently at RPA level [32]; and which, for finite d and any $U > 0$, is naturally gapless. Physically, the single spin-flip pole at ω_p reflects the general suppression of spatial fluctuations for d^∞: it corresponds simply to the energy cost of flipping a spin in the AF background. This is particularly clear in strong coupling, where the Stoner bands are eliminated entirely. Here, as is well known [20, 33], the RPA transverse spin spectrum reduces (for any d) to the linear spin-wave spectrum of the nearest-neighbour AF Heisenberg model, with exchange coupling $J_\infty = 4t^2/U = 2t_*^2/ZU$, onto which the $\frac{1}{2}$-filled Hubbard model maps rigorously. And for d^∞ it is straightforward to show that the resultant linear spin-wave spectrum collapses to an Ising-like spin-flip pole at $\omega_p^\infty = ZJ_\infty/2 = t_*^2/U$. Further, since linear spin-wave theory for the Heisenberg model is exact for d^∞ [34], it follows that in strong coupling UHF + RPA gives the exact spin excitation spectrum of the $\frac{1}{2}$-filled Hubbard model.

The occurrence of the single ω_p-pole is robust to further renormalization of particle–hole lines in $\Pi_{ii}^{+-}(\omega)$, as discussed in section 3.3. We stress further that to capture it requires the full ladder sum of repeated p–h interactions shown in figure 2: retention solely of the 'bare' polarization bubble diagram will clearly not suffice.
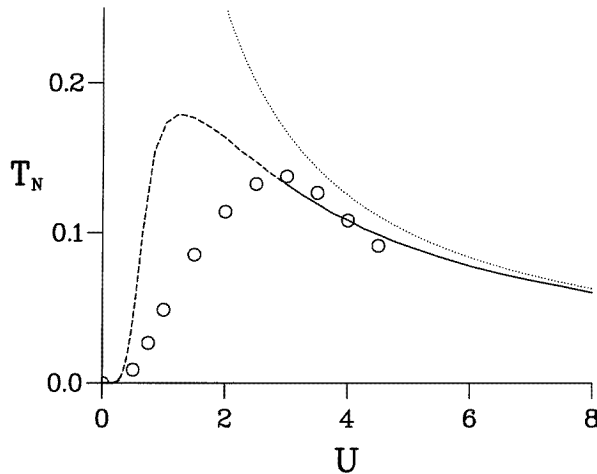

Figure 4. The QMC Néel temperature versus U/t_* (open circles) for the d^∞ hypercubic lattice [24, 25]. The simple estimate $T_N \simeq \frac{1}{2}\omega_p(U)$, argued to be valid for $U/t_* \gtrsim 3$, is also shown (the solid line for $U/t_* > 3$). The strong-coupling asymptote $T_N^\infty = t_*^2/2U$ is indicated by the dotted line.

The necessity of including the AF spin-flip scale will be evident when discussing the $T = 0$ single-particle spectra; sections 3.2, 3.3. Here, we illustrate briefly its importance at finite temperature, as reflected in the Néel temperature $T_N(U)$. Molecular-field theory is exact for the Heisenberg model in d^∞ [35]; thus, in strong coupling, $T_N = ZJ_\infty/4 = \frac{1}{2}\omega_p^\infty$. At finite U, we expect $T_N(U) \simeq \frac{1}{2}\omega_p(U)$ to yield a good estimate of the Néel temperature in a U-regime where thermal properties are dominated by the low-lying spin-flip excitations. Jarrell and Pruschke [36, 37] have obtained the finite-T phase diagram for the d^∞ hypercubic

lattice via quantum Monte Carlo (QMC). The thermal paramagnetic phase above $T_N(U)$ is found to be metallic for $U/t_* \lesssim 3$ and insulating for $U/t_* \gtrsim 3$ (with a small 'crossover' regime); it is thus in the latter region that we expect $T_N \simeq \frac{1}{2}\omega_p$. This is borne out. Figure 4 shows the QMC $T_N(U)$ for the d^∞ hypercubic lattice, together with the corresponding $\frac{1}{2}\omega_p(U)$ and the exact strong-coupling asymptote $T_N = \frac{1}{2}\omega_p^\infty$. The QMC Néel temperature is indeed well described by $\frac{1}{2}\omega_p(U)$ down to $U/t_* \sim 3$.

2.2.2. The P phase. For the $T = 0$ P phase, figure 3(b) shows $\mathrm{Im}\,\Pi_{\mathrm{AA}}^{+-}(\omega)$ at $U/t_* = 4$ for the d^∞ Bethe lattice. Compared to its AF counterpart (figure 3(a)) the key difference is that the spin-flip pole occurs at $\omega = 0$, reflecting the fact that the energy cost for a spin flip is zero in the paramagnetic insulator, as argued physically in section 2.1.2. The formal origin of this at RPA level is seen readily by noting that the bare transverse spin-polarization bubble (figure 2, diagram (a)) is given by

$$^0\Pi_{\mathrm{AA}}^{+-}(\omega) = \mathrm{i} \int_{-\infty}^{\infty} \frac{\mathrm{d}\tilde{\omega}'}{2\pi}\, G_{\mathrm{A}\downarrow}^0(\tilde{\omega}')G_{\mathrm{A}\uparrow}^0(\tilde{\omega}' - \omega). \tag{2.12}$$

From equation (2.4) for the UHF Green functions, using $G_{\mathrm{A}\downarrow}^0 = G_{\mathrm{B}\uparrow}^0$, it follows that

$$G_{\mathrm{A}\downarrow}^0(\tilde{\omega})G_{\mathrm{A}\uparrow}^0(\tilde{\omega}) = -\frac{1}{U|\mu|}(G_{\mathrm{A}\uparrow}^0(\tilde{\omega}) - G_{\mathrm{A}\downarrow}^0(\tilde{\omega})). \tag{2.13}$$

Hence, using the spectral representation of $G_{\mathrm{A}\sigma}^0(\tilde{\omega})$,

$$^0\Pi_{\mathrm{AA}}^{+-}(\omega = 0) = \frac{1}{U|\mu|} \int_{-\infty}^{0} \mathrm{d}\tilde{\omega}\, \left[D_{\mathrm{A}\uparrow}^0(\tilde{\omega}) - D_{\mathrm{A}\downarrow}^0(\tilde{\omega})\right]. \tag{2.14}$$

Since the UHF local moment $|\mu| = |\mu_0|$ is given by equation (2.5), $^0\Pi_{\mathrm{AA}}^{+-}(\omega = 0) = 1/U$; and thus from equation (2.11b) the RPA $\Pi_{\mathrm{AA}}^{+-}(\omega)$ has a spin-flip pole at $\omega = 0$.

Note again, as for the AF case, that the full ladder sum of particle–hole interactions in the transverse spin channel is required to capture the $\omega = 0$ spin-flip pole. Further, although we have shown explicitly its existence within RPA, the occurrence of the zero-frequency spin-flip scale is naturally a general feature of the d^∞ paramagnetic insulating phase where, locally, the ground state is a doubly degenerate local moment (as for the single-impurity Anderson model embedded in an insulating host) [5].

For both phases, the evident virtues of the RPA for excitations of the $\frac{1}{2}$-filled state contrast sharply with the deficiencies of the single-particle spectra at UHF level; section 2.1. This itself hints at what is necessary to describe the single-particle spectra successfully: single-particle processes must be coupled dynamically to the transverse spin excitations, reflected in the frequency dependence of the self-energy. This is now considered.

3. Green functions

It is helpful to separate the full interaction self-energies $\tilde{\Sigma}_{\alpha\sigma}(\tilde{\omega})$ as

$$\begin{aligned}
\tilde{\Sigma}_{\mathrm{A}\sigma}(\tilde{\omega}) &= -\frac{\sigma}{2}U|\mu| + \Sigma_{\mathrm{A}\sigma}(\tilde{\omega}) \\
\tilde{\Sigma}_{\mathrm{B}\sigma}(\tilde{\omega}) &= \frac{\sigma}{2}U|\mu| + \Sigma_{\mathrm{B}\sigma}(\tilde{\omega})
\end{aligned} \tag{3.1}$$

where $\Sigma_{\alpha\sigma}(\tilde{\omega})$ ($\alpha = \mathrm{A}$ or B) excludes the first-order UHF-type contribution, and contains the dynamics on which we want to focus. From equations (2.2), (2.3) for the Bethe lattice,

the exact site-diagonal Green functions are thus given formally by

$$G_{A\sigma}(\tilde{\omega}) = \left[\tilde{\omega} + \frac{\sigma}{2}U|\mu| - S_{A\sigma}(\tilde{\omega}) - \Sigma_{A\sigma}(\tilde{\omega})\right]^{-1} \tag{3.2a}$$

$$G_{B\sigma}(\tilde{\omega}) = \left[\tilde{\omega} - \frac{\sigma}{2}U|\mu| - S_{B\sigma}(\tilde{\omega}) - \Sigma_{B\sigma}(\tilde{\omega})\right]^{-1}. \tag{3.2b}$$

Here, the medium self-energy is given for the AF and P phases by

$$S_{\alpha\sigma}(\tilde{\omega}) = \begin{cases} \dfrac{1}{2}t_*^2 G_{\bar{\alpha}\sigma}(\tilde{\omega}) & \text{AF} \\[2mm] \dfrac{1}{2}t_*^2 G(\tilde{\omega}) & \text{P} \end{cases} \tag{3.2c}$$

where the site index $\bar{\alpha} = \text{B}$ or A for $\alpha = \text{A}$ or B respectively; and

$$G(\tilde{\omega}) = \frac{1}{2}[G_{A\sigma}(\tilde{\omega}) + G_{B\sigma}(\tilde{\omega})] \tag{3.3}$$

is the total Green function.

As at UHF level, $\uparrow/\downarrow$-spin symmetry and particle–hole symmetry for the corresponding spectral densities imply

$$D_{A\sigma}(\tilde{\omega}) = D_{B-\sigma}(\tilde{\omega}) \tag{3.4a}$$

and

$$D_{A\sigma}(\tilde{\omega}) = D_{A-\sigma}(-\tilde{\omega}) \tag{3.4b}$$

respectively. For the associated Green functions $G_{\alpha\sigma}(\tilde{\omega}) = G_{\alpha\sigma}^+(\tilde{\omega}) + G_{\alpha\sigma}^-(\tilde{\omega})$, a Hilbert transform of equations (3.4) gives directly

$$G_{A\sigma}^{\pm}(\tilde{\omega}) = G_{B-\sigma}^{\pm}(\tilde{\omega}) \tag{3.5a}$$

and

$$G_{A\sigma}^{\pm}(\tilde{\omega}) = -G_{A-\sigma}^{\mp}(-\tilde{\omega}). \tag{3.5b}$$

Thus, from equation (3.3),

$$G(\tilde{\omega}) = -G(-\tilde{\omega}) \tag{3.5c}$$

while from equations (3.2), (3.5)

$$\Sigma_{B\sigma}(\tilde{\omega}) = \Sigma_{A-\sigma}(\tilde{\omega}) \tag{3.6a}$$

and

$$\Sigma_{B\sigma}(\tilde{\omega}) = -\Sigma_{A\sigma}(-\tilde{\omega}) \tag{3.6b}$$

and likewise for the $\tilde{\Sigma}_{\alpha\sigma}$. Equations (3.5a) with (3.3) show also that $G(\tilde{\omega})$ is correctly independent of spin.

The symmetries reflected in equations (3.5), (3.6) play an important role in the following analysis. For the P phase, note also the physical interpretation of equation (3.3) for $G(\tilde{\omega})$: viewing the paramagnet in terms of randomly oriented local moments, where a site is equally likely to likely to be A type as B type, we can consider equation (3.3) as a configurationally averaged Green function. This is a natural alloy analogy interpretation but, unlike the static approximation to such inherent in UHF or HIII, it is formally exact since no approximation to the interaction self-energies has thus far been made.

3.1. Self-consistent renormalization

Our aim now is to develop a specific approximation to the self-energy which in particular (a) becomes exact in strong coupling, ensuring thereby a controlled limit; and (b) is constructed in renormalized form, enabling a self-consistent solution for the single-particle Green functions.
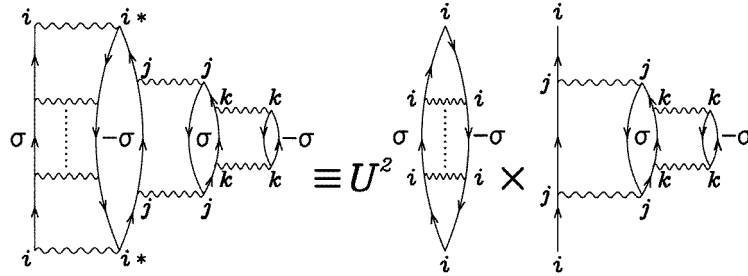

Figure 5. A diagram contributing to single-particle self-energy $\Sigma_{i\sigma}$, with the same notation as in figure 2; for a full discussion, see the text.

A relevant diagram contributing to $\Sigma_{i\sigma}$ is shown in figure 5, employing the same diagrammatic notation as figure 2. Using deliberately a strong-coupling terminology, its physical interpretation is as follows (with $t > 0$ for convenience): at $t = 0$ a ($\sigma =$) $\uparrow$-spin electron, say, is added to site i, thus creating a 'doublon'; at $t_1 > 0$ the ($-\sigma =$) $\downarrow$-spin electron already present on site i hops from i to j, and at $t_2 > t_1$ an $\uparrow$-spin hops from j to k; the entire path is then retraced. The diagram thus describes motion of the doublon (or hole for $t < 0$) from $i \rightarrow j \rightarrow k$ via a correlated sequence of alternating spin hops, creating behind it a string of flipped spins. All ladder interactions of the resultant on-site particle–hole pair—which reflect the on-site spin-flip created by motion of the doublon/hole—are shown explicitly for site i in figure 5; from which it is seen that their sum is exactly $U^2\Pi_{ii}^{-+}$, with $\Pi_{ii}^{-+}(\omega)$ ($=\Pi_{ii}^{+-}(-\omega)$) the RPA transverse spin propagator discussed in section 2.2 (cf. figure 2).

It is precisely correlated dynamics of the sort exemplified by figure 5 that we seek to include and generalize in the frequency-dependent $\Sigma_{\alpha\sigma}(\tilde{\omega})$. To this end we first define an undressed (or self-consistent host) Green function by

$$\mathcal{G}_{ii;\sigma}(\tilde{\omega}) = [G_{ii;\sigma}^{-1}(\tilde{\omega}) + \Sigma_{i\sigma}(\tilde{\omega})]^{-1}. \tag{3.7}$$

This is shown diagrammatically in figure 6(a), as obtained simply from equation (3.7) using the Dyson equation for the full Green function $G_{ii;\sigma}$ expressed in terms of the UHF propagators and self-energy insertions. As seen from the figure, the implicit sum over intermediate sites j, k, etc is thus restricted to exclude site i itself (unlike the full $G_{ii;\sigma}$ where the site sums are free). While including all interactions on sites $j \neq i$, $\mathcal{G}_{ii;\sigma}$ thus excludes all interactions on site i beyond the simple first-order UHF contribution to $\tilde{\Sigma}_{i\sigma}$. The latter is of course subsumed into the UHF Green functions (as in section 2.1), which here constitute the 'bare' propagators; and in this important sense the above definition, equation (3.7), of the host propagator differs from that of e.g. references [21, 36–40] (which would be recovered if we set the site local moment $|\mu| = 0$).

To generalize the processes contained in figure 5, we renormalized the self-energy as shown in figure 6(b), replacing the $-\sigma$-spin particle lines connecting the starred vertices i in figure 5 by the self-consistent host Green function $\mathcal{G}_{ii;-\sigma}$; the infinite set of diagrams
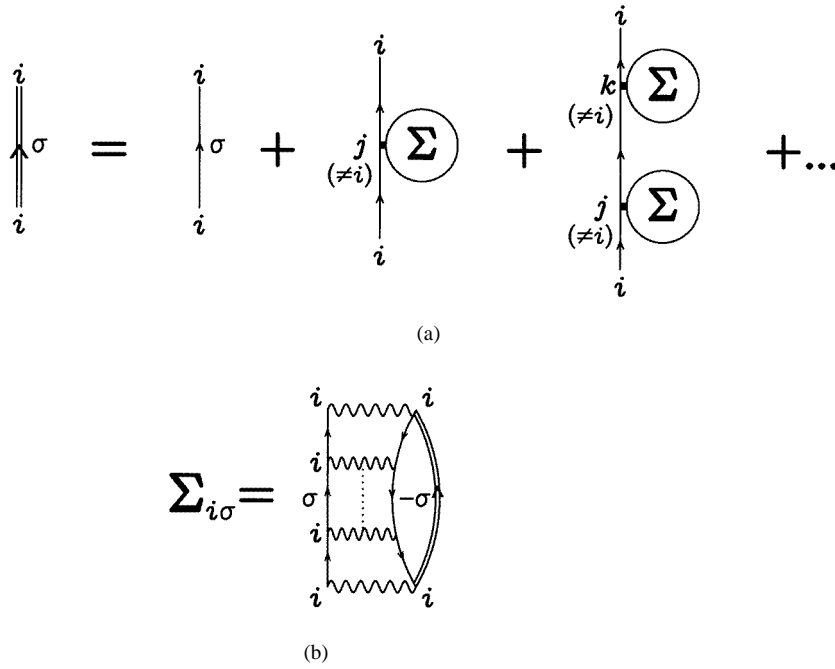
(a)

(b)

Figure 6. (a) The undressed (or self-consistent host) Green function $\mathcal{G}_{ii;\sigma}$, expressed in terms of bare (UHF) propagators and site-diagonal self-energy insertions $\Sigma_{i\sigma}$. Note the restrictions on intermediate site sums. (b) The basic approximation to $\Sigma_{i\sigma}$ used in the present work, from which the full set of diagrams retained follows by iteration using (a).
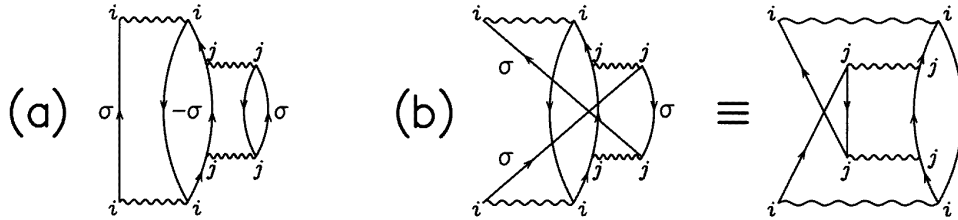

Figure 7. A 'direct' diagram (a), and its exchange counterpart (b). With site $j = i$ excluded from the implicit sum over j, the direct diagram is O(1) while the exchange diagram is O(1/d).

thus retained in $\Sigma_{i\sigma}$ follows simply by direct iteration of figure 6(b) using figure 6(a) for $\mathcal{G}_{ii;-\sigma}$. This renormalization is adopted for the following reasons.

(i) It ensures that an on-site spin flip occurs only when the doublon/hole hops off a site, and that its outward path is self-avoiding. With reference to figure 5 for example, site $j \neq i$ is guaranteed, likewise $k \neq j$; while terms with $k = i$ vanish for d^∞, being at least O(1/d) since $G^0_{ij;\sigma} \sim$ O($d^{-m/2}$) for sites i and j mth-nearest neighbours.

(ii) In addition, the resultant site restrictions further prevent the need to include a class of partially cancelling exchange diagrams, as illustrated simply in figure 7 (where a sum over $j \neq i$ is implicit). Since $j \neq i$ is guaranteed, the exchange diagram (figure 7(b)) is at least O(1/d) and thus vanishes for d^∞, while the 'direct' diagram (figure 7(a)) is O(1). If,

however, $j = i$ was included in the direct diagram, its exchange counterpart would also be O(1) and would thus need to be retained.

Our basic approximation to $\Sigma_{i\sigma}(\tilde{\omega})$ is thus figure 6(b), namely

$$\Sigma_{A\uparrow}(\tilde{\omega}) = U^2 \int \frac{d\Omega}{2\pi i} \, \Pi_{AA}^{-+}(\Omega) \mathcal{G}_{A\downarrow}(\tilde{\omega} - \Omega) \tag{3.8a}$$

and the remaining $\Sigma_{\alpha\sigma}$ follow by symmetry; equation (3.6). From section 2.2 the RPA $\Pi_{AA}^{-+}(\Omega)$ ($=\Pi_{AA}^{+-}(-\Omega)$) may be separated into the spin-flip pole contribution, $Q[\Omega + \omega_s - i\eta]^{-1}$ (with pole weight Q), plus the Stoner contribution; whence equation (3.8a) may be cast as

$$\Sigma_{A\uparrow}(\tilde{\omega}) = QU^2 \mathcal{G}_{A\downarrow}^-(\tilde{\omega} + \omega_s) + \Sigma_{A\uparrow}^{\text{Stoner}}(\tilde{\omega}) \tag{3.8b}$$

with spin-flip frequency

$$\omega_s = \begin{cases} \omega_p & \text{AF} \\ 0 & \text{P.} \end{cases} \tag{3.8c}$$

By symmetry, $\mathcal{G}_{A\downarrow}(\tilde{\omega}) = \mathcal{G}_{B\uparrow}(\tilde{\omega})$ using equations (3.5)–(3.7); and from equations (3.7), (3.2):

$$\mathcal{G}_{A\downarrow}(\tilde{\omega}) = \begin{cases} \left[\tilde{\omega} - \frac{1}{2}U|\mu| - \frac{1}{2}t_*^2 G_{A\uparrow}(\tilde{\omega}) \right]^{-1} & \text{AF} \\[3mm] \left[\tilde{\omega} - \frac{1}{2}U|\mu| - \frac{1}{2}t_*^2 G(\tilde{\omega}) \right]^{-1} & \text{P.} \end{cases} \tag{3.9}$$

From equation (3.8a), $\Sigma_{A\uparrow}(\tilde{\omega})$ is thus a functional of the Green functions, the basic equations for which (equations (3.2)) must therefore be solved self-consistently, as now described.

3.2. Strong coupling

We consider first the behaviour in strong coupling, as this can be extracted analytically. Since $|\mu| = 1 - O(t_*^2/U^2)$, equations (3.2) reduce in strong coupling to

$$G_{A\uparrow}(\omega) = [\omega - S_{A\uparrow}(\omega) - \Sigma_{A\uparrow}(\omega)]^{-1} \tag{3.10a}$$

$$G_{B\uparrow}(\omega) = [\omega - U - S_{B\uparrow}(\omega) - \Sigma_{B\uparrow}(\omega)]^{-1}. \tag{3.10b}$$

The interaction self-energy, equation (3.8), likewise simplifies in strong coupling, since the Stoner contribution vanishes (see section 2.2) and the pole weight $Q \to 1$. Hence from equation (3.8b),

$$U^{-2}\Sigma_{A\uparrow}(\omega) = \mathcal{G}_{A\downarrow}^-(\omega + \omega_s) = \mathcal{G}_{B\uparrow}^-(\omega + \omega_s). \tag{3.11}$$

This may be reduced further, noting that $\mathcal{G}_{B\uparrow}^-(\omega)$ is given by

$$\mathcal{G}_{B\uparrow}^-(\omega) = \int_{-\infty}^{U/2} d\omega_1 \, \frac{\mathcal{D}_{B\uparrow}(\omega_1)}{\omega - \omega_1 - i\eta} \tag{3.12}$$

as a one-sided Hilbert transform of the corresponding lower-Hubbard-band spectral density, $\mathcal{D}_{B\uparrow}(\omega_1) = -\pi^{-1} \operatorname{sgn}(\omega_1 - U/2) \operatorname{Im} \mathcal{G}_{B\uparrow}(\omega_1)$; and from equations (3.7), (3.10b)

$$\mathcal{G}_{B\uparrow}(\omega_1) = [\omega_1' - U - S_{B\uparrow}(\omega_1)]^{-1} = [\omega_1' - U]^{-1} + [\omega_1' - U]^{-2}S_{B\uparrow}(\omega_1) + O(U^{-3}) \tag{3.13}$$

where $\omega_1' = \omega_1 + i\eta\, \mathrm{sgn}(\omega_1 - U/2)$. From equation (3.12), the leading large-U contribution to $\mathcal{G}_{\mathrm{B}\uparrow}^-(\omega)$ thus arises from the second term in equation (3.13), yielding $\mathcal{G}_{\mathrm{B}\uparrow}^-(\omega) = U^{-2} S_{\mathrm{B}\uparrow}^-(\omega)$; hence from equations (3.11), (3.2c):

$$\Sigma_{\mathrm{A}\uparrow}(\omega) = S_{\mathrm{B}\uparrow}^-(\omega + \omega_s) = \begin{cases} \dfrac{1}{2}t_*^2 G_{\mathrm{A}\uparrow}^-(\omega + \omega_p) & \text{AF} \\[2mm] \dfrac{1}{2}t_*^2 G^-(\omega) & \text{P.} \end{cases} \tag{3.14}$$

We focus now on the lower Hubbard band (LHB) in strong coupling, i.e. for $\omega \approx 0 \ll U \to \infty$; the upper Hubbard band follows trivially by symmetry. Since $\Sigma_{\mathrm{B}\uparrow}(\omega) = -\Sigma_{\mathrm{A}\uparrow}(U - \omega)$, it follows that for ω in the LHB $\Sigma_{\mathrm{B}\uparrow}(\omega)$ is pure real and $O(1/U)$; it can thus be neglected. The $G_{\alpha\uparrow}^+(\omega \approx 0)$ are likewise pure real, with $G_{\mathrm{B}\uparrow}^+(\omega) \sim O(1/U)$ and $G_{\mathrm{A}\uparrow}^+(\omega) \sim O(1/U^3)$ (as may be shown using equation (3.10) together with the analogue of equation (3.12) for $G_{\alpha\sigma}^+(\omega)$); together with $G^+(\omega) = \frac{1}{2}[G_{\mathrm{A}\uparrow}^+ + G_{\mathrm{B}\uparrow}^+]$, they too may be neglected. And from equation (3.10b), $G_{\mathrm{B}\uparrow}^-(\omega) \equiv \mathcal{G}_{\mathrm{B}\uparrow}^-(\omega) = U^{-2} S_{\mathrm{B}\uparrow}^-(\omega)$ which can also be neglected asymptotically. Hence in total, $G_{\mathrm{A}\uparrow}(\omega) \equiv G_{\mathrm{A}\uparrow}^-(\omega)$ and $G(\omega) = \frac{1}{2}[G_{\mathrm{A}\uparrow}(\omega) + G_{\mathrm{B}\uparrow}(\omega)] \equiv \frac{1}{2}G_{\mathrm{A}\uparrow}(\omega)$. For the LHB in strong coupling, equation (3.10a) thus reduces to

$$G_{\mathrm{A}\uparrow}(\omega) = \left[\omega - \frac{1}{2}t_*^2 G_{\mathrm{A}\uparrow}(\omega + \omega_p)\right]^{-1} \qquad \text{AF} \tag{3.15a}$$

$$G_{\mathrm{A}\uparrow}(\omega) = \left[\omega - \frac{1}{2}t_*^2 G_{\mathrm{A}\uparrow}(\omega)\right]^{-1} \qquad \text{P.} \tag{3.15b}$$

These are the equations for the corresponding t–J_z model on the Bethe lattice, for an AF and random spin background respectively (see e.g. [11, 41]); the t–J_z model itself is naturally equivalent for d^∞ to the t–J model since the spin excitations are purely Ising-like. We add in passing that a much more detailed asymptotic analysis, picking up constant terms $O(1/U)$, leads to the 'bare' ω in the denominators of equations (3.15) being replaced by $\omega + \omega_p$ and $\omega + \frac{1}{2}\omega_p$ respectively for the AF and P phases. These shifts, neglected in the brief analysis above, reflect simply the presence of the trivial charge terms in the t–J model (see equation (2.9)); they are irrelevant to our subsequent discussion.

Since the t–J limit emerges correctly in strong coupling, the present theory is thus asymptotically exact. Consider for example the P phase, noting that for the $U = 0$ non-interacting limit the Green function $G_0(\omega) = \mathrm{Re}\, G_0(\omega) - i\pi\, \mathrm{sgn}(\omega) D_0(\omega)$ is given by

$$G_0(\omega) = \left[\omega - \frac{1}{2}t_*^2 G_0(\omega)\right]^{-1} \qquad U = 0 \tag{3.16a}$$

whence the non-interacting spectrum

$$D_0(\omega) = \frac{1}{\pi t_*}[2 - (\omega/t_*)^2]^{1/2} \qquad |\omega| \leqslant \sqrt{2}t_* \tag{3.16b}$$

is a semi-ellipse with full width $2\sqrt{2}t_*$. From equation (3.15b) this is also precisely the spectral density for $G_{\mathrm{A}\uparrow}(\omega)$ in the lower Hubbard band. And since $G(\omega \simeq 0) = \frac{1}{2}G_{\mathrm{A}\uparrow}(\omega)$ as above, the total lower-Hubbard-band spectrum in strong coupling is $D_L(\omega) = \frac{1}{2}D_0(\omega)$; see also section 2.1.2 (the normalization factor of $\frac{1}{2}$ naturally reflects the fact that the remaining half of the single-particle spectrum occurs in the upper Hubbard band centred on $\omega = U$, namely $D_U(\omega) = \frac{1}{2}D_0(U - \omega)$). Note further that the Feenberg ('medium') and interaction self-energies contribute equally to the $\frac{1}{2}t_*^2 G_{\mathrm{A}\uparrow}(\omega)$ denominator in equation

(3.15*b*) for the P phase. Physically, this reflects the fact discussed in section 2.1.2 that, upon adding a σ-spin electron to a site, it is equally probable for either the added σ-spin or the $-\sigma$-spin electron already present to hop off the site. At the UHF/HIII level, in contrast, only the former can by construct occur ($\Sigma_{\alpha\sigma} = 0$): the analogue of equation (3.15*b*) is then $G^0_{A\uparrow}(\omega) = [\omega - \frac{1}{4}t_*^2 G^0_{A\uparrow}(\omega)]^{-1}$, producing an incorrect strong-coupling bandwidth of $2t_*$ as argued physically in section 2.1.2.

The AF case itself is discussed further in the following section since, in contrast to the P phase, the *approach* to strong coupling is subtle and physically revealing. Here we simply add the following points. (i) In contrast to in the P phase, the $\frac{1}{2}t_*^2 G_{A\uparrow}(\omega + \omega_p)$ denominator in equation (3.15*a*) for the AF stems *solely* from the interaction self-energy $\Sigma_{A\uparrow}(\omega)$. Thus at the UHF/HIII level atomic limit behaviour arises (incorrectly)—that is, $G^0_{A\uparrow}(\omega) = 1/\omega$—as argued physically in section 2.1.1. (ii) Although obtained explicitly for the Bethe lattice, equation (3.15*a*) holds equally for the hypercubic lattice in strong coupling. This is because retraceable paths, which by construct are the only self-energy paths for a Bethe lattice, are for the d^∞ hypercube also the only paths which restore the Néel spin configuration; see also [29].

3.3. Results

At finite U the basic self-consistency equations, (3.2) and (3.8*a*), are solved numerically. We consider first the AF phase.
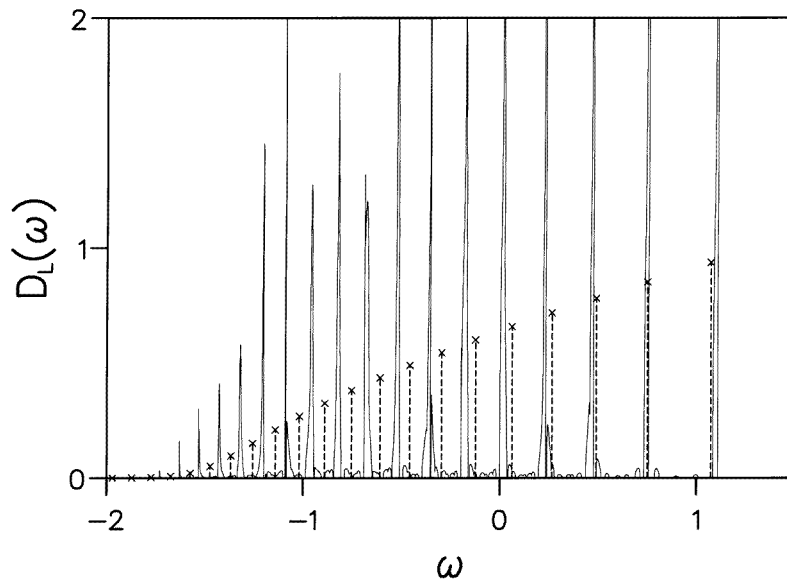

Figure 8. The lower-Hubbard-band spectrum $D_L(\omega)$ versus ω (in units of t_*) for the AF phase (Bethe lattice) at $U/t_* = 10$; the Fermi level lies at $U/2 = 5$. The corresponding t–J_z limit spectrum is also shown, as discussed in the text.

3.3.1. Antiferromagnet. For $U/t_* = 10$, figure 8 shows the resultant lower Hubbard band, $D_L(\omega) = \pi^{-1} \operatorname{Im} G(\omega)$; from particle–hole symmetry the upper band follows by reflection about the Fermi level, $D_U(\omega) = D_L(U - \omega)$.

For the same ω_p-value (figure 3(a), inset), figure 8 shows also the corresponding t–J_z limit spectrum from equation (3.15a). As is well known [11] the t–J_z spectrum is discrete (and to illustrate the relative intensities is thus shown with the height proportional to the integrated weight). Physically, this reflects the fact that the hole is pinned by the string of spin flips that its motion creates, leading therefore to spatially localized single-particle excitations and hence a discrete spectrum; mathematically, it is reflected in the convergence of the continued fraction implicit by iteration of equation (3.15a).

Although the $U/t_* = 10$ spectrum evidently bears a close resemblance to its t–J_z counterpart, it is by contrast continuous. This persists for any finite U: with increasing interaction strength the individual sub-bands in $D_L(\omega)$ centre ever more closely on their t–J_z counterparts, and their integrated spectral weights tend to those of the t–J_z limit; but they retain a finite width, reflecting delocalization of the hole. The peculiarities for $U \to \infty$ are further evident in the t–J_z model itself, equation (3.15a). For any $\omega_p > 0$ the t–J_z spectrum is discrete, while for $\omega_p = 0$ (as in equation (3.15b) for the P phase) the spectrum is continuous: the point $\omega_p = 0$ thus corresponds to a transition from localized to extended single-particle excitations, and since $\omega_p \to t_*^2/U$ as $U \to \infty$ it is clear that $U = \infty$ is a singular point.

While the physical mechanism leading to delocalization of the hole at any finite U is not of course inherent in the t–J_z model equation (3.15a) itself, it is readily inferred. Consider the Néel spin configuration and imagine removing, say, an ↑-spin electron from an A-type site, i. The nearest neighbours (NN) to any ↑-spin site all all ↓-spins. Hence to leading order in U—the t–J_z limit—the hole initially moves via a NN ↓-spin electron hopping onto site i, creating thereon a spin flip (with an associated exchange energy penalty); and the subsequent motion of the hole via such a correlated sequence of alternating NN spin hops, in leaving behind a string of upturned spins, would by itself render the hole spatially confined.

At large but finite U there is however a small but non-vanishing probability amplitude, of order t_*^2/U, for an ↑-spin electron on a *second* NN site, also A-type, to hop to site i via an intervening ↓-spin site: the hole thus moves two lattice spacings, to the second NN A-type site. Unlike the 't–J_z processes' above, this does not entail a spin flip with concomitant exchange penalty: the hole moves freely.

This mechanism evidently leads to hole delocalization and, in tandem with the t–J_z processes, produces the strong-coupling spectrum. Its formal origins reside in the passage from equation (3.10a) to equation (3.15a) for the AF lower Hubbard band in strong coupling, where the Feenberg part of the self-energy $S_{A\uparrow}(\omega) = \frac{1}{2}t_*^2 G_{B\uparrow}(\omega)$ was neglected. As seen readily from the asymptotics of section 3.2, the leading corrections to Im $S_{A\uparrow}(\omega \approx 0)$ are Im $S_{A\uparrow}(\omega) = (t_*^2/2U)^2$ Im $G_{A\uparrow}(\omega)$. It is these that embody the delocalization described above, and lead to spectral broadening (contributions to Re $S_{A\uparrow}(\omega)$ are O(1/U) and lead simply to residual energy shifts). Further, note that since the energetic width of the spectral broadening is naturally the smallest energy scale in strong coupling, the principal effect on the 'bare' t–J_z spectrum is a small resonant broadening of the individual t–J_z lines. This is seen in figure 8, and becomes clearer still with further increasing U.

To our knowledge, the above mechanism is the only one which can lead to hole delocalization for the d^∞ AF in strong coupling; and for the reasons already given in section 3.2 this applies to the hypercubic as well as the Bethe lattice. For finite d it is for example well known that Trugman paths [42] lead to hole delocalization for the hypercubic lattice, but such processes are O(d^{-4}) and do not therefore contribute in d^∞ [29].

As U is decreased, the spectra continue to exhibit essentially strong-coupling behaviour down to modest interaction strengths of $U/t_* \sim 2$–3, and can thus be understood

qualitatively starting from the t–J_z limit. This is shown in reference [7] (see e.g. figure 3(b) therein).
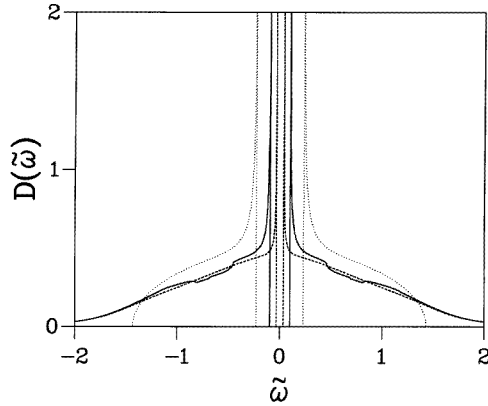

Figure 9. The full spectrum $D(\tilde{\omega})$ versus $\tilde{\omega} = \omega - U/2$ (in units of t_*) for the AF phase (Bethe lattice) at $U/t_* = 1$. Dotted line: the UHF spectrum; full line: from the present theory; dashed line: with $|\mu|$ and Π_{AA}^{-+} further renormalized as described in the text.

With further decreasing U however, the spectra evolve continuously to a weak-coupling form that shows no trace of remnant t–J_z-like behaviour. The spectral gap closes only in the non-interacting limit whence, correctly, the system is an AF insulator for all $U > 0$. The full spectrum $D(\tilde{\omega}) = D_L + D_U$ is shown in figure 9 for $U/t_* = 1$, together with the corresponding UHF spectrum to which (as one expects) it is qualitatively closer, although the single-particle gap Δ_g is reduced to 0.42 of the UHF gap $\Delta = U|\mu_0|$.

Two further renormalizations have been performed to check the veracity of the above results. First, note that although the Green functions have been obtained self-consistently via equations (3.2), (3.8), the single-particle propagators occurring in the RPA Π_{AA}^{-+} that enters the self-energy kernel equation (3.8a) are themselves bare UHF propagators; see figure 2. To ensure that the theory is robust, we have thus additionally renormalized the single-particle lines entering Π_{AA}^{-+} in terms of both the (self-consistent) full Green functions $G_{\alpha\sigma}$ and the host Green functions $\mathcal{G}_{\alpha\sigma}$. The results in either case differ only quantitatively, and at low U, from those just described; see also below.

The second renormalization concerns the local moment $|\mu|$ which, in the calculations above, has been set to its UHF value $|\mu_0|$. In weak coupling, van Dongen [43] has examined perturbatively the Néel temperature and the moment magnitude $|\mu|$ (the order parameter) for the d^∞ hypercubic lattice, and has shown that even for $U \to 0+$ these are reduced by a factor q of order unity ($q \simeq 0.28$ [43]) below their corresponding UHF values. The present theory is not of course perturbative (e.g. the emergence of the AF spin-flip scale is intrinsically non-perturbative), but it is certainly closer in spirit to the work of van Dongen to renormalize the moment beyond the UHF level. This is quantitatively important at low U, and is achieved by requiring that $|\mu|$ be determined fully self-consistently via (cf. equation (2.5))

$$|\mu| = \int_{-\infty}^{0} d\tilde{\omega} \, \left[D_{A\uparrow}(\tilde{\omega}) - D_{A\downarrow}(\tilde{\omega}) \right] \tag{3.17}$$

where $D_{A\sigma}$ is the full (as opposed to the UHF) spectral density.

For illustration, figure 9 shows the Bethe lattice spectrum at $U/t_* = 1$, obtained with both $|\mu|$ and Π_{AA}^{-+} renormalized (the latter in terms of the full Green functions). The gap Δ_g is further diminished, the ratio $g = \Delta_g/\Delta$ being ~ 0.15; while the local moment $|\mu|$ is likewise reduced below its UHF counterpart, such that $m = |\mu|/|\mu_0| \sim 0.39$. It is unfortunately not feasible to obtain numerically accurate estimates of g and m as $U \rightarrow 0$ (since $|\mu|$ and Δ rapidly become exponentially small). But for $U/t_* = 1$ the UHF moment itself is accurately represented by its asymptotic $U \rightarrow 0$ limit, $|\mu_0| = 8\sqrt{2}\exp[-\pi t_*/\sqrt{2}U - 1]$, so the above result for m may be reasonably close to its limiting value.

3.3.2. Paramagnet. To obtain correctly the strong-coupling limit for either phase is, as has been shown, fairly subtle. But in contrast to the AF, the *approach* to strong coupling for the paramagnetic phase is not. Figure 10 shows the full spectrum $D(\tilde{\omega}) = -\pi^{-1}\,\text{sgn}(\tilde{\omega})\,\text{Im}\,G(\tilde{\omega})$ ($\tilde{\omega} = \omega - U/2$) for the P phase at $U/t_* = 8$, 6 and 4, compared to the strong-coupling t–J_z limit from equation (3.15*b*). For $U/t_* = 8$, the strong-coupling limit has in practical terms been reached: the Hubbard bands are essentially symmetrically centred on $\tilde{\omega} = \pm U/2$ respectively, with widths $W \sim W_\infty = 2\sqrt{2}t_*$ and a band gap of $\Delta_g \sim \Delta_g^\infty = U - 2\sqrt{2}t_*$; even for $U/t_* = 6$ the departure from the asymptotic spectrum is relatively minor. With further decreasing U, however, the individual bands become increasingly asymmetric; and the gap tends to zero more rapidly than Δ_g^∞, signalling the collapse of the insulating phase. This we now discuss, adding that throughout the insulating regime the local moments are well developed ($|\mu| \gtrsim 0.95$), as in Mott's conception of a Mott insulator [44].

4. Destruction of the Mott insulator

Figure 11 shows the resultant band gap, $\Delta_g(U)$, for the paramagnetic insulator as a function of U/t_*. $\Delta_g(U)$ is found to vanish continuously at a critical $U_c = 3.41t_*$. Detailed numerical analysis shows the corresponding exponent to be unity:

$$\Delta_g(U) \sim (U - U_c)^\nu \qquad \nu = 1 \tag{4.1}$$

and we note that the width of the critical regime is quite narrow: the behaviour equation (4.1) is seen clearly for $(U - U_c) \lesssim 0.05t_*$, corresponding to gaps $\Delta_g(U) \lesssim 0.1t_*$.

The continuous closure of the gap is intimately connected to the divergence of the low-frequency dynamical characteristics of the system. Consider first the self-energy $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega})$. At frequencies $\tilde{\omega} \in [-\tilde{\omega}_+, \tilde{\omega}_+]$ inside the spectral gap ($\Delta_g = 2\tilde{\omega}_+$), $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega}) \equiv \tilde{\Sigma}_{A\uparrow}^R(\tilde{\omega})$ is pure real with a leading low-$\tilde{\omega}$ expansion:

$$\tilde{\Sigma}_{A\uparrow}^R(\tilde{\omega}) - A = B\tilde{\omega} \qquad \tilde{\omega} \rightarrow 0. \tag{4.2}$$

Here, $A \equiv \tilde{\Sigma}_{A\uparrow}(\tilde{\omega} = 0)$ ($= -\frac{1}{2}U|\mu| + \Sigma_{A\uparrow}(0)$, see equations (3.1), (3.8)), and is finite for all $U > U_c$ (see also below). We wish to find the behaviour of $B = -|B|$ as $U \rightarrow U_c$.

This is obtained by a scaling analysis. Defining $y = \tilde{\omega}/\tilde{\omega}_+$, it is found that as the gap closes ($\tilde{\omega}_+ \rightarrow 0$), $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega}) - A$ obeys the scaling form

$$\tilde{\Sigma}_{A\uparrow}^R(\tilde{\omega}) - A = \tilde{\omega}_+^\alpha f(y) \qquad \alpha = \frac{1}{2} \tag{4.3}$$

with exponent $\alpha = \frac{1}{2}$; i.e. for different values of U close to U_c, with correspondingly different gaps $\Delta_g(U) = 2\tilde{\omega}_+(U)$, the $\tilde{\omega}$-dependent functions $[\tilde{\Sigma}_{A\uparrow}^R(\tilde{\omega}) - A]/\tilde{\omega}_+^{1/2}$ plotted in terms of $y = \tilde{\omega}/\tilde{\omega}_+$, collapse to a 'universal' function $f(y)$. Four points should be noted
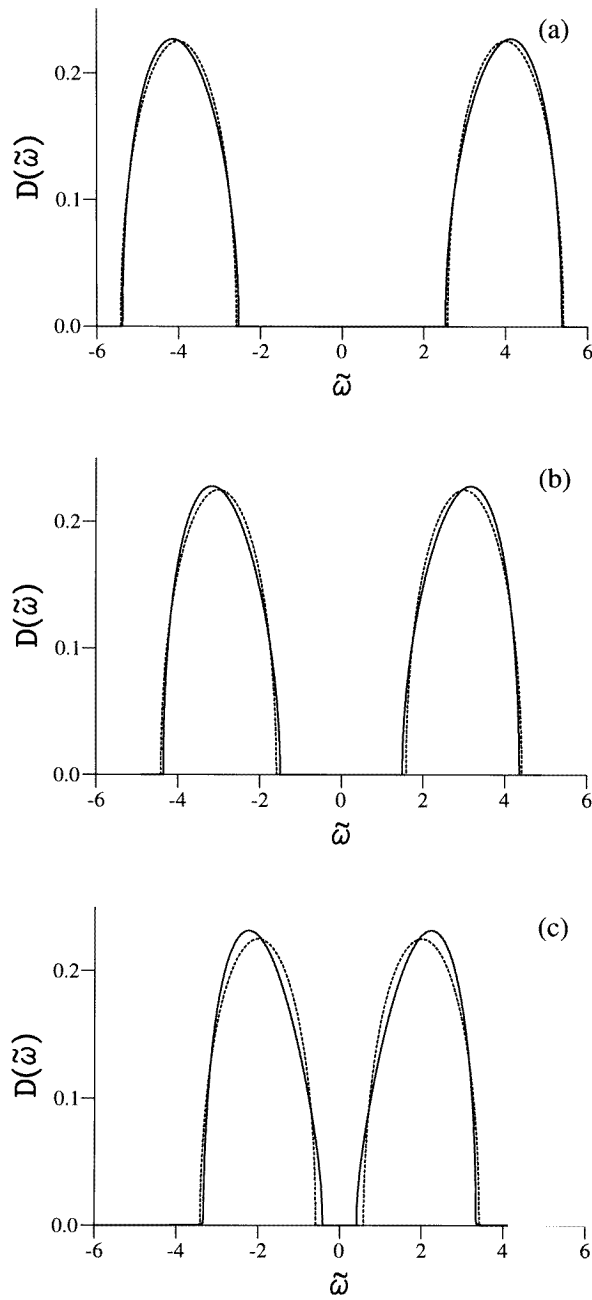
Figure 10. Full spectra $D(\tilde{\omega})$ versus $\tilde{\omega} = \omega - U/2$ (in units of t_*) for the P phase (Bethe lattice) at $U/t_* = 8$ (a), 6 (b) and 4 (c). The corresponding strong-coupling spectra are shown as dashed lines.

about the scaling behaviour. (i) Good scaling is found in practice for gaps $\Delta_g \lesssim 0.1t_*$, consistent with the critical regime found above for closure of the gap. (ii) The scaling is not confined to frequencies $y \ll 1$ well inside the spectral gap, but encompasses the region
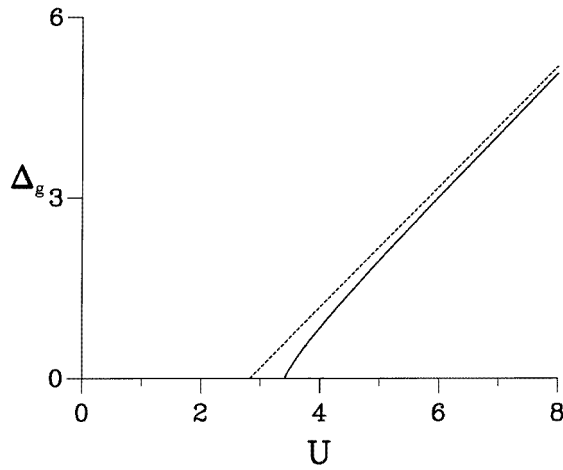
Figure 11. The resultant spectral gap $\Delta_g(U)$ versus U/t_* for the P insulator (Bethe lattice). The strong-coupling asymptote $\Delta_g^\infty = U - 2\sqrt{2}t_*$ is also shown (dashed line).

of non-zero spectral density ($|y| > 1$), certainly up to $|y| \sim 2$. (Similar scaling with $\alpha = \frac{1}{2}$ naturally occurs for $\text{Im}\,\tilde{\Sigma}_{A\uparrow}(\tilde{\omega})$, as follows from Kramers–Kronig; see also below.) (iii) In numerical terms the scaling analysis is sufficiently accurate to distinguish readily between an exponent of $\alpha = \frac{1}{2}$ and, e.g., $\alpha = \frac{1}{3}$. (iv) The scaling function $f(y)$ is a finite, well behaved function of $y = \tilde{\omega}/\tilde{\omega}_+$, with $f(y) \sim y$ for $y \to 0$ as is evident from equation (4.2).

From equations (4.3) and (4.2) it follows immediately that $|B| \sim \tilde{\omega}_+^{-1/2}$; i.e.

$$|B| \sim \Delta_g^{-1/2} \qquad \Delta_g \to 0 \tag{4.4}$$

or $|B| \sim (U - U_c)^{-1/2}$ from equation (4.1) (which we have confirmed by direct calculation of $B = (\partial \tilde{\Sigma}_{A\uparrow}(\tilde{\omega})/\partial \tilde{\omega})_0$).

The divergence of $|B|$ controls additionally the low-frequency behaviour of $\text{Re}\,G(\tilde{\omega}) = X(\tilde{\omega})$. From equation (3.5c), $X(\tilde{\omega}) = -X(-\tilde{\omega})$, whence its leading low-$\tilde{\omega}$ behaviour is

$$X(\tilde{\omega}) = \gamma_1 \tilde{\omega} \qquad \tilde{\omega} \to 0 \tag{4.5}$$

(with $\gamma_1 = -|\gamma_1|$). From equations (3.1)–(3.3) and (3.6b), $G(\tilde{\omega})$ may be written generally as

$$G(\tilde{\omega}) = \frac{1}{2}\left\{\left[\tilde{\omega} - \frac{1}{2}t_*^2 G(\tilde{\omega}) - \tilde{\Sigma}_{A\uparrow}(\tilde{\omega})\right]^{-1} + \left[\tilde{\omega} - \frac{1}{2}t_*^2 G(\tilde{\omega}) + \tilde{\Sigma}_{A\uparrow}(-\tilde{\omega})\right]^{-1}\right\}. \tag{4.6}$$

Using equations (4.2) and (4.5) on either side of (4.6) enables $|\gamma_1|$ to be related to $|B|$; the result is

$$|\gamma_1| = \frac{1 + |B|}{A^2 - \frac{1}{2}t_*^2}. \tag{4.7}$$

We find that $A^2 > \frac{1}{2}t_*^2$ for all $U \geqslant U_c$, whence the divergence of $|B|$ as $U \to U_c$ controls that of $|\gamma_1|$:

$$|\gamma_1| \sim \Delta_g^{-1/2} \qquad \Delta_g \to 0. \tag{4.8}$$

This is further confirmed by a scaling analysis of $X(\tilde{\omega})$ itself. In direct analogy to that for $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega})$ above, $X(\tilde{\omega})$ is found to satisfy the scaling form

$$X(\tilde{\omega}) = \tilde{\omega}_+^\alpha \, x(y) \qquad \alpha = \frac{1}{2} \tag{4.9}$$

with $x(y) = -x(-y)$, from which equation (4.8) in particular follows. For $|y| > 1$ the corresponding spectrum $D(\tilde{\omega})$ likewise shows the same scaling form as expected, with $D(\tilde{\omega}) \sim \tilde{\omega}_+^{1/2} [y-1]^{1/2} = [\tilde{\omega} - \tilde{\omega}_+]^{1/2}$ for $y \sim 1$ close to the lower edge of the upper Hubbard band; this, combined with the spectral representation of $|\gamma_1|$,

$$|\gamma_1| = 2 \int_{\tilde{\omega}_+}^{\infty} d\tilde{\omega} \, \frac{D(\tilde{\omega})}{\tilde{\omega}^2} \tag{4.10}$$

leads again to equation (4.8).

It is instructive to compare the above results with those obtained from both the simple HIII approximation discussed in section 2.1.2, and with the resonance broadening contributions [13] additionally included, which we refer to as HIII'. For HIII, $A = -\frac{1}{2}U$ and $|B| = 0$—the approximation is purely static. Equation (4.6) becomes a cubic for $G(\tilde{\omega})$, leading as is well known to $\Delta_g \sim (U - U_c)^{3/2}$ [13]. As is clear from equation (4.7) with $|B| = 0$, the transition occurs when $A^2(U_c) = \frac{1}{2}t_*^2$, i.e. $U_c = \sqrt{2}t_*$, and $|\gamma_1| \sim (U - U_c)^{-1} \sim \Delta_g^{-2/3}$. The HIII' approximation can also be shown to be of the form of equation (4.6), but with a $\tilde{\omega}$-dependent $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega})$ given by $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega}) = -U/2 + t_*^2 G(\tilde{\omega})$ at low frequencies (which is sufficient for analysing the critical behaviour); so $A = -U/2$ and $|B| = t_*^2|\gamma_1|$. Since $\tilde{\Sigma}_{A\uparrow}(\tilde{\omega})$ is a simple linear function of $G(\tilde{\omega})$, equation (4.6) again becomes a cubic for $G(\tilde{\omega})$; and, as for HIII, the gap exponent $\nu = \frac{3}{2}$ [13]. Equation (4.7) with $|B| = t_*^2|\gamma_1|$ yields $|\gamma_1| = (A^2 - \frac{1}{2}t_*^2)/(A^2 - \frac{3}{2}t_*^2)$. The transition thus occurs when $A^2(U_c) = \frac{3}{2}t_*^2$, i.e. $U_c = \sqrt{6}t_*$ as is well known [13]; and, again, $|\gamma_1| \sim (U-U_c)^{-1} \sim \Delta_g^{-2/3}$.

HIII and HIII' are thus both in the same universality class, reflected more generally in the fact that in either case scaling of the form of equation (4.9) can be shown to hold, but with an exponent of $\alpha = \frac{1}{3}$. Gros [45] has recently extended Hubbard's hierarchical equation-of-motion decoupling scheme to higher order. The critical exponents are unchanged from those of HIII/HIII'; and the value of U_c itself is barely changed from its HIII' value of $U_c/t_* \simeq 2.45$. From the above discussion it is apparent that the present theory belongs to a different universality class from that of HIII or its extensions.

In direct analogy to the AF phase discussed in section 3.3, we have tested the robustness of our results by further self-consistently renormalizing single-particle lines in the polarization propagator $\Pi_{AA}^{-+}(\omega) = {}^0\Pi_{AA}^{-+}/(1 - U\,{}^0\Pi_{AA}^{-+})$ that enters the self-energy kernel, equation (3.8a). To illustrate what this involves, consider renormalizing ${}^0\Pi_{AA}^{-+}$ (and hence Π_{AA}^{-+}) in terms of the self-consistent host Green functions $\mathcal{G}_{\alpha\sigma}$. The resultant ${}^0\Pi_{AA}^{-+}(\omega)$ is then given generally by equation (2.12), with the bare (UHF) Green functions $G_{A\sigma}^0$ now replaced by $\mathcal{G}_{A\sigma}$. For $\omega = 0$ in particular, equation (2.14) likewise holds, but with the bare $D_{A\sigma}^0(\tilde{\omega})$ replaced by the renormalized spectral densities $\mathcal{D}_{A\sigma}(\tilde{\omega}) = -\pi^{-1}\,\mathrm{sgn}(\tilde{\omega})\,\mathrm{Im}\,\mathcal{G}_{A\sigma}(\tilde{\omega})$; i.e.

$$U\,{}^0\Pi_{AA}^{-+}(\omega = 0) = \frac{1}{|\mu|} \int_{-\infty}^{0} d\tilde{\omega} \, [\mathcal{D}_{A\uparrow}(\tilde{\omega}) - \mathcal{D}_{A\downarrow}(\tilde{\omega})]. \tag{4.11}$$

As discussed in sections 2, 3 the key feature of the paramagnetic insulator is the zero-frequency spin-flip scale. To preserve this, the local moment $|\mu|$ in equation (4.11) is itself renormalized to ensure that at each step of the self-consistent iteration scheme $U\,{}^0\Pi_{AA}^{-+}(\omega = 0) = 1$ (and we note that throughout the entire insulating regime, the resultant

moment $|\mu|$ is also self-consistent in the sense of equation (3.17) to <1% accuracy). The results of this further renormalization are found to differ negligibly from those that we have reported above.

Finally, to demonstrate the importance of the $\omega_s = 0$ spin-flip scale, we have eliminated it: both by (a) neglecting its contribution to $\Sigma_{A\uparrow}(\tilde{\omega})$ in equation (3.8b), retaining only $\Sigma_{A\uparrow}^{\text{Stoner}}(\tilde{\omega})$; and (b) replacing Π_{AA}^{-+} by $^0\Pi_{AA}^{-+}$ in the self-energy kernel equation (3.8a). Results obtained from (a) and (b) are very similar, but differ qualitatively from those reported above. In particular, although the self-energy remains ω-dependent, the resultant critical behaviour is found to be that of HIII/HIII′—the gap closes continuously, but with an exponent $\nu = \frac{3}{2}$. This points clearly to the necessity of including the $\omega_s = 0$ spin-flip scale throughout the entire insulating phase—not only in achieving the correct strong-coupling limit (as in section 3.2), but also in describing the destruction of the insulating state.

5. Discussion

We now discuss the present work, particularly in relation to the iterated perturbation theory (IPT) approach [21 38–40, 46], use and application of which has been extensive [5]. Although our theory of the Mott–Hubbard insulating phases, with its explicit emphasis on local moments, is conceptually and technically distinct from IPT, some general points of marked contrast are evident.

For the antiferromagnetic phase we have emphasized the importance of the ω_p spin-flip scale, inclusion of which is necessary to obtain even qualitatively reasonable results throughout essentially the entire range of interaction strengths, and in particular to recover exact strong-coupling asymptotics. However, IPT does not appear to capture the AF spin-flip scale, presumably because it omits repeated particle–hole interactions of the sort shown in figure 2 (which, as in section 2.2.1, are required to pick up the spin flip). This is seen, for example, from the known inability of IPT to describe correctly the U-dependence of the Néel temperature [5], particularly in the 'Heisenberg' regime.

For the paramagnetic insulator, the results of section 4 also disagree qualitatively with those obtained from IPT; see in particular [40] and the review [5]. Within IPT the paramagnetic insulating solution is found to break down *discontinuously* (at a critical $U_{c1} = 3.67t_*$, where the IPT gap $\Delta_g(U_{c1}) \sim 0.3t_*$), and $|\gamma_1|$ (equation (4.10)) remains *finite* at the transition.

The same authors [12] have recently examined the insulator via exact diagonalization (ED) on clusters of $n_s = 3$, 5 and 7 sites, extrapolated to $n_s \to \infty$ assuming $1/n_s$ scaling behaviour. The resultant data suggest a continuous closure of the gap at a $U_{c1}/t_* = 3.04 \pm 0.35$ and are consistent with $\Delta_g(U) \sim (U - U_c)$; see also [5]. Further, and independently of the gap analysis, the behaviour of $|\gamma_1|$ has also been examined by ED [12], noting (see equation (4.10)) that a divergence in $|\gamma_1|$ implies a continuous closure of the gap: $1/|\gamma_1|$ is found to show good scaling behaviour, and to scale to zero when n_s is extrapolated to ∞.

The present theory evidently agrees with the inferences drawn from ED. These concur with our predictions (section 4) that the gap closes continuously and with an exponent $\nu = 1$, that $|\gamma_1|$ diverges, and (less importantly) the value of U_c itself; note moreover that the ED gap [12] is in rather good agreement with the present work over a wide range of U. As described in section 4, inclusion of the $\omega_s = 0$ spin-flip scale is central in describing the destruction of the Mott insulator. That IPT appears unreliable close to U_c [5] thus suggests an incomplete inclusion of the effects of this spin scale—which cannot be entirely absent since IPT does give the correct strong-coupling spectrum [47]—although in physical terms

the origin of the spin-flip scale within IPT is not transparent.

To conclude, we have developed in this paper a theory for the $T = 0$ Mott–Hubbard insulating phases of the d^∞ Hubbard model, encompassing both the antiferromagnetic and paramagnetic insulators. The microscopic perspective that it affords hinges on the importance of low-energy scales for insulating spin-flip excitations. Their existence is physically natural within the explicit local moment picture intrinsic to the theory, and inclusion of them is required not only to obtain the strong-coupling limits of the single-particle spectra—which are captured exactly—but more generally to describe the entire insulating regimes, including for the paramagnetic phase in particular the destruction of the Mott insulator.
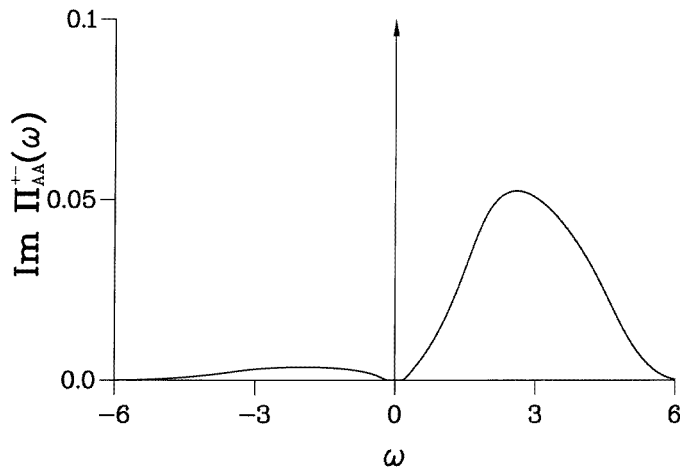

Figure 12. Im $\Pi_{AA}^{+-}(\omega)$ versus ω (in units of t_*) at $U/t_* = 3.5$ close to the boundary of the P insulating state, with renormalization as described in the text.

Let us also note what we have not considered: the metallic state of the paramagnetic phase. But a glimpse of what is required to describe the metal within the present framework is evident from figure 12. For $U/t_* = 3.5$, close to the critical U_c of section 4, this shows the spectral density of transverse spin excitations Im $\Pi_{AA}^{+-}(\omega)$ (obtained here, as described in section 4, with $^0\Pi_{AA}^{+-}$ renormalized in terms of the $\mathcal{G}_{A\sigma}$). The $\omega = 0$ spin-flip pole characteristic of the paramagnetic insulator is evident, and persists down to U_c. Clearly, however, the spectral edges of the Stoner-like bands are themselves approaching $\omega = 0$. This they do at $U = U_c$, and for $U < U_c$ in the metallic phase the insulating spin-flip *pole* at $\omega = 0$ is replaced by a *resonance* at a small non-zero frequency $\omega = \omega_K$, indicative of the Kondo-like physics known to dominate the correlated metal [5, 6]. Extension of the present approach for describing the metal, encompassing the Kondo spin scale in such a manner that the correlated state is correctly a Fermi liquid, will be described in a subsequent paper.

Acknowledgments

References

[1] Hubbard J 1963 *Proc. R. Soc.* A **276** 238
Gutzwiller M C 1963 *Phys. Rev. Lett.* **10** 59
Kanamori J 1963 *Prog. Theor. Phys.* **30** 275
[2] Lieb E H and Wu F Y 1968 *Phys. Rev. Lett.* **20** 1445
[3] Metzner W and Vollhardt D 1989 *Phys. Rev. Lett.* **62** 324
[4] Vollhardt D 1993 *Correlated Electron Systems* ed V J Emery (Singapore: World Scientific)
[5] Georges A, Kotliar G, Krauth W and Rozenberg M J 1996 *Rev. Mod. Phys.* **68** 13
[6] Pruschke T, Jarrell M and Freericks J K 1995 *Adv. Phys.* **44** 187
[7] A preliminary account of part of this work has been given in a recent letter:
Logan D E, Eastwood M P and Tusch M A 1996 *Phys. Rev. Lett.* **76** 4785
[8] Janiš V and Vollhardt D 1993 *Z. Phys.* B **91** 317
Janiš V, Mašek J and Vollhardt D 1993 *Z. Phys.* B **91** 325
[9] Li Y M and d'Ambrumenil N 1992 *Mod. Phys. Lett.* B **6** 1827
Li Y M and d'Ambrumenil N 1994 *Phys. Rev.* B **49** 6058
[10] Wermbter S and Czycholl G 1995 *J. Phys.: Condens. Matter* **7** 7335
[11] For a review with references, see
Dagotto E 1994 *Rev. Mod. Phys.* **66** 763
[12] Rozenberg M J, Kotliar G and Kajueter H 1996 *Phys. Rev.* B **54** 8452
[13] Hubbard J 1964 *Proc. R. Soc.* A **281** 401
[14] Falicov L M and Kimball J C 1969 *Phys. Rev. Lett.* **22** 997
[15] Müller-Hartmann E 1989 *Z. Phys.* B **74** 507
Müller-Hartmann E 1989 *Z. Phys.* B **76** 211
[16] Feenberg E 1948 *Phys. Rev.* **74** 206
[17] Economou E N 1983 *Green's Functions in Quantum Physics* (Berlin: Springer)
[18] Penn D R 1966 *Phys. Rev.* **142** 350
[19] Cyrot M 1972 *Phil. Mag.* **25** 1031
[20] Singh A and Tešanović Z 1990 *Phys. Rev.* B **41** 614
Singh A and Tešanović Z 1990 *Phys. Rev.* B **41** 11 457
[21] Georges A and Krauth W 1993 *Phys. Rev.* B **48** 7167
[22] Edwards D M 1993 *J. Phys.: Condens. Matter* **5** 161
[23] Brandt U and Mielsch C 1989 *Z. Phys.* B **75** 365
Brandt U and Mielsch C 1990 *Z. Phys.* B **79** 295
[24] van Dongen P G J and Vollhardt D 1990 *Phys. Rev. Lett.* **65** 1663
[25] van Dongen P G J 1992 *Phys. Rev.* B **45** 2267
Note that van Dongen's hopping matrix element, t, is related to t_* used here by $t_* = \sqrt{2}t$.
[26] Edwards D M and Hewson A C 1968 *Rev. Mod. Phys.* **40** 810
[27] Kawabata A 1975 *Prog. Theor. Phys.* **54** 45
[28] Brinkman W F and Rice T M 1970 *Phys. Rev.* B **2** 1324
[29] Metzner W, Schmit P and Vollhardt D 1992 *Phys. Rev.* B **45** 2237
[30] This holds for the Bethe lattice under consideration because all hole/doublon self-energy paths are retraceable; it does not hold for the d^∞ hypercube; see reference [29].
[31] Rozenberg M J, Zhang X Y and Kotliar G 1992 *Phys. Rev. Lett.* **69** 1236
The quoted UHF/HIII strong-coupling spectral width arising from equation (7) is in error, and should be $\sqrt{2}D$ (with $D = \sqrt{2}t_*$) as in equation (2.10b) of the present work.
[32] Tusch M A, Szczech Y H and Logan D E 1996 *Phys. Rev.* B **53** 5505
[33] Schrieffer J R, Wen X-G and Zhang S-C 1989 *Phys. Rev.* B **39** 11 663
[34] Kennedy T, Lieb E H and Shastry B S 1988 *Phys. Rev. Lett.* **61** 2582
Wang Y L, Shtrikman S and Callen H 1966 *Phys. Rev.* **148** 419
[35] Stanley H E 1971 *Introduction to Phase Transitions and Critical Phenomena* (Oxford: Oxford University Press)
[36] Jarrell M 1992 *Phys. Rev. Lett.* **69** 168
[37] Jarrell M and Pruschke T 1993 *Z. Phys.* B **90** 187

[38] Georges A and Krauth W 1993 *Phys. Rev. Lett.* **69** 1240
[39] Zhang X Y, Rozenberg M J and Kotliar G 1993 *Phys. Rev. Lett.* **70** 1666
[40] Rozenberg M J, Kotliar G and Zhang X Y 1994 *Phys. Rev.* B **49** 10 181
[41] Kane C L, Lee P A and Read N 1989 *Phys. Rev.* B **39** 6880
[42] Trugman S 1990 *Phys. Rev.* B **41** 892
[43] van Dongen P G J 1991 *Phys. Rev. Lett.* **67** 757
        van Dongen P G J 1994 *Phys. Rev.* B **50** 14 016
[44] Mott N F 1990 *Metal–Insulator Transitions* (London: Taylor and Francis)
[45] Gros C 1994 *Phys. Rev.* B **50** 7295
[46] Georges A and Kotliar G 1992 *Phys. Rev.* B **45** 6479
[47] This is a subtle point that can be proven analytically. It bears little relation to the fact [5] that IPT—like
        UHF/HIII—is exact in the atomic limit.