Springer Series in Solid-State Sciences 171

Adolfo Avella
Ferdinando Mancini *Editors*

Strongly Correlated Systems

Theoretical Methods

Springer

Springer Series in

SOLID-STATE SCIENCES 171

Springer Series in

SOLID-STATE SCIENCES

The Springer Series in Solid-State Sciences consists of fundamental scientific books prepared by leading researchers in the field. They strive to communicate, in a systematic and comprehensive way, the basic principles as well as new developments in theoretical and experimental solid-state physics.

Adolfo Avella
Ferdinando Mancini
Editors

Strongly Correlated Systems

Theoretical Methods

With 147 Figures

Springer

Editors
Dr. Adolfo Avella
Professor Dr. Ferdinando Mancini
Università degli Studi di Salerno, Dipartimento di Fisica "E.R. Caianiello"
Via Ponte don Melillo, 84084 Fisciano (SA), Italy
E-mail: avella@physics.unisa.it, mancini@physics.unisa.it

Springer Series in Solid-State Sciences ISSN 0171-1873
ISBN 978-3-642-21830-9 e-ISBN 978-3-642-21831-6
DOI 10.1007/978-3-642-21831-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011940501

To Maria Marinaro, who introduced us to the wonderful field of many-body physics

Preface

This volume is the first of three volumes reviewing the main analytical, numerical and experimental techniques specifically devised to study Strongly Correlated Systems. This editorial project builds upon the long-standing experience we have acquired in organizing the "Training Course in the Physics of Strongly Correlated Systems" in Vietri sul Mare (Salerno, Italy) since 1996 and our working scientific experience in the field. Running a school for advanced graduate students and junior post-docs, we realized that this field of condensed matter and solid state physics was missing adequate textbooks and that the whole Strongly Correlated Systems community would benefit by a systematic exposition of the field. The present volume consists of a series of monographs on the most relevant analytical methods currently used to tackle the hoary problem of correlations. Authors have been selected, consulted major experts in the field, among the most world-wide famous scientists who invented or greatly helped to improve/spread the specific method in the community. Each chapter presents the method in a pedagogical way and contains at least one case study where the method has proved to give a substantial leap forward in the knowledge and a very rich bibliography. The book is mainly intended for neophytes, who will find in one single volume all pieces of information necessary to choose and start learning an analytical method. Also more experienced researchers would benefit from this volume as they would gain a deeper understanding of what any single technique can really tell them and what cannot. Accordingly, the accent is more on the ideas behind (origins, pros/cons, perspectives, etc.) than on the technical details, which are left to the comprehensive bibliography.

We wish to thank all authors of this volume as they all joined this editorial project with enthusiasm and provided the whole community with what we hope will become a relevant resource for any researcher in the field as comprehensive and extended reference.

Salerno, October 2011

Adolfo Avella
Ferdinando Mancini

Contents

Foreword

Peter Fulde

It required the discovery of the high-temperature superconducting cuprates before the field of strongly correlated electrons obtained the attention it deserves. Shortly after their discovery it became clear that the normal state properties of these materials and moreover the appearance of superconductivity cannot be understood without better insight into the strong correlations prevailing in these systems [1]. Even before that time, a sizeable amount of research had been devoted to strong correlations. The Kondo effect, intermediate valence systems, and materials with heavy low-energy quasiparticles are areas which required to deal with them. But these efforts were small when compared with those which set in after high-T_c superconductivity was discovered. In a series of three volumes, of which the present is the first, an overall survey is made of the present status of the theory of strongly correlated electrons. Some aspects, like numerical methods, are the subject of the second volume in this series. The last one includes experimental techniques, which are used to study strongly correlated electrons.

When speaking about strongly correlated electrons, the first question which comes to one's mind is, how can we decide whether or not electrons in a given material are strongly correlated. In other words, how can we determine the strength of correlations provided that we know the full many-electron wavefunction $|\psi_0\rangle$ to a reasonable approximation? For that, a simple measure can be introduced, which is based on the degree of suppression of electronic charge fluctuations on different atomic sites compared with those which are obtained when instead of $|\psi_0\rangle$ the uncorrelated or Hartree–Fock counterpart $|\Phi_0\rangle$ is used [2]. This gives rise to a quantity $0 \leq \Sigma_A \leq 1$. It defines the strength of correlations at a site A, so that $\Sigma_A = 0$ is the limit of uncorrelated electrons, while $\Sigma_A = 1$ describes the strong

P. Fulde (✉)
Asia Pacific Center for Theoretical Physics, Pohang, Korea

Max-Planck-Institut für Physik komplexer Systeme, Dresden, Germany
e-mail: fulde@apctp.org; fulde@pks.mpg.de

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6,

correlation limit. More specifically, we define Σ_A through

$$\Sigma_A = \frac{\left\langle \Phi_0 \left| (\delta n_A)^2 \right| \Phi_0 \right\rangle - \left\langle \psi_0 \left| (\delta n_A)^2 \right| \psi_0 \right\rangle}{\left\langle \Phi_0 \left| (\delta n_A)^2 \right| \Phi_0 \right\rangle}, \tag{1.1}$$

where $\delta n_A = n_A^2 - \bar{n}_A^2$ with n_A denoting the electron number operator, e.g., of d electrons on site A and $\bar{n}_A$ its average value with respect to $|\Phi_0\rangle$ or $|\psi_0\rangle$. Note that definition (1.1) deviates slightly from the one given in [2]. Applied to the simple example of a H_2 molecule, one finds $\Sigma_{\mathrm{H}} \approx 0.1$, i.e., rather weak correlations. A wavefunction with a doubly occupied molecular orbital would give $\Sigma_{\mathrm{H}}(\mathrm{MO}) = 0$ while the Heitler-London wavefunction yields $\Sigma_{\mathrm{H}}(\mathrm{HL}) = 1$, since all ionic configurations are suppressed. It is interesting that the π electrons in graphene have $\Sigma_\pi \simeq 0.5$, showing that correlations are fairly strong in that system, while the d electrons on a Cu site in the Cu–O planes of the cuprates have $\Sigma_{\mathrm{Cu}} \simeq 0.8$. Thus, correlations are indeed strong here. But it is also apparent that theories which apply, e.g., a Gutzwiller projector to a wavefunction of electrons in the Cu–O planes overestimate the correlation strength. That projector eliminates charge fluctuations completely and is equivalent to setting $\Sigma_{\mathrm{Cu}} = 1$.

Electronic charge fluctuations at a given atomic site interfere with the formation of intra-atomic Hund's rule correlations. When $\Sigma_A \simeq 1$ the latter can fully form, while with decreasing values of Σ_A they become less important. This shows up in the behavior of $\langle \psi_0 | \mathbf{S}_A^2 | \psi_0 \rangle$ with changing Σ_A where $\mathbf{S}_A$ is the total spin at site A.

1 Ab-Initio Calculations

Ab initio electronic structure calculations are dominated by density functional theory (DFT). They have revolutionized that field. The subject is reviewed in this volume by Jones (see Chap. 8), one of the pioneers in the field. But DFT is a ground-state theory and therefore it is of little surprise that it fails, in particular for strongly correlated systems, when low-energy excitations are calculated from it. The failure is inherent and independent of any approximations which are made for the potential in the Kohn-Sham equation. This is seen by considering the simplest possible system of strongly correlated electrons, i.e., two electrons in two orbitals. As seen below, this model also shows a characteristic feature of strong correlations, namely the appearance of new low-energy scales. In weakly correlated systems the characteristic energy scale is given by the Fermi energy ϵ_{F}, or alternatively, by the electronic hopping matrix elements t_{ij} between neighboring sites i and j. Strong correlations cause additional, much lower energy scales.

The two orbitals are denoted by L (for ligand) and F (for $4f$, for example) and we assume the corresponding orbital energies to be ϵ_l and ϵ_f with $\epsilon_f < \epsilon_l$. Two electrons in the F orbital are expected to repel each other with an energy $U \gg \Delta\epsilon$ with $\Delta\epsilon = \epsilon_l - \epsilon_f$. When both electrons are in the L orbital, or when one electron is in the L and the other in the F orbital, we neglect their Coulomb interaction.

This is justified, if the ligand orbital has a large spatial extent. It applies when, for example, the ligand orbital is that of a large molecule or when it is a Bloch state. We assume that the hybridization t between the two orbitals is small, i.e., $t \ll \Delta\epsilon$. The Hamiltonian of the system is

$$H = \epsilon_l \sum_\sigma l_\sigma^+ l_\sigma + \epsilon_f \sum_\sigma f_\sigma^+ f_\sigma + t \sum_\sigma \left(l_\sigma^+ f_\sigma + f_\sigma^+ l_\sigma \right) + U n_\uparrow^f n_\downarrow^f . \tag{1.2}$$

The $l_\sigma^+(l_\sigma)$, $f_\sigma^+(f_\sigma)$ create (annihilate) electrons with spin σ in the L and F orbitals, respectively; furthermore $n_\sigma^f = f_\sigma^+ f_\sigma$. The Hamiltonian is so simple that it can be easily diagonalized. When $t = 0$, the ground state of the system has energy $E_0 = \epsilon_l + \epsilon_f$ and is fourfold degenerate. One electron is in the F orbital, while the other is in the L orbital. The four states are eigenstates of the total spin S and consist of a singlet $|\Phi_{S=0}\rangle$ and a triplet $|\Phi_{S=1}\rangle$. The system has one excited state of the form

$$|\Phi_{\text{ex}}\rangle = l_\uparrow^+ l_\downarrow^+ |0\rangle . \tag{1.3}$$

The energy of that state is $E_{\text{ex}} = 2\epsilon_l$. The state $f_\uparrow^+ f_\downarrow^+ |0\rangle$ is excluded from further consideration, since its energy is of order U and we assume $U \to \infty$.

When the hybridization is turned on, the singlets $|\Phi_{S=0}\rangle$ and $|\Phi_{\text{ex}}\rangle$ are coupled, while the $S = 1$ states $|\Phi_{S=1}\rangle$ remain unchanged. The coupling leads to the two eigenvalues

$$\begin{aligned} \tilde{E}_0 &= E_0 - \frac{2t^2}{\Delta\epsilon}, \\ \tilde{E}_{\text{ex}} &= E_{\text{ex}} + \frac{2t^2}{\Delta\epsilon}. \end{aligned} \tag{1.4}$$

For small values of t there is a low-lying triplet excitation above the singlet ground state. One can attach a characteristic temperature $T^* = 2t^2/(k_B \Delta\epsilon)$ to the energy gain associated with the singlet formation. It is by a factor $t/\Delta\epsilon$ smaller than the energy scale set by the hopping matrix element t and is an example of the new low-energy scales, which are generated by strong correlations. The same system can be treated using the density functional theory. In fact, that is quite interesting to do so, because one can derive explicitly the exact exchange-correlation potential $v_{\text{xc}}[\rho]$ as function of the density. When the 2×2 Kohn–Sham equation is solved, one finds that the energy difference between the two eigenvalues should *not* be interpreted as the excitation energy, since it is of order t instead of T^*. These findings hold irrespective of approximations to the functional [3].

With the above pointed out, it is clear that treatments of low-energy excitations by Kohn–Sham equations have no sound basis. Nevertheless, semi-empirical methods based on density functional theory have had numerous successes. An excellent example is the renormalized band-structure method [4]. It is the only one which has been able to make detailed predictions for the effective mass anisotropies at

the Fermi surface of heavy-quasiparticle systems. The idea hereby is to formulate the quasiparticle dispersions in terms of phase shifts like in an independent electron approach. For these phase shifts, one is using the ones obtained from an LDA, i.e., a local approximation to the density functional, except for those of the strongly correlated electrons. For Ce intermetallic compounds like $CeRu_2Si_2$, $CeSn_3$, $CeCoIn_5$, etc., these are the f electron phase shifts at the Ce sites. For the latter, a phenomenological ansatz is made. It has the effect of reducing the bare bandwidth to a renormalized one of order $k_B T^*$. Take $CeRu_2Si_2$ as an example. The phase shifts near the Fermi energy ϵ_F are

$$\{\eta_l^A(\epsilon)\} = \left\{\eta_l^{\mathrm{Ce}}(\epsilon), \eta_{l'}^{(\mathrm{Ru})_\nu}(\epsilon), \eta_{l''}^{(\mathrm{Si})_\mu}(\epsilon)\right\}, \quad \nu;\mu = 1,2, \tag{1.5}$$

where l, l', l'' are angular momenta and the indices ν, μ count different atoms within the unit cell. Except for $\eta_{l=3}^{\mathrm{Ce}}(\epsilon)$ all other phase shifts are assumed to be given by the LDA. Regarding the f phase shifts at a Ce site, only the one with $j = 5/2$ (Hund's rule multiplet) is relevant and, more specifically, with the symmetry of the crystalline field ground-state doublet, i.e., $\eta_\tau^{\mathrm{Ce}}(\epsilon)$ with $\tau = 1,2$. It can be parameterized by the resonant form

$$\eta_\tau^{\mathrm{Ce}}(\epsilon) = \arctan \frac{\Gamma}{\tilde{\epsilon} - \epsilon} \tag{1.6}$$

with two parameters Γ and $\tilde{\epsilon}$. One of them is fixed by the f-electron number at a Ce site, while the remaining one is fixed by requiring that the large γ coefficient of the low temperature specific heat $C = \gamma T$ is reproduced. As mentioned earlier, with these phase shifts not only a Fermi surface but also the strongly anisotropic quasiparticle masses can be determined. They agree well with experiments.

Another extension of density functional theory which is often used is the LDA+U method. Here the LDA is supplemented by adding an on-site Coulomb interaction U and exchange interaction J term to the LDA energy. For example, for d electrons it is

$$E = E_{\mathrm{LDA}} + \frac{U}{2}\sum_{lij\sigma} \delta n_{i\sigma}(l)\delta n_{j-\sigma}(l) + \frac{(U-J)}{2}\sum_{li(\neq j)\sigma} \delta n_{i\sigma}(l)\delta n_{j\sigma'}(l), \tag{1.7}$$

where l is a site index while i and j are d orbital indices. Furthermore, $\delta n_{i\sigma}(l) = n_{i\sigma}(l) - n_0(l)$ where $n_0(l) = n_d(l)/10$ and $n_d(l)$ is the total d electron number at site l. The LDA is an orbital-independent molecular-field approximation and therefore an inclusion of the deviations $\delta n_{i\sigma}(l)$ allows for an improved treatment of correlations. The potential entering the Kohn-Sham equation is obtained from $\delta E/\delta n_{i\sigma}(l)$ as

$$V_{i\sigma}^{\mathrm{eff}}(l) = V_{\mathrm{LDA}} + U\sum_j \delta n_{j-\sigma}(l) + (U-J)\sum_{j\neq i} \delta n_{j\sigma}(l), \tag{1.8}$$

which shows that results different from LDA are obtained if the spin orbitals are differently populated, i.e., when $\delta n_{i\sigma}(l) \neq 0$. An unequal population is favored by a large Coulomb interaction U like in any unrestricted mean-field approximation. As explained earlier, this way charge fluctuations are suppressed and correlation energies (not wavefunctions) are improved. There has also been an approach developed in which the LDA is used for computing Wannier functions and Coulomb parameters as input for a multiband Hubbard Hamiltonian. The latter is treated by a generalized tight-binding method, i.e., one which combines the exact diagonalization of an isolated cluster, i.e., a unit cell, with a perturbation treatment of the intercluster hopping and interactions [5] (see Chap. 4.4). The aim has been to find, e.g., the size of the gap in La_2CuO_4, which in LDA is absent.

Another hybrid method is the LDA+DMFT [6]. The dynamical mean-field theory (DMFT) [7–9] is a dynamical coherent potential approximation (DCPA) [10] which was stimulated by work on the Hubbard model in infinite dimensions [11, 12] (see Chap. 6.5). While the standard coherent potential approximation (CPA) introduced by Hubbard in connection with his Hamiltonian (see below) reduces at temperature $T = 0$ to a self-consistent field (SCA), i.e., Hartree–Fock theory, the DMFT (or DCPA) contain correlation effects in that limit. A site of an infinite lattice with, e.g., d electrons is treated as an impurity in a medium for which an LDA calculation has been done. The electronic self-energy $\Sigma(\omega)$ at the impurity site is computed and the medium is self-consistently modified until the self-energy of the impurity coincides with that of the medium. A shortcoming of the DMFT and DCPA is that only on-site correlations are treated, i.e., any $\mathbf{k}$ dependence of the self-energy is neglected. From quantum chemical calculations it is well known that intersite correlations are important and must be treated if one is interested in quantitative results.

Extensions of the original DMFT to a cluster DMFT [13] go in the right direction, but the clusters one can treat are rather small. Therefore, one cannot distinguish, e.g., between short-range AF correlations and long-range AF order. One possible way of including properly the short-ranged correlations is by using the method of increments in connection with a self-consistent projection operator method. This allows $\Sigma(\mathbf{k}, \omega)$ to be calculated with rather high accuracy. These approaches are still at their infancy as far as realistic calculations for specific materials are concerned, but some promising results have been obtained.

A rather different approach to strongly correlated electron systems is based on wavefunction methods [2]. They are combined with quantum chemical techniques and provide a rigorous theoretical framework for addressing the correlation problem which avoids any uncontrolled approximations. Many-body wavefunctions can be explicitly constructed at levels of increasing sophistication and accuracy. But in order to ensure size extensivity of the modifications induced by correlations the wavefunctions have to be formulated using cumulants. Standard quantum chemical methods thus offer a systematic path to converged results. They provide the right framework for coping with issues like a rigorous treatment of the ubiquitous short-range correlations and of a realistic representation of the crystalline environment. The way to proceed is to cut out a finite atomic cluster C from the infinite solid which is large enough to describe the crucial short-range correlations. Partially

filled d-electron shells require a multiconfiguration representation of the correlated many-electron wavefunction, which is achieved using the complete-active-space (CAS) self-consistent field (CASSCF) method. This way, strong correlations can be very well accounted for. The crystalline environment is described by an effective one-electron potential which is extracted from prior Hartree–Fock calculations for the periodic system. Remember that Hartree–Fock calculations describe charge distributions quite well. They are robust against correlation corrections, even when the latter are strong. With this approach the ground state of such strongly correlated systems as $LaCoO_3$ and LiFeAs was determined as well as the Zhang-Rice-like electron removal band for CuO_2 planes in La_2CuO_4 [14].

Wavefunction-based quantum chemical calculations for strongly correlated electrons are a field with high potential for the future. They have been somewhat neglected, because they require investments in program development and time, features not particularly favored by research funding. However, concerning actual calculations, they are the best one can do to obtain insight into the most important microscopic processes.

2 Model Hamiltonians

From Sect. 1 it is apparent that ab-initio calculations for strongly correlated electron systems are still at their beginning. Therefore, simplifying model Hamiltonians are very helpful in order to unravel physical effects caused by them. The one studied most is the Hubbard Hamiltonian [15]

$$H = -t \sum_{\langle ij \rangle \sigma} c^{+}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}. \tag{1.9}$$

It has one orbital associated with each site i and the Hamiltonian contains a hopping term between nearest-neighbor sites and an on-site Coulomb repulsion term U. The simplifications made by this ansatz as compared with a quantum chemical ab-initio Hamiltonian are enormous. But nevertheless, important insight can be gained by studying (1.9), in particular, when a system is at or close to half filling. Various techniques have been applied to study the Hubbard model with emphasis on two dimensions (2D). This is, because claims have been made that a 2D Hubbard model contains all of the important physics of high-T_c superconductivity.

A large role is being played by cluster approximations. They replace the infinite lattice by a finite cluster which is often embedded in an effective medium. In view of the short range of the correlation hole of an electron, that seems fine even for an infinite system, provided the correlation hole is properly constructed. In order not to limit for computational reasons the cluster to a too-small size, the method of increments can be used as a tool. It can be looked at as an expansion of a (cumulant) scattering matrix in terms of one, two, three etc. site clusters. But that is often not done and in that case translational invariance within a cluster is

missing. Established approaches within quantum cluster theory are the dynamical cluster approximation [16, 17] (see Chap. 8.8), the cluster DMFT mentioned earlier [18] (see Chap. 10.7) and the cluster perturbation theory [19] (see Chap. 7.9). In the latter case a perturbation expansion in terms of the hopping matrix element t is combined with the exact diagonalization of small clusters. The dynamical cluster approximation has the advantage of ensuring translational invariance within a cluster. In that respect it differs from the cluster DMFT in which this symmetry is violated. Otherwise, both methods are similar in many aspects. An incremental cluster expansion of the self-energy or scattering matrix has been applied in a fully self-consistent projection operator approach [20].

One can try to find a formalism from which all the different cluster methods derive as special cases, at least as long as they are based on Feynman diagrams. The hope is that this allows for the development of improved cluster methods. Such an attempt has been made with the Self-Energy Functional Theory [21] (see Chap. 9.5). It starts from the generating functional $\Phi[G, U]$ of the Green's function G. The self-energy is required to be given by the derivative of Φ with respect to G, i.e., $\Sigma = \delta\Phi[G]/\delta G$. From the work of Baym and Kadanoff it is known that conservation laws are obeyed if $\Phi[G, U]$ is calculated from all distinct, connected, and closed skeleton diagrams expressed in G and U. One may then specify to which approximations for Φ a given cluster method corresponds to.

Quantum cluster theories are one possibility to find approximate solutions for the Hubbard model. Of course, other methods have been developed too, which treat the infinite system. Thereby special attention is paid to conserving approximations, i.e., approximations which do not violate conservation laws [22] (see Chap. 12.6). Also the functional renormalization group has been applied to the two-dimensional Hubbard model. It works like a microscope with variable resolution. For understanding the low excitation energy sector of the Hubbard model, one eliminates all degrees of freedom of the system which are irrelevant for its behavior in that limit. For that purpose differential equations for the one-particle Green's function are derived which describe the flow of parameters as the degrees of freedom are reduced. Since it is difficult to calculate the renormalization flow for a strongly interacting system, one is starting out from a weakly interacting system in which case the energy scale is given by the kinetics. But as the energy scale decreases, the coupling function increases. This is indicative of possible instabilities, e.g., of magnetic or pair forming origin. In that case one has to switch to a modified description of the system which accounts for these changes [23].

A much studied example is the Mott–Hubbard metal-to-insulator phase transition at half filling when $U \gg t$ [52]. But even now it is not known at which critical ratio of U_c/t this phase transition is taking place at $T = 0$. The value has been steadily increasing with improvements of the approximation schemes. At present the most accurate estimates are obtained when the self-energy is computed with a dynamical CPA in combination with self-consistent projection operator methods. That allows for the calculation of $\Sigma(\mathbf{k}, \omega)$ with an accuracy, which includes effects up to the 12th nearest neighbors. Despite considerable efforts, the topic itself is still wide open. There are several reasons for this. One is that in the Hubbard model

we include one orbital per site only. But in a realistic multiorbital system it is likely that criteria for localization are first fulfilled for a single orbital or for two of them (compare with the dual model of $5f$ electrons) before the system becomes an insulator. Also, there may be a redistribution of electrons among different orbitals as the phase transition is approached. The spectral density and in particular the low-energy peak as obtained, e.g., from a DMFT calculation behave quite different when the ground state is paramagnetic and when it has a long-range AF order. An open question is how AF correlations modify the spectral density as the correlation length increases continuously near a metal-insulator transition.

For small deviations from half filling the Hubbard Hamiltonian can be transferred into a $t-J$ model Hamiltonian of the form

$$\begin{aligned} H_{t-J} &= -t\sum_{\langle ij\rangle\sigma}\left(\hat{a}^{+}_{i\sigma}\hat{a}_{j\sigma}+h.c.\right)+J\sum_{\langle ij\rangle}\left(\mathbf{S}_i\cdot\mathbf{S}_j-\frac{\hat{n}_i\hat{n}_j}{4}\right)\\ &= H_t+H_J, \end{aligned} \tag{1.10}$$

where the operators $\hat{a}^{+}_{i\sigma}$, $\hat{a}_{i\sigma}$ are defined by

$$\begin{aligned} \hat{a}^{+}_{i\sigma} &= a^{+}_{i\sigma}(1-n_{i-\sigma}),\\ \hat{a}_{i\sigma} &= a_{i\sigma}(1-n_{i-\sigma}). \end{aligned} \tag{1.11}$$

The spin operators are $\mathbf{S}_i=(1/2)\sum_{\alpha\beta}\hat{a}^{+}_{i\alpha}\boldsymbol{\sigma}_{\alpha\beta}\hat{a}_{i\beta}$ and $\hat{n}_{i\sigma}=\hat{a}^{+}_{i\sigma}\hat{a}_{i\sigma}$. The coupling constant is $J=4t^2/U$ and defines a low-energy scale caused by the strong correlations. At half filling $H_t=0$ and we deal with an antiferromagnetic (AF) Heisenberg Hamiltonian. Its excitations involve only spin degrees of freedom and constitute the simplest example of spin-charge separation. The $t-J$ model has been very successful in understanding the motion of doped holes or electrons in an antiferromagnetic surroundings. The bandwidth of a coherent hole motion is of order J and therefore strongly renormalized as compared with the bare hopping matrix element t. The energy dispersion of the coherent quasihole motion strongly resembles that found by ab initio calculations using quantum chemical methods. With the $t-J$ model one can also show that two doped holes attract each other, a possible mechanism for high-T_c superconducting in hole-doped cuprates. One can also study the effects of small hole concentration back on the form of AF order, i.e., the development of spiral spin states [24].

Improved Hubbard Hamiltonians have been applied to the Cu–O planes of the high-T_c cuprates. For example, in addition to a $3d_{x^2-y^2}$ orbital for the Cu ions the $2p$ orbitals of the O ions which hybridize with the $3d_{x^2-y^2}$ Cu orbitals are included.This leads to three-band or five-band Hubbard Hamiltonian depending on whether one includes one or two of the $2p$ oxygen orbitals in the Cu–O plane. A new feature which is not contained in a one-band Hubbard description is the formation of a Zhang–Rice singlet state, when a hole is doped in an otherwise half-filled system. It

is mainly located on the oxygen sites and the singlet is formed with the hole on the Cu^{2+} ion [25].

Model Hamiltonians have also been widely used in order to study the microscopic origins of different systems with heavy quasiparticle excitations. The Anderson lattice Hamiltonian has played a prominent role. It is of the form

$$\begin{aligned} H = & \sum_{\mathbf{k}n\sigma} \epsilon_n(\mathbf{k}) a^+_{\mathbf{k}n\sigma} a_{\mathbf{k}n\sigma} + \sum_m \epsilon_{fm} f^+_m(i) f_m(i) \\ & + \frac{1}{\sqrt{N_0}} \sum_{im\mathbf{k}n\sigma} V_{m\sigma}(\mathbf{k}, n) \left[a^+_{\mathbf{k}n\sigma} f_m(i) \mathrm{e}^{\mathrm{i}\mathbf{k}\mathbf{R}_i} + h.c. \right] \\ & + \frac{U}{2} \sum_{im \neq m'} n^f_m(i) n^f_{m'}(i). \end{aligned} \tag{1.12}$$

There are conduction electrons with band index n which weakly hybridize with strongly correlated electrons. They are created (destroyed) by f^+, f operators with an orbital index m and are positioned at lattice sites i. Various approximation schemes have been applied to that Hamiltonian. They are of diagrammatic nature like the noncrossing approximation [26, 27], which describes the Anderson lattice as a system of independent impurities. Or they introduce auxiliary fields like slave boson fields [28–30] which allow for mean-field treatments. Also the so-called Gutzwiller approximation has been applied [31–33]. It renormalizes the hybridization matrix element and moves the orbital energy close to the Fermi energy.

The common feature of those approximations is to reproduce the low-energy scales, which strongly correlated electrons generate.

3 Systems with Heavy Quasiparticles

A sizeable number of intermetallic compounds have a high density of fermionic low-energy excitations. Experiments show that their low-temperature thermodynamic properties strongly resemble those of ordinary metals but with a large quasiparticle mass. The latter can become as large as a meson or even proton mass. It is a consequence of the low-energy scales which strong electron correlations create. Systems with heavy quasiparticles involve $4f$ or $5f$ electrons, in one case, i.e., LiV_2O_4 even $3d$ electrons. Much studied examples are $CeAl_3$, $CeRu_2Si_2$, $CeCu_2Si_2$ or Yb_3As_4 with $4f$ electrons and UPt_3, UBe_{13} and UPd_2Al_3 with $5f$ electrons.

After the discovery of the heavy quasiparticles [34] it was generally believed that the Kondo effect is responsible for those low-energy excitations. Indeed, breaking up a Kondo singlet formed between a $4f$ electron, e.g., of Ce, with the conduction electrons requires a small amount of energy only, i.e., of the order of a few meV. However, in the meantime we have learned that heavy quasiparticles may have quite different microscopic origins. In addition to the Kondo effect (e.g., $CeAl_3$,

$CeRu_2Si_2$, $CeCu_2Al_2$) also partial charge order (e.g., Yb_4As_3), partial localization (e.g., UPt_3, UPd_2Al_3) or the Zeeman effect (e.g., $Nd_{2-x}Ce_xCuO_4$) may cause heavy quasiparticles [35]. In LiV_2O_4, a spinel, the frustrated pyrochlore sublattice of the V ions plays an important role in the formation of the large number of low-energy excitations. While the low temperature properties of heavy quasiparticles are reasonably well understood, the transition to higher temperatures $T > T^*$ at which the quasiparticles lose their large effective mass is less understood. Related with it is a transition from a large Fermi surface, in which the strongly correlated f electrons take part to a small Fermi surface at $T > T^*$ where they are excluded. In the latter case, they behave like localized electrons.

A field of its own are quantum critical points and their neighborhood [36]. At a quantum critical point a system undergoes a phase transition at zero temperature as a function of an external parameter like pressure, impurity concentration etc. Clearly, there are many low-lying excitations in the vicinity of quantum critical point. In some cases, e.g., YMn_2 [37] heavy quasiparticle behavior is observed, but more common is a break down of Fermi-liquid theory near such a point.

4 Mean Field Approximations

Whenever mean-field approximations (MFA) can be made to a Hamiltonian, they not only simplify significantly its solution, but in most cases, also provide new physical insight. The simplest MFA to a Hamiltonian of interacting electrons is the Hartree–Fock or SCF approximation. It neglects totally electron correlations. The question arises whether MFAs can be made which take strong correlations reasonably well into account. Of special interest are clearly mean-field solutions with broken symmetries, i.e., with a ground state which has a lower symmetry than the Hamiltonian. However, care must be exercised here. Often only a symmetry breaking occurs, because correlation effects are simulated which otherwise are insufficiently taken into account.

For example, an AF ground state reduces on-site charge fluctuations as do correlations. In that case, one has to decide whether an AF mean-field solution describes the system correctly or whether it is merely favored because it suppresses charge fluctuations. The simplest example is a H_2 molecule with variable bond length. An unrestricted SCF calculation gives a symmetry broken ground state when the bond length exceeds a critical value. Near one proton the electrons have predominantly spin up while near the other proton they have predominantly spin down. In reality, there is, of course, no distinction between the spins on the two sites. But therefore, all ionic configurations are eliminated by the unrestricted MFA in the limit of large bond lengths.

A special object of mean-field investigations has been the Hubbard model on a cubic or square lattice at or near half filling. With hopping limited to nearest neighbors one finds for the half-filled case perfect nesting at the Fermi surface and a spin-density mean-field ground state for any value of $U > 0$.

An interesting variety of MF solutions is obtained when auxiliary slave boson or fermion fields are introduced [38] (see Chap. 3). An example is the replacement

$$a_{i\sigma} = f_{i\sigma} b_i^+, \quad a_{i\sigma}^+ = f_{i\sigma}^+ b_i, \tag{1.13}$$

where the fermion operator $f_{i\sigma}^+$ creates a spinon at site i while the boson b_i^+ creates an empty site (holon). Strong correlations can be taken into account by forbidding double occupancies via the subsidiary condition

$$\sum_\sigma f_{i\sigma}^+ f_{i\sigma} + b_i^+ b_i = 1, \tag{1.14}$$

i.e., a site is either empty or singly occupied. When the $t - J$ Hamiltonian (1.10) is re-expressed in terms of spinons and holons, different MFAs with different order parameters can be made [39]. One of them has the form of a BCS superconducting order parameter, but here in terms of spinons, not electrons, while another is of resonating valence bond (RVB) type (see Chap. 1.6). There are also other MFAs possible, in particular, since the decomposition (1.14) of electrons in form of spinons and holons is not the only possible one. Another decomposition is

$$a_{i\sigma}^+ = f_i b_{i\sigma}^+, \quad a_{i\sigma} = f_i^+ b_{i\sigma}, \tag{1.15}$$

where the spin degree of freedom is represented by a boson field. We want to find a representation of the Hamiltonian in terms of auxiliary fields in terms of which a MFA describes strong electron correlations as well as possible.

5 Deviations from Fermi Liquids

There is no a priori reason why strongly correlated metallic systems should depict Fermi-liquid behavior. But experiments show that in many cases they do. Systems with heavy quasiparticles are an example. However, in dimensions lower than three deviations may easily occur.

In one dimension (1D) a much discussed item is the separation of spin and charge degrees of freedom. An example is trans-polyacetylene where it is known that excitations involving only spin or charge degrees of freedom do exist. Moreover, when heavily doped, those systems can have excitations with fractional charges. Spin-charge separation is also found in Tomonaga-Luttinger liquids, i.e., one-dimensional interacting electron systems. These features are found independent of the strength of correlations.

In 2D, deviations from Fermi-liquid behavior have been observed not only in the high-T_c cuprates, but also most importantly in samples which exhibit a fractional quantum Hall effect [40]. In the latter case, excitations with fractional charges are widely spread and their statistics changes from fermionic to anyonic. We are able

to understand in the cuprates a number of deviations from standard Fermi-liquid behavior by a *marginal* Fermi-liquid description [41].

Even in three dimensions, deviations from a Fermi liquid are known. Examples are doped Mott–Hubbard systems, but also, e.g., Yb_4As_3 where due to partial charge order, spin and charge degrees of freedom are nearly separated. The break down of a Fermi liquid near a quantum critical point was mentioned before.

6 Superconductivity

Two features of superconductivity in strongly correlated electron systems need special attention. One is the form of the pair state and the other is the origin of electron-electron attraction resulting in a Cooper pair formation. Before the discovery of the high-T_c cuprates, s-wave pairing was assumed to be ubiquitous, although there were indications that in some of the systems with heavy quasiparticles the pair state was of a more complex form. However, it turned out that in the high-T_c materials, electrons pair in a d wave state. This state is preferred when the electron-electron attractions are caused by electronic excitations rather than phonons. Here, one has to distinguish between excitations within the electron system, which forms the Cooper pairs and those, which take place in another, i.e., localized electron system with which the itinerant electrons are interacting. The first case seems to be realized in the cuprates with spin fluctuations leading to electron attractions [42, 43] (see Chap. 5.8). Their form is though quite different in the underdoped and overdoped regimes. The second case is realized, e.g., in the filled skutterudite compound $PrOs_4Sb_{12}$ where intra-atomic, i.e., crystal field excitations of the $4f^2$ subshell of Pr^{3+} provide for an attraction between conduction electrons. The more recent FeAs superconducting compounds, which came as a complete surprise, re-emphasize the need for a better understanding of strongly correlated electrons.

A main topic of research is to describe the high-T_c cuprates for low and high hole doping concentrations, i.e., in the under- and overdoped regimes. While in the overdoped regime a Fermi liquid description applies, this is different in the underdoped regime, where the system is close to a Mott–Hubbard transition to an insulating state.

Understanding the form of the Fermi surface as function of hole doping is an important issue. Angular-resolved photon emission spectroscopy is an invaluable tool here [44]. For a satisfactory description of the normal state a broad spectrum of methods has been used. At one end, are MFA to the Hubbard Hamiltonian rewritten in terms of auxiliary fields [39], at the other end, are quantum chemical calculations with a basis set of respectable size for the energy dispersion of a hole in a Cu–O plane. A separate question is the form of the wavefunction in the superconducting state. An obvious choice is a BCS wavefunction multiplied by a Jastrow function in order to take the strong correlations into account. A simplified form of the Jastrow factor is a Gutzwiller projector which eliminates all configurations with doubly occupied sites [45] (see Chap. 3.7). This wavefunction relates to an RVB state as

mentioned earlier. It is used to estimate the transition temperature in the regime of small hole doping. In the overdoped regime, solving Eliashberg's equations with a spin fluctuation propagator determined within the FLEX approximation has given good results [46].

7 Composite Operator and Projection Techniques

One way of dealing with strongly correlated electrons is to limit the operator or Liouville space within which one wants to describe the many-body system. Stated differently, among all possible degrees of freedom of the system only those are kept, which are considered to be the most important ones. The composite operator method developed by Mancini and Avella [47] and expanded continuously [48–50] (see Chap. 3.7) and the projection operator method (see, e.g., [2]) share this goal. They differ in the way the relevant part of the operator space is chosen. The composite operator method starts from the equation of motion for the single particle Green's function, which generates a new Green's function, this time for composite operators instead of a single electron operator. This procedure is repeated until the sequence of Green's function is terminated at some stage. The projection operator method selects the relevant operator space simply by specifying the microscopic processes which one wants to include, independent of the order in which they appear in the hierarchy of Green's function. This has to do with the lack of a proper expansion parameter, when kinetic energy and Coulomb repulsion are of comparable size. After the space of composite or relevant operators has been chosen, one can diagonalize the matrix of Green's functions. However, there are constants appearing in the solution and different methods differ in the way they are computed They may or may not include couplings to the neglected part of the full operator space. Notwithstanding the possible derivation of low energy scales in certain special cases, like the Kondo problem [51], the selection of a limited number of composite operators is best suited for determining high-energy incoherent excitations of strongly correlated systems. The Zhang–Rice singlet in Cu–O planes or satellite peaks in transition metals serve as examples here. Usually, the lower the considered energy scales are, the more the relevant operator space, i.e., the number of composite operators has to be extended to deal with them.

8 Summary

We have tried to point out a number of features which are characteristic for systems with strongly correlated electrons. Needless to say that neither are they complete nor do they give a historical account of their development. For example, we have not mentioned the effects characteristic for transition metals. The interplay of Hund's rule correlations and the kinetic energy requires a description which interpolates

between a tight binding and a Heisenberg Hamiltonian. A multiband Hubbard Hamiltonian contains that physics. We have also not touched on the topic of orbital ordering, a phenomena widely spread in transition metal oxides. Here, Hund's rule correlations and crystalline electric field effects compete with each other. Other topics like strongly correlated electrons on frustrated lattices could be added. Some of them, like numerical methods, are the subject of the second volume in this series. Nevertheless, we hope that the introduction given here will help the reader of the present volume to view the different aspects of strong correlations which are discussed in the following chapters in a more general context.

Dresden, October 2011 *Peter Fulde*

References

1. P.W. Anderson, Science **235**, 1196 (1987)
2. P. Fulde, *Electron Correlations in Molecules and Solids*. Springer Series in Solid-State Sciences, vol. 100, 3rd edn. (Springer, Berlin, 1995)
3. E. Runge, G. Zwicknagl, Ann. Physik **5**, 333 (1996)
4. G. Zwicknagl, Adv. Phys. **41**, 203 (1992)
5. M.M. Korshunov, V.A. Gavrichkov, S.G. Ovchinnikov, I.A. Nekrasov, Z.V. Pchelkina, V.I. Anisimov, Phys. Rev. B **72**, 165104 (2005)
6. V.I. Anisimov, Y.A. Izyumov, *Electronic Structure of Strongly Correlated Materials*. Springer Series in Solid State Sciences (Springer, Berlin, 2009)
7. A. Georges, G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
8. M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
9. A. Georges, G. Kotliar, W. Krauth, M. J. Rosenberg, Rev. Mod. Phys. **68**, 13 (1996)
10. Y. Kakehashi, Adv. Phys. **53**, 497 (2004)
11. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987)
12. E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)
13. T. Maier, M. Jarrell, T. Pruschke, M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
14. L. Hozoi, U. Birkenheuer, H. Stoll, P. Fulde, New J. Phys. **11**, 023023 (2009)
15. M. Imada, A. Fujimori, Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
16. M.H. Hettler, A.N. Tahvilder-Zadeh, M. Jarrell, T. Pruschke, H.R. Krishnamurthy, Phys. Rev. B **58**, 7475 (1998)
17. M.H. Hettler, M. Mukherjee, M. Jarrell, H.R. Krishnamurthy, Phys. Rev. B **61**, 12739 (2000)
18. G. Kotliar, S. Savrasov, G. Pallson, G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
19. D. Senechal, D. Perez, D. Plouffe, Phys. Rev. B **66**, 075129 (2002)
20. Y. Kakehashi, T. Nakamura, P. Fulde, J. Phys. Soc. Jpn. **78**, 124710 (2009)
21. M. Potthoff, Euro. Phys. J. B **32**, 429 (2003)
22. Y.M. Vilk, A.M. S. Tremblay, J. de Phys. **7**, 1309 (1997)
23. M. Salmhofer, C. Hohnerkamp, Progr. Theor. Phys. **105**, 1 (2001)
24. B.I. Shraiman, E.D. Siggia, Phys. Rev. Lett. **60**, 740 (1988)
25. F.C. Zhang, T.M. Rice, Phys. Rev. B **37**, 3759 (1988)
26. N.E. Bickers, Rev. Mod. Phys. **59**, 845 (1987)
27. Y. Kuramoto, Z. Phys. B **57**, 95 (1984)
28. S.E. Barnes, J. Phys. F **6**, 1375 (1976)
29. S.E. Barnes, J. Phys. F **7**, 2637 (1977)
30. T. Li, P. Wölfle, P.J. Hirschfeld, Phys. Rev. B **40**, 6817 (1989)

31. M.C. Gutzwiller, Phys. Rev. A **134**, 923 (1964)
32. M.C. Gutzwiller, Phys. Rev. A **137**, 1726 (1965)
33. B. Edegger, V.N. Muthukumar, C. Gross, Adv. Phys. **56**, 927 (2007)
34. K. Andres, J.E. Graebner, H.R. Ott, Phys. Rev. Lett. **35**, 1779 (1975)
35. P. Fulde. P. Thalmeier, G. Zwicknagl, in *Solid State Physics*, vol. 60, ed. by H. Ehrenreich, N. Sneddon (Elsevier, Amsterdam, 2006)
36. S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, London, 1999)
37. C. Pinettes, C. Lacroix, J. Phys. Cond. Mat. **6**, 10093 (1994)
38. P. Wölfle, J. Low Temp. Phys. **99**, 625 (1995)
39. P. Lee, N. Nagaosa, X.-G. Wen, Rev. Mod. Phys. **78**, 17 (2006)
40. D. Yoshioka, *The Quantum Hall Effect*. Solid State Sciences, vol. 133 (Springer, Berlin, 2002)
41. P.B. Littlewood, C.M. Varma, J. Appl. Phys. **69**, 4979 (1991)
42. A.V. Chubukov, D. Pines, J. Schmalian, in *The Physics of Superconductors*, ed. by K. H. Bennemann, J.B. Ketterson (Springer, Berlin, 2008), p. 1349
43. N. Plakida, *High-Temperature Cuprate Superconductors*. Solid State Sciences (Springer, Berlin, 2010)
44. A. Damascelli, Z. Hussein, Z.-X. Shen, Rev. Mod. Phys. **75** 473 (2003)
45. P.W. Anderson, P.A. Lee, M. Randeria, T.M. Rice. N. Trivedi, F.C. Zhang, J. Phys. Cond. Matt. **16** R755 (2004)
46. T. Dahm, V. Hinkov, S.V. Borisenko, A.A. Kordyuk, V.B. Zabolotnyy, J. Fink, B. Büchner, D.J. Scalapino, W. Hanke, B. Keimer, Nat. Phys. **5** 217 (2009)
47. F. Mancini, A. Avella, Adv. Phys. **53** 537 (2004)
48. H. Matsumoto, T. Saikawa, F. Mancini, Phys. Rev. B **54**, 14445 (1996)
49. A. Avella, F. Mancini, Phys. Rev. B **75**, 134518 (2007)
50. F. Mancini, Europhys. Lett. **70** 485 (2005)
51. D. Villani, E. Lange, A. Avella, G. Kotliar, Phys. Rev. Lett. **85** 804 (2000)
52. F. Gebhard, *The Mott Metal-Insulator Transition* Springer Tracs in Modern Physics, vol. 137 (Springer, Berlin, 1997)

Chapter 1
Density Functional Theory: A Personal View

Robert O. Jones

Abstract A practical definition of "strongly correlated" systems covers those that cannot be described well by density functional (DF) theory. DF theory has become an indispensable part of computational condensed matter physics and chemistry, but its origins go back to the early years of quantum mechanics in the late 1920s, and this chapter is devoted to a historical view of its development. Thomas and Fermi recognized the electron density as a basic variable, and Dirac showed already in 1930 that the state of an atom is completely determined by its density; it is not necessary to specify the wave function. We follow the development of these ideas in the following decades to the single-particle equations of Kohn and Sham in 1965. Many details of the history are not well known, even to specialists in the field. The single application discussed is the Be_2, which was perhaps the first unexpected DF prediction on small molecules that proved to be correct.

1.1 Introduction

Many seminars and publications on "strongly correlated" systems mention at the outset the widespread use of density functional (DF) theory in condensed matter physics and chemistry and point out the physical insight that often results. The second sentence, however, often lists the systems where DF results are disastrous (an insulator is found to conduct, almost anything to do with rare earth elements, etc.), emphasizing the importance of describing strongly correlated materials correctly

R.O. Jones (✉)
Institut für Festkörperforschung, Forschungszentrum Jülich, 52425 Jülich, Germany

German Research School for Simulation Sciences, RWTH Aachen University and FZ Jülich, 52425 Jülich, Germany
e-mail: r.jones@fz-juelich.de

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_1,

(see, e.g., the Foreword of this volume). At first sight, one might wonder about the relevance of an article on DF theory in a book devoted to the areas where it fails. Further reflection, however, suggests that we should have a feel for the areas where it does a good job, i.e., cases where physical insight can be obtained *without* recourse to the methods described in other articles in this volume. It is also important to know *why* approximations used in DF calculations should give sensible answers far from their regions of obvious validity. Yes, the DF approach should be understood by all who are interested in systems where it fails, and the editors have asked me to help you.

The DF formalism shows that ground state properties of a system of electrons in an external field can be determined from a knowledge of the *density distribution* $n(r)$ alone; one does not need to know the much more complicated many-electron wave function. We shall see below that this remarkable result was proved by Dirac [1] already in 1930. We focus in this article on a property for which DF calculations are particularly valuable: the total energy E of a system of electrons in the presence of ions located at $R_{\rm I}$. Accurate calculations of the entire energy surface $E(R_{\rm I})$ are possible only for systems with very few atoms, and this function generally has vast number of maxima and minima at unknown locations. The lowest energy, however, corresponds to the ground state *structure*, and paths between minima are essential to our studies of chemical reactions, including their activation energies. When I read the autobiography of Francis Crick [2] some 15 years ago, I was taken by his observation

> If you want to study *function*, study *structure*.

This may be self-evident to molecular biologists, but it is also true in many other areas. The DF approach allows us to calculate $E(R_{\rm I})$, and hence the structure and many related properties, without using experimental input. If you are more interested in "real materials" than in mathematical models, this is a crucial advantage in the area of strongly correlated materials as well.

Twenty years ago, Olle Gunnarsson and I reviewed the DF formalism, its history, and its prospects in an article that has been cited more than 2,000 times [3]. This is very gratifying, particularly if all who cited it had actually read it. I have given many seminars that took a similar perspective on the history of the method. Perhaps colloquium organizers (editors?) think that someone who has been active in DF research for over 30 years might have something to say. Some 4 years ago, I looked in detail into the original literature, some of which many (including me) simply cite. This has resulted in a substantial change in format that has been reflected in several seminars. Many listeners, even (or particularly) DF specialists, were surprised at what they heard. In these talks, and in the present chapter, I have traced DF history from the first years after the development of quantum mechanics. I have quoted the original texts in numerous places; some are in journals that are not readily accessible, but it is important in all cases that the authors speak for themselves.

1.2 The Early Years: The Density as a Basic Variable

The recent book by Gino Segrè, *Faust in Copenhagen* [4], gives a fascinating account of the development of quantum mechanics in the years following 1926. Methods for finding approximate solutions of the Schrödinger equation were developed almost immediately after the equations were published, and these methods have had a profound impact on chemistry and condensed matter physics ever since.

A method for calculating the wave function of an atom was developed by Hartree [5]. The "Hartree approximation" to the many-electron wave function is a product of single-particle functions:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, ...) = \psi_1(\mathbf{r}_1) \cdots \psi_N(\mathbf{r}_N), \tag{1.1}$$

where each $\psi_i(\mathbf{r}_i)$ satisfies a one-electron Schrödinger equation with a potential term arising from the average field of the other electrons. Although the papers of Hartree [5] introduced the idea of such a "self-consistent field," with specific reference to the core and valence electrons, there is no mention of the approximation (1.1). Slater [6] and Fock [7] noted shortly afterwards that the product wave function (1.1) in conjunction with the variational principle led to a generalization of the method that would apply to systems more complex than atoms. Both authors noted that replacing (1.1) by a "linear combination of such products with permuted indices" [6] or a determinant of such functions [7] led to equations that were not much more complicated than those of Hartree, but satisfied important symmetry requirements such as the Pauli exclusion principle. These determinantal functions, which had been used previously in discussions of atoms [8] and ferromagnetism [9], are known today as "Slater determinants," and the resulting "Hartree–Fock equations" have formed the basis of most discussions of atomic and molecular structure since.

This was the view of Dirac in 1929 [10]:

> The general theory of quantum mechanics is now almost complete, the imperfections that still remain being in connection with the exact fitting of the theory with relativity ideas. These give rise to difficulties only when high-speed particles are involved, and are therefore of no importance in the consideration of atomic and molecular structure and ordinary chemical reactions, in which it is, indeed, usually sufficiently accurate if one neglects relativity variation of mass with velocity and assumes only Coulomb forces between the various electrons and atomic nuclei. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

Dirac emphasizes the difficulty of solving the equations of quantum mechanics and the desirability of developing "approximate practical methods of applying quantum

mechanics" to explain complex systems. Density functional theory seems to me to be just such a method.

Dirac did not stop at this statement, but set about finding a way to improve the newly developed Thomas–Fermi (TF) model [11, 12] for calculating atomic properties based purely on the electron density $n(r)$ [10]. Here it is assumed that the electrons form a perfect gas satisfying Fermi statistics and the kinetic energy has a simple dependence on the $n(r)$. The TF equations

$$\frac{5}{3}C_k n(\boldsymbol{r})^{\frac{2}{3}} + e^2 \int \mathrm{d}\boldsymbol{r}' \, \frac{n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} + V_{\mathrm{ext}}(\boldsymbol{r}) + \lambda = 0, \tag{1.2}$$

where $C_k = 3\hbar^2(3\pi^2)^{\frac{2}{3}}\,10\,\mathrm{m}^{-1}$ and λ is the Lagrange multiplier related to the constraint of constant particle number, are derived in Appendix 1. Dirac noted the necessity of incorporating "exchange" phenomena, as in the Hartree–Fock approach [10]:

> For dealing with atoms involving many electrons the accurate quantum theory, involving the solution of the wave equation in many-dimensional space, is far too complicated to be practicable. One must therefore resort to approximate methods. The best of these is Hartree's method of the self-consistent field. Even this, however, is hardly practicable, ... so that one then requires a still simpler and rougher method. Such a method is provided by Thomas' atomic model, ...
>
> The method of the self-consistent field has recently been established on a very much better theoretical basis in a paper by Fock, which shows how one can take into account the exchange phenomena between equivalent electrons. Fock shows that if one takes the best approximation to the many-dimensional wave function that is of the form of a product of a number of three-dimensional wave functions, one for each electron, then the three-dimensional wave functions will satisfy just Hartree's equations. .... The exclusion principle of Pauli, however, requires that the wave function representing a number of electrons shall always be antisymmetrical. One would therefore expect to get a better approximation if one first made the product of a number of three-dimensional wave functions antisymmetrical, by applying perturbations and taking a linear combination, and then made it approach as closely as possible to the accurate many-dimensional wave function. The three-dimensional wave functions will then, as found by Fock, satisfy equations somewhat different from Hartree's, containing extra terms that may be considered as representing the exchange phenomena.

Dirac included exchange effects in the "Thomas atom" [1] by means of the potential

$$V_{\mathrm{x}}^{\mathrm{Dirac}} = -\left(\frac{1}{\pi}\right)\left(3\pi^2 n(\boldsymbol{r})\right)^{\frac{1}{3}}. \tag{1.3}$$

This approximation should be valid for weak spatial variations of the density $n(\boldsymbol{r})$, and the correction to the TF equation is often referred to as the "Thomas–Fermi–Dirac" equation.

The Thomas–Fermi method and its extensions give a rough description of some properties of atoms (e.g. the charge density and the electrostatic potential), and it is a well-defined model whose mathematical properties have attracted considerable attention [13, 15]. The TF scheme is exact, for example, in the limit of infinite

nuclear charge. However, the model has severe deficiencies. The charge density is infinite at the nucleus and decays as r^{-6}, not exponentially, far from it. It has also been shown by Teller [16] and others that TF theory does not result in atoms binding to form molecules or solids. There is also no shell structure in the TF atom, so that the periodic variation of many properties with changing atomic number cannot be reproduced, and no ferromagnetism [3]. Moreover, atoms *shrink* with increasing atomic number Z (as $Z^{-1/3}$) [17].

On rereading [1], however, I was struck particularly by the following lines, where Dirac states quite clearly the basic premise of all involved in DF calculations:

> Each three-dimensional wave function will give rise to a certain electric density. This electric density is really a matrix, like all dynamical variables in the quantum theory. By adding the electric densities from all the wave functions we can obtain the total electric density for the atom. If we adopt the equations of the self-consistent field as amended for exchange, then this total electric density (the matrix) has one important property, namely, if the value of the total electric density at any time is given, then its value at any later time is determined by the equations of motion. This means that *the whole state of the atom is completely determined by this electric density; it is not necessary to specify the individual three-dimensional wave functions that make up the total electric density.* Thus one can deal with any number of electrons by working with just one matrix density function.

The italics are in the original. The observation that the density follows the equations of motion is much in the spirit of Ehrenfest's theorem [18]. Ehrenfest had proved in 1927 what I have seen referred to as the "time-dependent Hellmann-Feynman theorem," namely that the acceleration of a quantum wave packet that does not spread satisfied Newton's equations of motion.

The central role of the density in the present discussion means that we must have a clear picture of its nature. In Fig. 1.1, we show the spherically averaged density in the ground state of the carbon atom. The density falls monotonically from the

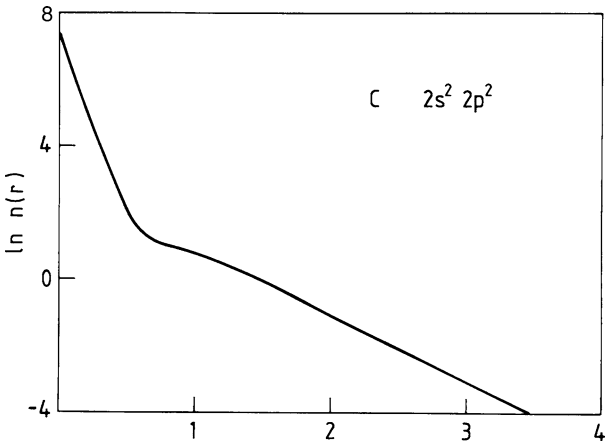

Fig. 1.1 Logarithm of spherical average of density in ground state of C atom as a function of the distance from the nucleus [3]

nucleus and does not show the radial oscillations that occur if we plot $4\pi r^2 n(r)$. The charge density in small molecules is also relatively featureless, with maxima at the nuclei, saddle points along the bonds, and a generally monotonic decay from both. It is essential also to recognize that the electron density in molecules and solids shows relatively small departures from the overlapped densities of the constituent atoms. Energy differences, including binding, ionization, and cohesive energies, are the focus of much DF work and result from subtle changes in relatively featureless density distributions. It still amazes me that this is all we need to know to determine ground state properties.

1.3 Towards an "Approximate Practical Method"

The basis of a quantum theory of atoms, molecules, and solids was already in place at the beginning of the 1930s. Linear combinations of atomic orbitals formed molecular orbitals, from which determinantal functions could be constructed, and linear combinations of determinants would provide approximations to the complete wave function. Dirac had noted, however, that this procedure could not be implemented in practice, so that approximations are essential. Furthermore, numerical techniques for solving the Schrödinger equation in extended systems needed to be developed.

Wigner and Seitz [19] developed a cellular method for treating the self-consistent problems in crystals, and the "Wigner–Seitz cell" is a construction that is known to all students of condensed matter physics. The first application was to metallic sodium using a pseudopotential for the Na ion, and calculations of the lattice constant, cohesive energy, and compressibility gave satisfactory results. Of particular interest for our purposes, however, is the calculation of the probability of finding electrons with parallel spins a distance r apart (Fig. 1.2). This function obtains its half-value for $r = 1.79\,d'$ or $0.460\,d$ for a body-centered cubic lattice with cube edge d, which is close to the radius of the "Wigner–Seitz sphere" $(\frac{3}{8\pi})^{\frac{1}{3}}\,d = 0.492\,d$. It follows that two electrons with parallel spins will very rarely be at the same ion, simply as a consequence of the exclusion principle. Wigner and Seitz noted that this argument does not depend significantly on the potential and should apply to a Fermi gas subject to periodic boundary conditions. The corresponding curves for spin up and spin down electrons, as well as for both spins combined, were discussed in the 1934 review article of Slater [20].

The picture that results is simple and appealing: The exclusion principle means that an electron with a given spin produces a surrounding region where there is a deficiency of charge of the same spin. This region contains unit charge and is referred to as the "Fermi" [19] or "exchange hole" [21]. In the Hartree–Fock scheme, the exchange hole is different for each wave function, but Slater [21] developed a simplified "exchange potential":

$$V_{\mathrm{x}}^{\mathrm{Slater}} = -\left(\frac{3}{2\pi}\right)\left(3\pi^2 n(r)\right)^{\frac{1}{3}}\,. \tag{1.4}$$

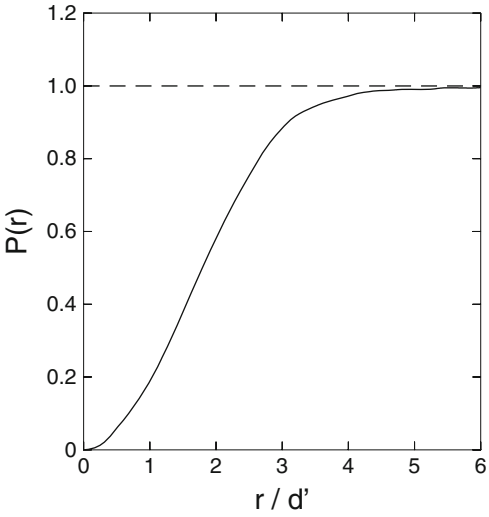

Fig. 1.2 Probability function that electrons in Na with parallel spins are a distance r/d' apart. $d'^3 = V_0/(3\pi^2)$, with V_0 the atomic volume. After Wigner and Seitz [19]

This approximation was proposed at the time that electronic computers were becoming available for electronic structure calculations and proved to be very useful in practice. Methods for solving the Schrödinger equation had been developed around this time, including the augmented plane wave (APW) [22] and Korringa–Kohn–Rostoker approaches [23, 24].

The exchange potential of Slater (1.4) is 3/2 times that of Dirac (1.3), but Slater [25] pointed out that an effective potential proportional to the cube root of the density could be obtained by arguments based on the exchange hole that are quite independent of the free electron gas arguments used in the original derivation [21]. The exchange hole discussed above for a spin up electron contains a single electron. If we assume that it can be approximated by a sphere of radius $R_\uparrow$, then

$$\left(\frac{4\pi}{3}\right) R_\uparrow^3 n_\uparrow = 1; \quad R_\uparrow = \left(\frac{3}{4\pi n_\uparrow}\right)^{\frac{1}{3}}, \tag{1.5}$$

where $n_\uparrow$ is the density of spin up electrons. Since the electrostatic potential at the center of such a spherical charge is proportional to $1/R_\uparrow$, the exchange potential will be proportional to $n_\uparrow^{\frac{1}{3}}$. This argument was used by Slater to counter a common misconception that local density approximations were only valid if the density was close to that in a uniform electron gas.

In 1954, Gáspár [26] pointed out that the prefactor of the effective exchange potential (1.4) was questionable. If one varies the spin orbitals to minimize the total energy in the Thomas–Fermi–Dirac form, one obtains a coefficient just $\frac{2}{3}$ as large. Gáspár applied this approximation to the Cu^+ ion and found good agreement with Hartree–Fock eigenfunctions and eigenvalues. Slater noted that Gáspár's method was "more reasonable than mine" [27], but most calculations continued to use the larger value.

1.4 Density Functional Formalism

The variational principle on the energy was the basis of the formulation of the density functional formalism given by Hohenberg and Kohn [28]. The ground state (GS) properties of a system of electrons in an external field can be expressed as functionals of the GS electron density, i.e., they are determined by a knowledge of the density alone. The total energy E can be expressed in terms of such a functional, and $E[n]$ satisfies a variational principle. These basic theorems were proved by Hohenberg and Kohn [28] for densities that can be derived from some potential V ("V-representable"). A simpler and more general proof for ("N-representable") densities that can be derived from some antisymmetric wave function was given by Levy [29] and is given in Appendix 2. These proofs leave open, of course, how we write the functional relationship between energy E and density n in practice.

1.4.1 Single-Particle Description of a Many-Particle System

The task of finding good approximations to the energy functional $E(n)$ is greatly simplified if we use the decomposition introduced by Kohn and Sham [30]:

$$E[n] = T_0[n] + \int \mathrm{d}\boldsymbol{r}\; n(\boldsymbol{r})\Big(V_{\mathrm{ext}}(\boldsymbol{r}) + \frac{1}{2}\Phi(\boldsymbol{r})\Big) + E_{\mathrm{xc}}[n]. \tag{1.6}$$

T_0 is the kinetic energy that a system with density n would have if there were no electron–electron interactions, Φ is the classical Coulomb potential for electrons, and E_{xc} defines the exchange-correlation energy. Although T_0 differs from the true kinetic energy T, it is of comparable magnitude and is treated here without approximation. This removes many of the deficiencies of the Thomas–Fermi approximation, such as the lack of a shell structure of atoms or the absence of chemical bonding in molecules and solids. In the expression (1.6) all terms other than the exchange-correlation energy E_{xc} can be evaluated exactly, so that approximations for this term play a central role in the following discussion.

The variational principle applied to (1.6) yields

$$\frac{\delta E[n]}{\delta n(\boldsymbol{r})} = \frac{\delta T_0}{\delta n(\boldsymbol{r})} + V_{\mathrm{ext}}(\boldsymbol{r}) + \Phi(\boldsymbol{r}) + \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\boldsymbol{r})} = \mu, \tag{1.7}$$

where μ is the Lagrange multiplier associated with the requirement of constant particle number. If we compare this with the corresponding equation for a system with an effective potential $V(\boldsymbol{r})$ but *without* electron-electron interactions,

$$\frac{\delta E[n]}{\delta n(\boldsymbol{r})} = \frac{\delta T_0}{\delta n(\boldsymbol{r})} + V(\boldsymbol{r}) = \mu, \tag{1.8}$$

we see that the mathematical problems are identical, provided that

$$V(\boldsymbol{r}) = V_{\text{ext}}(\boldsymbol{r}) + \Phi(\boldsymbol{r}) + \frac{\delta E_{\text{xc}}[n]}{\delta n(\boldsymbol{r})}. \tag{1.9}$$

The solution of (1.8) can be found by solving the Schrödinger equation for *non-interacting* particles,

$$\left(-\frac{1}{2}\nabla^2 + V(\boldsymbol{r})\right)\ \psi_i(\boldsymbol{r}) = \epsilon_i \psi_i(\boldsymbol{r}), \tag{1.10}$$

yielding

$$n(\boldsymbol{r}) = \sum_{i=1}^{N} |\psi_i(\boldsymbol{r})|^2. \tag{1.11}$$

Condition (1.9) can be satisfied in a self-consistent procedure.

The solution of this system of equations leads then to the energy and density of the lowest state and all quantities derivable from them. The formalism can be generalized to the lowest state with a given set of quantum numbers [31]. Instead of seeking these quantities by determining the wave function of the system of interacting electrons, the DF method reduces the problem to the solution of a single-particle equation of Hartree form. In contrast to the Hartree–Fock potential,

$$V_{\text{HF}}\ \psi(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}'\ V_{\text{HF}}(\boldsymbol{r}, \boldsymbol{r}')\psi(\boldsymbol{r}'), \tag{1.12}$$

the effective potential, $V(\boldsymbol{r})$ is *local*.

The numerical advantages of solving the Kohn-Sham equations [30] are obvious. Efficient methods exist for solving single-particle Schrödinger-like equations with a local effective potential, and there is no restriction to small systems. With a local approximation to E_{xc}, the equations are no more complicated to solve than Hartree's equations. In contrast to the Thomas–Fermi method, where the large kinetic energy term is approximated, the valence kinetic energy $T_{0,\text{v}}$, the electrostatic interaction between core and valence electrons E_{cv}, and the electrostatic interaction between valence electrons E_{vv} are treated exactly. However, the exchange-correlation energy E_{xc} is the difference between the exact energy and the terms we can evaluate numerically exactly, and approximations are unavoidable.

1.4.2 Approximations to E_{xc}

Kohn and Sham [30] proposed using the "local density (LD) approximation"

$$E_{\text{xc}}^{\text{LD}} = \int \mathrm{d}\boldsymbol{r}\ n(\boldsymbol{r})\ \varepsilon_{\text{xc}}[n(\boldsymbol{r})], \tag{1.13}$$

where $\varepsilon_{\rm xc}[n]$ is the exchange and correlation energy per particle of a homogeneous electron gas with density n. This approximation is exact in the limits of slowly varying densities and very high densities. The authors noted that this approximation "has no validity" in the "surface" of atoms and in the overlap regions of molecules and concluded [30]:

> We do not expect an accurate description of chemical bonding.

The generalization to spin-polarized systems is

$$E_{\rm xc}^{\rm LSD} = \int {\rm d}\boldsymbol{r}\; n(\boldsymbol{r})\; \varepsilon_{\rm xc}[n_\uparrow(\boldsymbol{r}), n_\downarrow(\boldsymbol{r})], \tag{1.14}$$

where $\varepsilon_{\rm xc}[n_\uparrow, n_\downarrow]$ is the exchange and correlation energy per particle of a homogeneous, spin-polarized electron gas with spin up and spin down densities $n_\uparrow$ and $n_\downarrow$, respectively.[1] The "Xα" approximation

$$E_{\rm x}^{\rm X\alpha} = -\frac{3}{2}\alpha C \int {\rm d}\boldsymbol{r}\; \{(n_\uparrow(\boldsymbol{r}))^{4/3} + (n_\downarrow(\boldsymbol{r}))^{4/3}\}, \tag{1.15}$$

where $C = 3(3/4\pi)^{1/3}$, was used in numerous calculations in the late 1960s and 1970s. The α-dependence of energy differences for a given atom or molecule is weak for values near $2/3$, the value of Dirac [1], Gáspár [26], and Kohn and Sham [30].

We have noted that the electron density in molecules and solids is generally far from that of a homogeneous electron gas, and the validity of calculations based on properties of a gas of constant density has often been questioned. The many improved approximations for $E_{\rm xc}$ that have been developed in the past 20 years are beyond the scope of this article, but an example of a rung on the "Jacob's ladder" of approximations is that of TPSS [32], where $n_\uparrow$ and $n_\downarrow$ are joined by their gradients and the kinetic energy density of the occupied Kohn–Sham orbitals. In recent years, there has been much focus on incorporating "exact exchange" into DF calculations. Olle Gunnarsson and I showed [33] that LD and LSD approximations could yield very large errors in exchange energy differences if there was a change in the nodal structure of the orbitals involved. The exact inclusion of exchange would circumvent this problem, but the computation cost is substantial. We now discuss some general properties of $E_{\rm xc}$ using arguments closely related to the "exchange hole" picture of Wigner and Seitz [19] and Slater [21, 25].

[1] The calculation by Bloch [9] of ferromagnetism in a free-electron model of a metal was the first where the exchange-correlation energy was expressed as the sum of terms proportional to $n_\uparrow^{4/3}$ and $n_\downarrow^{4/3}$.

1.4.3 Exchange-Correlation Energy, E_{xc}

The crucial simplification in the density functional scheme is the relationship between the interacting system, whose energy and density we seek, and the fictitious, noninteracting system for which we solve (1.10) and (1.11). This can be studied by considering the interaction $\lambda/|\boldsymbol{r}-\boldsymbol{r}'|$ and varying λ from 0 (noninteracting system) to 1 (physical system). This is done in the presence of an external potential, V_λ [34], such that the ground state of the Hamiltonian

$$H_\lambda = -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\boldsymbol{r}) + V_\lambda + \lambda V_{\text{ee}} \tag{1.16}$$

has density $n(\boldsymbol{r})$ for all λ. The exchange-correlation energy of the interacting system can then be expressed *exactly* as an integral over the coupling constant λ [31]:

$$E_{\text{xc}} = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\; n(\boldsymbol{r}) \int \mathrm{d}\boldsymbol{r}'\; \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} n_{\text{xc}}(\boldsymbol{r}, \boldsymbol{r}'-\boldsymbol{r}), \tag{1.17}$$

with

$$n_{\text{xc}}(\boldsymbol{r}, \boldsymbol{r}'-\boldsymbol{r}) \equiv n(\boldsymbol{r}') \int_0^1 \mathrm{d}\lambda \left(g(\boldsymbol{r}, \boldsymbol{r}', \lambda) - 1\right). \tag{1.18}$$

The function $g(\boldsymbol{r}, r', \lambda)$ is the pair correlation function of the system with density $n(\boldsymbol{r})$ and Coulomb interaction λV_{ee}. The exchange-correlation hole, n_{xc}, describes the fact that an electron at point $\boldsymbol{r}$ reduces the probability of finding one at $\boldsymbol{r}'$. The xc-energy may then be viewed as the energy resulting from the interaction between an electron and its exchange-correlation hole. This is a straightforward generalization of the work of Wigner and Seitz [19] and Slater [21] discussed above.

Second, the isotropic nature of the Coulomb interaction V_{ee} has important consequences. A variable substitution $\boldsymbol{R} \equiv \boldsymbol{r}'-\boldsymbol{r}$ in (1.17) yields

$$E_{\text{xc}} = \frac{1}{2}\int \mathrm{d}\boldsymbol{r}\; n(\boldsymbol{r}) \int_0^\infty \mathrm{d}R\; R^2\, \frac{1}{R} \int \mathrm{d}\Omega\; n_{\text{xc}}(\boldsymbol{r}, \boldsymbol{R}). \tag{1.19}$$

Equation (1.19) shows that the xc-energy depends only on the spherical average of $n_{\text{xc}}(\boldsymbol{r}, \boldsymbol{R})$, so that approximations for E_{xc} can give an *exact* value, even if the description of the nonspherical parts of n_{xc} is quite inaccurate. Third, the definition of the pair-correlation function leads to a sum-rule requiring that the xc-hole contains one electron, i.e., for all $\boldsymbol{r}$,

$$\int \mathrm{d}\boldsymbol{r}'\; n_{\text{xc}}(\boldsymbol{r}, \boldsymbol{r}'-\boldsymbol{r}) = -1. \tag{1.20}$$

This means that we can consider $-n_{\rm xc}(\boldsymbol{r}, \boldsymbol{r}' - \boldsymbol{r})$ as a normalized weight factor and define locally the radius of the xc-hole

$$\left\langle \frac{1}{\boldsymbol{R}} \right\rangle_r = -\int \mathrm{d}\boldsymbol{r}\ \frac{n_{\rm xc}(\boldsymbol{r}, \boldsymbol{R})}{|\boldsymbol{R}|}. \tag{1.21}$$

This leads to

$$E_{\rm xc} = -\frac{1}{2}\int \mathrm{d}\boldsymbol{r}\ n(\boldsymbol{r})\left\langle \frac{1}{\boldsymbol{R}} \right\rangle_r. \tag{1.22}$$

Provided (1.20) is satisfied, $E_{\rm xc}$ is determined by the first moment of a function whose second moment we know exactly and depends only weakly on the details of $n_{\rm xc}$ [31]. It is therefore possible that approximations to the exchange and correlation energy can lead to good total energies even if the details of the exchange-correlation hole are described very poorly. This is shown in Fig. 1.3, where the exchange hole in a nitrogen atom is shown for a representative value of $\boldsymbol{r}$ for both the local density and exact (Hartree–Fock) cases. The holes are qualitatively different: The LD hole is

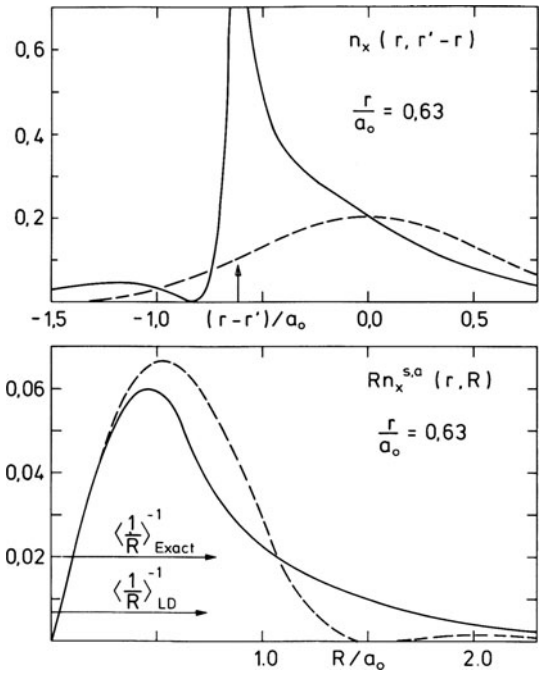

Fig. 1.3 Exact (*solid*) and approximate (*dashed*) exchange hole $n_{\rm xc}(\boldsymbol{r}, \boldsymbol{r}' - \boldsymbol{r})$for spin up electron in a nitrogen atom for 0.63 a.u. The *top figure* shows the hole along a line through the nucleus and the electron. The *arrow* shows the electron position, and $\boldsymbol{r} - \boldsymbol{r}' = 0$ gives the electron position. The *lower figure* shows the spherical average of the hole around the electron, as well as the value of $\langle 1/\boldsymbol{R} \rangle$ (1.21) [3]

spherically symmetric and centered on the electron, while the exact hole has a large weight at the nucleus and is very asymmetric. Nevertheless, the spherical averages are very similar, and the exchange energies differ by only a few percent.

1.4.4 DF Theory in Retrospect

Density functional calculations are so well established in condensed matter physics and chemistry that one might assume that they have always found general acceptance. This was not the case. Condensed matter physicists were generally pleased to have justification for the "local density" calculations they had been performing for decades, and they were sometimes less than critical of the approximations involved. A well-known condensed matter theorist, Volker Heine, looked back on the 1960s in this way [35]:

> Of course at the beginning of the 1960s the big event was the Kohn Hohenberg Sham reformulation of quantum mechanics in terms of density functional theory (DFT). Well, we recognize it now as a big event, but it did not seem so at the time. That was the second big mistake of my life, not to see its importance, but then neither did the authors judging from the talks they gave, nor anyone else. Did you ever wonder why they never did any calculations with it?

Nevertheless, large number of electronic structure theorists and a few theoretical chemists moved seamlessly from performing "Xα" or "Hartree–Fock–Slater" calculations into the density functional world.

It took many years for DF calculations to be taken seriously by most chemists. The reasons given were often quite convincing: (1) Unlike the TF theory, the Kohn-Sham expression for the energy is not really a "functional" of the density, since the kinetic energy term is treated exactly and is defined by an effective potential that leads to the density. (2) The original functional of Hohenberg and Kohn [28] is not even *defined* for all n, because not all densities can be derived from the ground state of some single-particle potential [13]. (3) Approximations to the exchange-correlation energy are essential, and their usefulness can only be determined by performing calculations. (4) There is no *systematic* way to approach the exact solution of the Schrödinger equation.

This last point was emphasized by many. The Hartree–Fock method could be extended to multiple determinants ("configuration interaction") and, coupled with a large basis set, we could obtain, in principle, the exact wave function and all properties obtainable from it. This is a very attractive proposition, and new generations of computers might make the reservations of Dirac [10] less formidable. Nevertheless, obtaining numerically exact total energies from calculations of the wave function remains a major challenge to this day, and it is not surprising that several groups looked at alternatives.

Hartree–Fock–Slater calculations were carried out in the early 1970s, particularly by Evert Jan Baerends and collaborators in Amsterdam, and some of the first DF calculations on small molecules were carried out during the thesis work of Olle Gunnarsson in Göteborg [31]. John Harris and I were very surprised that the LD

and LSD approximations gave reasonable results for small molecules, and we decided to expand the effort when Olle joined our group in 1975. We invested man-years in the development of a full-potential LMTO code for small molecules and clusters, and we performed calculations on many small systems. The local density approximations led to good geometries and reasonable binding energies in most cases. Most condensed matter physicists were not surprised, but theoretical chemists remained sceptical or critical, and this situation continued throughout the 1980s and into the 1990s.

Two developments changed this. Axel Becke developed a nonlocal exchange functional that promised improvements over local approximations [36], and application of this functional to the atomization energies of small molecules by John Pople and coworkers gave very promising results [37]:

> In summary, these initial results indicate that DFT is a promising means of obtaining quantum mechanical atomization energies; here, the DFT methods B-VWN and B-LYP outperformed correlated *ab initio* methods, which are computationally more expensive.

In addition to the growing body of results on molecules and clusters that were beyond the scope of calculations of correlated wave functions, this change in attitude by one of the most prominent theoretical chemists led to a dramatically new attitude towards the DF method in chemistry. One sign of this was the incorporation of DF options into molecular program packages.

One willing convert seems to have been Michael Frisch, first author of the Gaussian program package, who gave a seminar at the ACS National Meeting in San Francisco (13 April 1997) on "Ab initio calculations of vibrational circular dichroism and infrared spectra using SCF, MP2, and density functional theories for a series of molecules." At the end of the talk, an unknown (to me) member of the audience asked:

> What about Hartree–Fock?

The answer was unambiguous and so much to my liking that I wrote it down immediately:

> It does not matter what you want to calculate, and it does not matter what functional you use; density functional results are always better than Hartree–Fock.

Fritz Schaefer, a prolific and much-cited theoretical chemist, gave his perspective on this transitional period when summarizing a conference in 1996 [38]:

> This conference has displayed a remarkable synthesis of density functional methods with *ab initio* quantum mechanical methods. In light of this successful synthesis, one might fairly ask the question, why did it take some of us so long to accept DFT?
>
> As one of those who resisted DFT until a few years ago, I would suggest two reasons: (i) The overselling of the Xα method during the 1970s. (ii) Grave reservations concerning the numerical precision of early density functional computations. I distinctly remember a discussion with Professor Handy five or six years ago in which we juggled the question, has *any* density functional calculation ever given a correct molecular energy to ± 0.01 hartree, other than fortuitously?

The first point is certainly valid, as there had been many promises to solve long-standing problems. It is often overlooked, however, that some of the bizarre results that were found using the Xα method (such as linear water or an unbound carbon dimer) were a consequence of the use of a "muffin tin" potential *throughout* the "self-consistent" cycle, not from the use of approximation (1.15). There were very few people performing DF calculations on molecules in the 1970s, so the second remark must also apply to our work. Two comments are in order: (a) There is no reason to think that the LMTO basis functions we were using were inferior in any way to the very limited bases then favored in the Gaussian world. (b) The variational principle indicates that wave function-based calculations will indeed give better total energies than DF theory. However, we focused on energy *differences*, such as binding energies, for which there is no variational principle. *All* calculations of energy differences rely on cancellation of errors.

1.5 The Beryllium Dimer

The editors have asked that I "present just one case study where the method has proved to give a substantial leap forward in knowledge." It would be surprising if I did not choose this single case study from my own work. I have chosen the binding energy of the beryllium dimer Be_2, and we shall see that its story also says something about scientists.

1.5.1 The Story to Late 1979

The beryllium dimer (Be_2) is my favorite molecule. With two identical atoms and just four valence electrons ($2\sigma_g^2 2\sigma_u^2$) it is not exactly pretty. However, for many years it resisted discovery, and much theoretical work indicated that this situation would not change. I was astonished when my DF calculations showed that the binding energy should be substantially greater than in its group 12 neighbors in the periodic table (He_2, Mg_2). Either I had an unexpected prediction that could be checked by other methods, or I had a case where the DF method led to a qualitatively incorrect result. An interesting situation, or so it seemed.

My interest in group 12 dimers arose from discussions with a colleague in the Institut für Chemie in Jülich, Chung Wu. He had used a Knudsen source to evaporate metals and study the concentration of clusters of different sizes as a function of temperature. If one makes assumptions about the bond length and the vibration frequency, it is possible to estimate the dissociation energy of the ground state. Chung had performed measurements on Mg_2 and Ca_2, for which other spectroscopic data were available, but his estimates of the binding energies were significantly larger than earlier results of infrared spectroscopy. I decided to carry out calculations on all group 12 dimers.

The results for Mg_2 were consistent with measured values of the bond length, but the vibration frequency and the well depth were both overestimated, a feature we now know follows from the local density (LD) approximation used in these calculations. Absorption spectra had also been measured for the calcium dimer, and the results of the calculations showed similar parallels. I was confident enough in the DF description of the binding in these molecules to extend the calculations to the heavier dimers (Sr_2, Ba_2, and Ra_2), even to the lightest, He_2. I did not expect a good description of the last of these, a prototype van der Waals molecule with a minimum in the binding energy curve that is so shallow that it does not support vibrations (this means that the minimum energy associated with vibrations – the "zero point energy" – is above that of two separated atoms). The DF calculations led, as expected, to a minimum that was too deep.

There were good reasons why I came last to the beryllium dimer. The opinions of many eminent theoretical chemists were overwhelmingly against its existence, and our first applications of the DF formalism to "first-row" molecules [39] omitted Be_2 for precisely this reason! Here are quotes from some of the most cited theoretical chemists of the past decades:

> One would expect that two normal beryllium atoms would behave towards each other as do two normal helium atoms. This expectation is given support by our calculations, which show that the resulting molecular state is repulsive.

J.H. Bartlett, Jr, W.H. Furry [40]

> There is no evidence from this calculation that the ground state of Be–Be is bound ... The CI results are just as repulsive as the SCF results ....

C.F. Bender, E.R. Davidson [41]

> .. the van der Waals bound molecule Be_2, which should have a dissociation energy less than 1.2 kcal/mol, the experimental D_0 for Mg_2 .. To obtain a realistic result, say 0.7 kcal/mol, for the dissociation energy ...

E. Dykstra, H.F. Schaefer III, W. Meyer [42]

> Of course, the ground states of analogous diatomics (Mg_2, Hg_2, Cd_2) are bound due to the much larger dispersion interaction which occurs in these much more polarizable species. ... The second order perturbation theory calculations indicate that the ground state $^1\Sigma_g^+$ potential curve of Be_2 is indeed repulsive.

K.D. Jordan, J. Simons [43]

> Because the ground state of Be_2 has not been detected experimentally, it appears that the van der Waals minimum is very shallow ($\lesssim$ 1 kcal/mol), it may even be too shallow to support vibrational levels.

K.D. Jordan, J. Simons [44]

> Since a system of two Be atoms, each with closed shell ($1s^2 2s^2$) electronic configurations, does not show appreciable bonding, substantial changes must occur in the nature of bonding if beryllium metal is to be formed. ... At the highest level of theory used, RMP4(SDQ)/6-31G*, the bond length is 3.999 Å and the binding is 0.3 kcal/mol.

R.A. Whiteside, R. Krishnan, J.A. Pople, M.-B. Krogh-Jespersen, P. von R. Schleyer, G. Wenke [45]

These articles paint a clear picture: The beryllium dimer has an equal occupancy of bonding and antibonding orbitals and will be unbound in the absence of long ranged "van der Waals forces." These forces are proportional to the atomic polarizability, which lies in Be between the values of He and Mg. As noted in [45], this implies that the bonding in larger Be clusters must be substantially different from that in the dimer, or beryllium metal would not form.

I carried out the calculations in the last quarter of 1978, completing the last (Be_2) in early December. The results were quite disconcerting: Not only should Be_2 exist, but its equilibrium separation (2.57 Å) would be shorter than in Mg_2 and about half of the anticipated minimum of the binding energy curve. Moreover, the binding energy should be significantly *greater* than that of Mg_2. The results are shown in Fig. 1.4, where DF calculations of the cohesive energies of the bulk elements and experimental values are shown where available. The obvious similarities between the two curves suggest that binding in the diatomic molecules and the bulk has the same origin. A shorter bond length in Be_2 than in Mg_2 would be consistent with the measured lattice constants in the bulk materials.

I planned to submit the results for publication, but I was uncertain enough to seek the reaction of theoretical chemists before doing so. Just before Christmas (21 December 1978) I wrote to Prof. Werner Kutzelnigg in Bochum, who headed one of the leading theoretical groups in Germany. Bochum is just 100 km from Jülich, and I had given a seminar in this group in February 1978. We arranged

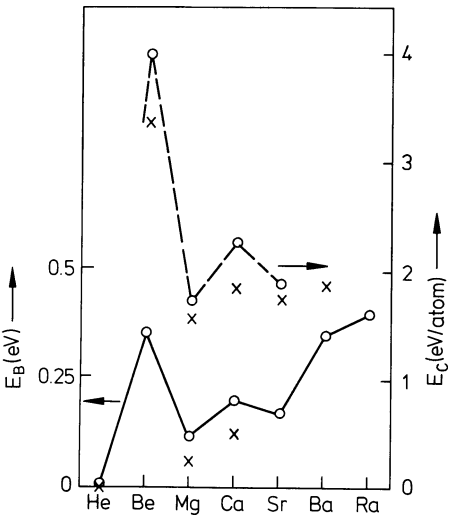

Fig. 1.4 Well depths calculated for $^1\Sigma_g^+$ state of group 12 dimers (*solid line*, *left* scale [46]) and cohesive energies of bulk materials (*dashed line*, *right* scale [47]). Experimental values (*crosses*) are given were known [46]

a second seminar on 7 February 1979. In the meantime I began analyzing and writing up the results. I also tried several times to convince Chung Wu to perform measurements on Be_2, but that is another story.

I have vivid memories of my second Bochum seminar. It was obvious that nobody believed my results, and the discussion after my presentation focused solely on *which* mistake I must have made, not on whether or not there was one. I asked what I had done wrong in Be_2 that had apparently been right in Mg_2, but there were ready answers. Herbert Kollmar agreed to perform additional calculations on Be_2, and I returned to Jülich. Prof. Kutzelnigg wrote to me shortly afterwards (19 February 1979) with the results of these calculations for an interatomic separation of 5.0 a.u. (2.65 Å) (a little longer than the minimum in my binding energy curve), informing me that[2]:

> The molecule is certainly not bound at this separation, but (by ca. 0.5 kcal/mol) repulsive. The minimum is then at greater distances and is obviously much weaker than you imagine, but in agreement with the pseudopotential calculations that Herr Schwartmann has shown you, where Mg_2 has a clearly deeper minimum.

In the following 2 weeks I completed the manuscript. The calculations indicated that *sp*-polarization is the origin of binding in these molecules, and I showed that the spatial overlap between the valence *s*- and *p*-orbitals was greatest in beryllium. I sent the manuscript to the *Journal of Chemical Physics* at the University of Chicago on 13 March 1979, and I left for Australia with my family shortly afterwards. I included several minor improvements suggested by the referee, and the manuscript was accepted on 2 May 1979 and published on 1 August 1979 [46].

The absence of experimental evidence for the existence of the beryllium dimer was very frustrating, of course, and I tried again to convince Chung Wu of the interest that the identification of the molecule would cause. I had no success, and I understood his reluctance to work with a metal whose vapor is highly toxic. Shortly afterwards, Olle Gunnarsson returned from the 46th Nobel Symposium in Aspenäsgården (near Göteborg, Sweden from 11 to 16 June 1979). Rod Bartlett had told him there that correlated wave function calculations on Be_2 by Bowen Liu (IBM San Jose) had led to a minimum between 4 and 5 a.u. with a binding energy of several kcal/mol. On 26 June 1979 I sent copies of my manuscript to Bowen Liu, with a request for more information, and to Walter Balfour (University of Victoria, BC, Canada), who had observed Mg_2 while at the National Research Council laboratories in Ottawa. I asked whether he could do similar measurements on Be_2.

---

[2] W. Kutzelnigg to ROJ (19 February 1979)
"Herr Kollmar hat inzwischen Be_2 bei 5 a.u. (2.65 Å) gerechnet. Das Molekül ist bei diesem Abstand sicher nicht bindend, sondern (mit ca. 0.5 kcal/mol) abstoßend. Das Minimum liegt wohl bei größeren Abständen und ist offenbar viel schwächer als Sie vermuten, aber in Einklang mit den Pseudopotentialrechnungen, die Ihnen Herr Schwartmann gezeigt hat, wonach Mg_2 ein deutlich tieferes Minimum hat."

Balfour responded on 4 July 1979 with more information on Mg_2 and Ca_2 and the news of unsuccessful attempts by Reginald Colin at NRC to find the beryllium dimer. Colin (who had moved to the Université Libre in Brussels) provided me subsequently with more details of a difficult experiment (18 July 1979). Bowen Liu did not answer my letter and he was not particularly forthcoming when I phoned him, so I wrote on 7 August 1979 to his collaborator Douglas McLean. Doug came originally from Western Australia, and he had spent a year (1962) in the Chemistry Department of the University of Western Australia in Perth when I was an Honours student in Physics there. He responded at length (27 August 1979) with full details of the calculations and the results. Liu and McLean had found a minimum at 4.75 a.u. and a well depth of $810\,\mathrm{cm}^{-1}$, but the calculation (ICF, interacting correlated fragments) left several questions open.

I had not received the promised details of Herbert Kollmar's calculations, and I wrote to him on 30 July. He responded (1 August) that he found no minimum in CEPA (coupled electron pair approximation) calculations for a Be–Be separation near 5 a.u. On 7 August Prof. Kutzelnigg wrote a letter with an unambiguous message[3]:

> Herr Kollmar has shown me your letter and his response. In order to avoid misunderstanding, I should like to add the following remarks. (....) This indicates that the CEPA-curve should be very similar to the exact curve - even if you like the IEPA-curve better, because it is similar to yours. Perhaps you neglect exactly those inter-intra correlation effects that are ignored by IEPA.

Several weeks later (21 September), however, Herbert Kollmar informed me of a presentation by Paul Bagus from IBM San Jose at a meeting held in Bad Neuenahr (18 September), where the results of Liu and McLean were presented (Be_2: $r_e = 2.49\,\text{Å}$, $D_e = 814\,\mathrm{cm}^{-1}$, $\nu_1 = 202\,\mathrm{cm}^{-1}$, $\nu_2 = 132\,\mathrm{cm}^{-1}$). He wrote:

> The results confirm your calculations; the discrepancy to our calculations is unclear.[4]

The results of Liu and McLean were submitted to *Journal of Chemical Physics* as a Note on 10 December 1979 and published on 1 March 1980 [48]. This article referred to my *JCP* publication.

[3] W. Kutzelnigg to ROJ (7 August 1979)
"Herr Kollmar hat mir Ihren Brief und seine Antwort gezeigt. Damit keine Mißverständnisse auftreten, möchte ich noch folgende ergänzende Bemerkungen machen. (.....) Dies spricht schon dafür, dass die CEPA-Kurve der exakten Kurve sehr ähnlich sein muss - auch wenn Ihnen die IEPA-Kurve sympatischer sein sollte, weil sie Ihrer ähnlich ist. Vielleicht vernachlässigen Sie auch genau jene inter-intra Korrelationseffekte, die man bei IEPA vernachlässigt."

[4] H. Kollmar to ROJ (21 September 1979)
"Die Ergebnisse bestätigen Ihre Rechnungen; unklar ist die Diskrepanz zu unseren Rechnungen."

1.5.2 1980–1984

The density functional results for Be_2 were so different from all preceding work that they were greeted with much scepticism. This was also true to some extent for the work of Liu and McLean. A notable exception was Volker Heine, the adviser of my Ph.D. work, who told me:

> It is perfectly obvious to any solid-state physicist that Be_2 must be more strongly bound than Mg_2.

He knew immediately that this was a consequence of the relative compactness of the $2p$-orbitals in the first-row elements, but few other condensed matter physicists were interested.

It was by no means obvious that the beryllium dimer could be identified using the mass-spectrometric methods in use at the time. The melting and boiling points of Be are far higher than Mg, and – even if one could develop an appropriate furnace – such high temperatures lead to very unstable clusters. Other DF calculations led to similar results to mine, and Lengsfield et al. [49] extended the correlated wave function work of Liu and McLean to larger basis sets and more extensive methods of including correlation effects. The location of the minimum of the Be_2 binding energy curve changed very little (the revised best estimate was $r_e = 4.73 \pm 0.03$ a.u, $D_e = 2.04 \pm 0.21$ kcal mol^{-1}), and the work was published on 4 May 1983.

I was surprised that this paper did not refer to my *JCP* article 4 years earlier, and I thought that a "Comment" to the journal was appropriate. I added some newer results on Be_2, as well as those of Painter and Averill [50]. The Editor sent it to Bowen Liu, who noted the lack of new results and suggested asking the opinion of a neutral referee. Liu added, however:

> In the event that you decide to publish the comment, we would like an opportunity to reply to the claim that 'density functional calculations are simpler to interpret than CI calculations'.

The "neutral" referee wrote:

> No new results or analyses are presented here. This contribution is mainly in the nature of a polemic in favor of more serious consideration of density functional results in studies of chemical binding. All the data and arguments on this point are already in the literature, and no useful purpose is served by publishing this discussion.
>
> While I agree that it would have been appropriate for Lengsfield et al. to mention the density functional results, their omission is not a sufficient reason for the publication of this comment. In any case, the ab initio calculation is a definitive calculation which establishes a standard of accuracy. I do not think that any density functional calculation can be regarded in the same light.

It is interesting to read these remarks over 20 years later, when most theoretical chemists value density functional calculations for opening up the study of much larger systems than possible with more traditional (wave function based) methods *and* the single-particle picture that eases the interpretation of the results.

Nevertheless, it is not surprising that the Editor, John Light, rejected the article and equally unsurprising that I responded to these comments.[5]

I had heard so often about the difficulties facing the experimental identification of Be_2 that it came as a real surprise when it was carried out [51] (published 31 August 1984). Vladimir Bondybey at Bell Laboratories used liquid N_2 to cool Be vapor, after which laser-induced fluorescence from Be_2 was measured. The $^1\Sigma_g^+$ ground state was found to be characterized by $r_e = 2.45$ Å, $D_e = 790 \pm 30\,\mathrm{cm}^{-1}$, $\omega_e = 275.8\,\mathrm{cm}^{-1}$. I am sure that Liu and McLean were as pleased with this result as I was. The minimum in the binding energy curve *is* much shorter than in all work prior to 1979, and I had no doubt that my picture of the nature of the bond was right.

1.5.3 After 1984

The experimental confirmation of the DF prediction that the bond is stronger in Be_2 than in Mg_2, so that the binding trends were similar to those in the solids, might have helped the acceptance of density functional methods by chemists. It did not. I met many who remained convinced by the weight of literature favoring a very weak bond, and others focused on the overbinding that resulted from our use of the local density approximation in the DF calculations. A representative view came from John Light, Editor of the *Journal of Chemical Physics* and *ex officio* a member of the chemical establishment. During the refereeing process of a paper I

[5] ROJ to J. C. Light, Editor, *Journal of Chemical Physics* (29 November 1983)
Thank you for the author's and referee's response to the above comment. Given the content of these remarks, I would probably have acted as you did. Nevertheless, it would be surprising if I had nothing further to say on the matter.

I think that you should make it quite clear in your Announcement that you are less interested in "discussion and comments" than in controversy. The first draft of my comment would have been much closer to your requirements, and I shall bear your change of policy in mind for future occasions.

Over the past few years, I have had numerous opportunities to discuss the density functional method and its applications with quantum chemists. I have found that most chemists of a particular generation are incapable of giving the method serious consideration. My visit to San Jose in August indicated that Dr. Liu is no exception and his letter comes as no surprise. The last sentence of the referee's remarks indicates that he, too, subscribes to the prevailing view of the CI fraternity.

For some years, I have been attempting to convince chemists that density functional calculations are very useful in certain contexts and solid state physicists that the approximations they use in DF calculations can lead to unreliable results. I have not yet had much success in either direction. Quantum chemists and solid state physicists can learn a lot from each other, but the communication is poor and is not helped by editorial decisions to throttle scientific discussion. Given the amount of rubbish generated by earlier CI calculations on Be_2, it would have been interesting to hear Dr. Liu's arguments for the relative simplicity of their interpretation. Making such a comparison would also have provided him and his colleagues with the opportunity to discover what the density functional formalism is.

submitted on the carbon trimer C_3, I asked him about his view of the DF method and its appropriateness for his journal. His response is given below.[6]

I had spent the first 4 months of 1984 at the Max-Planck-Institut für Festkörperforschung in Stuttgart, after which Olle Gunnarsson and I had a clear picture of a major source of error in DF calculations using the local density approximation [33]. There is no doubt that LD calculations for Be_2 will lead to overbinding, and the experimental result for Be_2 told us by how much. It comes as no surprise that other efforts were made to improve on the initial calculations of Liu and McLean. One of the most extensive was that of Petersson and Shirley [52], who studied the convergence of wave function-based calculations for Be_2. After extrapolation to the complete basis set and full CI limits, they found good agreement with measured values. My 1979 paper was not cited.

The DF calculation and the prediction that Be_2 should be more stable than Mg_2 were pieces in the mosaic that led to the breakthrough of DF calculations in chemistry. As noted by Nicholas Handy and coworkers, it was one of the first successes in DF applications to small molecules [53]:

> Jones was indeed the first theoretician to suggest that Be_2 has a potential minimum near 2.45 Å, and although his well depth at 8 kcal/mol was too deep, his frequency was good at 300 cm^{-1}. All other *ab initio* calculations to that date (1979) had predicted a minimum near 4.5 Å, principally because of deficient basis sets. Jones result was one of the first successes of DFT to the study of small molecules.

I still like my *JCP* article of 1979, where my tone was appropriately cautious, and I am pleased that I did not yield to the weight of opinion of the chemical establishment concerning binding in the beryllium dimer. Chemistry is a conservative field, and I cannot help but wonder whether established opinion is as wrong in other areas as it was in this one.

One postscript might also be interesting: I have not been to the Theoretical Chemistry group in Bochum since 1979. However, my colleague John Harris gave a seminar at the 29th Symposium for Theoretical Chemistry, held at Oberwiesenthal

---

[6] J.C. Light, Editor, Journal of Chemical Physics to ROJ (16 November 1984)

"Thank you for your letter of 2 November 1984 concerning the above manuscript. In order to try to evaluate the problem better, I have read through your manuscript. The problem, as I see it, is that despite the large number of manuscripts published in the Journal of Chemical Physics on this method, many chemists remain to be convinced of its value.

There appear to be two reasons for this. First, the local density approximation introduces an uncertainty about the reliability of the results which is compounded by the use of muffin-tin orbitals. Second, many chemists prefer the more rigorous and familiar approaches even if they be computationally more computer intensive. It is also clear, from looking at some of the published papers listed, that the approach has continued to evolve (i.e., different parameterizations of E^{xc}).

This leaves me, and probably the reviewer, with an unhappily ambiguous feeling toward this manuscript. it does not treat the methodology problem (clearly stated in JCP **79**, 1874, part IV. D) but gives only one example where the method apparently works well, whereas for the NH_3 inversion barrier it doesn't.

I'm not sure that I have answered your questions as to criteria, which are probably decided empirically by averaging over referees."

(Erzgebirge) from 27 September to 1 October 1993. In the discussion following the talk, which considered the role of density functional calculations in chemistry, Prof. Kutzelnigg said to John that he acknowledged our "triumph." He had noticed, after all.

1.6 Concluding Remarks

The personal view I referred to in the title should now be clear; the density functional approach is the culmination of a sequence of contributions dating back to the earliest days of quantum mechanics. Thomas and Fermi developed the first theory where the electron density is the basic variable, and Dirac (1930) not only incorporated exchange effects into this theory, but made the remarkable observation that the density determines completely the state of an atom. It is not necessary to specify the wave function. The self-consistent field of Hartree and the use of determinantal wave functions by Slater and Fock (the Hartree–Fock method) and by Bloch (ferromagnetism) were followed by the calculations of Wigner and Seitz, who showed that the exchange hole in Na metal was localized to a single ion. This picture was developed by Slater in 1951, and an effective exchange potential of Dirac form was derived and tested by Gáspár in 1954. The work of Hohenberg and Kohn [28] and of Kohn and Sham [30] is known to everybody.

Kohn and Sham [30] did not expect the local density approximation for exchange and correlation to give "an accurate description of chemical binding," and they were not alone. Nevertheless, LD (LSD) and Hartree–Fock–Slater calculations on molecules were carried out by several groups – including our own in Jülich – from the mid-1970s. In fact, Mel Levy told me in Santa Barbara in August 1983 that one reason for the success of our project was that we had not understood that such calculations could not possibly work! More than 15 years were to pass after we began our work before DF methods found widespread acceptance in chemistry, but it is now known that the LD and LSD approximations and their many improvements have provided us with a scheme that gives valuable information about large molecules and extended systems. It does appear to satisfy the 80-year-old goal of Dirac [1] to find "approximate practical methods of applying quantum mechanics to [explain] the main features of complex atomic systems without too much computation." Those who know how many DF calculations I have done in the past 30 years know that I did not take the last goal all that seriously.

Those readers working on "strongly correlated systems" know that this is not the whole story. There are many cases where DF calculations using standard approximations lead to errors that are quite unacceptable. Striking examples are found in atoms, where excitation energies between different states are often poor. The energy to transfer a d- to an f-electron in the rare earth and lanthanide atoms, for example, differs by around 2 eV from measured values (Fig. 1.5). There are smaller, but substantial, errors in s–d transfer energies in transition element atoms. Many such problems can be traced to changes in the nodal structure of the wave

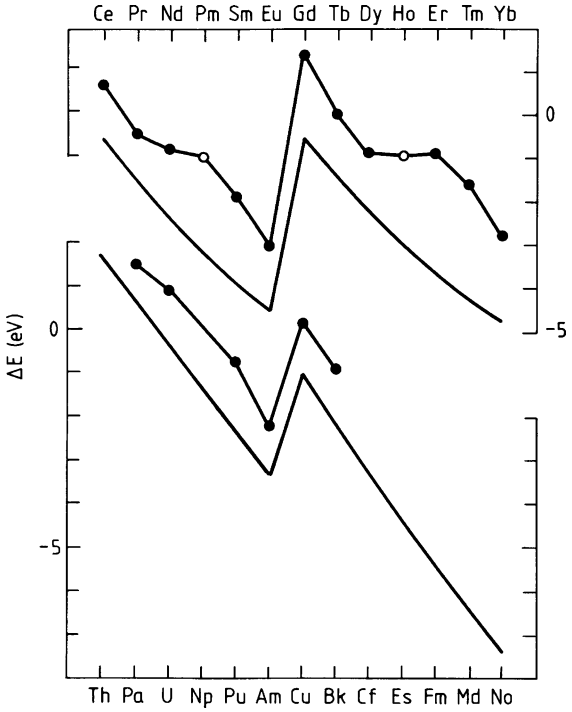

Fig. 1.5 The df-transfer energies Δ_{df} $[E(f^{n-3}ds^2) - E(f^{n-2}s^2)]$ for rare earth and lanthanide atoms [3] [D. Glötzel, private communication (1980). The solid lines are calculations for spherically averaged densities]

function of the system, which can lead to large changes in the exchange energy that are very difficult to describe with local density approximations [33].

I conclude with a word of caution for the "strongly correlated" community. Very few theoretical chemists thought that DF calculations were relevant to understanding the electronic structure of molecules, but local density approximations (and their modifications) have given far better results than *anyone* expected. It was shown afterwards (see, e.g., Sect. 1.4.3) why approximations to $E_{\rm xc}$ could give good results for density distributions far from those where they are obviously valid. Recent modifications to such approximations are leading to small but significant improvements in the density functional description of rather complicated systems. Perhaps DF theory has some real surprises in store for the "strongly correlated" world.

Acknowledgments The historical part of this article developed out of seminars given in recent years at NIST (Gaithersburg), Cornell University, Forschungszentrum Jülich, and the Universities of Jyväskylä, Bayreuth, and the Saarland (Saarbrücken). I thank members of these audiences for their comments and Richard Jones for suggestions on the manuscript. Our work on DF problems has been the effort of many collaborators over decades, but Olle Gunnarsson and John Harris

deserve special mention for essential contributions in those exciting, early years. Any successes we have had also reflect the spirit of free scientific enquiry in the research group of Gert Eilenberger, whose focus on curiosity-driven research was not deterred by many pressures to the contrary. Funding agencies and others define today what research is "useful," and we were fortunate indeed to experience a different world.

Appendix 1: Thomas–Fermi (TF) Equations

The electrons are treated as independent particles, and the electron-electron interaction energy arises solely from the electrostatic energy,

$$E_{\rm es}[n] = \frac{e^2}{2}\int {\rm d}\boldsymbol{r}\int {\rm d}\boldsymbol{r}'\,\frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{1.23}$$

It is also assumed that the kinetic energy is

$$T[n] = \int {\rm d}\boldsymbol{r}\; t[n(\boldsymbol{r})], \tag{1.24}$$

where $t[n]$ is the kinetic energy density for a system of noninteracting electrons with density n. We have

$$t[n] = 2\frac{1}{(2\pi)^3}\int_{|k|\le k_{\rm F}} {\rm d}\boldsymbol{k}\;\frac{\hbar^2k^2}{2m}, \tag{1.25}$$

where $2(4\pi/3){k_{\rm F}}^3/(2\pi)^3 = n$. This gives

$$T_0[n] = C_k\int {\rm d}\boldsymbol{r}\;\{n(\boldsymbol{r})\}^{\frac{5}{3}}, \tag{1.26}$$

where $C_k = 3\hbar^2(3\pi^2)^{\frac{2}{3}}\;10{\rm m}^{-1}$.

If $E[n]$ is minimized under the constraint that the number of electrons is constant,

$$\int {\rm d}\boldsymbol{r}\; n(\boldsymbol{r}) = N. \tag{1.27}$$

Using the method of Lagrange multipliers, we require that

$$E[n] + \lambda N = T[n] + E_{\rm es}[n] + \int {\rm d}\boldsymbol{r}\; n(\boldsymbol{r})\{V_{\rm ext}(\boldsymbol{r}) + \lambda\} \tag{1.28}$$

has a minimum. The corresponding Euler equation is the Thomas–Fermi equation,

$$\frac{5}{3}C_k n(\boldsymbol{r})^{\frac{2}{3}} + e^2\int {\rm d}\boldsymbol{r}'\,\frac{n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} + V_{\rm ext}(\boldsymbol{r}) + \lambda = 0. \tag{1.29}$$

Appendix 2: The Density Functional Formalism

We follow the approach of Levy [29], who provided a simpler and more general derivation of the DF formalism than that given by Hohenberg and Kohn [28]. The Hamiltonian for N electrons moving in an external potential $V_{\rm ext}(\boldsymbol{r})$ is

$$H = T + V_{\rm ee} + \sum_{i=1}^{N} V_{\rm ext}(\boldsymbol{r}_i), \tag{1.30}$$

where T and $V_{\rm ee}$ are the kinetic and electron-electron interaction operators, respectively. For all "N-representable" densities $n(\boldsymbol{r})$, i.e., those obtainable from some antisymmetric wave function, $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N)$, Levy defined the functional

$$F[n] = \min_{\psi \to n} \langle \psi | T + V_{\rm ee} | \psi \rangle, \tag{1.31}$$

where the minimum is taken over all ψ that give the density n. $F[n]$ refers neither to a specific system nor to the external potential $V_{\rm ext}(r)$, so may be viewed as "universal." If $E_{\rm GS}$, $\psi_{\rm GS}$, and $n_{\rm GS}(\boldsymbol{r})$ are the ground state energy, wave function, and density, respectively, then the basic theorems of DF theory are

$$E[n] \equiv \int {\rm d}\boldsymbol{r}\ V_{\rm ext}(\boldsymbol{r}) n(\boldsymbol{r}) + F[n] \geq E_{\rm GS} \tag{1.32}$$

for all N-representable $n(\boldsymbol{r})$, and

$$\int {\rm d}\boldsymbol{r}\ V_{\rm ext}(r) n_{\rm GS}(\boldsymbol{r}) + F[n_{\rm GS}] = E_{\rm GS}\ . \tag{1.33}$$

To prove the variational principle (1.32) we write $\psi^n_{\rm min}(\boldsymbol{r})$ for a wave function that minimizes (1.31), so that

$$F[n] = \langle \psi^n_{\rm min} | T + V_{\rm ee} | \psi^n_{\rm min} \rangle\ . \tag{1.34}$$

Writing $V = \sum_i V_{\rm ext}(\boldsymbol{r}_i)$, we have

$$\int {\rm d}\boldsymbol{r}\ V_{\rm ext}(\boldsymbol{r}) n(\boldsymbol{r}) + F[n] = \langle \psi^n_{\rm min} | V + T + V_{\rm ee} | \psi^n_{\rm min} \rangle \geq E_{\rm GS}, \tag{1.35}$$

according to the minimum property of the GS. This proves (1.32). A further use of the minimum property leads to

$$E_{\rm GS} = \langle \psi_{\rm GS} | V + T + V_{\rm ee} | \psi_{\rm GS} \rangle \leq \langle \psi^{n_{\rm GS}}_{\rm min} | V + T + V_{\rm ee} | \psi^{n_{\rm GS}}_{\rm min} \rangle. \tag{1.36}$$

We subtract the interaction with the external potential and obtain

$$\langle \psi_{\rm GS}|T+V_{\rm ee}|\psi_{\rm GS}\rangle \le \langle \psi_{\rm min}^{n_{\rm GS}}|T+V_{\rm ee}|\psi_{\rm min}^{n_{\rm GS}}\rangle. \tag{1.37}$$

On the other hand, the definition of $\psi_{\rm min}^{n_{\rm GS}}$ yields the reverse relation between the two sides of (1.37). This is possible only if

$$\langle \psi_{\rm GS}|T+V_{\rm ee}|\psi_{\rm GS}\rangle = \langle \psi_{\rm min}^{n_{\rm GS}}|T+V_{\rm ee}|\psi_{\rm min}^{n_{\rm GS}}\rangle. \tag{1.38}$$

Then we have

$$\begin{aligned} E_{\rm GS} &= \int {\rm d}\boldsymbol{r}\ V_{\rm ext}(\boldsymbol{r})n_{\rm GS}(\boldsymbol{r}) + \langle \psi_{\rm GS}|T+V_{\rm ee}|\psi_{\rm GS}\rangle \\ &= \int {\rm d}\boldsymbol{r}\ V_{\rm ext}(\boldsymbol{r})n_{\rm GS}(\boldsymbol{r}) + \langle \psi_{\rm min}^{n_{\rm GS}}|T+V_{\rm ee}|\psi_{\rm min}^{n_{\rm GS}}\rangle \\ &= \int {\rm d}\boldsymbol{r}\ V_{\rm ext}(\boldsymbol{r})n_{\rm GS}(\boldsymbol{r}) + F[n_{\rm GS}]\,. \end{aligned} \tag{1.39}$$

Another result follows from (1.39). If the ground state GS is nondegenerate, $\psi_{\rm min}^{n_{\rm GS}} = \psi_{\rm GS}$. If the ground state is degenerate, $\psi_{\rm min}^{n_{\rm GS}}$ is equal to one of the GS wave functions, and the others can be obtained by symmetry. The GS charge density then determines the GS wave function(s), from which all GS properties can be calculated. These properties are therefore functionals of the density. These theorems provide the formal justification for working with the density instead of the wave function and a general method for calculating GS properties.

References

1. P.A.M. Dirac, Proc. Camb. Phil. Soc. **26**, 376 (1930)
2. F.H.C. Crick, *What Mad Pursuit* (Penguin, London, 1988), p. 150
3. R.O. Jones, O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
4. G. Segrè, *Faust in Copenhagen: A Struggle for the Soul of Physics* (Jonathan Cape, London, 2007)
5. D.R. Hartree, Proc. Camb. Phil Soc. **24**, 89, 111 (1928)
6. J.C. Slater, Phys. Rev. **35**, 210 (1930)
7. V. Fock, Z. Physik **61**, 126 (1930)
8. J.C. Slater, Phys. Rev. **34**, 1293 (1929)
9. F. Bloch, Z. Physik **57**, 545 (1929)
10. P.A.M. Dirac, Proc. Roy. Soc. (London) A **123**, 714 (1929)
11. L.H. Thomas, Proc. Camb. Phil Soc. **23**, 542 (1927)
12. E. Fermi, Z. Physik **48**, 73 (1928)
13. E.H. Lieb, Rev. Mod. Phys. **53**, 601 (1981)
14. E.H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983)
15. L. Spruch, Rev. Mod. Phys. **63**, 151 (1991)
16. E. Teller, Rev. Mod. Phys. **34**, 627 (1962)

17. E.H. Lieb, B. Simon, Phys. Rev. Lett. **31**, 681 (1973)
18. P. Ehrenfest, Z. Physik A **45**, 455 (1927)
19. E. Wigner, F. Seitz, Phys. Rev. **43**, 804 (1933); E. Wigner, F. Seitz, Phys. Rev. **46**, 509 (1934)
20. J.C. Slater, Rev. Mod. Phys. **6**, 209 (1934)
21. J.C. Slater, Phys. Rev. **81**, 385 (1951)
22. J.C. Slater, Phys. Rev. **51**, 846 (1937)
23. J. Korringa, Physica **13**, 392 (1947)
24. W. Kohn, N. Rostoker, Phys. Rev. **94**, 111 (1954)
25. J.C. Slater, Phys. rev. **165**, 658 (1968)
26. R. Gáspár, Acta Phys. Hung. **3**, 263 (1954)
27. J.C. Slater, in *The World of Quantum Chemistry*, ed. by R. Daudel, B. Pullman. (D. Reidel, Dordrecht, 1974), pp. 3–15
28. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964)
29. M. Levy, Proc. Nat. Acad. Sci. (USA), **76**, 6062 (1979)
30. W. Kohn, L. J. Sham, Phys. Rev. **140**, A1133 (1965)
31. O. Gunnarsson, B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976)
32. J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003)
33. O. Gunnarsson, R.O. Jones, Phys. Rev. B **31**, 7588 (1985)
34. J. Harris, R.O. Jones, J. Phys. F **4**, 1170 (1974)
35. V. Heine, Ψ_k Newsletter, Ab Initio (from Electronic Structure) Calculations of Complex Processes in Materials, **50**, 7 (2002). www.psi-k.org/newsletters/News_50/newsletter_50.pdf
36. A.D. Becke, Phys. Rev. A **38**, 3098 (1988)
37. B.G. Johnson, P.M.W. Gill, J.A. Pople, J. Chem. Phys. **97**, 8846 (1992)
38. H.F. Schaefer III, J. Phys. Chem. **100**, 6003 (1996).
39. O. Gunnarsson, J. Harris, R.O. Jones, J. Chem. Phys. **67**, 3970 (1977)
40. J.H. Bartlett Jr, W.H. Furry, Phys. Rev. **38**, 1615 (1931)
41. C.F. Bender, E.R. Davidson, J. Chem. Phys. **47**, 4792 (1967)
42. E. Dykstra, H.F. Schaefer III, W. Meyer, J. Chem. Phys. **65**, 5141 (1976)
43. K.D. Jordan, J. Simons, J. Chem. Phys. **65**, 1601 (1976)
44. K.D. Jordan, J. Simons, J. Chem. Phys. **67**, 4027 (1977)
45. R.A. Whiteside, R. Krishnan, J.A. Pople, M.-B. Krogh-Jespersen, P. von R. Schleyer, G. Wenke, J. Comput. Chem. **1**, 307 (1980)
46. R.O. Jones, J. Chem. Phys. **71**,1300 (1979)
47. V.L. Moruzzi, J.F. Janak, A.R. Williams, *Calculated Electronic Properties of Metals* (Pergamon, New York, 1978)
48. A.D. McLean, B. Liu, J. Chem. Phys. **72**, 3418 (1980) [Received 10 December 1979, accepted 17 December 1979, published 1 March 1980]
49. B.H. Lengsfield III, A.D. McLean, M. Yoshimine, B. Liu, J. Chem. Phys. **79**, 1891 (1983) [Received 16 March 1983, accepted 4 May 1983, published 15 August 1983]
50. G.S. Painter, F.W. Averill, Phys. Rev. B **26**, 1781 (1982)
51. V.E. Bondybey, Chem. Phys. Lett. **109**, 436 (1984) [Received 9 July 1984, published 31 August 1984]
52. G.A. Petersson, W.A. Shirley, Chem. Phys. Lett. **160**, 494 (1989)
53. C.W. Murray, N.C. Handy, R.D. Amos, J. Chem. Phys. **98**, 7145 (1993)

Chapter 2
Projected Wavefunctions and High T_c Superconductivity in Doped Mott Insulators

Mohit Randeria, Rajdeep Sensarma, and Nandini Trivedi

Abstract We review the use of projected wavefunctions to gain insight into the strongly correlated d-wave superconducting state of high T_c cuprates within the framework of the large U Hubbard model. Using sum rules, we show that doped Mott insulators exhibit a strong particle-hole asymmetry in their single-particle spectral function. We calculate the doping dependence of a variety of observables using a simple approximation scheme, the Gutzwiller approximation, and compare the results with variational Monte Carlo results and with experimental data on the cuprates. We gain detailed insights into the superconducting dome, the energy gap, nodal excitations, their quasiparticle weight and dispersion, momentum distribution, superfluid stiffness, and optical spectral weight. We show that strong correlations make the d-wave state robust against disorder-induced pair breaking. Finally, we discuss the competition between antiferromagnetism and superconductivity, the difference between hole and electron-doped Mott insulators, and how the range of hopping enhances superconductivity for hole-doped materials.

2.1 Introduction

In 1986, Bednorz and Müller [1] discovered high-temperature superconductivity in the cuprates by adding mobile carriers to an antiferromagnetic Mott insulator. A quarter of a century later, we are still debating the physics of these materials. The problem has turned out to be nontrivial because the phase diagram of the

M. Randeria (✉) · N. Trivedi
Physics Department, The Ohio State University, Columbus, OH 43210, USA
e-mail: randeria@mps.ohio-state.edu; trivedi@mps.ohio-state.edu

R. Sensarma
Condensed Matter Theory Center, University of Maryland, College Park, MD 20742, USA
e-mail: sensarma@umd.edu

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_2,

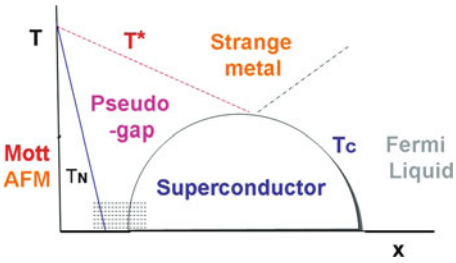

Fig. 2.1 Schematic phase diagram of the hole-doped high-temperature superconductors in the doping-temperature plane

cuprates, a cartoon of which is shown in Fig. 2.1, represents a challenge for three of the central paradigms of twentieth century condensed matter physics. First, the parent (undoped) insulator represents a failure of band theory. Coulomb correlations give rise to a Mott insulator. Second, the "normal" (i.e., nonsuperconducting) phase is highly abnormal. Landau's Fermi liquid theory, the underpinning of our understanding of conventional metals, fails in the strange metal and pseudogap regimes of the phase diagram. The electron is not a sharply defined excitation above T_c. Third, even the superconducting phase, now known to have d-wave pairing symmetry and supporting well-defined electronic excitations, cannot be described by BCS theory. Its very short coherence length, a superfluid phase stiffness much smaller than the energy gap (for low doping), and its insensitivity to disorder, all make the superconducting state different from a standard BCS superconductor. To add to all this, there are a plethora of possible competing order parameters in the vicinity of the Mott insulator and possible quantum critical points.

In a seminal paper written within months of the experimental discovery, Anderson [2] proposed that the basic physics of the copper–oxygen planes in the cuprates could be captured by the single-band Hubbard model on a 2D square lattice. He further argued that these materials are in the large U limit, where the on-site Coulomb repulsion is much larger than the kinetic energy, and thus one is restricted to a projected Hilbert space in which no double occupancy is allowed. Implementing this projection has turned out to be hard, since perturbative methods are not applicable. In this review, we describe one approach, that of projected wavefunctions, and an approximate calculation scheme, the Gutzwiller approximation, which have turned out to give enormous insight into the strongly correlated d-wave superconducting state at $T = 0$, its low-lying excitations and its response functions.

Quite apart from its usefulness in understanding the physics of high T_c cuprates, the Hubbard model is the simplest lattice model of interacting fermions, occupying a place in many-body theory analogous to the Ising model in classical statistical mechanics. However, not much is known definitively about the Hubbard model outside of one dimension (where one can use Bethe ansatz and other techniques), infinite dimensions (where dynamical mean field theory (MFT) becomes exact), and weak coupling. Thus, nonperturbative progress on this problem is very important, and comparison with experiments provides a useful check. One must keep in mind that the Hubbard model is a minimal model for real materials, and it is actually quite

surprising how much of the phenomenology of the superconducting state of the cuprates it can capture. Progress on the anomalous normal states has been perhaps less successful, in part due to a lack of technique. Although new developments like cluster dynamical MFT are beginning to provide very interesting numerical results.

In the remainder of this section, we briefly describe early work which provided the foundation for our studies, and then end with an outline of the rest of the review. Projected states have been associated with the idea of *resonating valence bond* (RVB) wavefunctions [2]. Any singlet state can be written as a superposition of configurations with singlet pairs, or valence bonds, on a lattice. (Note that these are *not*, in general, near-neighbor pairs). In the half-filled case, this representation may be most useful for spin-liquid states that do not break any symmetries, but can also describe states with long range antiferromagnetic order [3]. Here, our primary focus is on systems doped away from half-filling, where the wavefunctions for the projected d-wave superconductor involve both singlet bonds and holes. The valence bonds are in a sense the "Cooper pairs," but it is the holes that permit transport and hence superconductivity of pairs in a doped Mott insulator.

These ideas were first explored in the framework of slave-boson MFT [4, 5], and implementing the no double-occupancy constraint beyond mean field leads to gauge theories [6]. A parallel development used variational wavefunctions [7] and the Gutzwiller approximation [8]. It was realized early on in these strong coupling approaches [5, 7] that d-wave pairing is the most natural symmetry for the repulsive Hubbard model. This had already been recognized from weak coupling studies of spin fluctuations in the Hubbard model [9]. From the large U perspective, the superexchange J acts as the near-neighbor attraction to bind a pair, while the Hubbard U suppresses the on-site amplitude, thus leading to d-wave pairing.

The projected wavefunction approach was revived a decade later by Paramekanti et al. [10, 11] who calculated a variety of experimentally relevant quantities and made predictions and comparisons with data on high T_c cuprates. These results, and later developments, which give a remarkably complete description of the strongly correlated superconducting state will be described in detail below. Related reviews may be found in [6, 12, 13]. Note that we give references to a very small fraction of the extensive experimental literature on high T_c cuprates, referring only to those papers with which we make direct contact.

The rest of the article is organized as follows. In Sect. 2.2, we introduce the Hubbard model and the idea of projection onto a Hilbert space with no double occupancy. We also describe a canonical transformation that builds in the effects of a finite but large local repulsion U and discuss the relation between the large U Hubbard and tJ models. We focus exclusively on hole-doped systems (except at the end in Sect. 2.10).

In Sect. 2.3, we derive exact sum rules for the spectral function in projected systems and show that there is a large particle-hole asymmetry in the tunneling spectrum that increases with underdoping. The observation of this asymmetry in scanning tunneling microscopy (STM) experiments on cuprates implies that Coulomb correlations are very important even in doped materials.

We introduce the Gutzwiller approximation in Sect. 2.4, which permits us to deal with the projection in a simple, analytically tractable approach. We show (in later sections) that the results of this approximation are in good qualitative agreement with those of a computationally intensive method, variational Monte Carlo (VMC), that imposes projection exactly.

We focus on the d-wave superconducting state and its properties in Sects. 2.5–2.9, turning briefly to the competition between superconductivity and antiferromagnetism in Sect. 2.10. We begin in Sect. 2.5 with a variational analysis of the superconducting (SC) ground state. We show that the SC order parameter naturally exhibits a dome as a function of doping, qualitatively similar to T_c, while the pairing amplitude follows the antinodal energy gap measured by angle-resolved photoemission spectroscopy (ARPES) of cuprates. Section 2.6 describes the momentum distribution and the nodal quasiparticle spectral weight, which vanishes as one approaches the Mott insulator.

In Sect. 2.7, we describe the electronic excitations above the d-wave SC state, particularly the nodal quasiparticle dispersion. We also address properties of the the single-particle spectral function measured in ARPES, including the notion of an "underlying Fermi surface" in the SC state. In Sect. 2.8, we look at the response to electromagnetic fields. Using optical conductivity sum-rules, we bound the superfluid stiffness and we describe how nodal quasiparticles affect its T-dependence.

The problem of why d-wave SC in the cuprates is so insensitive to disorder is addressed in Sect. 2.9, where we develop an inhomogeneous generalization of the Gutzwiller approximation. Our results give new insights into how the low-energy excitations in strongly correlated SCs are protected against disorder.

Section 2.10 contains a brief discussion of the competition between superconductivity and antiferromagnetism at very low doping and gives insights into the role of longer range in-plane hopping in aiding SC, and the differences between electron and hole doping a Mott insulator. Finally, we conclude with open problems related to the non-Fermi liquid normal states in Sect. 2.11.

2.2 Hubbard Model and Projected Wavefunctions

The one band Hubbard model is described by the Hamiltonian

$$H = -\sum_{\langle ij \rangle} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}, \tag{2.1}$$

where $c^{\dagger}_{i\sigma}$ creates an electron of spin σ at site i and $n_{i\sigma} = c^{\dagger}_{i\sigma} c_{i\sigma}$ is the number operator. (We assume summation over repeated spin indices.) The first term is the kinetic energy of electrons hopping between sites i and j of a square lattice with a hopping matrix element t_{ij} that is short ranged, usually connecting nearest and

next nearest neighbors. There is a local repulsion U which penalizes configurations with double occupancy, i.e., both $\uparrow$ and $\downarrow$ electrons on the same site. In the large U limit ($U/|t_{ij}| \gg 1$), this model is expected to capture the essential physics of strong electronic correlations relevant to the high T_c materials.

In the infinite U limit, the low-energy states of the system with density $\langle n \rangle \leq 1$ have no double occupancy. These states can be written as

$$\mathcal{P}|\psi_0\rangle = \prod_i (1 - n_{i\uparrow} n_{i\downarrow})|\psi_0\rangle, \tag{2.2}$$

where $|\psi_0\rangle$ is any state. $\mathcal{P}$ is the Gutzwiller projection operator [14],[1] which projects out configurations with double occupancy in $|\psi_0\rangle$. The Hilbert space at each site is thus reduced from four states to three: $|\uparrow\rangle$, $|\downarrow\rangle$, and $|0\rangle$ (or a vacancy), with $\mathcal{P}$ eliminating the doubly occupied state $|\uparrow\downarrow\rangle$. The fraction of sites that have a vacancy is equal to the *hole doping* $x = 1 - \langle n \rangle$.

The effects of having a large but finite U are then taken into account through a canonical transformation on these projected wavefunctions, so that the final wavefunction can be written as

$$|\psi\rangle = \mathrm{e}^{-\mathrm{i}\mathcal{S}} \mathcal{P}|\psi_0\rangle \tag{2.3}$$

where the operator $\mathcal{S}$ is determined as described below [15, 16]. The expectation value of any operator A in the states of the form (2.3) can be written as

$$\langle\psi|A|\psi\rangle = \langle\psi_0|\mathcal{P}\tilde{A}\mathcal{P}|\psi_0\rangle, \qquad \tilde{A} = \mathrm{e}^{\mathrm{i}\mathcal{S}} A \mathrm{e}^{-\mathrm{i}\mathcal{S}}. \tag{2.4}$$

Operationally, expectation values of various operators in a state $|\psi\rangle$ are obtained by evaluating the expectation of the transformed operators $\tilde{A}$ in projected states $\mathcal{P}|\psi_0\rangle$.

We determine $\mathcal{S}$ by demanding that the transformed Hamiltonian does not have any terms connecting projected states to states with nonzero double occupancy. This can be achieved order by order in t/U, leading to a series expansion of $\mathcal{S}$ in t/U. There are three types of terms in the kinetic energy. The first, $K_0 = \sum_{\langle ij \rangle} t_{ij}(1 - n_{i\bar{\sigma}}) c^\dagger_{i\sigma} c_{j\sigma} (1 - n_{j\bar{\sigma}})$, does not change the double occupancy of the state it acts on, the second $K_1 = \sum_{\langle ij \rangle} t_{ij} n_{i\bar{\sigma}} c^\dagger_{i\sigma} c_{j\sigma} (1 - n_{j\bar{\sigma}})$, increases the double occupancy by one, and the third, $K_{-1} = \sum_{\langle ij \rangle} t_{ij}(1 - n_{i\bar{\sigma}}) c^\dagger_{i\sigma} c_{j\sigma} n_{j\bar{\sigma}}$, decreases the double occupancy by one. To second order in t/U, we find $\mathrm{i}\mathcal{S} = (K_1 - K_{-1})/U + ([K_1, K_0] + [K_{-1}, K_0])/U^2$.

Let us now look at the transformed Hamiltonian, which acts on projected states. Dropping the "tilde" on H for simplicity, to leading order in t/U, we find that

[1] Gutzwiller's original proposal is $\mathcal{P}_\alpha = \prod_i (1 - \alpha n_{i\uparrow} n_{i\downarrow})$, where $\alpha < 1$ is treated as a variational parameter. This is often called the partial Gutzwiller projector. The one we use is the full projector given by $\alpha = 1$.

$$H = \sum_{\langle ij \rangle} t_{ij} h_{i\bar{\sigma}} c^{\dagger}_{i\sigma} c_{j\sigma} h_{j\bar{\sigma}} + \frac{1}{2} \sum_{\langle ij \rangle} J_{ij} \left(\boldsymbol{S}_i \cdot \boldsymbol{S}_j - \frac{1}{4} n_i n_j \right)$$
$$- \frac{1}{U} \sum_{\langle ij \rangle \langle jl \rangle i \neq l} t_{ij} t_{jl} h_{i\bar{\sigma}} c^{\dagger}_{i\sigma} c_{j\sigma} n_{j\bar{\sigma}} c^{\dagger}_{j\sigma'} c_{l\sigma'} h_{l\bar{\sigma}'}. \tag{2.5}$$

Here, $h_{i\bar{\sigma}} = (1 - n_{i\bar{\sigma}})$, the spin $\boldsymbol{S}_i = c^{\dagger}_{i\alpha} \boldsymbol{\sigma}_{\alpha\beta} c_{i\beta}$, and $J_{ij} = 4t^2_{ij}/U$ is the *antiferromagnetic superexchange* coupling.

The transformed Hamiltonian consists of the three sets of terms depicted in Fig. 2.2 : (1) Kinetic energy terms of $\mathcal{O}(t)$, where electrons of either spin hop to a neighboring site as long as a hole is present there. This process can alternatively be described as the hopping of the holes with a backflow of spins. (2) The spin–spin interaction, with its associated super-exchange scale J. This is a process where the electrons hop to a neighboring site, creating a double occupancy and then hop back to the site they started from, relaxing the double occupancy. Note that this process does not lead to motion of electrons. (3) Three-site terms, in which the electrons create a double occupancy by hopping on to a neighboring site and then relax the double occupancy by hopping on to a second neighboring site which has a hole. This is alternatively a second order hole hopping process. Now, the two hops in this process can be made by electrons of the same spin ($\sigma = \sigma'$ in the third term), in which case the spin configuration of the mediating site remains unchanged, or by electrons of different spin ($\sigma' = \bar{\sigma}$ in the third term), in which case the spin of the mediating site is flipped. Note that the transformed Hamiltonian connects projected states to projected states, even though it may involve double-occupancy in an intermediate state.

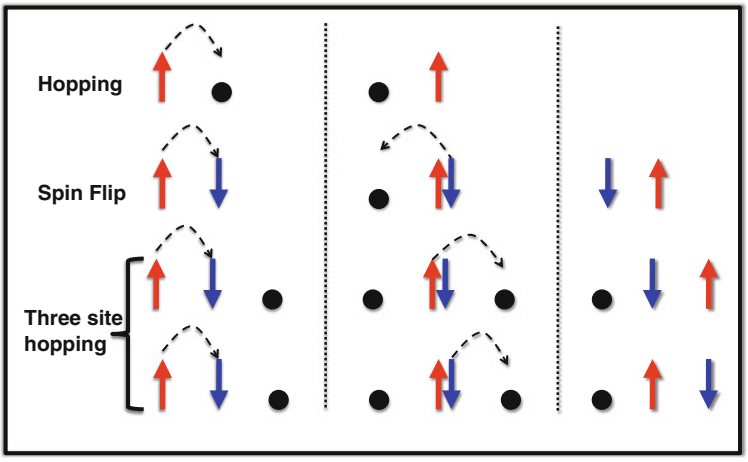

Fig. 2.2 The different processes represented by the terms in the canonically transformed Hamiltonian of (2.5)

It is useful to clarify the relation between the large U Hubbard and the tJ model, widely used in the literature [6]. There are two differences. First, the tJ Hamiltonian is obtained from the canonically transformed Hubbard Hamiltonian by simply *neglecting* the three site-hopping terms. The tJ model thus describes hopping of projected electrons together with the antiferromagnetic superexchange interactions. The second difference is that the canonical transformation is used only to determine the tJ Hamiltonian, but *not* applied to any other operators when their expectation values are obtained in the projected states. Thus for operators other than the Hamiltonian, the tJ model retains terms which are only zeroth order in t/U.

To summarize: The effects of the strong electron–electron interactions have been taken into account in a nonperturbative way by the use of the Gutzwiller projector $\mathcal{P}$, which imposes the *no double occupancy* constraint. The effects of a noninfinite U are then included perturbatively in powers of t/U via the canonical transformation e^{-iS}. (A different way to incorporate finite U physics is to use a "partial" projection operator that penalizes, but does not forbid, double-occupancy (See footnote 3). We will not discuss partial projection because it cannot describe the Mott insulator at half-filling).

2.3 Particle-hole Asymmetry in Doped Mott Insulators

At half filling (one electron per lattice site) there are no holes, and thus no conduction in the large U Hubbard model of (2.5) at $T = 0$. This is the Mott insulating state whose low-energy physics is described by the antiferromagnetic $S = 1/2$ Heisenberg Hamiltonian on a square lattice. It is by now very well established that the undoped parent compounds of the high T_c superconductors are antiferromagnetically ordered Mott insulators (actually charge transfer insulators, but that distinction is not important for our present purposes).

The question then arises: Is the strong Hubbard U that leads to Mott insulating behavior at half filling also relevant for the doped materials? Or, could the screening in the doped materials effectively lead to a weakly correlated system? We show here that the answer is unequivocal: the strong Hubbard U leads to a characteristic doping-dependent particle-hole asymmetry [17–22] in the tunneling spectra of the doped Mott insulators. Theoretically, this is a rigorous consequence of Gutzwiller projection, requiring no assumptions about the nature of the ground state. Experimentally, this characteristic p–h asymmetry has been observed [23–26] by STM experiments on several hole-doped cuprates including Bi2212 and oxychlorides. This establishes unambiguously that correlations *are* important in the doped materials.

We discuss the asymmetry between the electron and hole density of states (DOS) in the lightly doped Mott insulators in a very general setting that does not depend on the details of the Hamiltonian, e.g., presence of disorder, presence of longer range interaction or longer range hopping etc., as long as the local Hubbard repulsion U is the largest energy scale in the model. Our results do not make any assumptions

about the broken symmetry (if any) in the projected state and are valid even in the presence of quenched disorder as we do not assume translational invariance.

We focus, for the sake of simplicity, on the case of $T = 0$ and small but finite hole doping $x = (1 - n) \ll 1$. The STM spectrum at a location $\mathbf{r}$ and a bias V is proportional to the local density of states (LDOS) $N(\mathbf{r}, \omega = eV)$. (The question of the STM tunneling matrix element, which can itself be a function of $\mathbf{r}$, is discussed below.) The LDOS is given by $N(\mathbf{r}, \omega) = A(\mathbf{r}, \mathbf{r}, \omega)$, where the spectral function $A(\mathbf{r}, \mathbf{r}', \omega) = -(1/\pi)\mathrm{Im}G(\mathbf{r}, \mathbf{r}', \omega + i\eta)$ is related to the imaginary part of the retarded (real space) Green's function of the electrons. It has the spectral decomposition

$$A(\mathbf{r}, \mathbf{r}', \omega) = \sum_{m\sigma} \langle 0|c^\dagger_{\mathbf{r}'\sigma}|m\rangle\langle m|c_{\mathbf{r}\sigma}|0\rangle\delta(\omega + E_m - E_0) + \langle 0|c_{\mathbf{r}\sigma}|m\rangle\langle m|c^\dagger_{\mathbf{r}'\sigma}|0\rangle\delta(\omega - E_m + E_0), \tag{2.6}$$

where $|0\rangle$ is the ground state with energy E_0 and $|m\rangle$ are excited states with energy E_m and an implicit sum over σ is assumed. To look at the asymmetry we want to derive sum rules for the frequency integrals of $A(\mathbf{r}, \mathbf{r}', \omega)$. The first sum rule is that $\int_{-\infty}^{\infty} \mathrm{d}\omega A(\mathbf{r}, \mathbf{r}', \omega) = 2\delta_{\mathbf{r},\mathbf{r}'}$, which is just a statement of probability conservation and the factor of 2 comes from the spin sum. We can also obtain

$$\int_{-\infty}^{0} \mathrm{d}\omega N(\mathbf{r}, \omega) = n(\mathbf{r}) = 1 - x(\mathbf{r}), \tag{2.7}$$

where $n(\mathbf{r})$ is the local electron density and $x(\mathbf{r})$ is the corresponding local hole doping. (At finite temperatures, the integration would be cut off by a Fermi function in the integrand. In the $T = 0$ case discussed here, we get a hard cutoff at the chemical potential $\omega = 0$). This sum rule just implies that the probability of *extracting* an electron from the system at point $\mathbf{r}$ is proportional to the local density. The above two sum rules also lead to an obvious third one, $\int_0^{\infty} \mathrm{d}\omega N(\mathbf{r}, \omega) = 1 + x(\mathbf{r})$. However, a nontrivial and more interesting sum rule is obtained if the positive frequency integral is cutoff at a scale Ω, where $t, J, \ldots \ll \Omega \ll U$. Physically, this means that we are restricting attention to the "lower Hubbard band" in which there is no double occupancy. Thus, the states $|m\rangle$ that contribute to the sum rule all lie in the projected low-energy subspace, and are of the form $|m\rangle = \mathrm{e}^{-\mathrm{i}\mathcal{S}}\mathcal{P}|\phi_m\rangle$ with $|\phi_m\rangle$ forming a complete basis in the full Hilbert space; i.e., $\sum_m |\phi_m\rangle\langle\phi_m| = 1$. Using the form of the low-energy states and the spectral decomposition of the DOS, we find that

$$\int_0^{\Omega} \mathrm{d}\omega N(\mathbf{r}, \omega) = \sum_m \langle 0|\mathcal{P}\tilde{c}_{\mathbf{r}\sigma}\mathcal{P}|\phi_m\rangle\langle\phi_m|\mathcal{P}\tilde{c}^\dagger_{\mathbf{r}\sigma}\mathcal{P}|0\rangle = \langle 0|\mathcal{P}\tilde{c}_{\mathbf{r}\sigma}\mathcal{P}\tilde{c}^\dagger_{\mathbf{r}\sigma}\mathcal{P}|0\rangle, \tag{2.8}$$

where $\tilde{c}^\dagger$ is the canonically transformed form of the creation operator and we have used the completeness of $|\phi_m\rangle$ to write the final form. We have

$$\mathcal{P}\tilde{c}^{\dagger}_{\mathbf{r}\sigma}\mathcal{P} = c^{\dagger}_{\mathbf{r}\sigma}h_{\mathbf{r}\bar{\sigma}} + (1/U)\sum_{\langle \mathbf{r}\mathbf{r}'\rangle} t_{\mathbf{r}\mathbf{r}'}h_{\mathbf{r}'\bar{\sigma}'}c^{\dagger}_{\mathbf{r}'\sigma'}c_{\mathbf{r}\sigma'}n_{\mathbf{r}\bar{\sigma}}c^{\dagger}_{\mathbf{r}\sigma}, \tag{2.9}$$

where the first term in (2.9) simply ensures that the creation of the electron does not create a double occupancy, while the second term represents a process where the double occupancy formed by the creation operator is relaxed through a hopping process. Using the above form for the creation operator, we finally arrive at the sum rule for low-energy electron injection:

$$\int_0^{\Omega} \mathrm{d}\omega N(\mathbf{r},\omega) = 2x(\mathbf{r}) + 2\frac{\langle K(\mathbf{r})\rangle}{U}. \tag{2.10}$$

Here, $t, J, \ldots \ll \Omega \ll U$ and $\langle K(\mathbf{r})\rangle$ is the local kinetic energy of the electrons at $\mathbf{r}$. The first term in (2.10) simply says that one can inject an electron into any of the x empty sites, with two for spin degeneracy. The second term gives an order xt/U correction since the injected electron can create a temporary double occupancy and then hop off to a neighboring empty site. Note that while the probability to add a low-energy electron (2.10) is obtained to order t/U, the corresponding result (2.7) to extract an electron is exact to all orders in t/U.

Thus, we conclude that the integrated weight for extracting an electron (2.7) is much larger that the integrated low-energy weight to add an electron (2.10) in a doped Mott insulator, and this asymmetry increases with underdoping.

The spectral weight asymmetry in doped Mott insulators can be understood by the following argument [18–20], which also shows why these systems are completely different from doped band insulators; see Fig. 2.3. Upon doping a band insulator with a density x of holes, the chemical potential just moves into the

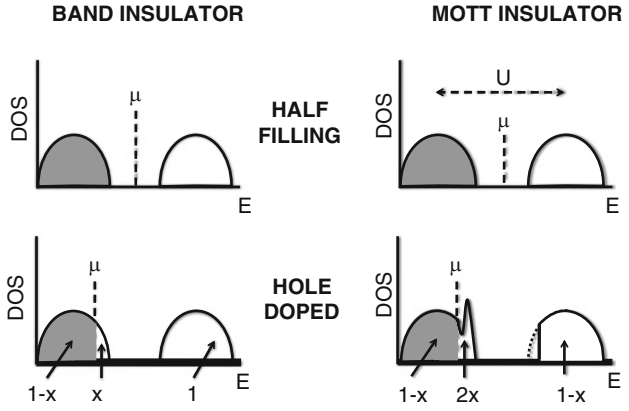

Fig. 2.3 Schematic figure showing the difference between doped band insulators and doped Mott insulators, based on [18–20]. Upon doping a band insulator, the chemical potential moves with the bands remaining "rigid." However, there is a transfer of spectral weight across the Mott–Hubbard gap upon doping a Mott insulator, with the creation of new low-energy states

valence band, with the bands remaining "rigid." The integrated spectral weight for electron extraction is then $(1 - x)$, whereas that for addition at low energies is x. In marked contrast, the upper and lower Hubbard bands are not "rigid." While they each have spectral weight of 1 at half-filling, upon hole doping, there is a transfer of spectral weight x from the upper to the lower Hubbard band. Thus, in the doped system the spectral weight for electron extraction is still $(1 - x)$ ((2.7) with translation invariance), while that for addition at low energies is now $2x$ (large-U limit of (2.10)). This transfer of spectral weight from high energies to the low-energy subspace with doping is a hallmark of doped Mott insulators [18–20, 27].

As already noted, the p–h asymmetry and its doping dependence are in excellent qualitative agreement with all available STM data on cuprates [23–26], where the measured conductance $g(\mathbf{r}; V) = [\mathrm{d}I/\mathrm{d}V](\mathbf{r})$ shows strong asymmetry between positive and negative bias V. The conductance is related to the LDOS via $g(\mathbf{r}, V) = M(\mathbf{r})N(\mathbf{r}; \omega = eV)$ where $M(\mathbf{r})$ is a spatially varying tunneling matrix element. It was proposed in [21, 22] that local ratios of $V > 0$ and $V < 0$ conductances would lead to a cancellation of the unknown matrix elements. This has certainly proved very useful in the analysis of STM data [28] and the extracted LDOS ratios are very similar from one material to another, even though the matrix elements are very different. However, it has not been possible to quantitatively test the ratio of sum rules (2.10) to (2.7) and extract the local charge density. The difficulty stems from a choice of a suitable (negative) lower cutoff in (2.7) for experimental data.

2.4 Gutzwiller Approximation

Having looked at general consequences of projection, we now turn to the calculation of detailed properties of projected wavefunctions. We should make clear the distinction between *Gutzwiller projection* and *Gutzwiller approximation*. Projection $\mathcal{P}|\psi_0\rangle$ imposes the "no double occupancy" constraint acting on an arbitrary state $|\psi_0\rangle$, where $\mathcal{P}$ was defined in (2.2). The Gutzwiller approximation, to be described below, is a simple approximation scheme to estimate expectation values in projected states.

The key difficulty in analyzing projected wavefunctions is that the projection operator $\mathcal{P}$ of (2.2) is a product of single-site operators in real space, while most of the interesting wavefunctions $|\psi_0\rangle$ (e.g., a BCS superconducting state) that one would like to project are naturally written in momentum space. Thus, it is impossible to analytically calculate even the norm of $\mathcal{P}|\psi_0\rangle$, let alone expectation values in such a state. One can, however, use the VMC method [7, 10, 11] to compute properties of projected wavefunctions without any approximation. Here, one works in real space and uses Monte Carlo methods to compute the multiple integrals involved in computing expectation values. We will mention some results of the VMC approach here but mostly for the sake of comparison.

Our main focus will be on the Gutzwiller approximation (GA) [8, 14, 29], which permits us to evaluate matrix elements of operators in projected states using a simple

approximation scheme. The resulting equations have the structure of usual MFT, but with a renormalization of parameters that impacts the results in important ways. Hence, the Gutzwiller approximation scheme has also been called "renormalized MFT" [8]; for recent reviews, see [12, 13]. As a check on its validity, we note that GA scheme gives results in good qualitative agreement with the exact VMC results. Further, given the simplicity of the GA/RMFT equations one can use them to gain analytical insights into important doping dependent trends, and also generalize them to more complex situations, such as inhomogeneous states [30–36]; see Sect. 2.9. (There have also been attempts to include effects of short range correlations [37, 38].)

The essential idea of Gutzwiller approximation is to express the expectation value of an operator A in a projected state as the expectation value in the corresponding *unprojected* state times a renormalization factor g_A called the Gutzwiller factor for that operator:

$$\frac{\langle\psi_0|\mathcal{P}A\mathcal{P}|\psi_0\rangle}{\langle\psi_0|\mathcal{P}|\psi_0\rangle} \simeq g_A \frac{\langle\psi_0|A|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} \tag{2.11}$$

The approximation provides a prescription for evaluating the Gutzwiller factors for different operators, which depend only on the average electron density in the underlying state. (Here, we focus only on the simplest case of homogeneous systems without spin order.) This seemingly ad-hoc prescription may be rationalized in terms of combinatorial factors for finding different configurations in the projected and unprojected states [39].

The easiest way to understand the Gutzwiller approximation is to give some explicit examples. Here, we will explain in detail the renormalization of the kinetic energy and superexchange. For a detailed derivation of various other terms, the reader is referred to [40].

Consider a homogeneous system with density n. In the unprojected states, a single site can be in one of four configurations: (a) $|\uparrow\rangle$, (b) $|\downarrow\rangle$, each of which occurs with probability $(n/2)(1-n/2)$, (c) doubly occupied with probability $n^2/4$ and (d) empty with probability $(1-n/2)^2$. In the projected states, each site is in one of three configurations: (a′) $|\uparrow\rangle$, (b′) $|\downarrow\rangle$, each with probability $n/2$ and (c′) empty with probability $(1-n)$.

Let us consider the hopping term $c^\dagger_{i\uparrow}c_{j\uparrow}$, and look at a single hopping event where an up electron hops from i to j. In the unprojected states, there should be an $\uparrow$ spin on site i and no $\uparrow$ spin on site j (Pauli exclusion) for this hop to occur. After the event one has an $\uparrow$ spin on site j and no $\uparrow$ spin on site i. The probability of occurrence of the initial configuration is $n_{i\uparrow}(1-n_{j\uparrow})$ and that of the final configuration is $n_{j\uparrow}(1-n_{i\uparrow})$. In projected states, the same hopping event would occur if we have an $\uparrow$ spin on site i and a vacancy on site j to begin with and finally lead to an $\uparrow$ spin on site j and a vacancy on site i. The probability of these configurations are respectively $n_{i\uparrow}(1-n_j)$ and $n_{j\uparrow}(1-n_i)$. The initial and final configurations for projected and unprojected states are shown in Fig. 2.4. The Gutzwiller factor for the hopping term is then given by

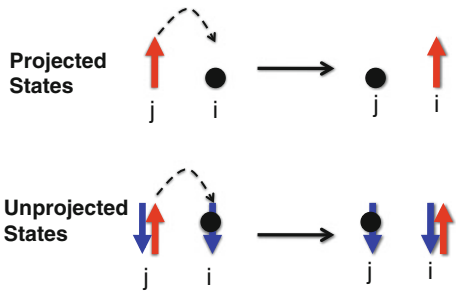

Fig. 2.4 The initial and final configurations for a spin-↑ electron to hop from site j to site i in projected and in unprojected states. Initially, site i must be empty for projected states, but can have either a hole or a ↓ spin in unprojected states

$$g_t = \left[\frac{n_{i\uparrow}(1-n_j)n_{j\uparrow}(1-n_i)}{n_{i\uparrow}(1-n_{j\uparrow})n_{j\uparrow}(1-n_{i\uparrow})}\right]^{\frac{1}{2}}. \tag{2.12}$$

We thus have a clear prescription for obtaining the Gutzwiller factors. For a given operator, first neglect the configurations on sites that are not connected by the operator, i.e., ignore the effects of longer-range electronic correlations that would make this probability dependent on the full wavefunction. Then multiply the probabilities of the configurations before and after the event has occurred and take a square root, since we need an amplitude. The ratio in the projected and unprojected states gives the Gutzwiller factor for the corresponding operator. The final answer depends only on the density. The Gutzwiller factor for hopping is given by

$$g_t = \frac{2(1-n)}{2-n} = \frac{2x}{1+x}, \tag{2.13}$$

where $x = 1 - n$ is the hole doping of the system away from half-filling. Thus, the kinetic energy in projected states is suppressed by a factor of g_t, and vanishes $\sim x$ as one approaches the Mott insulator at half filling.

Next, consider the spin-flip terms $S_i^+ S_j^-$. In both projected and unprojected states, spin flips require an initial configuration of a single ↑ spin (no double occupancy) on site j and a single ↓ spin on site i. The final configuration has the same requirements with spins interchanged. In projected states, the probability of the initial and final configurations are $n_{i\downarrow}n_{j\uparrow}$ and $n_{i\uparrow}n_{j\downarrow}$, respectively. In the unprojected states, one has to explicitly take care of the fact that configurations should avoid double occupancy for the spin flip to occur. The probability of initial and final configurations are then given by $n_{i\downarrow}n_{j\uparrow}(1-n_{i\uparrow})(1-n_{j\downarrow})$ and $n_{i\uparrow}n_{j\downarrow}(1-n_{i\downarrow})(1-n_{j\uparrow})$ respectively, leading to the Gutzwiller factor for the spin–spin interactions

$$g_s = \frac{4}{(2-n)^2} = \frac{4}{(1+x)^2}. \tag{2.14}$$

The x-dependence of g_s reflects the increasing importance of spin fluctuations as one goes toward the Mott insulator within the Gutzwiller approximation.

Similar considerations can be used to derive the Gutzwiller factors for other operators and we simply state the results without derivation (see [40]). The renormalization of the remaining terms in the transformed Hamiltonian (2.5) are as follows. The $n_i n_j$ term is not renormalized, i.e., it has a Gutzwiller factor of 1. For the three-site hopping term with spin flip the Gutzwiller factor is

$$g_{st} = 4x/(1+x)^2, \tag{2.15}$$

while that for three-site term without spin flip is g_t. In certain cases, we will need additional Gutzwiller factors which will be described below as they are needed.

2.5 Superconducting Ground State

In this section, we will focus on the strongly correlated d-wave superconducting phase that arises upon doping a Mott insulator. We work with the 2D one-band Hubbard model (2.1) whose parameters are known from a combination of theory and experiment. The near-neighbor hopping ($j = i \pm \hat{x}(\hat{y})$) is $t_{ij} = t = 300\,\text{meV}$ and the next-near-neighbor hopping ($j = i \pm \hat{x} + \pm\hat{y}$) is $t_{ij} = t' = -t/4$, consistent with electronic structure calculations and ARPES measurements [41, 42]. We will see below that the t' term determines the shape of the Fermi surface (end of Sect. 2.7) and can be important for superconductivity (Sect. 2.10). The bare dispersion is then given by

$$\epsilon_{\mathbf{k}} = -t\gamma_{\mathbf{k}} - t'\lambda_{\mathbf{k}}, \quad \gamma_{\mathbf{k}} \equiv 2(\cos k_x + \cos k_y), \quad \lambda_{\mathbf{k}} \equiv 4\cos k_x \cos k_y. \tag{2.16}$$

We choose the on-site Hubbard $U = 12t$, which gives the superexchange scale $J = 4t^2/U = 100\,\text{meV}$ consistent with neutron scattering experiments [43]. Although J ($\ll t \ll U$) is the smallest energy scale in the large-U Hubbard model, in reality $J \simeq 100$–$150\,\text{meV}$ found in the high T_c cuprates is amongst the largest superexchange interaction in all known transition metal oxides!

For the superconducting state, we consider a variational wavefunction of the form

$$|\psi\rangle = \mathrm{e}^{-\mathrm{i}\mathcal{S}}\mathcal{P}|\text{dBCS}\rangle \;\text{ with }\; |\text{dBCS}\rangle = \prod_{\mathbf{k}}(u_{\mathbf{k}} + v_{\mathbf{k}} c^\dagger_{\mathbf{k}\uparrow} c^\dagger_{-\mathbf{k}\downarrow})|0\rangle, \tag{2.17}$$

where $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are variational parameters with $u^2_{\mathbf{k}} + v^2_{\mathbf{k}} = 1$. The d-wave symmetry requires that the pairing amplitude $u_{\mathbf{k}} v_{\mathbf{k}}$ change sign under $\pi/2$ rotation. We can also study states with other broken symmetries such as antiferromagnetism; see Sect. 2.10.

The variational parameters are determined by minimizing the expectation value of $H - \mu N$ in $|\psi\rangle$, where μ is the chemical potential and N the total number operator. We calculate ground state energy by evaluating the expectation value of the canonically transformed Hamiltonian (2.5) in projected states using Gutzwiller

approximation. The expectation value of N is the same as in the unprojected BCS wavefunction. (This will be seen explicitly when we calculate the momentum distribution.) One then obtains equations that look like BCS theory albeit with various Gutzwiller factors renormalizing the different terms in the Hamiltonian [8].

Within this renormalized MFT, one obtains $v_{\mathbf{k}}^2 = 1 - u_{\mathbf{k}}^2 = (1/2)(1 - \xi_{\mathbf{k}}/E_{\mathbf{k}})$ and $u_{\mathbf{k}} v_{\mathbf{k}} = \Delta_{\mathbf{k}}/(2E_{\mathbf{k}})$, where $\xi_{\mathbf{k}}$ is the renormalized dispersion, $\Delta_{\mathbf{k}} = \Delta(\cos k_x - \cos k_y)$ is the d-wave pairing function and $E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}$. The renormalized dispersion $\xi_{\mathbf{k}}$ has the form

$$\xi_{\mathbf{k}} = g_t \epsilon_{\mathbf{k}} - \chi_{\mathbf{k}} - \overline{\mu}, \qquad \chi_{\mathbf{k}} = \chi_0 + \chi_1 \gamma_{\mathbf{k}} + \chi_2 \lambda_{\mathbf{k}} + \chi_3 \gamma_{\mathbf{k}}^2 + \chi_4 \lambda_{\mathbf{k}}^2 + \chi_5 \gamma_{\mathbf{k}} \lambda_{\mathbf{k}}. \quad (2.18)$$

Here, $\chi_{\mathbf{k}}$ comes from a mean field decomposition of the interaction terms in the Hamiltonian (2.5); χ_0 is a Hartree shift and χ_n's (for $n \geq 1$) determine the Fock shift. The chemical potential has an additional renormalization coming from the fact that the expectation value of $H - \mu N$ depends on the density through the Gutzwiller factors. Following [8], we get $\overline{\mu} = \mu + (1/N_s)\partial\langle H\rangle/\partial x$, where the derivative is taken only with respect to the Gutzwiller factors and N_s is the number of sites.

The variational minimization leads to self-consistent equations for the pairing amplitude Δ, the chemical potential μ and different Fock shift amplitudes. The gap equation for Δ and the number equation are given by

$$1 = \frac{g_0(x) J}{2} \sum_{\mathbf{k}} \frac{\phi_{\mathbf{k}}^2}{E_{\mathbf{k}}} \quad \text{and} \quad n = \sum_{\mathbf{k}} (1 - \xi_{\mathbf{k}}/E_{\mathbf{k}}) \quad (2.19)$$

with $\phi_{\mathbf{k}} = \cos k_x - \cos k_y$ and $g_0(x) = 3g_s/4 + 1/4 - g_t/2 - g_{st}/2$, where we use the Gutzwiller factors of (2.13), (2.14), and (2.15). These equations look like BCS theory, but with renormalized coefficients that have profound consequences on the results.

The self-consistency equations for the Fock shifts χ_n are omitted for simplicity and can be found in [40]. We should note that the higher harmonics entering the Fock shift (see (2.16,2.18)) are necessarily demanded by the form of the self-consistency equation. In contrast, the gap function has the simple $(\cos k_x - \cos k_y)$ form. Higher harmonics in $\Delta_{\mathbf{k}}$ can only arise from higher order terms in t/U, like four-spin terms, that we have not retained.

2.5.1 Energy Gap

The solution of these equations leads to a $\Delta(x)$, which is a monotonically decreasing function of x and goes to zero around $x = 0.35$. In Fig. 2.5a, we compare the Gutzwiller approximation $\Delta(x)$ with that obtained from a VMC calculation that implements projection exactly. Although the agreement is not quantitative, both

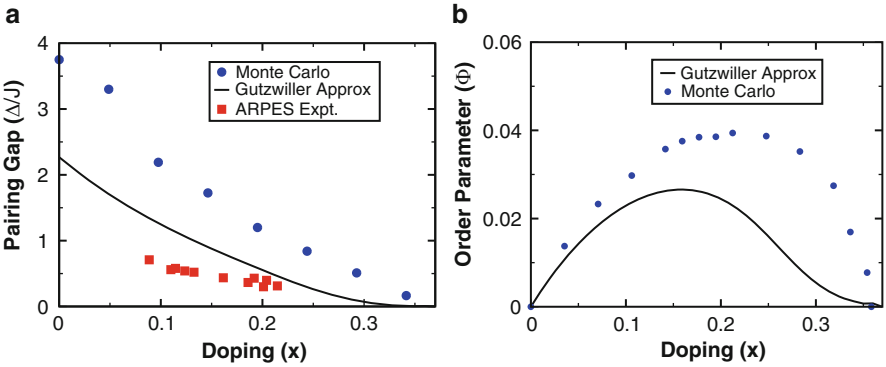

Fig. 2.5 *Left panel* (**a**): The pairing amplitude Δ in units of J as a function of doping as obtained from Gutzwiller approximation and variational Monte Carlo [11]. Also plotted is the antinodal quasiparticle gap obtained from ARPES experiments [44]. *Right panel* (**b**): The superconducting order parameter Φ as a function of doping from Gutzwiller approximation and variational Monte Carlo [10]

have the same qualitative behavior. We also emphasize that it is the superexchange J that sets the scale for superconductivity in the gap equation.

The physical meaning of $\Delta(x)$ in a projected superconductor is not immediately clear. We will see that although Δ is related to the energy gap in the spectrum, it is *not* the ODLRO order parameter associated with superconductivity (Fig. 2.5b). In the next section, we will treat quasiparticle excitations about the correlated superconducting state and show that Δ is the spectral gap of the antinodal quasiparticles. We also compare in Fig. 2.5(a) the experimentally observed antinodal gap from ARPES measurements on cuprates [44], with our calculated $\Delta(x)$. The overall doping trends are very similar, but quantitatively, the Gutzwiller approximation overestimates the experimental energy gap by about a factor of two. This is not unusual for simple variational excited state calculations (e.g., Feynman's results for the roton minimum in Helium-4).

2.5.2 *Order Parameter*

We now show that the results for the doping dependence of the superconducting order parameter Φ are qualitatively different from simple BCS theory (where Φ and Δ are essentially the same). In a projected superconductor, $\mathcal{P}$ suppresses number fluctuations as $x \to 0$ leading to a strong enhancement in the (conjugate) phase fluctuations. Thus, local quantum phase fluctuations lead to the destruction of superconductivity and $\Phi(x)$ vanishes as $x \to 0$ despite the presence of pairing $\Delta(x) \neq 0$.

To see this physical picture emerging from our calculations, we examine off-diagonal long range order (ODLRO) in the two-particle reduced density matrix. We define the near-neighbor singlet creation operator on a bond $B^{\dagger}_{\mathbf{r},\alpha} = (1/\sqrt{2})(c^{\dagger}_{\mathbf{r}\uparrow}c^{\dagger}_{\mathbf{r}+\alpha,\downarrow} - c^{\dagger}_{\mathbf{r}\downarrow}c^{\dagger}_{\mathbf{r}+\alpha,\uparrow})$, where $\alpha = \pm\hat{x}$ or $\pm(\hat{y})$. The two-particle density matrix is then given by $F_{\alpha\beta}(\mathbf{r}-\mathbf{r}') = \langle B^{\dagger}_{\mathbf{r}\alpha}B_{\mathbf{r}'\beta}\rangle$. From its long distance limit $\mathrm{Lim}_{|\mathbf{r}-\mathbf{r}'|\to\infty}F_{\alpha\beta}(\mathbf{r}-\mathbf{r}') = \pm\Phi^2$, we obtain the superconducting order parameter Φ. The $\pm$ sign in the last equation reflects d-wave pairing, and depends on whether the bonds α and β are parallel or perpendicular (sign change under $\pi/2$ rotation).

To zeroth order in t/U, one can set $\mathrm{e}^{-\mathrm{i}\mathcal{S}} \simeq 1$ and simply evaluate $F_{\alpha\beta}$ in the projected states. The easiest way to obtain the Gutzwiller factor for this correlator is to note that $F_{\alpha\beta}(\mathbf{r}-\mathbf{r}')$ takes two electrons from neighboring sites near $\mathbf{r}'$ and places them on neighboring sites around $\mathbf{r}$; i.e., this is equivalent to hopping of two electrons in the projected states. The simple Gutzwiller approximation neglects intersite correlation in evaluating Gutzwiller factors. Thus, the Gutzwiller factor for two independent hops is g_t^2. Straightforward algebra then leads to

$$\Phi(x) = \frac{g_t}{2}\sum_{\mathbf{k}}\phi_{\mathbf{k}}u_{\mathbf{k}}v_{\mathbf{k}} = g_t(x)\frac{\Delta(x)}{4Jg_0(x)}, \tag{2.20}$$

where $\phi_{\mathbf{k}}$ and $g_0(x)$ were defined just below (2.19).

We see that the superconducting order parameter Φ vanishes: (a) when Δ vanishes (for large hole doping $x > 0.35$) so that there is no pairing in the system, and (b) as $x \to 0$ due to the $g_t \sim x$ factor. In the second case, even though there is electron pairing ($\Delta(x) \neq 0$), there is no mobility as one approaches the Mott insulator and, as explained above, local quantum phase fluctuations destroy superconductivity. We note that it is the large coulomb U that suppresses superconductivity on the underdoped side and *not* some competing order parameter, which can only be associated with a much lower energy scale J.

Thus, $\Phi(x)$ is a nonmonotonic function of hole doping x as shown in Fig. 2.5b. It mimics the superconducting dome [45] in the phase diagram of high T_c superconductors, in contrast to Δ, which is a monotonically decreasing function of x. The doping dependence of the order parameter as obtained from Gutzwiller approximation and Monte Carlo studies is plotted in Fig. 2.5b. The top of the dome can be identified with optimal doping around $x = 0.15$.

We note that the projected superconductor naturally leads to a superconducting dome. Projection leads to a very non-BCS characteristic of the system, i.e., the superconducting order parameter does not follow the energy gap due to pairing as a function of doping. On the overdoped side, Φ follows Δ as in usual BCS theory, while on the underdoped side the two quantities show diametrically opposite behavior as one approaches the Mott insulator.

2.6 Momentum Distribution

We next describe the calculation of the momentum distribution $n_{\mathbf{k}}$. This gives very useful insights into the low-energy excitations of the system (nodal quasiparticles), and also serves to illustrate important technical points in Gutzwiller approximation calculations. Before plunging into technical details, let us explain how a ground state correlation function like $n_{\mathbf{k}}$ can give information about low-energy excitations. Recall that $n_{\mathbf{k}}$ is the energy integral of the spectral function $A(\mathbf{k},\omega)$ over the occupied states. Thus, gapless quasiparticles at a momentum $\mathbf{k}_{\mathrm{F}}$ lead to sharp discontinuities in $n_{\mathbf{k}}$ at $T = 0$. The magnitude of the jump discontinuity is the quasiparticle spectral weight Z of these low-energy excitations. In a d-wave superconductor, the superconducting gap vanishes along the zone diagonals in the Brillouin zone ($k_x = \pm k_y$) and leads to four nodes in the excitation spectrum. We will use the location and magnitude of the singularities in $n_{\mathbf{k}}$ to learn about the doping dependence of the nodal quasiparticle excitations.

Since Gutzwiller projection is defined in real space, we calculate $n_{\mathbf{k}}$ by taking the Fourier transform of the ground state expectation value of the operator $G(\mathbf{r}-\mathbf{r}')_\sigma = c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}'\sigma}$. Its canonically transformed form (see (2.4)) can be written as

$$\tilde{G}_\sigma(\mathbf{r}-\mathbf{r}') = h_{\mathbf{r}\bar{\sigma}} c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}'\sigma} h_{\mathbf{r}'\bar{\sigma}} + \sum_{\langle \mathbf{R}\mathbf{r}\rangle} t_{\mathbf{R}\mathbf{r}} h_{\mathbf{R}\bar{\sigma}'} c^\dagger_{\mathbf{R}\sigma'} c_{\mathbf{r}\sigma'} n_{\mathbf{r}\bar{\sigma}'} c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}'\sigma} h_{\mathbf{r}'\bar{\sigma}'}/U + \sum_{\langle \mathbf{R}\mathbf{r}'\rangle} t_{\mathbf{R}\mathbf{r}'} h_{\mathbf{r}\bar{\sigma}} c^\dagger_{\mathbf{r}\sigma} c_{\mathbf{r}'\sigma} n_{\mathbf{r}'\bar{\sigma}} c^\dagger_{\mathbf{r}'\sigma'} c_{\mathbf{R}\sigma'} h_{\mathbf{R}\bar{\sigma}'}/U. \tag{2.21}$$

$\langle \tilde{G} \rangle$ can be evaluated in projected states using Gutzwiller factors. After Fourier transforming, we finally get

$$n_{\mathbf{k}} = n_0 + n_{1\mathbf{k}} + Z_{\mathbf{k}} v^2_{\mathbf{k}}, \tag{2.22}$$

where n_0 is a constant: $n_0 = (1-x)^2/[2(1+x)] + (g_t + g_{st})(1-x)\sum_{\mathbf{k}'} \epsilon_{\mathbf{k}'} v^2_{\mathbf{k}'}/U$ and $n_{1\mathbf{k}}$ is a smooth function of $\mathbf{k}$ given by $n_{1\mathbf{k}} = -(1-g_t)(1-x)^2\epsilon_{\mathbf{k}}/2U - 4g_0(x)\sum_{\mathbf{k}'\mathbf{k}''} \epsilon_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''} u_{\mathbf{k}'} v_{\mathbf{k}'} u_{\mathbf{k}''} v_{\mathbf{k}''}/U - 4g_1(x)\sum_{\mathbf{k}'\mathbf{k}''} \epsilon_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''} v^2_{\mathbf{k}'} v^2_{\mathbf{k}''}/U$. The term $Z_{\mathbf{k}} v^2_{\mathbf{k}}$ in (2.22) leads to a jump discontinuity in $n_{\mathbf{k}}$ along the zone diagonal. Along this direction the superconducting gap $\Delta_{\mathbf{k}}$ vanishes, $v_{\mathbf{k}} = \Theta(-\xi_{\mathbf{k}})$ and hence the jump discontinuity in $n_{\mathbf{k}}$. The quasiparticle weight of the nodal excitations is then given by

$$Z_{\mathbf{k}} = g_t - 2g_{st}\sum_{\mathbf{k}'} \epsilon_{\mathbf{k}'} v^2_{\mathbf{k}'}/U - g_t(1-x)\epsilon_{\mathbf{k}}/U. \tag{2.23}$$

Although the momentum distribution only gives information about the gapless nodal quasiparticles, an explicit calculation of the coherent spectral function shows that $Z_{\mathbf{k}}$ is the spectral weight of the coherent quasiparticles all over the Brillouin zone. The nodal quasiparticle weight increases with doping and for very low doping vanishes like x.

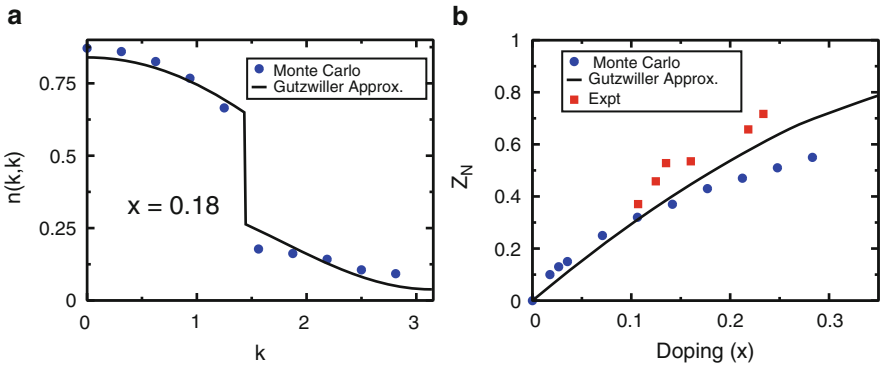

Fig. 2.6 *Left panel* (**a**): Momentum distribution along the zone diagonal for a doping of $x = 0.18$. The *blue circles* are the Monte Carlo results [10] while the thick line is the Gutzwiller approximation result. The jump discontinuity shows the presence of gapless quasiparticles. *Right panel* (**b**): The nodal quasiparticle weight as a function of doping. The *blue circles* are Monte Carlo results [10], the *thick line* is the Gutzwiller approximation results and the *red squares* are experimental ARPES results [46]

Figure 2.6a shows a plot of $n_{\mathbf{k}}$ along the zone diagonal as calculated from VMC and Gutzwiller approximation for a doping of $x = 0.18$, and the two are in very good agreement. Figure 2.6b shows a plot of the nodal quasiparticle weight $Z_N(x)$ as a function of doping. The Monte Carlo and the Gutzwiller approximation are both in very good agreement with the doping dependence of the nodal quasiparticle weight, which is obtained from ARPES experiments [46]. Note that there is no adjustable parameter in this comparison with experiment.

The suppression of the nodal quasiparticle weight with underdoping is a profound consequence of projection and proximity to the Mott insulating state. Projection hinders the coherent propagation of electrons. But, as we shall see in the next section, a surprise awaits us when we compute the doping dependence of the effective mass.

2.7 Electronic Excitations and Spectral Properties

We now turn our attention to single particle excitations above the superconducting ground state. The single particle spectral properties of the high T_c cuprates have been measured by ARPES [41, 42] and STM [47, 48] and have provided a wealth of information which has been crucial for understanding these systems. We have already compared our projected wavefunction results with the experimentally measured magnitudes and doping dependences of p–h asymmetry (Sect. 2.3), the antinodal energy gap (Fig. 2.5) and the nodal quasiparticle weight (Fig. 2.6).

In this section, we consider a variational ansatz for the fermionic excitations of projected superconductors. Among other things, this will allow us to examine the nodal quasiparticle dispersion, justify the identification of $\Delta(x)$ with the antinodal gap, and discuss the coherent contributions to the single particle spectral function $A(\mathbf{k}, \omega)$ and the notion of an "underlying Fermi surface" in the superconducting state.

We make a variational ansatz for the single-particle excitation

$$|\mathbf{k}, \sigma\rangle = \exp(-\mathrm{i}\mathcal{S})\mathcal{P}\gamma^{\dagger}_{\mathbf{k}\sigma}|\mathrm{dBCS}\rangle, \tag{2.24}$$

where $\gamma^{\dagger}$ is the Bogoliubov quasiparticle creation operator given by $\gamma^{\dagger}_{\mathbf{k}\sigma} = u_{\mathbf{k}}c^{\dagger}_{\mathbf{k}\sigma} - \sigma v_{\mathbf{k}}c_{-\mathbf{k}\bar{\sigma}}$. The excitation spectrum is obtained by subtracting the ground state energy from that of the variational quasiparticle states. Within the renormalized MFT [8], the spectrum has the usual BCS form

$$E_{\mathbf{k}} = \sqrt{\xi^2_{\mathbf{k}} + \Delta^2_{\mathbf{k}}}, \tag{2.25}$$

where the renormalized dispersion $\xi_{\mathbf{k}}$ and the gap function $\Delta_{\mathbf{k}}$ were defined in and above (2.18). It thus follows that the pairing amplitude Δ, plotted in Fig. 2.5a, is the antinodal spectral gap.

Linearizing the spectrum near the nodes, the spectrum of nodal quasiparticles is given by

$$E_{\mathbf{k}} \approx \sqrt{v_{\mathrm{F}}k^2_{\perp} + v_{\Delta}k^2_{\parallel}}, \tag{2.26}$$

where $k_{\perp}$ and $k_{\parallel}$ are momenta, perpendicular and parallel to the Fermi surface, measured from the nodal $\mathbf{k}_{\mathrm{F}}$. The nodal Fermi velocity v_{F} and the "gap slope" v_{Δ} determine the low-energy quasiparticle DOS $N(\omega) \sim |\omega|/(v_{\mathrm{F}}v_{\Delta})$ and, hence, various low-temperature properties of the system. Within the renormalized mean-field theory, the nodal k_{F} is weakly dependent on doping and $v_{\Delta}(x)$ follows the pairing amplitude $\Delta(x)$.

2.7.1 Nodal Fermi Velocity

We predicted that v_{F} is essentially independent of doping from VMC calculations [10] of projected wavefunctions. (This required the use of higher moment sum rules of the spectral function [10, 11] not discussed here). Our prediction was confirmed by later ARPES experiments [49] (see Fig. 2.7). Optical measurements [50] found a mass renormalization of 2 consistent with our results. The Gutzwiller approximation v_{F}, also shown in Fig. 2.7, is not in quantitative agreement with the VMC and experimental results, in that it has considerable x-dependence. However, it does capture an important qualitative aspect of the VMC results in that v_{F} goes to a constant as $x \to 0$.

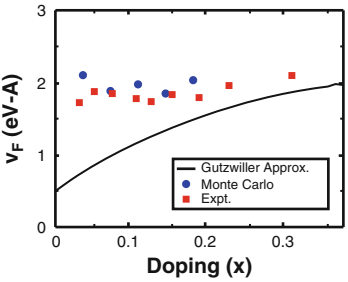

Fig. 2.7 The nodal Fermi velocity from Gutzwiller approximation, variational Monte Carlo [10] and ARPES experiments [49]

We thus find the remarkable result that although the weight $Z \to 0$ as $x \to 0$, the nodal v_F, or effective mass $m^* = k_F/v_F$, remains finite. This is very unusual and different from well-studied examples such as electron–phonon interactions in metals and conventional superconductors, or electron–electron interactions in three-dimensional transition metal oxides or the heavy fermions. In all of these cases, the electron self-energy has a weak k-dependence and thus a vanishing Z goes hand in hand with a divergent m^*.

Along the nodal direction, where the gap vanishes, one can relate the nodal quasiparticle weight Z and the nodal Fermi velocity v_F to the real part of self-energy Σ' using

$$Z^{-1} = 1 - \partial\Sigma'(\mathbf{k},\omega)/\partial\omega \quad \text{and} \quad v_F = Z\left(v_F^0 + \partial\Sigma'(\mathbf{k},\omega)/\partial k\right), \tag{2.27}$$

where all derivatives are evaluated at the node ($k = k_F, \omega = 0$) and v_F^0 is the bare (noninteracting) Fermi velocity. The vanishing of $Z \sim x$ seen from (2.23) implies a singularity in $\partial\Sigma(\mathbf{k},\omega)/\partial\omega \sim 1/x$. The only way in which v_F can remain nonzero in this limit is for $\partial\Sigma(\mathbf{k},\omega)/\partial\mathbf{k}$ to have a compensating $\sim 1/x$ singularity! The vanishing of Z upon approaching a Mott insulator implies loss of coherence due to strong correlations. The finite v_F as $x \to 0$ arises from the fact that the Fock shift contribution χ_1 to the renormalized dispersion $\xi_{\mathbf{k}}$ in (2.18) remains finite (of order J) as $x \to 0$; (all other $\mathbf{k}$-dependent contributions to $\xi_{\mathbf{k}}$ vanish in this limit).

2.7.2 *Spectral Function*

In Sect. 2.3, we discussed sum rules for the single-particle spectral function in a possibly inhomogeneous doped Mott insulator. Now, we turn to a more detailed study of $A(\mathbf{k},\omega)$ for a (homogeneous) projected superconductor. It has the spectral decomposition

$$A_\sigma(\mathbf{k},\omega) = \sum_m |\langle m|c^\dagger_{\mathbf{k}\sigma}|0\rangle|^2\delta(\omega - E_m + E_0) + |\langle m|c_{\mathbf{k}\sigma}|0\rangle|^2\delta(\omega + E_m - E_0). \tag{2.28}$$

In a strongly correlated system, $A(\mathbf{k},\omega)$ can be separated into two parts: a contribution $A^{\rm coh}(\mathbf{k},\omega)$ coming from *coherent* quasiparticles (poles of the Green function) and the remaining *incoherent* part $A^{\rm inc}(\mathbf{k},\omega)$. Only states $|m\rangle = |\mathbf{k},\sigma\rangle$ with excitation energy $E_{\mathbf{k}}$ contribute to the coherent piece. Using $\gamma^{\dagger}_{\mathbf{k}\sigma}|{\rm dBCS}\rangle = (1/u_{\mathbf{k}})c^{\dagger}_{\mathbf{k}\sigma}|{\rm dBCS}\rangle$, we may write the matrix element required for the first term of (2.28) as

$$M_{\mathbf{k}} = \frac{1}{u_{\mathbf{k}}}\sum_{ij} {\rm e}^{{\rm i}\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)}\langle {\rm dBCS}|c_{i\sigma}\mathcal{P}\tilde{c}^{\dagger}_{j\sigma}\mathcal{P}|{\rm dBCS}\rangle. \tag{2.29}$$

Here, $\tilde{c}^{\dagger}$ is the canonically transformed creation operator (see (2.4)). In addition to the creation of a particle on a site having a hole, $\tilde{c}^{\dagger}$ also contains a term where a double occupancy is created and then relaxed by hopping onto neighboring sites.

The Gutzwiller factors for the different terms can be calculated as before. Note that the $\mathcal{O}(1/N_s)$ change in density between the two states is irrelevant in calculating the Gutzwiller factors. For the first term, the initial configuration in a projected state is a hole at site i and the final configuration is a spin σ. In the unprojected states, we need an absence of spin σ on site i in the initial configuration. Following the procedure outlined in Sect. 2.5, the Gutzwiller factor for this term is $\sqrt{g_t}$. Similar considerations show that the Gutzwiller factor for the $1/U$ term where the created spin hops is also $\sqrt{g_t}$, whereas the renormalization for the term in which the opposite spin hops is $\sqrt{g_t g_s}$. Using these renormalization factors, we get

$$M_{\mathbf{k}} = \sqrt{g_t}u_{\mathbf{k}} - \frac{\sqrt{g_t}(1-x)}{U}\epsilon_{\mathbf{k}}u_{\mathbf{k}} - \frac{\sqrt{g_t g_s}}{U}u_{\mathbf{k}}\sum_{\mathbf{k}'}\epsilon_{\mathbf{k}'}v^2_{\mathbf{k}'}. \tag{2.30}$$

After a similar calculation for the second term in (2.28), we find that the coherent part of the spectral function can be written as

$$A^{\rm coh}(\mathbf{k},\omega) = Z_{\mathbf{k}}[u^2_{\mathbf{k}}\delta(\omega - E_{\mathbf{k}}) + v^2_{\mathbf{k}}\delta(\omega + E_{\mathbf{k}})], \tag{2.31}$$

where $Z_{\mathbf{k}}$ is found to be given by the same expression as that shown in (2.23). The fact that the quasiparticle weight calculated from the excited states exactly matches that calculated for nodal quasiparticles from $n_{\mathbf{k}}$ (which is a ground state property) serves as a useful consistency check on our results. To summarize: the coherent part of the spectral function, (2.31) has a BCS-like form, but with a reduced weight $Z_{\mathbf{k}}$. The remainder of the spectral weight $(1 - Z_{\mathbf{k}})$ is in the incoherent part.

2.7.3 Sum Rules

We can break up the spectral function sum rules into coherent and incoherent parts and see how each part contributes to the particle-hole asymmetry discussed in general in Sect. 2.3. For the occupied part of the spectral function, we have

$\int_{-\infty}^{0} d\omega A(\mathbf{k},\omega) = n_{\mathbf{k}}$. Separating it into coherent and incoherent parts, we get

$$\int_{-\infty}^{0} A^{\rm coh}(\mathbf{k},\omega) = Z_{\mathbf{k}} v_{\mathbf{k}}^2 \quad \text{and} \quad \int_{-\infty}^{0} A^{\rm inc}(\mathbf{k},\omega) = n_0 + n_{1\mathbf{k}}, \tag{2.32}$$

where n_0 and $n_{1\mathbf{k}}$ were defined below (2.22). Note that the coherent contribution vanishes as $x \to 0$ and the sum rule is completely dominated by the incoherent response.

On the unoccupied side, the interesting low-energy sum rule is obtained by putting an upper cutoff Ω, where $U \gg \Omega \gg t, J$. This ensures that only projected states in the "lower Hubbard band" contribute to the sum rule. Following Sect. 2.3 and using Gutzwiller approximation, we get $\int_0^{\Omega} A^{\rm coh}(\mathbf{k},\omega) = Z_{\mathbf{k}} u_{\mathbf{k}}^2$ and $\int_0^{\Omega} A^{\rm inc}(\mathbf{k},\omega) = 2g_{st} \sum_{\mathbf{k}'\mathbf{k}''} \epsilon_{\mathbf{k}+\mathbf{k}'-\mathbf{k}''} [v_{\mathbf{k}'}^2 u_{\mathbf{k}''}^2 - u_{\mathbf{k}'} v_{\mathbf{k}'} u_{\mathbf{k}''} v_{\mathbf{k}''}]/U - g_t(1-x) \sum_{\mathbf{k}'} \epsilon_{\mathbf{k}'} v_{\mathbf{k}'}^2/U$. We see that the incoherent contribution to the sum rule is $\mathcal{O}(t/U)$. Using the nonnegativity of spectral functions, we reach the very interesting conclusion that $A^{\rm inc}(\mathbf{k},\omega) = \mathcal{O}(xt/U)$ for $\omega > 0$. The unoccupied low-energy sum rule is much smaller than the occupied part, but it is mainly dominated by the coherent contributions. This also shows that the particle-hole asymmetry in the spectral weight comes chiefly from the incoherent part of the spectral function. Figure 2.8 shows the spectral function sum rules for two different dopings $x = 0.15$ and $x = 0.35$ along the zone diagonal. On the occupied side, the sum rule is dominated by incoherent spectral weight at low doping while the coherent part starts to dominate at higher doping. For the unoccupied part, the sum rule is always dominated by the coherent contribution.

2.7.4 Fermi Surface

We conclude this section by mentioning progress on the question [51, 52] of how one can determine the "underlying" Fermi surface in the superconducting ground state of strongly correlated Fermi systems. In a metallic state, following Luttinger, the change in sign of the zero energy Green's function with $G^{-1}(\mathbf{k},0) = 0$ can be used to define the Fermi surface. It has been suggested that in a superconductor one might be able to use the criterion $G(\mathbf{k},0) = 0$. However, this contour (in 2D, or surface in 3D) does *not* enclose the Luttinger count of n electrons; in fact, there is no analog of Luttinger's theorem in the superconducting state at $T = 0$ [51].

Nevertheless, one can look at various criteria that give information about where the low-energy excitations live in $\mathbf{k}$-space and compare them with each other and with ARPES experiments. These include criteria based on (1) the renormalized dispersion $\xi_{\mathbf{k}} = 0$, (2) the "minimum gap" locus from $E_{\mathbf{k}}$, (3) the momentum distribution $n_{\mathbf{k}}$, and (4) the low-energy spectral function $A(\mathbf{k},0)$. All these give similar, though not identical, results. The Luttinger count is violated in the superconducting state in an interesting way. Its deviation from the electron density is positive if the

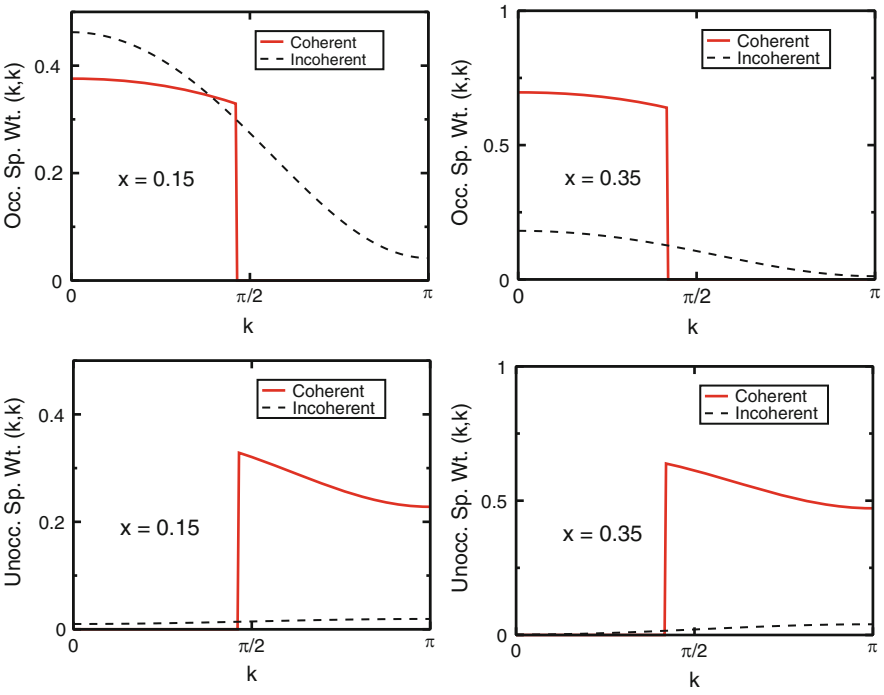

Fig. 2.8 The integrated spectral weight along the zone diagonal showing coherent and incoherent contributions on the occupied (electron extraction) and unoccupied (electron injection) sides for hole doping $x = 0.15$ and $x = 0.35$

Fermi surface is hole like and negative if the Fermi surface is electron like with a sign change when the Fermi surface topology changes. A similar result has been experimentally seen in cuprates through ARPES measurements [53]. For a detailed discussion, see [51, 52].

2.8 Superfluid Density

An important defining property of the superconducting state is its superfluid stiffness D_s (also called the superfluid density ρ_s) that is related to the energy cost of twisting the phase of the condensate. It is directly related to the measured London penetration depth λ_L via $\lambda_L^{-2} = 4\pi e^2 D_s/(\hbar^2 c^2 d)$, where c is the speed of light, e the electronic charge, and d the mean inter layer spacing.

In ordinary BCS superconductors, the energy scale D_s is of order the Fermi energy, and thus much larger than the pairing gap Δ. The smaller of the two scales controls T_c. Loss of superconductivity upon raising the temperature is controlled by

pair breaking, or thermally excited quasiparticles, resulting in the well-known BCS relation between T_c and the energy gap, $2\Delta \simeq 3.5 k_B T_c$.

In contrast, the underdoped high T_c superconductors show a completely different behavior. The antinodal gap increases with underdoping (decreasing x) while the T_c decreases. In addition, it is well established that the superfluid stiffness D_s also decreases upon underdoping so that it is much smaller than the gap for small enough x. The natural conclusion is that fluctuations of the phase of the order parameter destroy superconductivity on the underdoped side with T_c related to D_s [54] and not to the gap (The precise form of this relationship is not so important here, but is now understood in detail [55–58]).

In this section, we will first describe the doping dependence of the superfluid stiffness within the projected wavefunction framework, which is in good agreement with experiments. We will then look at the temperature dependence of D_s in a strongly correlated system and relate it to the current carried by the quasiparticle states.

2.8.1 Doping Dependence of D_s

For a superconductor, the real part of the frequency dependent conductivity has the form

$$\sigma(\omega) = \pi D_s \delta(\omega) + \sigma_{\text{reg}}(\omega), \tag{2.33}$$

where the superfluid stiffness D_s controls the strength of the delta-function and σ_{reg} is the regular part of the conductivity (in units where $\hbar = c = e = 1$). For the Hubbard model, there is the well-known optical conductivity sum-rule

$$\int_0^\infty d\omega Re\sigma(\omega) = 2\pi \sum_{\mathbf{k}} m_{\mathbf{k}}^{-1} n_{\mathbf{k}} \equiv \pi D_{\text{tot}}/2, \tag{2.34}$$

where $m_{\mathbf{k}}^{-1} = (\partial^2 \epsilon_{\mathbf{k}}/\partial k_x^2)$ is the noninteracting inverse mass. However, for projected systems, the more useful sum rule is restricted to the lower Hubbard band:

$$\int_0^\Omega d\omega \text{Re}\sigma(\omega) = \pi D_{\text{low}}/2, \tag{2.35}$$

where the cutoff Ω is chosen so that $U \gg \Omega \gg t, J$. It is evident from (2.33) that the low-energy integrated optical weight places an upper bound on the superfluid stiffness [11] so that: $D_s \leq D_{\text{low}} \leq D_{\text{tot}}$. The low-energy optical weight D_{low} is also called the "Drude weight."

We next describe how we can calculate D_{low} for projected superconductors. We determine the response of the system to an external vector potential $\mathbf{A}$, which acts

as a phase twist on the superconducting order parameter. The hopping amplitudes t_{ij} acquire a phase $t_{ij} \rightarrow t_{ij} e^{iA_{ij}}$, where $A_{ij} = \int_{\mathbf{r}_i}^{\mathbf{r}_j} \boldsymbol{A} \cdot d\boldsymbol{l} = \boldsymbol{A} \cdot (\mathbf{r}_i - \mathbf{r}_j)$. The canonically transformed Hamiltonian in the presence of the vector potential takes the form

$$H_A = \sum_{\langle ij \rangle} t_{ij} e^{A_{ij}} h_{i\bar{\sigma}} c^{\dagger}_{i\sigma} c_{j\sigma} h_{j\bar{\sigma}} - \frac{1}{U} \sum_{\langle il \rangle \langle lj \rangle} t_{il} t_{lj} e^{iA_{ij}} h_{i\bar{\sigma}'} c^{\dagger}_{i\sigma'} c_{l\sigma'} n_{l\bar{\sigma}} c^{\dagger}_{l\sigma} c_{j\sigma} h_{j\bar{\sigma}}. \tag{2.36}$$

We then evaluate $D_{\text{dia}} = \partial^2 H_A / \partial A^2 (A = 0)$. Using standard Kubo formula analysis, we find that $\langle D_{\text{dia}} \rangle$ gives both the diamagnetic response to a static $\mathbf{q} = 0$ vector potential as well as the low-energy optical spectral weight D_{low}. Assuming the vector potential to be along the x axis, one can write $D_{\text{dia}} = \sum_{\langle ij \rangle} t_{ij} (x_i - x_j)^2 h_{i\bar{\sigma}} c^{\dagger}_{i\sigma} c_{j\sigma} h_{j\bar{\sigma}} - \frac{1}{U} \sum_{\langle il \rangle \langle lj \rangle} t_{il} t_{lj} (x_i - x_j)^2 h_{i\bar{\sigma}'} c^{\dagger}_{i\sigma'} c_{l\sigma'} n_{l\bar{\sigma}} c^{\dagger}_{l\sigma} c_{j\sigma} h_{j\bar{\sigma}}$. Note that D_{dia} contains terms which represent hopping of holes ($\sim t$) and terms representing second-order processes where a double occupancy is created by hopping and then annihilated by a second hop. However, in this second-order process, terms which return the electrons to the point where they started ($i = j$ in $1/U$ terms, corresponding to spin fluctuations) do not contribute due to the $(x_i - x_j)^2$ factor.

Using Gutzwiller factors to take care of projection effects, we find that the final result can be written as

$$\langle D_{\text{dia}} \rangle = D_{\text{low}} = 2 \sum_{\mathbf{k}} Z_{\mathbf{k}} m^{-1}_{\mathbf{k}} v^2_{\mathbf{k}} + \frac{2 g_t (1 - x)}{U} \sum_{\mathbf{k}} \left(\frac{\partial \epsilon_{\mathbf{k}}}{\partial k_x} \right)^2 v^2_{\mathbf{k}}, \tag{2.37}$$

where $Z_{\mathbf{k}}$ is the quasiparticle weight given by (2.23). It is interesting to compare the low-energy sum rule D_{low} with the full sum rule D_{tot} of (2.34). The same (bare) inverse mass occurs in both, but the momentum distribution in D_{tot} has been replaced in the first term of D_{low} by the part of $n_{\mathbf{k}}$ coming from the coherent part of the spectral function. There is an extra term in D_{low} which is proportional to the square of the bare group velocity, but it is also multiplied by a factor of g_t.

For low doping $D_{\text{low}} \sim x$ and vanishes at half-filling. We thus conclude that (a) the low-energy charge response is suppressed in the underdoped limit leading to insulating behavior at $x = 0$, and (b) the superfluid stiffness, bounded from above by D_{low}, must vanish as $x \rightarrow 0$.

The variation of $D_{\text{low}}$with doping is shown in Fig. 2.9. The Gutzwiller approximation results and the Monte Carlo results closely match each other. The experimental data for optical conductivity is usually presented in terms of the plasma frequency $\omega_p^2 = 4\pi D_{\text{low}} K e^2 / d$, where d is the layer separation of the 2D planes and K is the number of planes. Using calculated $D_{\text{low}} \simeq 90\,\text{meV}$ for $x \sim 0.2$ (optimal doping), $d = 11.68\,\text{A}$ and $K = 2$, the estimated $\omega_p = 1.67\,\text{eV}$ for YBCO, while the experimentally measured value is $2.12\,\text{eV}$ [59]. We next use $\lambda_L^{-2} = 4\pi e^2 D_s / (\hbar^2 c^2 d)$ with the mean inter layer spacing $d = 7.5\,\text{A}$ for BSCCO,

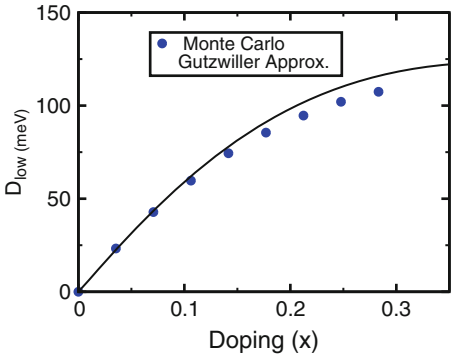

Fig. 2.9 The variation of the low energy integrated spectral weight or Drude weight D_{low} with hole doping. D_{low} is an upper bound for the superfluid stiffness. The *filled circles* are Monte Carlo results [10] while the *thick line* is obtained from the Gutzwiller approximation

to get a lower bound on the penetration depth $\lambda_L \geq 1350A$ at optimality, whereas the measured value is 2,100 A [60]. However, the overestimate of the superfluid stiffness is to be anticipated as the theory leaves out effects of impurities and of long-wavelength quantum phase fluctuations [61], both of which degrade the superfluid stiffness at $T = 0$.

2.8.2 Temperature Dependence of D_s

At finite temperatures $T \ll T_c$, thermally excited nodal quasiparticles lead to a reduction in the superfluid stiffness. For d-wave superconductors, in the absence of disorder, the low-energy DOS $N(\omega) \sim |\omega|$ leads to a linear decrease:

$$D_s(T) = D_s(0) - CT + \ldots \tag{2.38}$$

Experimentally, the zero temperature value $D_s(0)$ decreases with underdoping (as already discussed above), while the slope $\mathrm{d}D_s/\mathrm{d}T(0)$ shows interesting doping dependence: it is fairly constant near optimal doping and decreases in the highly underdoped samples.

We now address the question of the extent to which we can understand the doping dependence of the slope within the projected wavefunction framework. $\mathrm{d}D_s/\mathrm{d}T(0)$ can be related to the quasiparticle spectrum through [62]

$$\frac{\mathrm{d}D_s}{\mathrm{d}T}(0) = \frac{2\ln 2}{\pi}\alpha^2\frac{v_{\mathrm{F}}}{v_\Delta}, \tag{2.39}$$

where v_{F} and v_Δ were defined in (2.26) and αe is the effective charge defined by the current carried by the nodal quasiparticles

$$j_N = \alpha e v_{\mathrm{F}}. \tag{2.40}$$

Our goal is to determine the doping dependence of the charge renormalization $\alpha(x)$ arising from many-body backflow corrections [63].

The paramagnetic current operator is given by $j_x = -e\partial H_A/\partial A_x|_{A=0}$. Using (2.36), we find

$$j_x = ie\sum_{\langle ij\rangle} t_{ij} h_{i\bar{\sigma}} c^\dagger_{i\sigma} c_{j\sigma} h_{j\bar{\sigma}} (x_i - x_j) + ie\frac{1}{U}\sum_{\langle il\rangle\langle lj\rangle} t_{il}t_{lj}(x_i - x_j) h_{i\bar{\sigma}'} c^\dagger_{i\sigma'} c_{l\sigma'} n_{l\bar{\sigma}} c^\dagger_{l\sigma} c_{j\sigma} h_{j\bar{\sigma}}. \tag{2.41}$$

The current has the usual hopping term and a second-order term involving two hops, with the double occupancy created by one relaxed by the other. We now compute the expectation value of the current in the variational quasiparticle state $|\mathbf{q}\sigma\rangle = \exp(-\mathrm{i}\mathcal{S})\mathcal{P}\gamma^\dagger_{\mathbf{q}\sigma}|\mathrm{dBCS}\rangle$. This will be called $j^{\mathbf{q}}$ and should not be confused with the Fourier mode of the current.

Using the Gutzwiller approximation, we find $j^{\mathbf{q}}_x = eZ_{\mathbf{q}}(u^2_{\mathbf{q}} - v^2_{\mathbf{q}})\partial\epsilon_{\mathbf{q}}/\partial\mathbf{q}_x$. Note that the current carried by the quasiparticle is proportional to the bare velocity of the band electrons [62]. The $u^2_{\mathbf{q}}$ and $v^2_{\mathbf{q}}$ terms are the hole and particle contributions, respectively, which enter with opposite signs given their opposite charges. The overall current is renormalized by the quasiparticle weight Z of the excitations.

Evaluating this result at the node, we find that the current carried by the nodal quasiparticles is $j_N = eZ_N v^0_{\mathrm{F}}$, where v^0_{F} is the bare Fermi velocity. Comparing this with (2.40), we find the charge renormalization

$$\alpha = Z_N v^0_{\mathrm{F}}/v_{\mathrm{F}} = \left[1 + (1/v^0_{\mathrm{F}})\partial\Sigma'/\partial k\right]^{-1}, \tag{2.42}$$

where we have used (2.27). The results for $\alpha(x)$ are plotted in Fig. 2.10.

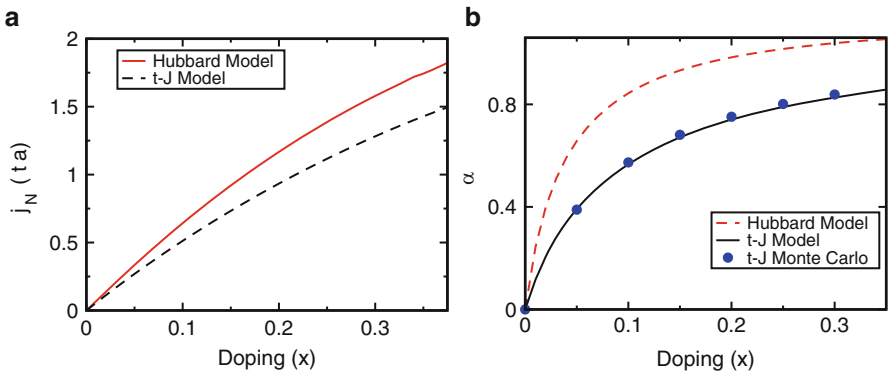

Fig. 2.10 *Left panel* (**a**): The current carried by the nodal quasiparticles as a function of doping for Hubbard and tJ models. *Right panel* (**b**): The effective charge α of the nodal quasiparticles as a function of doping for Hubbard and tJ models. Gutzwiller approximation results are compared with variational Monte Carlo results for the tJ model from [64]

To see how the Gutzwiller approximation gives semianalytical insights into VMC results, let us compare with Monte Carlo data on the tJ model [64]. In this case, the Gutzwiller approximation gives $Z_N = g_t$ and $\xi_{\mathbf{k}} = -(g_t t + \chi_1)\gamma_{\mathbf{k}}$, which leads to $v_{\mathrm{F}}^0/v_{\mathrm{F}} = t(g_t + \chi_1)$. The only numerical result we need is that the self-consistent solution yields $\chi_1 \approx J/2 \approx t/6$ for small x and the result is very weakly dependent on x. Thus, we find $\alpha(x) \simeq x/(1.08x + 0.08)$, which is in excellent agreement with the VMC data (see Fig. 2.10) and also gives insight into the large slope of $\alpha(x)$ at small x.

We thus see from (2.39) that although $\mathrm{d}D_s/\mathrm{d}T(0) \sim \alpha^2$, its x^2-behavior is restricted to a tiny range of doping values, where superconductivity gives way to other competing orders like antiferromagnetism (see below). Although theoretically very important, this analysis of $\mathrm{d}D_s/\mathrm{d}T(0)$ is hard to directly compare with experiments. In many materials, disorder changes the linear T behavior in $D_s(T)$ qualitatively, and in some experiments $\mathrm{d}\lambda_L/\mathrm{d}T$ is measured very accurately but not the absolute value of $\lambda_L(0)$ needed to determine $\mathrm{d}D_s/\mathrm{d}T \sim \mathrm{d}\lambda_L^{-2}/\mathrm{d}T$.

2.9 Disorder and Strong Correlations

We have thus far looked at the effects of strong correlations in a clean d-wave superconductor. In experiments, the process of doping the Mott insulator introduces intrinsic disorder in the system. Yet, superconductivity in the cuprates seems surprisingly robust against impurities. This is a remarkable finding in view of the standard Abrikosov–Gorkov theory [65], which implies that disorder is a pair-breaking perturbation for d-wave pairing symmetry. It has been confirmed that unconventional superconductors such as Sr_2RuO_4 with p-wave triplet pairing are indeed highly sensitive to disorder [66] as expected from the conventional theory.

The insensitivity of the cuprates to disorder stems in part from the fact that the dopants are far from the CuO_2 planes and also from the short coherence length. However, we point out here that there is another, more profound reason for the robustness of the cuprates against disorder. We present below a generalization of the Gutzwiller approximation to the inhomogeneous system to obtain the combined effects of strong correlation (projection) and disorder. We find the remarkable result that correlations play a central role in making the superconducting state and its low-energy excitations robust against disorder. The nodal quasiparticles, which are the low-energy electronic excitations, are found to be protected against disorder leading to characteristic signatures that are confirmed in scanning tunneling spectroscopy [67–71] and angle-resolved photoemission [41, 42, 72, 73] experiments.

We briefly describe our formalism. For simplicity, we work with the $t - t' - J$ model rather than the large U Hubbard model (see the discussion at the end of Sect. 2.2). We include a random impurity potential $V(\mathbf{r})$ that takes on values $V_0 > 0$ at a fraction n_{imp} of the sites and is zero elsewhere. Our Hamiltonian, acting on projected wavefunctions with no-double occupancy, is

$$H = -\sum_{\mathbf{r},\mathbf{r}',\sigma} t_{\mathbf{r}\mathbf{r}'}(c^{\dagger}_{\mathbf{r}\sigma}c_{\mathbf{r}'\sigma} + h.c.) + J\sum_{<\mathbf{r},\mathbf{r}'>}(\mathbf{S}_{\mathbf{r}}\cdot\mathbf{S}_{\mathbf{r}'} - n_{\mathbf{r}}n_{\mathbf{r}'}/4) + \sum_{\mathbf{r}}(V(\mathbf{r})-\mu)n_{\mathbf{r}}. \quad (2.43)$$

We focus on weak potentials $V_0 = t$ (Born scattering) as a simple model of intrinsic disorder in cuprates.

To clearly see the role of strong correlations, we compare the results of two different $T = 0$ calculations. (1) The first is a Bogoliubov deGennes (BdG) MFT that includes the inhomogeneity from the disorder, but treats interactions only at the simplest Hartree–Fock level. In other words, we make the gross simplification of ignoring the no-double occupancy constraint. In fact, most of the theoretical work [65, 74–76] in the field of disordered high T_c superconductors, barring a few exceptions [77, 78], uses just such an approximation. In this approach, the interaction in (2.43) is decomposed into *local* bond pairing amplitudes $\Delta_{\mathbf{r}\mathbf{r}'} = J\langle c_{\mathbf{r}\uparrow}c_{\mathbf{r}'\downarrow} - c_{\mathbf{r}\downarrow}c_{\mathbf{r}'\uparrow}\rangle/2$, and Fock shifts $W_{\mathbf{r}\mathbf{r}'} = J\langle c^{\dagger}_{\mathbf{r}'\sigma}c_{\mathbf{r}\sigma}\rangle/2$ with the density $n(\mathbf{r}) = \langle\sum_{\sigma} c^{\dagger}_{\mathbf{r}\sigma}c_{\mathbf{r}\sigma}\rangle$. The effective BdG Hamiltonian is then diagonalized and the fields defined above calculated self-consistently.

(2) The second calculation [35] includes *both* strong correlations *and* disorder, using a combination of the Gutzwiller approximation and the BdG method. Given the short coherence length, it is known from both theory and experiments that one has to deal with a highly inhomogeneous superconducting state. Thus, the treatment of strong correlations using the Gutzwiller approximation needs to be generalized [79] to take into account this spatial inhomogeneity. We discussed in Sect. 2.4 the standard Gutzwiller factors for translationally invariant systems and how they depend upon the hole density x. Those results must now be generalized to keep track of the *local* variations in the hole density $x(\mathbf{r}) = 1 - n(\mathbf{r})$.

Using the absence (or presence) of a subscript 0 to denote expectation values in projected (or unprojected) states, we obtain the following results. For the kinetic energy, we find $\langle c^{\dagger}_{\mathbf{r}\sigma}c_{\mathbf{r}'\sigma}\rangle \approx g_t(\mathbf{r},\mathbf{r}')\langle c^{\dagger}_{\mathbf{r}\sigma}c_{\mathbf{r}'\sigma}\rangle_0$ and for the spin correlation we get $\langle\mathbf{S}_{\mathbf{r}}\cdot\mathbf{S}_{\mathbf{r}'}\rangle \approx g_s(\mathbf{r},\mathbf{r}')\langle\mathbf{S}_{\mathbf{r}}\cdot\mathbf{S}_{\mathbf{r}'}\rangle_0$. The *local* Gutzwiller factors are given by $g_t(\mathbf{r},\mathbf{r}') = g_t(\mathbf{r})g_t(\mathbf{r}')$ with $g_t(\mathbf{r}) = [2x(\mathbf{r})/(1+x(\mathbf{r}))]^{1/2}$ and $g_s(\mathbf{r},\mathbf{r}') = 4/[(1+x(\mathbf{r}))(1+x(\mathbf{r}'))]$.

We can derive the renormalized BdG equations and solve them numerically to find the eigenvalues E_n and the corresponding eigenvectors $(u_n(\mathbf{r}), v_n(\mathbf{r}))$. We impose the *local* self-consistency equations for the density $n(\mathbf{r}) = 2\sum_n |v_n(\mathbf{r})|^2$, the pairing amplitude $\Delta_{\mathbf{r},\mathbf{r}'} = J_1(\mathbf{r},\mathbf{r}')\sum_n[u_n(\mathbf{r}')v^{\star}_n(\mathbf{r}) + u_n(\mathbf{r})v^{\star}_n(\mathbf{r}')]$ and the Fock shift $W_{\mathbf{r},\mathbf{r}'} = J_2(\mathbf{r},\mathbf{r}')\sum_n v_n(\mathbf{r}')v^{\star}_n(\mathbf{r})$, where the renormalized couplings $J_{1,2}(\mathbf{r},\mathbf{r}') = J(3g_s(\mathbf{r},\mathbf{r}')\pm 1)/4$.

In a disordered system, the one-particle spectral function $A(\mathbf{r},\mathbf{R};\omega) = -\mathrm{Im}G(\mathbf{r}_1,\mathbf{r}_2;\omega + i0^+)/\pi$ depends on both the center-of-mass $\mathbf{R} = (\mathbf{r}_1+\mathbf{r}_2)/2$ and the relative coordinate $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Within the Gutzwiller approximation, we obtain $A(\mathbf{r},\mathbf{R};\omega) = g_t(\mathbf{r}_1,\mathbf{r}_2)A^0(\mathbf{r},\mathbf{R};\omega)$ with $A^0 = \sum_n u_n(\mathbf{r}_1)u^{\star}_n(\mathbf{r}_2)\delta(\omega - E_n) +$

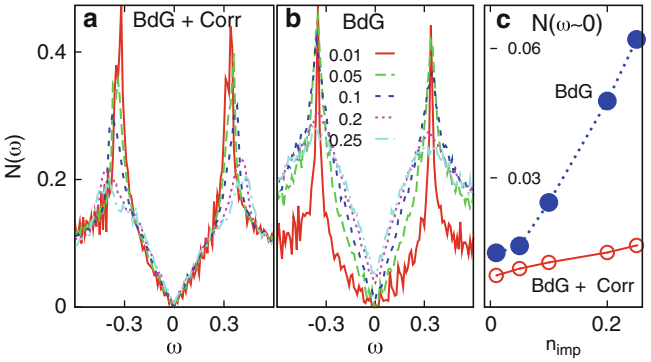

Fig. 2.11 DOS $N(\omega)$ for various impurity concentrations [35]. The results in panel (**a**) include both strong correlations using Gutzwiller projection and disorder using BdG. The *curves* correspond to different values of n_{imp} indicated in (**b**). Note the robust V-shaped density of states in (**a**), which implies that strong correlations protect the nodes. Panel (**b**) shows the results of a BdG calculation with disorder (but only a mean field treatment of interactions). Here, the low-energy DOS fills up with increasing disorder. This is further illustrated in panel (**c**), which shows the low-energy DOS integrated over a small window $\mid \omega \mid \leq 0.02t$ as a function of n_{imp}. Results with correlations (*open symbols* (*red*)) and plain BdG (*closed symbols* (*blue*)). The average electronic density $n = 0.8$, impurity potential $V_0 = 1.0t$

$v_n(\mathbf{r}_1)v_n^\star(\mathbf{r}_2)\delta(\omega + E_n)$. The Gutzwiller approximation describes the coherent part of the spectral function which dominates A for ω's smaller than or comparable to the gap. We will focus exclusively on these low-energy excitations here. The higher energy incoherent spectral weight, not accessible to the Gutzwiller approximation, can be constrained by exact sum rules and leads to $p - h$ asymmetry, as discussed in earlier sections.

A detailed discussion of the order parameter and density inhomogeneity in the presence of disorder and its implications for STM and ARPES observables may be found in [35]. Here, we show only one result: the DOS $N(\omega)$ as a function of impurity concentration in Fig. 2.11. This is calculated from the spectral function by first defining the LDOS $N(\mathbf{R}, \omega) = A(0, \mathbf{R}; \omega)$, and then spatially averaging it over the disordered system: $N(\omega) = \langle N(\mathbf{R}, \omega)\rangle_{\mathbf{R}}$.

The results of Fig. 2.11 clearly show that there is much less pair breaking in a strongly correlated superconductor compared with its uncorrelated counterpart, even if we take into account the short coherence length and inhomogeneity in both calculations. (We note that a weak coupling d-wave BCS theory [65] would predict an even larger impurity-induced density of states at low energies than that found in the plain BdG calculation).

The striking feature of the strongly correlated superconductor in panel (a) is that the low-energy DOS is highly insensitive to disorder, the only disorder-induced changes occur near the gap edge. These qualitative features are observed in all the STM studies of the cuprates [67–71]. Even though there are large inhomogeneous variations of the superconducting gap on the nanoscale, nevertheless the LDOS

at low energies is remarkably homogeneous. The "protected" nature of nodal quasiparticles is also seen in the ARPES spectral functions calculated from inhomogeneous Gutzwiller approximation [35] and gives insight into the nodal/antinodal dichotomy observed in ARPES experiments.

Let us conclude this section with a brief discussion of the reasons why correlations lead to this unusual effect. First, correlations lead to a mass-renormalization of about 2–3 and the lower v_F leads to a shorter healing length. Second, interactions screen the impurity potential. Finally, as shown in [35], the electron density decreases near a repulsive impurity potential leading to an increase in the local g_t. The resulting increase in the local hopping helps heal the loss of pairing near the impurity. The interplay of strong correlations and disorder raises many profound unsolved questions in condensed matter physics, and this is but one example where the Gutzwiller method has led to some new insights.

2.10 Competing Orders: Antiferromagnetism

The question of competing orders in lightly doped Mott insulators has received a lot of attention. The undoped parent compound of the high-T_c cuprates has long range antiferromagnetic (AF) order in its ground state. Upon doping the AF, T_{Neel} drops rapidly and d-wave superconductivity (SC) develops around $x \approx 0.05$. In addition, there is evidence for weak incommensurate spin density wave AF and also stripe order in lightly doped cuprates.

We have thus far discussed only the d-wave superconducting state in the doped system. We now briefly describe how the Neel AF state, known to be the ground state of the Heisenberg model at half-filling, competes and coexists with d-wave SC upon doping. These results are obtained from VMC calculations [80] of the tJ model with longer range hopping t' and t''.

We use a projected variational wavefunction that has the possibility of both d-wave SC and AF order parameters, and minimize its energy to determine the optimal state. Our approach builds on earlier studies of competition and coexistence of AF and SC [81–84]. The behavior of the two-order parameters as a function of electron and hole doping and of the next-neighbor hopping t' is shown in the "phase diagram" of Fig. 2.12. As explained in [80], key features of the $T = 0$ phase diagram, such as critical doping for SC-AF coexistence and the maximum value of the SC order parameter, are all determined by a single parameter that characterizes the curvature and topology of the "Fermi surface" at half filling, defined by the bare tight-binding parameters.

We emphasize two important conclusions that can be drawn from Fig. 2.12. First, t' introduces a striking asymmetry between electron and hole doping [85, 86]. As a result, AF is the stronger instability on the electron-doped side that dominates over SC, while just the opposite is true on the hole-doped side. The reasons for which order dominates are elucidated in [80]. Second, there is a very important trend on the hole-doped side: the larger the magnitude of t', the stronger the SC. In

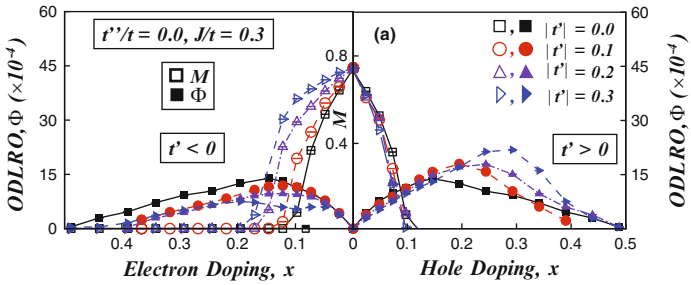

Fig. 2.12 Phase diagram showing the d-wave superconducting (SC) and antiferromagnetic (AF) order parameters as a function of doping, based on variational Monte Carlo results of Pathak et al. [80] for the t-t′-J model with various values of the next-neighbor hopping t′. Note the asymmetry between hole-doping (with SC dominating) and electron-doping (with AF dominating). Note also the systematic t′ dependence of SC on the hole-doped side

effect, our analysis constitutes a microscopic derivation of the connection between the range of hopping and superconductivity, which had been first emphasized by Pavarini et al. [87] on empirical grounds. For a detailed discussion of these issues, we refer the reader to [80].

2.11 Conclusion

We have described the insights one can get into the strongly correlated d-wave superconducting (SC) state in doped Mott insulators using projected wavefunctions. Of course, one cannot hope to *prove* using variational methods that the 2D large-U Hubbard model has a superconducting ground state. However, we *can* show that a SC state is energetically favored for a range of densities, and then compare the predictions of its properties with those observed in the high T_c cuprates. Our main goal was to calculate experimentally observed quantities and compare our results with experiments. Three strategies were used for the calculations: exact sum rules for projected fermions, VMC calculations that impose the no double-occupancy constraint exactly, and the Gutzwiller approximation – our primary focus – which gives detailed analytical insights.

To summarize, we have obtained a variety of detailed insights into the d-wave superconducting state in doped Mott insulators.

1. There is a striking $p - h$ asymmetry in the spectral function of doped Mott insulators that shows that strong correlations continue to play a crucial role in the doped systems.
2. The projected wavefunction approach leads naturally to a superconducting dome with its nonmonotonic dependence on hole density. The suppression of SC on the underdoped side of the dome is controlled by the high energy scale of U and the

approach to the Mott insulator, and *not* by a competing order whose characteristic scale can only be J.
3. The energy gap and the SC order parameter behave in a qualitatively different manner as a function of doping.
4. We have elucidated in detail the very interesting evolution of the excitations (nodal quasiparticles, their spectral weight, and dispersion), Fermi surface and response functions (optical weight, superfluid density) from the overdoped, weakly correlated SC with the underdoped, strongly correlated SC in the vicinity of the Mott insulator.
5. We showed that disorder effects are suppressed in the presence of strong correlations, thus giving insight into the robustness of d-wave SC in the cuprates against impurities and the nodal-antinodal dichotomy.
6. Finally, we have elucidated the asymmetry between electron and hole doping a Mott insulator, and the competition between AF and SC order. We have shown that longer range in-plane hopping promotes higher T_c on the hole-doped side.

What implications do our SC state results have for the anomalous normal states? It is clear from the results presented above that the phase transition to the normal state on the under- and overdoped sides are quite different. On the overdoped side, the energy gap Δ is much smaller than the superfluid stiffness D_s (as in usual BCS superconductors) and SC is destroyed by pair breaking. In contrast, on the underdoped side of the SC dome, $D_s \ll \Delta$ and phase fluctuations destroy SC order [88]. Thus, pairing must survive above T_c [89, 90] on the underdoped side. There is clear experimental evidence for fluctuation diamagnetism and fluctuating vortices just above T_c [91, 92], and for spin pairing from NMR [93] and a $\mathbf{k}$-dependent ARPES gap over a broad range of temperatures extending up to T^* [41, 42]. However, the pseudogap region in underdoped cuprates also shows evidence for other competing orders, including spin-density wave magnetism, stripes, and broken time reversal. Competing orders become more prominent in low-temperature experiments where SC is destroyed by external magnetic fields [94]. A complete understanding of the strongly fluctuating pseudogap regime close to the Mott insulator is not available at this time.

The strange metal normal state, above optimal doping, with its characteristic linear resistivity and marginal Fermi liquid phenomenology [95], also poses a challenge for theory. It has been interpreted as the quantum critical regime [96, 97] of a critical point hidden under the dome, but the order parameter for this transition is not definitively established, nor the unusual scaling in frequency, without any obvious spatial scaling, persisting up to very high temperatures. An alternative point of view is that the strange metal is a consequence of strong correlations leading to non-Fermi liquid behavior [6, 98–100].

In conclusion, while the strongly correlated superconducting state and its low-energy excitations and response functions are reasonably well understood, there is much less agreement on the non-Fermi liquid normal state properties of the large U Hubbard and tJ models.

Acknowledgments We thank Arun Paramekanti, Juan Carlos Campuzano, Seamus Davis, Arti Garg, and Fu-Chun Zhang for many discussions, and acknowledge inspiring interactions with P. W. Anderson. Our work was supported by NSF DMR-1006532 and DOE DE-SC0005035 (MR), DARPA QuEST and DARPA OLE (RS), and NSF DMR-0907275 (NT).

References

1. J.C. Bednorz, K.A. Muller, Z. Phys B **64**, 189 (1986)
2. P.W. Anderson, Science **235**, 1196 (1987)
3. S. Liang, B. Ducot, P.W. Anderson, Phys. Rev. Lett. **61**, 365 (1988)
4. G. Baskaran, Z. Zhou, P.W. Anderson, Solid State Comm. **63**, 973 (1987)
5. G. Kotliar, J. Liu, Phys. Rev. B **38**, 5142 (1988)
6. P.A. Lee, N. Nagaosa, X.G. Wen, Rev. Mod. Phys. **78**, 17 (2006)
7. C. Gros, Phys. Rev. B **38**, 931 (1988)
8. F.C. Zhang, C. Gros, T.M. Rice, H. Shiba, Supercond. Sci. Tech. **1**, 36 (1988)
9. D. J. Scalapino, J.E. Loh, J.E. Hirsch, Phys. Rev. B **34**, 8190 (1986)
10. A. Paramekanti, M. Randeria, N. Trivedi, Phys. Rev. Lett. **87**, 217002 (2001)
11. A. Paramekanti, M. Randeria, N. Trivedi, Phys. Rev. B **70**, 054504 (2004)
12. P.W. Anderson, P.A. Lee, M. Randeria, T.M. Rice, N. Trivedi, F.C. Zhang, J Phys. Condens. Matter **16**, R755 (2004)
13. B. Edegger, C. Gros, V.N. Muthukumar, Adv. Phys. **56**, 927 (2007)
14. M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
15. C. Gros, R. Joynt, T.M. Rice, Phys. Rev. B **36**, 381 (1987)
16. A.H. Mcdonald, S.M. Girvin, D. Yoshioka, Phys. Rev. B **37**, 9753 (1988)
17. A.B. Harris, R.V. Lange, Phys. Rev. **157**, 295 (1967)
18. M.B.J. Meinders, H. Eskes, G. Sawatzky, Phys. Rev. B **48**, 3916 (1993)
19. H. Eskes, A.M. Oles, Phys. Rev. Lett. **73**, 1279 (1993)
20. H. Eskes et al., Phys. Rev. B **50**, 17980 (1994)
21. M. Randeria, R. Sensarma, N. Trivedi, F.C. Zhang, Phys. Rev. Lett. **95**, 137001 (2005)
22. P.W. Anderson, N.P. Ong, J. Phys. Chem. Solids **67**, 1 (2006)
23. S.H. Pan et al., Nature **413**, 282 (2001)
24. J.E. Hoffman et al., Science **297**, 1148 (2002)
25. K. McKelroy et al., Nature **422**, 592 (2003)
26. H. Hanaguri et al., Nature **430**, 1001 (2004)
27. P. Phillips, T-P. Choy, and R.G. Leigh, Rep. Prog. Phys. **72**, 036501 (2009)
28. Y. Kohsaka et al., Science **315**, 1380 (2007)
29. D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984)
30. J.Y. Gan, Y. Chen, Z.B. Su, F.C. Zhang, Phys. Rev. Lett. **94**, 067005 (2005)
31. J. Y. Gan, F.C. Zhang, Z.B. Su, Phys. Rev. B **71**, 014508 (2005)
32. H.X. Huang, Y.Q. Li, F.C. Zhang, Phys. Rev. B **71**, 184514 (2005)
33. D. Poilblanc, Phys. Rev. B **72**, 060508(R) (2005)
34. C. Li, S. Zhou, Z. Wang, Phys. Rev. B **73**, 060501(R) (2006)
35. A. Garg, M. Randeria, N. Trivedi, Nat. Phys. **4**, 762 (2008)
36. W-H. Ko, C.P. Nave, P.A. Lee, Phys. Rev. B **76**, 245113 (2007)
37. M. Sigrist, T.M. Rice, F.C. Zhang, Phys. Rev. B **49**, 12058 (1994)
38. M. Ogata, A. Himeda, J. Phys. Soc. Jpn. **72**, 374 (2003)
39. T. Ogawa, K. Kanda, T. Matsubara, Prog. Theor. Phys. **53**, 614 (1975)
40. R. Sensarma, Ph.D. Thesis, The Ohio State University (2007)
41. J.C. Campuzano, M.R. Norman, M. Randeria, in *Physics of Superconductors*, vol. II, ed. by K.H. Bennemann, J.B. Ketterson (Springer, Berlin, 2004), pp. 167
42. A. Damasceli, Z. Hussain, Z.X. Shen, Rev. Mod. Phys. **75**, 473 (2003)

43. J.M. Tranquada et al., Phys. Rev. B **40**, 4503 (1989)
44. J.C. Campuzano et al., Phys. Rev. Lett. **83**, 3709 (1999)
45. M.R. Presland et al., Physica C **176**, 95 (1991)
46. P.D. Johnson et al., Phys. Rev. Lett. **87**, 177007 (2001)
47. K.M. Shen, J.C. Davis, Materials Today **11**, 14 (2008)
48. O. Fischer et al., Rev. Mod. Phys. **79**, 353 (2007)
49. A. Kaminski et al., Phys. Rev. Lett. **84**, 1788 (2000)
50. D.N. Basov, T. Timusk, Rev. Mod. Phys. **77**, 721 (2005)
51. R. Sensarma, M. Randeria, N. Trivedi, Phys. Rev. Lett. **98**, 027004 (2007)
52. C. Gros, B. Edegger, V.N. Muthukumar, P.W. Anderson, Proc. Nat. Acad. Sci. **103**, 14298 (2006)
53. T. Yoshida et al., Phys. Rev. B **74**, 224510 (2006)
54. Y.J. Uemura et al., Phys. Rev. Lett. **62**, 2317 (1989)
55. I. Hetel, T.R. Lemberger, M. Randeria, Nat. Phys. **3**, 700 (2007)
56. J.L. Tallon et al., Phys. Rev. B **68**, 180501 (2003)
57. T. Pereg-Bernea et al., Phys. Rev. B **69**, 184513 (2004)
58. D.M. Broun et al., Phys. Rev. Lett. **99**, 237003 (2007)
59. S.L. Cooper et al., Phys. Rev. B **47**, 8233 (1993)
60. S.F. Lee et al., Phys. Rev. Lett. **77**, 735 (1996)
61. A. Paramekanti, M. Randeria, T.V. Ramakrishnan, S.S. Mandal, Phys. Rev. B **62**, 6786 (2000)
62. A.C. Durst, P.A. Lee, Phys. Rev. B **62**, 1270 (2000)
63. A. Paramekanti, M. Randeria, Phys. Rev. B **66**, 214517 (2002)
64. C.P. Nave, D.A. Ivanov, P.A. Lee, Phys. Rev. B **73**, 104502 (2006)
65. A.V. Balatsky, I. Vekhter, J-X. Zhu, Rev. Mod. Phys. **78**, 373 (2006)
66. A.P. Mackenzie, Y. Maeno, Rev. Mod. Phys. **75**, 657 (2003)
67. S.H. Pan et al., Nature **413**, 282 (2001)
68. K. McElroy et al., Phys. Rev. Lett. **94**, 197005 (2005)
69. M. Vershinin et al., Science **303**, 1995 (2004)
70. K. McElroy et al., Science **309**, 1048 (2005)
71. A.N. Pasupathy et al., Science **320**, 196 (2008)
72. X.J. Zhou et al., Phys. Rev. Lett. **92**, 187001 (2004)
73. K. M. Shen et al., Science **307**, 901 (2005)
74. W.A. Atkinson, P.J. Hirschfeld, A.H. MacDonald, Phys. Rev. Lett. **85**, 3922 (2000)
75. A. Ghosal, M. Randeria, N. Trivedi, Phys. Rev. B **63**, 020505 (2000)
76. T.S. Nunner et al., Phys. Rev. Lett. **95**, 177003 (2005)
77. Z. Wang et al., Phys. Rev. B **65**, 064509 (2002)
78. Q-H. Wang et al., Phys. Rev. B **73**, 092507 (2006)
79. H. Tsuchiura et al., Phys. Rev. Lett. **84**, 3165 (2000)
80. S. Pathak, V. B. Shenoy, M. Randeria, N. Trivedi, Phys. Rev. Lett. **102**, 027002 (2009)
81. T. Giamarchi, C. Lhuillier, Phys. Rev. B **43**, 12943 (1991)
82. A. Himeda, M. Ogata, Phys. Rev. B **60**, R9935 (1999)
83. T.K. Lee, C-M. Ho, N. Nagaosa, Phys. Rev. Lett. **90**, 067001 (2003)
84. C.T. Shih et al., Phys. Rev. Lett. **92**, 227002 (2004)
85. G. Kotliar et al., Phys. Rev. Lett. **87**, 186401 (2001)
86. D. Senechal et al., Phys. Rev. Lett. **94**, 156404 (2005)
87. E. Pavarini et al., Phys. Rev. Lett. **87**, 047003 (2001)
88. V. Emery, S. Kivelson, Nature **374**, 434 (1995)
89. M. Randeria et al., Phys. Rev. Lett. **69**, 2001 (1992)
90. N. Trivedi, M. Randeria, Phys. Rev. Lett. **75**, 312 (1995)
91. L. Li et al., Phys. Rev. B **81**, 054510 (2010)
92. Y. Wang, L. Li, N.P. Ong, Phys. Rev. B **73**, 024510 (2006)
93. T. Timusk, B.W. Statt, Rep. Prog. Phys. **62**, 61 (1999)
94. L. Taillefer, Ann. Rev. Cond. Mat. Phys. **1**, 51 (2010)
95. C.M. Varma et al., Phys. Rev. Lett. **63**, 1996 (1989)

96. C.M. Varma, Phys. Rev. B **73**, 155113 (2006)
97. *The landscape of the Hubbard model* by S. Sachdev, TASI and Chandrasekhar Lectures, arXiv: 1012.0299
98. K-Y. Yang, T. M. Rice, F-C. Zhang, Phys. Rev. B **73**, 174501 (2006)
99. P. W. Anderson, Int. J. Mod. Phys. B **25**, 1 (2011)
100. P. A. Casey, P. W. Anderson, Phys. Rev. Lett. **106**, 097002 (2011)

Chapter 3
The Pseudoparticle Approach to Strongly Correlated Electron Systems

Raymond Frésard, Johann Kroha, and Peter Wölfle

Abstract The most prominent pseudoparticle representations and their applications to correlated spin and electronic models are reviewed, with approximate solution schemes ranging from saddle-point approximations with Gaussian fluctuations to conserving approximations and renormalization group (RG) techniques. Merits and shortcomings of these methods are described. In particular, the generic feature of radial slave boson fields to possess a finite expectation value is discussed, while pure fluctuation fields may best be treated by conserving approximations. We present applications to the magnetic phases of the Anderson lattice model and to the Hubbard model. The noncrossing approximation and the conserving T-matrix approximation are presented and discussed as the most important conserving approximations. Furthermore RG techniques for pseudoparticle representations, including "poor man's scaling" and functional RG for the Kondo model in and out of equilibrium, a novel RG approach to the Kondo model for strong coupling, and the functional RG approach to frustrated Heisenberg models are discussed.

R. Frésard (✉)
Laboratoire Crismat, UMR CNRS-ENSICAEN(ISMRA) 6508, 6, Bld. du Marèchal Juin, 14050 Caen CEDEX 4, France
e-mail: Raymond.Fresard@ensicaen.fr

J. Kroha
Physikalisches Institut, Universität Bonn, 53115 Bonn, Germany
e-mail: kroha@physik.uni-bonn.de

P. Wölfle
Institut für Theorie der Kondensierten Materie and Institut für Nanotechnologie, Karlsruher Institut für Technologie (KIT), 76049 Karlsruhe, Germany
e-mail: peter.woelfle@kit.edu

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_3,

3.1 Introduction

The immense and steadily increasing field of strongly correlated electrons has emerged as a central theme of many-body physics over the past 3 decades (for a review see [1]). Of particular interest are the so-called heavy-fermion metallic compounds [2] and copper-oxide superconductors [3]. While fully accounting for their properties remains a challenging task, it is believed that their key properties are embodied in the Anderson or Kondo lattice models in the former case, and in the Hubbard model and the t–J model in the latter one. The difficulty in solving these models is rooted in the fact that conventional many-body perturbation theory (including infinite resummations) does not work in these cases. This failure is obvious in lattice models with on-site repulsion U exceeding the band width D.

Take the Hubbard model with a large on-site repulsion U, where each lattice site can either be empty (state $|0\rangle$), singly occupied ($|\uparrow\rangle$ and $|\downarrow\rangle$) or doubly occupied ($|2\rangle$). The dynamics of an electron will be very different according to whether it resides on a singly or doubly occupied site. For large U the doubly occupied states will be pushed far up in energy, and will not contribute to the low-energy physics. This leads effectively to a projection of the Hilbert space onto a subspace devoid of doubly occupied states. It turns out to be difficult to effect the projection within conventional many-body theory, as has been realized early on in the context of the magnetic impurity problem. Indeed, this difficulty is at the heart of the single impurity Kondo problem, for which a sound physical picture and quantitative analytical and numerical methods of solution have been developed over a period of 40 years [4]. Over the past 15 years the Kondo problem has attracted renewed interest in the context of electron transport through nanostructures, in particular in situations out of equilibrium. We will discuss impurity models in Sects. 3.5 and 3.6.

A powerful technique for describing the projection in Hilbert space is the method of auxiliary particles (slave bosons and pseudofermions [5–10]): One assigns an auxiliary field or particle to each of the four states $|0\rangle, |\downarrow\rangle, |\uparrow\rangle$ and $|2\rangle$ at a given lattice site (considering one strongly correlated orbital per site). The Fermi character of the electrons requires that two of the auxiliary particles are fermions, e.g., the ones representing $|\downarrow\rangle, |\uparrow\rangle$ or equally well $|0\rangle, |2\rangle$, and the remaining two are bosons. Introducing new particles for the states $|0\rangle$ and $|2\rangle$ allows us to express the projection to the Hilbert space of states without double occupancy as $n_0 + n_\uparrow + n_\downarrow = 1$, where n_α are the occupation numbers of the states $|\alpha\rangle$; i.e., each lattice site is either empty or singly occupied. There are various ways of defining auxiliary particles for a given problem. It is wise to choose the one which is best adapted to the physical properties of the system.

Compared to alternative ways of effecting the projection, the auxiliary particle method has the advantage of allowing one to use the machinery of quantum field theory, i.e., Wick's theorem, diagrammatic perturbation theory and infinite resummations of diagrams, provided the constraint can be incorporated in a satisfactory way.

Historically, auxiliary particle representations have first been introduced in the context of spin models. Spin operators may be represented by Bose operators (Holstein–Primakoff [5] and Schwinger [6]) or in the case of spin 1/2 (and with additional complications for higher spins as well) by Fermi operators (Abrikosov [7] and Coqblin–Schrieffer [8]). Electron operators necessarily involve a combination of auxiliary fermion and boson operators. The simplest such representation has been proposed for the Anderson impurity problem by Barnes [9], and for lattice problems by Coleman [10]. A more complex representation of electron operators, incorporating the result of the Gutzwiller approximation [11] on the slave boson mean-field level has been developed by Kotliar and Ruckenstein [12]. Generalizations of the latter to manifestly spin rotation invariant form [13] and to particle–hole and spin rotation invariant form [14] have also been proposed. Generalizations to multiband Hubbard models have been introduced as well [15, 16].

Quite generally, auxiliary particle theories have to deal with two problems: the treatment of the constraint and the approximate description of the dynamics. An accurate control of the constraint alone does not yet make a good theory!

In Sect. 2 we review most of the available pseudoparticle representations for both spin operators and electron operators. Mean-field approximations of slave boson operators within several of the most prominent models are considered in Sect. 3. Section 4 is devoted to Gaussian fluctuation corrections on top of mean-field theories. Chapter 5 is devoted to a class of time-dependent mean-field type approximations, the noncrossing approximation (NCA) and the conserving T-matrix approximation. A powerful nonperturbative method allowing to capture collective many-body phenomena is the renormalization group method in its various forms. In Sect. 6 we present several examples of auxiliary particle representations employed to describe strong coupling problems within a RG treatment.

3.2 Pseudoparticle Representations of Quantum Operators

3.2.1 Spin Operators

As mentioned in the Introduction, the first auxiliary particle representations of quantum operators proposed and successfully applied are the bosonic representations of spin operators by Holstein-Primakoff and by Schwinger. Since both are well documented in the literature, we do not consider them here. Generally speaking, bosonic representations are useful to describe ordered states and fluctuations of collective variables. They are less useful if the fermionic character of spin 1/2 particles is of importance. In the following we will concentrate on fermionic representations.

3.2.1.1 Fermionic Representations of S=1/2 Spins

The spin 1/2 operator **S** has a faithful representation in terms of fermion operators f_σ ($\sigma = \uparrow, \downarrow$)

$$\mathbf{S} = \frac{1}{2} \sum_{\sigma,\sigma'=\uparrow,\downarrow} f_\sigma^\dagger \boldsymbol{\tau}_{\sigma\sigma'} f_{\sigma'}, \tag{3.1}$$

where $\boldsymbol{\tau}$ is the vector of Pauli matrices. The Hilbert space obtained by the creation operators $f_\sigma^\dagger$ acting on the vacuum state $|\mathrm{vac}\rangle$ is spanned by four states. The two unphysical states, the empty and the doubly occupied one, are eliminated by requiring that all states considered are eigenstates of the occupation number operator $Q = \sum_{\sigma=\uparrow,\downarrow} f_\sigma^\dagger f_\sigma$ with an eigenvalue $Q = 1$.

Abrikosov Projection

A first projection scheme involves adding a term λQ to the Hamiltonian, and taking the limit $\lambda \to \infty$ [7]. In that case double occupancy is forbidden, while empty states are not involved in expectation values of physical spin operators. The projection of the pseudofermion Green's function $G^<(\omega)$, for example, is effected by taking the following limit $G^<(\omega) = \lim_{\lambda\to\infty}[G_\lambda^<(\omega) Z_\lambda / \langle Q\rangle_\lambda]$. Here $G_\lambda^<(\omega)$ is the pseudofermion Green's function at a finite chemical potential λ, Z_λ is the partition function and $\langle Q\rangle_\lambda = \sum_{\sigma=\uparrow,\downarrow} \left\langle f_\sigma^\dagger f_\sigma \right\rangle_\lambda$ is the total number of pseudofermions at given λ. While taking the limit $\lambda \to \infty$ removes the contribution of doubly occupied states, the unphysical contribution of the empty level in the partition function Z_λ is removed in the limit $\lambda \to \infty$ by replacing Z_λ by $\langle Q\rangle_\lambda$. In equilibrium, we may now use that $-iG_\lambda^<(\omega) = f(\omega+\lambda)A(\omega) \overset{\lambda\to\infty}{\longrightarrow} \mathrm{e}^{-\lambda/T}\mathrm{e}^{-\omega/T}A(\omega)$, where $f(\omega)$ is the Fermi function, to express the projected Green's function as $G^<(\omega) = ie^{-\omega/T}A(\omega)/[2\int \mathrm{d}\omega \mathrm{e}^{-\omega/T}A(\omega)]$. Out of equilibrium the occupation function $f_\lambda(\omega)$ has to be calculated from the quantum Boltzmann equation.

The above projection scheme has the disadvantage that particle–hole symmetry is maximally broken. The spectral functions are, therefore, very unsymmetric under a sign change of ω, causing difficulty in numerical evaluations.

Particle–Hole Symmetric Projection

If one takes $\lambda = 0$ the pseudofermion spectral function is particle–hole symmetric, $A(-\omega) = A(\omega)$, which facilitates calculations considerably. In calculating any physical quantity, which only involves spin operators, unphysical states only come in through the partition function, which enters quantities described by diagrams with

closed pseudofermion loops. In the case of quantum impurity models, diagrams with more than one pseudofermion loop do not contribute, as each loop introduces a factor $1/N_{\mathrm{L}}$ with N_{L} the number of lattice sites in the system. The pseudofermion self-energy is not affected, because its relevant diagrams do not contain loops. By contrast, the diagrams for response functions like the conductance and the spin susceptibility necessarily contain one loop. The latter quantities have to be corrected by a normalization factor, as discussed by Larsen [17]. The correction amounts to multiplying the pseudofermion occupation factor $f_\lambda(\omega)$ by a factor

$$Y = Z/Z_{\mathrm{p}} = (Z/Z_0)/[(Z/Z_0) - 1/2]\,, \tag{3.2}$$

where $Z(\lambda) = \mathrm{Tr}[\mathrm{e}^{-(H+\lambda Q)/T}]$, $Z = Z(\lambda = 0)$, and $Z_0 = \mathrm{Tr}[\mathrm{e}^{-H_0/T}]$ are the partition functions of the (unprojected) interacting system and noninteracting system (taking $J \to 0$), respectively, and Z_{p} is the physical (projected) partition function. We now calculate (Z/Z_0) by using the relations $Z = \mathrm{Tr}[\mathrm{e}^{-H/T}] = \frac{1}{2}Z_0 + Z_{\mathrm{p}}$ and $Z\,\langle Q\rangle = \frac{1}{2}Z_0 + Z_{\mathrm{p}}$. From these two relations we conclude that $\langle Q\rangle = 1$ and that Y is indeed given by the above relation. One may calculate (Z/Z_0) by integrating the total pseudofermion occupation $\langle Q\rangle_\lambda = Z_\lambda^{-1}\mathrm{Tr}[Q\mathrm{e}^{-(H+\lambda Q)/T}] = -T\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln \mathrm{Tr}[\mathrm{e}^{-(H+\lambda Q)/T}]$ with respect to λ from 0 to ∞. As the spin levels are not occupied in the limit $\lambda \to \infty$ the system is, therefore, noninteracting in this limit. Hence $\lim_{\lambda\to\infty} Z_\lambda = Z_0/4$ (keeping in mind that the trace over the pseudofermion states at $\lambda = 0$ gives a factor of 4), leading to the result

$$Z/Z_0 = \frac{1}{4}\exp[\frac{1}{T}\int_0^\infty \mathrm{d}\lambda\,\langle Q\rangle_\lambda]\,. \tag{3.3}$$

In this setup, the total occupation number may be calculated approximately from $\int_0^\infty \mathrm{d}\lambda\,\langle Q\rangle_\lambda = \frac{1}{T}\int_0^\infty \mathrm{d}\lambda 2\int \mathrm{d}\omega f(\omega+\lambda)A(\omega) \approx 2\int \mathrm{d}\omega \ln(1+\mathrm{e}^{-\omega/T})A(\omega)$, neglecting the λ-dependence of the spectral function $A(\omega)$. In the case of the spin $1/2$ Kondo model the factor Y increases from $Y = 1$ at low temperatures ($T \ll T_{\mathrm{K}}$, the Kondo temperature) to $Y = 2$ at $T \gg T_{\mathrm{K}}$.

Popov–Fedotov Projection

A different approach allowing for an exact treatment of the constraint even for lattice systems has been proposed by Popov and Fedotov [18]. It amounts to applying a homogeneous, *imaginary-valued* chemical potential $\mu^{\mathrm{ppv}} = -\frac{i\pi T}{2}$, where T is the temperature. Thus, within this scheme, the Hamiltonian H is replaced by

$$H \longrightarrow H^{\mathrm{ppv}} = H - \mu^{\mathrm{ppv}}\sum_i Q_i\,. \tag{3.4}$$

Note that H denotes a given spin $1/2$ Hamiltonian using the fermionic representation of spin operators. Given a physical operator $\mathcal{O}$ (i.e., an arbitrary sum or product of spin operators) it can be shown[19] that the expectation value $\langle\mathcal{O}\rangle^{\text{ppv}}$, calculated with H^{ppv} and the *entire* Hilbert space, is identical to the physical expectation value $\langle\mathcal{O}\rangle$, where the average is performed with the original Hamiltonian H. The projection works by virtue of a mutual cancellation of the unphysical contributions of the sectors $Q_i = 0$ and $Q_i = 2$, at each site. It should be emphasized that, although the Hamiltonian H^{ppv} is no longer hermitian, the quantity $\langle\mathcal{O}\rangle^{\text{ppv}}$ comes out real-valued. If, on the other hand, $\mathcal{O}$ is unphysical in the sense that it is nonzero in the unphysical sector, e.g., the operator $\mathcal{O} = Q_i$, the expectation value $\langle Q_i\rangle^{\text{ppv}}$ is meaningless and one has $\langle Q_i\rangle \neq \langle Q_i\rangle^{\text{ppv}}$.

This approach is applicable to spin models [19–22], but unfortunately it cannot be extended to cases away from half-filling. Although μ^{ppv} vanishes in the limit $T \to 0$, in principle the exact projection with $\mu = \mu^{\text{ppv}}$ and the average projection with $\mu = 0$ are not equivalent at $T = 0$. This is due to the fact that the computation of an average $\langle\ldots\rangle^{\text{ppv}}$ does not necessarily commute with the limit $T \to 0$. Nevertheless it can be expected that in usual quantum spin models both projection schemes are equivalent at $T = 0$. This can be understood with the following argument: Starting from the physical ("true") ground state, a fluctuation of one fermion charge results in two sites with unphysical occupation numbers, one with no and one with two fermions. Since these sites carry spin zero the sector of the Hamiltonian with that occupation is identical to the physical Hamiltonian where the two sites are effectively *missing*. Thus, a fluctuation from the ground state into this sector costs the binding energy of the two sites which is of the order of the exchange coupling, even in the case of strong frustration. Consequently, at $T = 0$ charge fluctuations are not allowed and it is sufficient to use the simpler average projection with $\mu = 0$.

3.2.2 Electron Operators

The Hilbert space of electrons in a local orbital is spanned by four states: two with single occupancy (representing a local spin $1/2$) and the empty and doubly occupied states. Obviously, the singly occupied states have fermionic character, while the remaining two states have bosonic character. One may now envisage to create these states out of a vacuum state $|\text{vac}\rangle$, which is defined by the absence of any of the four occupation number states. These four states may then be created by fermionic or bosonic auxiliary operators. This may be done in a multitude of ways. We will concentrate here on the representations introduced by Barnes [9] and by Kotliar and Ruckenstein [12].

3.2.2.1 Barnes's Representation

The basic idea consists in locally decomposing the electronic excitations into spin and charge components. This can be achieved in many different ways. A suitable

Hubbard–Stratonovich decoupling of the interaction term could reach this goal, but would likely be limited to weak interaction. Instead, in the pioneering Barnes approach [9] the spin and charge degrees of freedom (DoF) are represented by fermionic and bosonic operators, respectively. Being more numerous than the original (physical) operators, the auxiliary operators span a Fock space that is larger than the physical one. Consequently they need to fulfill an appropriate set of constraints for such a representation to be faithful. Specifically Barnes considered the single impurity Anderson model (SIAM):

$$H = \sum_{\boldsymbol{k}\sigma} \varepsilon_{\boldsymbol{k}} c^{\dagger}_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \varepsilon_f \sum_{\sigma} a^{\dagger}_{\sigma} a_{\sigma} + V \sum_{\boldsymbol{k}\sigma} \left(c^{\dagger}_{\boldsymbol{k}\sigma} a_{\sigma} + a^{\dagger}_{\sigma} c_{\boldsymbol{k}\sigma} \right) + U a^{\dagger}_{\uparrow} a_{\uparrow} a^{\dagger}_{\downarrow} a_{\downarrow} . \quad (3.5)$$

Clearly, this problem may not be treated by means of perturbation theory in U, especially in the $U \to \infty$ limit. Instead, Barnes introduced the auxiliary fermionic (f_{σ}) and bosonic (e, d) operators in terms of which the physical electron operators a_{σ} read

$$a_{\sigma} = e^{\dagger} f_{\sigma} + \sigma f^{\dagger}_{-\sigma} d. \quad (3.6)$$

The a_{σ}-operators obey the ordinary Fermion anticommutation relations. This property is not automatically preserved when using the representation (3.6), even when the fermionic and bosonic auxiliary operators obey canonical commutation relations. In addition the constraint

$$Q \equiv e^{\dagger} e + \sum_{\sigma} f^{\dagger}_{\sigma} f_{\sigma} + d^{\dagger} d = 1 \quad (3.7)$$

must be satisfied. Equation (3.6) together with (3.7) constitutes a faithful representation of the physical electron operator in the sense that both have the same matrix elements in the physical Hilbert subspace with $Q = 1$. The above representation has been widely used, in particular, in the $U \to \infty$ limit where the operator d (linked to double occupancy) drops out. The constraint can be implemented by means of a functional integral representation. For example, for $U \to \infty$ the partition function, projected onto the $Q = 1$ subspace, reads

$$Z = \int_{-\pi/\beta}^{\pi/\beta} \frac{\beta d\lambda}{2\pi} \mathrm{e}^{\mathrm{i}\beta\lambda} \int \prod_{\sigma} D[f_{\sigma}, f^{\dagger}_{\sigma}] \int \prod_{\boldsymbol{k}\sigma} D[c_{\boldsymbol{k}\sigma}, c^{\dagger}_{\boldsymbol{k}\sigma}] \int D[e, e^{\dagger}] \mathrm{e}^{-\int_0^{\beta} \mathrm{d}\tau (\mathcal{L}_f(\tau) + \mathcal{L}_b(\tau))} \quad (3.8)$$

with the fermionic and bosonic Lagrangians

$$\mathcal{L}_f(\tau) = \sum_{\boldsymbol{k}\sigma} c^\dagger_{\boldsymbol{k}\sigma}(\tau)(\partial_\tau + \varepsilon_{\boldsymbol{k}} - \mu) c_{\boldsymbol{k}\sigma}(\tau) + \sum_\sigma f^\dagger_\sigma(\tau)(\partial_\tau + \varepsilon_c - \mu + i\lambda) f_\sigma(\tau)$$
$$+ V \sum_{\boldsymbol{k}\sigma} \left(c^\dagger_{\boldsymbol{k}\sigma}(\tau) f_\sigma(\tau) e^\dagger(\tau) + h.c. \right)$$
$$\mathcal{L}_b(\tau) = e^\dagger(\tau)(\partial_\tau + i\lambda) e(\tau)\,. \tag{3.9}$$

Here the λ integration enforces the constraint, and the Lagrangian is bilinear in the fermionic fields. Remarkably, this has been achieved without decoupling the interaction term. Besides, the correctness of the representation can be verified by carrying out all integrals in, e.g., the $V \to 0$ limit. By virtue of the substitution $z = \mathrm{e}^{-\mathrm{i}\beta\lambda}$, $\beta \mathrm{d}\lambda = i\,\mathrm{d}z/z$, the λ integral in (3.8) is transformed into a contour integral along the complex unit circle. Observing that this substitution implies exactly a second-order pole at $z = 0$ (i.e., at $i\lambda \to +\infty$, real), it is seen that the projection of Z amounts to calculating the grand canonical Q expectation value in the limit of infinite, real chemical potential, $Z = \lim_{i\lambda\to\infty} \langle Q \rangle_{i\lambda}$ [23], equivalent to the Abrikosov projection. Equation (3.8) may also be viewed as the projection of the non-interacting partition function onto the "$U = \infty$"-subspace. Indeed, (3.8) may be rewritten as

$$Z = P \prod_\sigma \det\left[S_\sigma[e(\tau), \lambda]\right] \tag{3.10}$$

with $\det[S_\sigma[e(\tau), \lambda]]$ the fermionic determinant for one spin species involving an effective time-dependent hybridization ($Ve^\dagger(\tau)$), and the projection operator

$$P = \int_{-\pi/\beta}^{\pi/\beta} \frac{\beta d\lambda}{2\pi} \mathrm{e}^{\mathrm{i}\beta\lambda} \int D[e, e^\dagger] e^{-\int_0^\beta \mathrm{d}\tau \mathcal{L}_b(\tau)}\,. \tag{3.11}$$

Yet, there is an asymmetry in the representation of spin and charge DoF. While the latter can be expressed in terms of bosons, this is not the case of the former, and may cause unnecessary errors in any approximate treatment (for details see [14]).

With this motivation Kotliar and Ruckenstein introduced a representation where spin and charge DoF may be expressed by bosons.

3.2.2.2 Kotliar and Ruckenstein Representation

In the Kotliar and Ruckenstein (KR) representation two additional Bose operators linked to the spin DoF are introduced, $p_\downarrow$ and $p_\uparrow$ [12]. In this approach the physical electron operators are represented as:

$$a_\sigma = \tilde{z}_\sigma f_\sigma \quad \text{with} \quad \tilde{z}_\sigma = e^\dagger p_\sigma + p^\dagger_{-\sigma} d, \tag{3.12}$$

where the first term corresponds to the transition from the singly occupied state to the empty one, and the second term to the transition from the doubly occupied state to the singly occupied one. Again the representation is faithful provided the auxiliary operators obey canonical commutation relations and satisfy constraints. They read

$$e^{\dagger}e + \sum_{\sigma} p_{\sigma}^{\dagger} p_{\sigma} + d^{\dagger}d = 1 \tag{3.13}$$

$$p_{\sigma}^{\dagger} p_{\sigma} + d^{\dagger}d = f_{\sigma}^{\dagger} f_{\sigma} \qquad \sigma = \uparrow, \downarrow . \tag{3.14}$$

They may be enforced in a functional integral representation with Lagrange multipliers in a fashion analogous to the one we encountered with the Barnes representation. Besides, the density operator ($\sum_{\sigma} p_{\sigma}^{\dagger} p_{\sigma} + 2d^{\dagger}d$) and the z-component of the spin operator ($\frac{1}{2}\sum_{\sigma=\pm} \sigma p_{\sigma}^{\dagger} p_{\sigma}$) may be expressed in terms of bosons. Spin and charge DoF may, therefore, be treated on equal footing. This procedure can be extended to multiband models [15].

3.2.2.3 Spin-Rotation Invariant Representation

Though faithful, the Kotliar and Ruckenstein representation is lacking spin rotational invariance as transverse components of the spin operator may not be simply represented in terms of auxiliary operators. Indeed, $S^{x,y}$ is neither related to $\frac{1}{2}\sum_{\sigma\sigma'} f_{\sigma}^{\dagger} \tau_{\sigma\sigma'}^{x(y)} f_{\sigma'}$ nor to $\frac{1}{2}\sum_{\sigma\sigma'} p_{\sigma}^{\dagger} \tau_{\sigma\sigma'}^{x(y)} p_{\sigma'}$. Hence fluctuations associated to the transverse modes are not treated on the same footing as the ones associated to the longitudinal mode. With this motivation a manifestly spin-rotation invariant (SRI) formulation has been introduced [13, 14]. In this setup the doublet p_{σ} [12] is replaced by a scalar ($S = 0$) field p_0 and a vector ($S = 1$) field $\boldsymbol{p} = (p_x, p_y, p_z)$, in terms of which the state $|\sigma\rangle = a_{\sigma}^{\dagger}|0\rangle$ may be represented as

$$|\sigma\rangle = \sum_{\sigma'} p_{\sigma\sigma'}^{\dagger} f_{\sigma'}^{\dagger} |\mathrm{vac}\rangle \quad \text{with} \quad p_{\sigma\sigma'}^{\dagger} = \frac{1}{2} \sum_{\mu=0,x,y,z} p_{\mu}^{\dagger} \tau_{\sigma\sigma'}^{\mu} . \tag{3.15}$$

The bosons p_{μ} obey canonical commutation relations, and all the auxiliary operators annihilate the vacuum ($f_{\sigma}|\mathrm{vac}\rangle = e|\mathrm{vac}\rangle = \ldots|\mathrm{vac}\rangle = 0$). With this at hand the electron operators may be written as

$$a_{\sigma} = \sum_{\sigma'} f_{\sigma'} \tilde{z}_{\sigma'\sigma} \quad \text{with} \quad \tilde{z}_{\sigma'\sigma} = e^{\dagger} p_{\sigma'\sigma} + \sigma'\sigma p_{-\sigma,-\sigma'}^{\dagger} d . \tag{3.16}$$

Again, the auxiliary operators need to satisfy constraints. They read

$$e^\dagger e + \sum_\mu p_\mu^\dagger p_\mu + d^\dagger d = 1 \tag{3.17}$$

$$\sum_\sigma f_\sigma^\dagger f_\sigma = \sum_\mu p_\mu^\dagger p_\mu + d^\dagger d \tag{3.18}$$

$$\sum_{\sigma,\sigma'} f_{\sigma'}^\dagger \boldsymbol{\tau}_{\sigma\sigma'} f_\sigma = p_0^\dagger \boldsymbol{p} + \boldsymbol{p}^\dagger p_0 - i\, \boldsymbol{p}^\dagger \times \boldsymbol{p}\,. \tag{3.19}$$

While the density operator ($n = \sum_\mu p_\mu^\dagger p_\mu + 2d^\dagger d$) and the density of doubly occupied sites operator ($D = d^\dagger d$) may be expressed in terms of bosons or fermions, the spin operator reads

$$\boldsymbol{S} = \sum_{\sigma\sigma'\sigma_1} \boldsymbol{\tau}_{\sigma\sigma'} p_{\sigma\sigma_1}^\dagger p_{\sigma_1\sigma'}\,. \tag{3.20}$$

This expression is especially useful in the context of the t–J model, in particular, because the spin DoF need not be expressed in terms of the original fermions. Using the above, one can tackle models of correlated electrons such as the SIAM, the Anderson lattice model, the t–J or the Hubbard model. However, while the spin and charge DoF have been mapped onto bosons, anomalous propagators necessarily vanish on a saddle-point level as the Lagrangian is bilinear in the fermionic fields, independent of the model. Here they are not treated on equal footing with the spin and charge DoF. This motivated two of us to introduce a manifestly spin- charge-rotation-invariant (SCRI) formulation [14].

3.2.2.4 Spin- and Charge-Rotation-Invariant Formulation

The SCRI representation is motivated both by the need to be able to account for anomalous expectation values (such as the ones arising when investigating excitonic states) and to satisfy the particle–hole rotational symmetry entailed in many models. The generators of these rotations are given by the components of the operators:

$$\boldsymbol{J} = \frac{1}{2}\left(a_\uparrow^\dagger, a_\downarrow\right) \boldsymbol{\tau} \begin{pmatrix} a_\uparrow \\ a_\downarrow^\dagger \end{pmatrix}, \tag{3.21}$$

which form a spin algebra with the usual commutation relations. One may then replace the doublet e, d by a scalar (vector) b_0 ($\boldsymbol{b}$) field (with respect to rotations in the particle–hole space), all of them satisfying canonical commutation relations. In terms of them the two local occupation number states $|2\rangle \equiv |+\rangle$ and $|0\rangle \equiv |-\rangle$ may be represented as

$$|\rho\rangle = \sum_{\rho'} b_{\rho\rho'}^\dagger \psi_{\rho'}^\dagger |\mathrm{vac}\rangle; \quad \rho = \pm \tag{3.22}$$

with $\psi_+^\dagger = f_\uparrow^\dagger f_\downarrow^\dagger$ and $\psi_-^\dagger = 1$. When considering the generalized $\tilde{z}$-operator $\tilde{z}_{\rho\sigma,\rho'\sigma'} = \rho\rho' b^\dagger_{-\rho,-\rho'} p_{\sigma'\sigma} + \sigma\sigma' p^\dagger_{-\sigma,-\sigma'} b_{\rho'\rho}$ (with $b^\dagger_{\rho\rho'} = \frac{1}{2}\sum_{\mu=0,x,y,z} b^\dagger_\mu \tau^\mu_{\rho\rho'}$) and the matrix operators:

$$A_{\rho,\sigma} = \begin{pmatrix} a_\uparrow & a_\downarrow \\ a_\downarrow^\dagger & -a_\uparrow^\dagger \end{pmatrix} \quad \text{and} \quad F_{\rho,\sigma} = \begin{pmatrix} f_\uparrow & f_\downarrow \\ f_\downarrow^\dagger & -f_\uparrow^\dagger \end{pmatrix} \tag{3.23}$$

one may write the physical electron operator as

$$A_{\rho\sigma} = \sum_{\sigma',\rho'} \tilde{z}_{\rho\sigma,\rho'\sigma'} F_{\rho'\sigma'} \,. \tag{3.24}$$

The constraints now read

$$\begin{aligned} f^\dagger_{\sigma'} f_\sigma &= 2\sum_{\sigma_1} p^\dagger_{\sigma_1\sigma} p_{\sigma'\sigma_1} + \frac{1}{2}\delta_{\sigma,\sigma'} \sum_{\mu=0,x,y,z} b^\dagger_\mu b_\mu \\ \sum_\sigma F^\dagger_{\rho\sigma} F_{\rho'\sigma} &= 2\sum_{\rho_1} b^\dagger_{\rho_1\rho} b_{\rho'\rho_1} + \frac{1}{2}\delta_{\rho,\rho'} \sum_{\mu=0,x,y,z} p^\dagger_\mu p_\mu \,. \end{aligned} \tag{3.25}$$

In particular, when performing the trace of (3.25), one obtains

$$\sum_\sigma f^\dagger_\sigma f_\sigma = 1; \qquad \sum_{\mu=0,x,y,z} \left(p^\dagger_\mu p_\mu + b^\dagger_\mu b_\mu \right) = 1 \,. \tag{3.26}$$

Therefore, both spin and charge DoF no longer possess a representation in terms of the auxiliary fermions. Instead, correcting (48) in [14], the density operator reads

$$n = \sum_\mu p^\dagger_\mu p_\mu + 2D \tag{3.27}$$

with the density of doubly occupied sites

$$D = \frac{1}{2}\sum_{\mu\mu'} b^\dagger_\mu b_{\mu'} \mathrm{Tr}\left[\left(\tau^0 + \tau^z\right) \tau^\mu \tau^{\mu'} \right] . \tag{3.28}$$

The spin operator is still given by (3.20). The SCRI representation of the Hubbard model is thus obtained using (3.24) and (3.28), together with the constraints (3.25).

3.2.2.5 Gauge Symmetry and Radial Slave Boson Fields

When representing the electron operators a_σ as $\tilde{z}_\sigma f_\sigma$, it is immediately clear that the latter expression is invariant under the group of transformations:

$$f_\sigma(\tau) \longrightarrow f_\sigma(\tau)\mathrm{e}^{\mathrm{i}\phi(\tau)} \quad \text{and} \quad \tilde{z}_\sigma(\tau) \longrightarrow \tilde{z}_\sigma(\tau)\mathrm{e}^{-\mathrm{i}\phi(\tau)}\,. \tag{3.29}$$

This local $U(1)$ gauge symmetry was first realized by Read and Newns [24–26] in the context of the $U \to \infty$ Barnes representation for the SIAM (with $\tilde{z}_\sigma = e^\dagger$). In that case this can be made use of to gauge away the phase of the slave boson, which remains as a purely radial field, while the constraint Lagrange parameter is promoted to a time-dependent field. Yet, standard textbooks do not mention representations of such radial fields that are set up on a discretized time mesh from the beginning.

A scheme specific to radial slave boson fields has been proposed by one of us [27]. In this scheme the partition function takes a form analogous to (3.8). However, the projection operator does not mix the N time steps, and may be written as

$$\begin{aligned} P &= \lim_{N\to\infty}\lim_{W\to\infty}\prod_{n=1}^{N} P_n\,, \quad \text{with} \\ P_n &= \int_{-\infty}^{\infty}\frac{\beta}{N}\frac{\mathrm{d}\lambda_n}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}x_n\mathrm{e}^{-\frac{\beta}{N}(i\lambda_n(x_n-1)+Wx_n(x_n-1))}\,. \end{aligned} \tag{3.30}$$

Here the constraint parameter λ_n is defined for each time step n, i.e., it is a time-dependent field, and x represents the radial slave boson field. In the discrete time step form, the fermionic part of the action reads

$$\begin{aligned} S_f = \sum_{n=1}^{N}\Bigg\{ &\sum_{\boldsymbol{k}\sigma} c^\dagger_{\boldsymbol{k},n,\sigma}\left(c_{\boldsymbol{k},n,\sigma} - e^{-\frac{\beta}{N}(\varepsilon_{\boldsymbol{k}}-\mu)}c_{\boldsymbol{k},n-1,\sigma}\right) \\ &+\sum_{\sigma} f^\dagger_{n,\sigma}\left(f_{n,\sigma} - e^{-\frac{\beta}{N}(\varepsilon_f+i\lambda_n-\mu)}f_{n-1,\sigma}\right) \\ &+\frac{\beta}{N}\sum_{\boldsymbol{k}\sigma} Vx_n\left(c^\dagger_{\boldsymbol{k},n,\sigma}f_{n-1,\sigma} + f^\dagger_{n,\sigma}c_{\boldsymbol{k},n-1,\sigma}\right)\Bigg\}\,. \end{aligned} \tag{3.31}$$

The integration over the fermionic fields can be carried out, and the partition function can be obtained by projecting the resulting fermionic determinant:

$$Z = P\prod_{\sigma}\det\left[S_\sigma\left[\{x_n\},\{\lambda_n\}\right]\right] \tag{3.32}$$

with the above projection operator (3.30). The expectation value of the hole density operator takes the simple form:

$$\langle n_h(\tau_m)\rangle = \langle x_m\rangle = \frac{1}{Z}P\left\{x_m\prod_{\sigma}\det\left[S_\sigma\left[\{x_n\},\{\lambda_n\}\right]\right]\right\}\,. \tag{3.33}$$

It is easily seen to be time-independent. In contrast to a Bose condensate $\langle x_m\rangle$ is generically finite, and may only vanish for zero hole concentration [28]. Another

specific feature of the radial slave boson field is that, for any power $a > 0$, one finds $\langle x_m^a \rangle = \langle x_m \rangle$, as the corresponding projections of the fermionic determinant all yield the same value.

The hole autocorrelation function may also be expressed as a projection of the fermionic determinant. It reads

$$\langle n_h(\tau_n) n_h(\tau_m) \rangle = \langle x_n x_m \rangle = \frac{1}{Z} P \left\{ x_n x_m \prod_{\sigma} \det \left[S_\sigma \left[\{x_n\}, \{\lambda_n\} \right] \right] \right\} . \tag{3.34}$$

Regarding the Kotliar and Ruckenstein representation the determination of the gauge symmetry group has been debated over several years [14, 29–34]. It was finally agreed that it reads $U(1) \times U(1) \times U(1)$. By promoting all three constraint parameters to fields one may gauge away the phase of three bosonic fields, the fourth one, for example, d, remaining complex. Therefore, in the $U \to \infty$ limit ($d \to 0$), the three remaining bosonic fields are radial slave boson fields. In functional integral language they may be handled in the same fashion as the above x-field.

3.3 Mean-Field Approximations

An economical and often physically reasonable way to determine observable quantities in the SB framework is provided by a saddle-point approximation (SPA) to the functional integral. This is equivalent to allowing for a finite expectation value of a Bose field amplitude. Strictly speaking, a finite expectation value of a Bose field operator violates gauge invariance and should not exist. In contrast, a finite saddle-point amplitude of the radial slave boson fields is compatible with Elitzur's theorem.

3.3.1 Saddle-Point Approximation to the Barnes Representation

In its simplest form the SPA consists of replacing the boson field operators b_i at each lattice site, or b at the impurity site, by the modulus of its expectation value, in accordance with the above. The remaining problem is a noninteracting model, which is easily solved. We will discuss the solution briefly for the Anderson impurity model and the Anderson lattice model.

3.3.1.1 Kondo Effect in the Anderson Impurity Model

In SPA the Anderson impurity Hamiltonian (3.5) takes the form

$$H = \sum_{\boldsymbol{k}\sigma} \varepsilon_{\boldsymbol{k}} c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} + \varepsilon_f \sum_{\sigma} f_\sigma^\dagger f_\sigma + V \sum_{\boldsymbol{k}\sigma} e_0 \left(c_{\boldsymbol{k}\sigma}^\dagger f_\sigma + f_\sigma^\dagger c_{\boldsymbol{k}\sigma} \right) + \lambda (Q - 1), \tag{3.35}$$

where the conserved charge is $Q = \sum_\sigma f_\sigma^\dagger f_\sigma + e_0^2 = 1$ and λ is the corresponding Lagrange multiplier. This is nothing but a resonant level model with renormalized parameters: $\widetilde{\varepsilon}_f = \varepsilon_f + \lambda$ and $\widetilde{V} = Ve_0$. At the stationary point of the free energy one finds that the level position $\widetilde{\varepsilon}_f$ and the level width $\widetilde{\Delta} = e_0^2\Delta = \pi N_0 \widetilde{V}^2$, where $\Delta = \pi N_0 V^2$ ($N_0 = 1/2D$ is the conduction electron DOS at the Fermi level), satisfy the equations

$$\widetilde{\varepsilon}_f = \varepsilon_f - \frac{2\Delta}{\pi} \ln \frac{\sqrt{\widetilde{\varepsilon}_f^2 + \widetilde{\Delta}^2}}{D} \tag{3.36}$$

$$\widetilde{\Delta} = \Delta - \frac{2\Delta}{\pi} \tan^{-1} \frac{\widetilde{\Delta}}{\widetilde{\varepsilon}_f}. \tag{3.37}$$

In the limit of $\Delta \ll |\varepsilon_f|$ the occupation of the local level, $n_f = \frac{2}{\pi} \tan^{-1} \frac{\widetilde{\Delta}}{\varepsilon_f}$, approaches unity, which means that a local moment forms at higher temperature. Below a characteristic temperature, the Kondo temperature T_K, the local moment gets screened by the conduction electron spins, which form a resonance state with the local moment, located close to the Fermi energy, at $\widetilde{\varepsilon}_f$, and of width $\widetilde{\Delta} \approx T_\mathrm{K} = D \exp \frac{-|\varepsilon_f|}{2N_0V^2} = D \exp \frac{-1}{2N_0 J}$, where $J = \frac{V^2}{|\varepsilon_f|}$ is the antiferromagnetic spin exchange coupling constant of the local spin and the local conduction electron spin density. The low-temperature behavior of Kondo systems is reasonably well described by slave boson mean-field theory. At higher temperatures one finds in this approximation a spurious first-order transition to the local moment regime, rather than a continuous crossover.

3.3.1.2 Heavy Fermions in the Anderson Lattice Model

The SB mean-field approximation to the Anderson lattice model in the limit $U \to \infty$ [24–26] leads to the following single-particle Hamiltonian of two hybridized bands

$$H = \sum_{\boldsymbol{k}\sigma} \varepsilon_{\boldsymbol{k}} c_{\boldsymbol{k}\sigma}^\dagger c_{\boldsymbol{k}\sigma} + \varepsilon_f \sum_{i,\sigma} f_{i\sigma}^\dagger f_{i\sigma} + V \sum_{i,\sigma} e_0 \left(c_{i\sigma}^\dagger f_{i\sigma} + f_{i\sigma}^\dagger c_{i\sigma} \right) + \sum_i \lambda_i (Q_i - 1). \tag{3.38}$$

The stationarity condition with respect to the fields λ_i leads to the condition $\langle Q_i \rangle = 1$,which for a translation invariant state is independent of the lattice position $\mathbf{R}_i$. As in the impurity problem the f-level position is shifted by the correlation effect to $\widetilde{\varepsilon}_f = \varepsilon_f + \lambda = \varepsilon_f - 2N_0V^2 \ln \frac{\widetilde{\varepsilon}}{D}$ and the square of the boson amplitude is related to the f-level occupation n_f

$$\widetilde{\varepsilon}_f = \varepsilon_f - 2N_0V^2 \ln \frac{\widetilde{\varepsilon}_f - \varepsilon_\mathrm{F}}{D} \tag{3.39}$$

$$e_0^2 = 1 - n_f = 1 - \frac{2N_0V^2e_0^2}{\widetilde{\varepsilon}_f}. \tag{3.40}$$

Here we have assumed $|\widetilde{\varepsilon}_f| \ll D$. We observe that, provided ε_f is sufficiently below the Fermi level $\varepsilon_{\rm F}$, we have $|\widetilde{\varepsilon}_f| \ll |\varepsilon_f|$ and it follows from (3.39) that $\widetilde{\varepsilon}_f - \varepsilon_{\rm F} = D\exp\frac{-|\varepsilon_f|}{2N_0V^2} = T_{\rm K}$, equal to the single impurity Kondo temperature. In this limit $e_0^2 \approx |\widetilde{\varepsilon}_f|/2N_0V^2 \ll 1$ and the hybridization amplitude is substantially reduced, leading to heavy quasiparticle bands of energy

$$E_{\boldsymbol{k}}^{\pm} = \frac{1}{2}\left[\varepsilon_{\boldsymbol{k}} + \widetilde{\varepsilon}_f \pm \sqrt{(\varepsilon_{\boldsymbol{k}} + \widetilde{\varepsilon}_f)^2 + V^2e_0^2}\right]. \tag{3.41}$$

3.3.2 *Saddle-Point Approximation to the KR Representation*

It is tempting to extend this approach to the Hubbard Model. Yet, at this stage of the formulation, the representation suffers from the following drawback: The noninteracting limit is not properly recovered on the SPA level (see [12] for the case of the Hubbard model), in contrast to more conventional approaches. This can be cured by noticing that there is no unique slave boson representation, but rather infinitely many different ones. They are all equivalent when the functional integral is exactly evaluated, but differ on saddle-point level. Fortunately enough, there is one representation of the kinetic energy which allows us to overcome the above drawback. For the KR representation it consists in replacing the operators $\tilde{z}_\sigma$ in (3.12) by

$$z_\sigma = e^\dagger L_\sigma R_\sigma p_\sigma + p_{-\sigma}^\dagger L_\sigma R_\sigma d \quad \text{with} \tag{3.42}$$

$$L_\sigma = \frac{1}{\sqrt{1 - p_\sigma^\dagger p_\sigma - d^\dagger d}} \quad \text{and} \quad R_\sigma = \frac{1}{\sqrt{1 - p_{-\sigma}^\dagger p_{-\sigma} - e^\dagger e}}, \tag{3.43}$$

and to consistently use $a_\sigma = z_\sigma f_\sigma$ in the representation of the kinetic energy operator. In this form the SPA to the KR representation is equivalent to the Gutzwiller approximation to the Gutzwiller wavefunction [12]. As the latter two are equivalent in the large d limit, the SPA to the KR representation turns variationally controlled in this limit. In addition it turns exact in several large N limits [14], or for particular toy models [35]. These properties are shared by the SRI formulation [14]. Indeed, introducing $\tilde{p}_{\sigma\sigma'} \equiv \sigma\sigma' p_{-\sigma',-\sigma}$ the z-operator reads

$$\underline{z} = e^\dagger \underline{L}\ M\ \underline{R}\ \underline{p}\ + \underline{\tilde{p}}^\dagger \underline{L}\ M\ \underline{R}\ d \tag{3.44}$$

with

$$M = \left[1 + e^\dagger e + \sum_\mu p_\mu^\dagger p_\mu + d^\dagger d\right]^{\frac{1}{2}} \quad \text{and} \tag{3.45}$$

$$\underline{L} = \left[\left(1 - d^\dagger d\right)\underline{1} - 2\underline{p}^\dagger \underline{p}\right]^{-\frac{1}{2}} \quad \text{and} \quad \underline{R} = \left[\left(1 - e^\dagger e\right)\underline{1} - 2\underline{\tilde{p}}^\dagger \underline{\tilde{p}}\right]^{-\frac{1}{2}}. \tag{3.46}$$

Equations (3.44) and (3.46) correct (22) in [14] and (3.46) corrects (3) in [36].

3.3.2.1 Mott–Hubbard Metal–Insulator Transition

The KR and SRI representations have been used to characterize a very broad range of phases of the Hubbard model [37–49]. In addition to the above discussed motivations, the popularity of the approach is fueled by its ability to describe a Mott metal-to-insulator transition as it encompasses the Brinkman–Rice mechanism [50–52] that we describe below. It arises when considering the paramagnetic saddle-point. In the SRI representation it corresponds to set the bosonic fields $\boldsymbol{p}_i(\tau)$ and the constraint fields enforcing (3.19) to zero, and to replace the remaining bosonic and constraint fields by their mean value. The free energy then reads

$$F = -T\sum_{\boldsymbol{k},\sigma} \ln\left(1 + \mathrm{e}^{-\frac{E_{\boldsymbol{k}\sigma}}{T}}\right) + Ud^2 + \alpha\left(e^2 + p_0^2 + d^2 - 1\right) - \beta_0\left(p_0^2 + 2d^2\right). \tag{3.47}$$

Here the Lagrange multiplier α (β_0) enforces the constraint (3.17) and (3.18). The quasiparticle dispersion relation is given by

$$E_{\boldsymbol{k}\sigma} = z_0^2 t_{\boldsymbol{k}} + \beta_0 - \mu \tag{3.48}$$

with

$$z_0 = \frac{1}{\sqrt{2}} \frac{p_0(e+d)}{\sqrt{1 - d^2 - \frac{1}{2}p_0^2}\sqrt{1 - e^2 - \frac{1}{2}p_0^2}}. \tag{3.49}$$

z_0^2 both plays the role of a mass renormalization factor and of a quasiparticle residue. Should it vanish in some parameter range, then a Mott insulating state would be realized. Solving the saddle-point equations at half-filling yields

$$z_0^2 = 1 - \left(\frac{U}{U_0}\right)^2 \tag{3.50}$$

with

$$U_0 = -4 \sum_{\boldsymbol{k},\sigma} t_{\boldsymbol{k}} f_{\mathrm{F}}(E_{\boldsymbol{k},\sigma}). \tag{3.51}$$

Therefore, the quasiparticle residue continuously varies from 1 down to 0 for $U \rightarrow U_0$. There, the quasiparticle mass diverges, its residue vanishes, and a Mott gap opens. Indeed, solving the equation for the chemical potential of the quasiparticles for $U > U_0$ and $n \rightarrow 1$ yields

$$\mu(n) = \frac{U}{2}\left[1 - \frac{1-n}{|1-n|}\sqrt{1-\frac{U_0}{U}}\right]. \tag{3.52}$$

The discontinuity in μ across $n = 1$ indicates a pair of first-order phase transitions from the metallic phase at $n < 1$ (with finite z_0) to the insulating phase at $n = 1$ (with a chemical potential $\mu = \frac{U}{2}$) and back to the metallic phase at $n > 1$ (with finite z_0). This discontinuity vanishes for $U \rightarrow U_0^+$, which is, therefore, a critical point. In the insulating phase the quasiparticle contribution to doubly occupied sites vanishes. This does not imply that the latter is predicted to be zero, but that it purely results from fluctuations, that we address in Sect. 3.4.

3.3.2.2 Magnetic Order in the Anderson Lattice Model

The Anderson lattice model is believed to describe the physics of many transition-metal, rare-earth and actinide compounds, including the so-called heavy fermion compounds. It is one of the archetypical models of correlated electrons on a lattice, consisting of a "light" conduction band hybridized with a strongly correlated narrow f-electron band. Depending on the strength of the on-site Coulomb repulsion on the f-orbital, the hybridization strength and the band filling, the model describes either localized moments interacting via spin exchange interaction (e.g., the RKKY interaction), which usually order at low temperature, or Kondo screened moments and heavy quasiparticles. The competition between these two ground states gives rise to a quantum phase transition [53, 54]. A qualitatively correct description (excluding the critical behavior at the quantum critical point, which requires a different approach) may be obtained within the SRI slave boson mean-field theory. The Hamiltonian of the Anderson lattice model reads

$$\begin{aligned} H = &\sum_{\boldsymbol{k}\sigma} \varepsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \varepsilon_a \sum_{i,\sigma} a^\dagger_{i\sigma} a_{i\sigma} + V \sum_{i,\sigma} \left(c^\dagger_{i\sigma} a_{i\sigma} + a^\dagger_{i\sigma} c_{i\sigma} \right) \\ &+ U \sum_i a^\dagger_{i\uparrow} a_{i\uparrow} a^\dagger_{i\downarrow} a_{i\downarrow}, \end{aligned} \tag{3.53}$$

where $c_{i\sigma} = \sum_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\mathbf{R}_i} c_{\boldsymbol{k}\sigma}$ and $\mathbf{R}_i$ is the lattice vector at site i. In terms of SRI slave boson operators H may be represented as

$$H = \sum_{\boldsymbol{k}\sigma} \varepsilon_{\boldsymbol{k}} c^{\dagger}_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \varepsilon_a \sum_i \left(\sum_\mu p^{\dagger}_{i\mu} p_{i\mu} + 2 d^{\dagger}_i d_i \right)$$
$$+ V \sum_{i,\sigma,\sigma'} \left(c^{\dagger}_{i\sigma} z_{i\sigma'\sigma} f_{i\sigma'} + h.c. \right) + U \sum_i d^{\dagger}_i d_i$$
$$+ \sum_i \{ \alpha_i (Q_b - 1) + \beta_{i0} Q_f + \boldsymbol{\beta}_i \cdot \mathbf{Q}_s \} \tag{3.54}$$

A mean-field approximation to this Hamiltonian describing spiral magnetic states has been considered in [55]. There the nonmagnetic boson mean-field amplitudes e, d, and p_0 and Lagrange parameters α and β_0 have been assumed to be spatially uniform, while the magnetic parameters $\boldsymbol{p}_i$ and $\boldsymbol{\beta}_i$ were taken to have the spatial dependence of a spiral vector field, $\boldsymbol{p}_i = p(\cos\phi_i, \sin\phi_i, 0)$ and $\boldsymbol{\beta}_i = \beta(\cos\phi_i, \sin\phi_i, 0)$ oriented perpendicular to the z-axis in spin space, and $\phi_i = \mathbf{q} \cdot \mathbf{R}_i$. The spatial periodicity characterized by the wave vector $\mathbf{q}$ leads to a coupling of Bloch states at wave vectors $\mathbf{k}$ and $\mathbf{k} + \mathbf{q}$. The energy matrix of the hybridized bands then takes the form

$$\epsilon_{\mathbf{k}} = \begin{pmatrix} \epsilon_{\mathbf{k}} - \mu & V z_+ & 0 & V z_- \\ V z_+ & \epsilon_a + \beta_0 - \mu & V z_- & \beta \\ 0 & V z_- & \epsilon_{\mathbf{k}+\mathbf{q}} - \mu & V z_+ \\ V z_- & \beta & V z_+ & \epsilon_a + \beta_0 - \mu, \end{pmatrix} \tag{3.55}$$

where the weight factors $z_{+,-}$ are defined by

$$z_\pm = \frac{e p_+ + d p_-}{\sqrt{1 - d^2 - p_+^2}\sqrt{1 - e^2 - p_-^2}} \pm [p_+ \Leftrightarrow p_-] \tag{3.56}$$

with $p_{+,-} = (p_0 \pm p)/\sqrt{2}$. The mean-field values are determined by requiring that the free energy

$$F = -T \sum_{\boldsymbol{k}\sigma\alpha} \ln [1 + \mathrm{e}^{\frac{-E_{\mathbf{k}\sigma\alpha}}{T}}] + N_L [U d^2 - \beta_0 (p_0^2 + p^2 + 2d^2) + 2\beta p_0 p] \tag{3.57}$$

be stationary. Here $E_{\mathbf{k}\sigma\alpha}$ are the eigenvalues of the energy matrix $\epsilon_{\mathbf{k}}$.

In Fig. 3.1 the zero-temperature phase diagram is shown in the $(t/U) - \delta$-plane (t is the nearest-neighbor hopping amplitude δ and filling factor of the conduction band). In a wide region a spiral magnetic state is found, with wave vector $\mathbf{q}$ approaching the edge of the Brillouin zone at $\delta = 1$ (antiferromagnetic order). Approaching the limit $\delta = 0$ one finds a ferromagnetic region, followed by another antiferromagnetic state very close to $\delta = 0$. These findings have been confirmed by quantum Monte Carlo simulations [56]. One should keep in mind that the spatial dimension enters only through the energy dispersion of the conduction

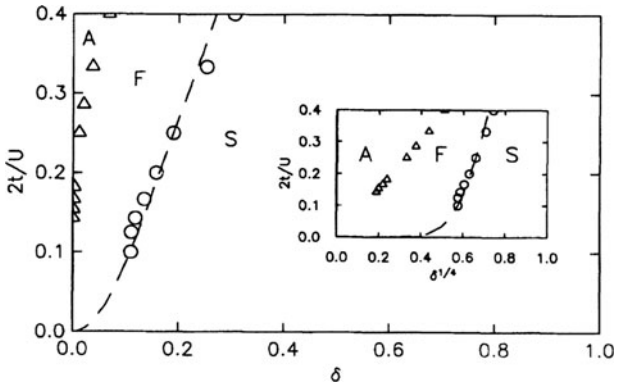

Fig. 3.1 Phase diagram in the $2t/U$ vs. δ plane at $T = 0$. Spiral (S), ferromagnetic (F), and antiferromagnetic (A) regions are indicated. The inset shows the behavior near $\delta = 0$

electrons. These results are, therefore, applicable in three or higher dimensions, where fluctuation effects are expected to be small.

3.4 Fluctuation Corrections to the Saddle-Point Approximation: SRI Representation of the Hubbard Model

Having mapped all DoF onto bosons allows us to directly evaluate the spin and charge response functions. Indeed, the spin and density fluctuations may be expressed as

$$\sum_{\sigma} \sigma \delta n_{\sigma} = \delta(p_0^{\dagger} p_3 + p_3^{\dagger} p_0) \equiv \delta S$$
$$\sum_{\sigma} \delta n_{\sigma} = \delta(d^{\dagger} d - e^{\dagger} e) \equiv \delta N \,, \tag{3.58}$$

in the SRI representation. The spin and charge autocorrelation functions can be written in terms of the slave boson correlation functions as

$$\chi_{\mathrm{s}}(k) = \sum_{\sigma,\sigma'} \sigma\sigma' \langle \delta n_{\sigma}(-k) \delta n_{\sigma'}(k) \rangle = \langle \delta S(-k) \delta S(k) \rangle$$
$$\chi_{\mathrm{c}}(k) = \sum_{\sigma\sigma'} \langle \delta n_{\sigma}(-k) \delta n_{\sigma'}(k) \rangle = \langle \delta N(-k) \delta N(k) \rangle. \tag{3.59}$$

Performing the calculation to one-loop order, one can make use of the propagators given in the Appendix of [36] to obtain:

$$\chi_s(k) = 2p_0^2 S_{77}^{-1}(k)$$
$$\chi_c(k) = 2e^2 S_{11} S^{-1}(k) - 4ed S_{12}^{-1}(k) + 2d^2 S_{22}^{-1}(k)\,. \tag{3.60}$$

As emphasized and analyzed by several authors, see for example, [57, 58], a Fermi liquid behavior is obtained when considering the above $\chi_s(k)$ and $\chi_c(k)$ in the long wavelength and low-frequency limit. However, in contrast to the conventional RPA results, the obtained Landau parameters involve effective interactions, which differ in the spin channel and in the charge channel. Performing the algebra at half-filling yields

$$F_0^a = -1 + \frac{1}{(1 + U/U_0)^2}$$
$$F_0^s = \frac{U(2U_0 - U)}{(U_0 - U)^2} \tag{3.61}$$

fulfilling the property $F_0^s(U) = F_0^a(-U)$ that can be derived on a more general ground [51, 52]. As can be seen in (3.61) F_0^a remains finite when reaching the Mott transition, while F_0^s diverges (for a recent manifestation of a related behavior see [59]).

Ferromagnetic instabilities and ferromagnetic phases have been investigated, too. In particular, in the limit $U \to \infty$, it could be shown analytically that the fully polarized ferromagnetic ground state and the paramagnetic ground state are degenerate at density $n = 2/3$, for any bipartite lattice [37]. For lower densities the ground state is paramagnetic.

Yet, in such an analysis, focus is put on a ferromagnetic instability only, while other commensurate or even incommensurate instabilities should be considered as well. This analysis has been carried out by two of us for the Hubbard model on the square lattice [60]. Off half-filling it turned out that the leading instabilities are systematically towards incommensurate states characterized by a wavevector (Q_x, π) for $U < 57t$ with Q_x smoothly varying from π for $U = 0^+$ down to 0 for $U = 57t$. For larger U the wavevector characterizing the instability is rather of the form $(0, Q_y)$, with $Q_y \simeq \pi$.

Charge instabilities have been looked for as well, in particular through the computation of the charge structure factor [36], though with a negative result even for the t–t'–U repulsive Hubbard model [61]. Instead, the charge structure factor quite systematically consists of one broad peak centered at (π, π). As an example we compare in Fig. 3.2 the slave boson result with quantum Monte Carlo simulations by Dzierzawa [62], for $U = 4t$ and $\delta = 0.275$ at a temperature $T = t/6$. The agreement between both approaches is excellent, as the difference does not exceed a few percent.

3.4.1 Magnetic and Stripe Phases

Having established that the leading instabilities of the paramagnetic phase are generically towards incommensurate phases, spiral and stripe phases have been

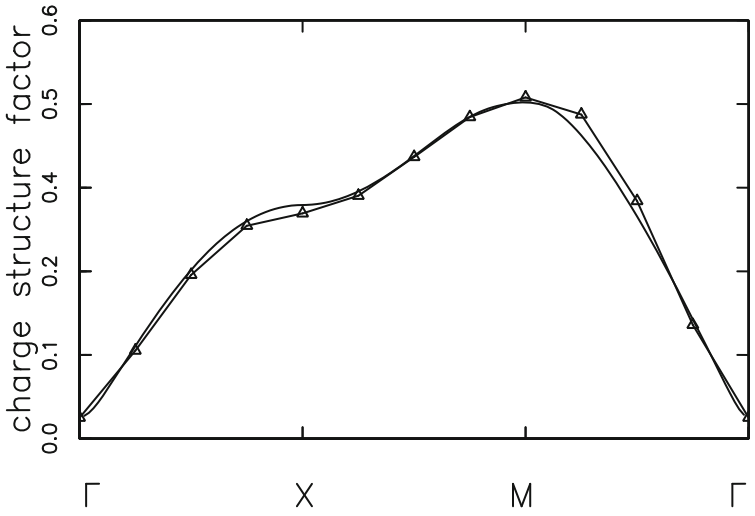

Fig. 3.2 Comparison of the quantum Monte Carlo (*triangles*) and slave boson (*full line*) charge structure factors for $U = 4t$, $\delta = 0.275$, and $\beta = 6$

thoroughly investigated [37–43, 49]. Comparison of ground state energies in spiral phases with numerical simulations showed a very good agreement [38, 40]. Regarding the pure Hubbard model it has been obtained that magnetic stripe phases are generically more stable than spiral phases. However, for the t–t'–U repulsive Hubbard model, the situation is more intricate. As shown in Fig. 3.3 for an intermediate value of t', a large number of phases compete. While the vertical site-centered stripe phases are generically lower in energy than the vertical bond-centered stripe phases at a low doping x, the opposite result is found at a larger x. For instance, for $U = 12t$, the transition occurs at $x \simeq 0.16$ for $t' = -0.15t$, and at $x \simeq 0.18$ for $t' = -0.3t$. Yet, in the latter case, the diagonal spiral phase is lower in energy for $x \geq 0.09$, in contrast to the former case [43].

3.5 Conserving Self-Consistent Approximations

3.5.1 General Properties

The KR representation and its SRI and spin–charge symmetric extensions have been remarkably successful in identifying and describing zero-temperature phases of correlated electron models within saddle-point and Gaussian approximations. However, in order to describe dynamical properties, like spectra or nonequilibrium transport, a more accurate treatment of the excitations is needed. This is especially important for understanding Fermi liquid or non-Fermi liquid signatures in these

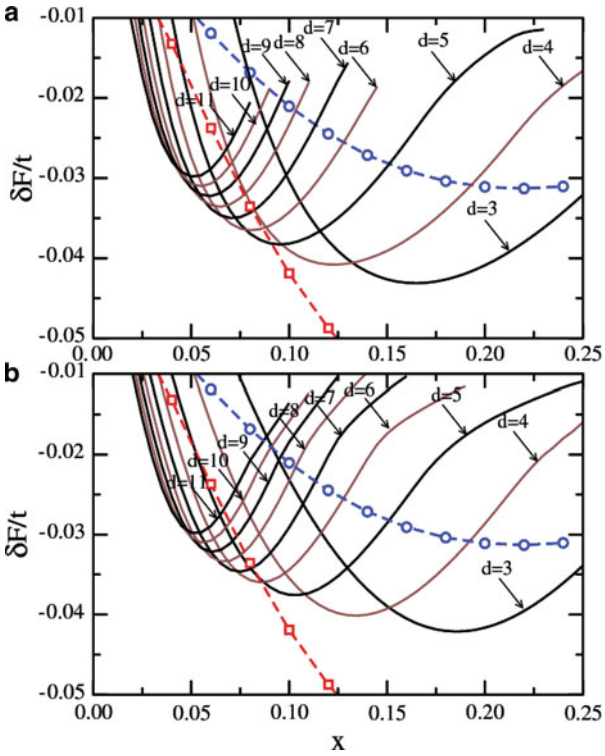

Fig. 3.3 Free energy gain δF per site with respect to the AF phase as a function of doping x, obtained for the t–t'–U Hubbard model with $U = 12t$ and $t' = -0.3t$ for: (**a**) vertical site-centered stripe phases and (**b**) vertical bond-centered stripe phases. Domain walls are separated by $d = 3, \ldots, 11$ lattice constants. *Circles* and *squares* show the corresponding data for vertical and diagonal spiral order, respectively

quantities. It requires, in particular, avoiding the spurious condensation transition of the auxiliary boson fields at a finite temperature or energy. That is, the auxiliary fields must be treated as pure fluctuation fields, preserving the local $U(1)$ gauge symmetries associated with the conservation of the local constraint charges, (3.7), (3.13), (3.14), or (3.17)–(3.19) for the Barnes, the KR, and the SRI representation, respectively. Gauge symmetric theories may be systematically constructed by means of conserving approximations [63,64], where all irreducible n-point vertex functions are derived from a generating functional Φ by means of appropriate functional derivatives. This implies that these quantities are calculated self-consistently, i.e., as functionals of the fully renormalized auxiliary particle propagators. Physical expectation values, calculated at first in the grand canonical ensemble of the auxiliary particle occupation numbers utilizing Wick's theorem, are projected onto the physical sector of Fock space at the end of the calculation, using the techniques described in Sect. 3.2.2. In order to capture the correct low-energy properties of a

given correlated electron model, the approximation for the generating functional must still be chosen appropriately. The conserving scheme has been extensively pursued within the Barnes representation, in order to keep the number of fluctuating auxiliary fields minimal.

3.5.2 Exact Infrared Properties of Pseudoparticle Propagators

The Fermi or non-Fermi liquid behavior is governed by the infrared asymptotics of the slave boson and pseudofermion propagators, $G_e(\omega)$, $G_d(\omega)$, and $G_{f\sigma}(\omega)$. By definition, the field operators appearing in these propagators create an impurity site which is initially empty, doubly occupied or singly occupied with spin $\sigma = \pm 1/2$, respectively, and which evolves in time to a final occupation number $n_a = \sum_\sigma n_{a\sigma}$. This process constitutes an orthogonality catastrophy [65] with characteristic infrared powerlaw singularities of the auxiliary particle spectral functions. In the case of a Fermi liquid, the infrared powerlaw exponents of the pseudofermion, empty and doubly occupied slave boson spectral functions, α_f, α_e, and α_d, respectively, are determined by the single-particle scattering phase shifts at the Fermi level, which may in turn be expressed in terms of the impurity occupation number n_a via the Friedel sum rule. The dependence of the infrared exponents on n_a is, therefore, an indicator of Fermi liquid or non-Fermi liquid behavior in a given system. For the single-channel Anderson model, (3.5), which has a Fermi liquid strong-coupling fixed point, one obtains in the absence of a magnetic field [66–68],

$$\alpha_{f\sigma} = n_a - n_a^2/2 \tag{3.62}$$

$$\alpha_e = 1 - n_a^2/2 \tag{3.63}$$

$$\alpha_d = -1 + 2n_a - n_a^2/2\,. \tag{3.64}$$

These expressions have been confirmed by direct numerical renormalization group calculations [69] and should be recovered by any approximation that is to describe a Fermi liquid fixed point.

3.5.3 Fock Space Projection in Saddle-Point Approximation

While the projection onto the physical Fock space with local charge $Q = 1$ (3.8) may easily be performed exactly for a single correlated site, it becomes cumbersome already for two sites, let alone for a lattice of correlated electrons. This is because by the exact projection the partition function, for instance, is transformed into the expectation value of the *product* of all the local charges Q_j on the correlated lattice sites j (compare the discussion after (3.9)), i.e., it becomes an N_{L}-point correlation function, where N_{L} is the number of correlated sites.

However, gauge symmetric, conserving approximations, which avoid any spurious condensate amplitudes and, hence, preserve the infrared properties, can still be

constructed when the Fock space projection is done in an approximate way. In this approach, first proposed in [23], the λ-integration of (3.8) is done in SPA, while all auxiliary particle Green's functions are derived as pure fluctuation propagators from a generating Luttinger–Ward functional. The λ saddle point can be shown to be equivalent to fixing $i\lambda$ as a real chemical potential for the thermodynamic average of the local charge, $\langle Q \rangle$. Spurious slave boson condensation is avoided by the fact that the fluctuation part of the Bose propagator acquires finite, *negative* spectral weight $A_{e,d}(\omega) \leq 0$ for negative frequencies, $\omega < 0$, such that the occupation number density, $b(\omega)\, A_{e,d}(\omega)$, remains nonnegative for all ω [23], with $b(\omega) = 1/(\mathrm{e}^{\beta\omega} - 1)$ the Bose–Einstein distribution. The structure of the self-consistent integral equations for the pseudoparticle propagators is not altered by this approximation. In fact, it may be shown explicitly and in a straightforward way along the lines of [70, 71] that the λ saddle-point projection preserves the infrared exponents on the level of the simplest conserving approximation, the NCA (to be discussed in the next section). It may be conjectured that this remains true also for more sophisticated conserving approximations. Since the λ SPA involves only the thermal average $\langle Q \rangle$, it is straightforwardly generalized to lattice problems, with a spatially homogeneous chemical potential $i\lambda \in \mathbb{R}$. The method has been applied to the Heisenberg lattice in pseudofermion representation in [72].

3.5.4 Noncrossing Approximation (NCA)

The rest of this chapter is concentrated on the Anderson impurity model.

3.5.4.1 Anderson Impurity Model for $U \to \infty$: NCA

The simplest conserving approximation in the limit $U \to \infty$ is obtained by choosing the generating functional to lowest, i.e., second order in the hybridization V, as shown as the first diagram of Fig. 3.4. Since the self-energies generated in this approximation do not contain any crossings of lines, it has been termed NCA. The NCA has been pioneered by Keiter and Kimball using the resolvent operator formalism [73, 74] and by Kuramoto, who first recognized the conserving nature of the NCA [75]. First numerical evaluations were performed by Kojima et al. [76–78] and by Bickers et al. [79, 80]. For an efficient and numerically stable algorithm for solving pseudoparticle integral equations, like the NCA and its extensions, see [81].

For $U \to \infty$ the NCA captures correctly the Kondo energy scale, and it provides a qualitative description of the formation of the Kondo resonance. It happens to describe also correctly the powerlaw dependence of physical properties at the non-Fermi liquid fixed point of the two-channel Anderson or Kondo impurity model [82]. For these reasons and for its technical simplicity the NCA has been applied to a wide variety of problems, including, as an impurity solver for dynamical mean-field theory (DMFT), to the t–J model [83, 84] and the non-Fermi liquid two-channel

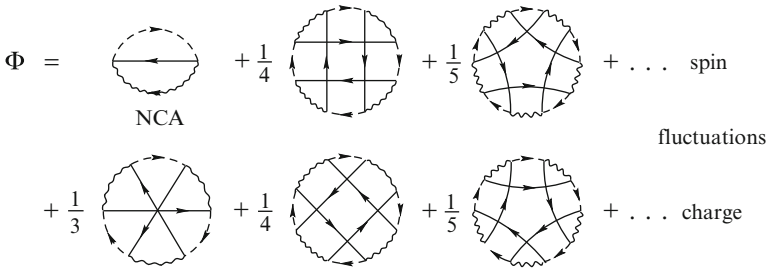

Fig. 3.4 Diagrammatic representation of the Luttinger–Ward functional generating the CTMA for $U \to \infty$. The first diagram constitutes the NCA. The two-loop diagram is excluded, since it is not a skeleton. *Solid*, *dashed*, and *wiggly lines* represent conduction electron, renormalized pseudofermion and auxiliary boson propagators, respectively. The terms with the conduction electron lines running clockwise generate the conduction electron-pseudofermion ladder vertex $T^{(cf)}$ with bosons as rungs (spin fluctuations), while the terms with the conduction electron lines running counter-clockwise generate the conduction electron-empty orbital ladder vertex $T^{(ce)}$ with pseudofermions as rungs (charge fluctuations)

Anderson lattice model [85, 86] as well as to phase transitions in dilute, magnetic semiconductors at not too low temperature [87]. The NCA has also been generalized to the case of multiple local orbitals, as in rare earth and transition metal ions, where the NCA correctly produces a distinct Kondo resonance for each crystal-field or spin-orbit split local orbital, each with a characteristic, logarithmic temperature dependence [88, 89].

However, in NCA the infrared exponents of the auxiliary particle propagators come out independent of n_a, $\alpha_f^{\rm NCA} = 1/(N+1)$, $\alpha_e^{\rm NCA} = N/(N+1)$ [70, 71, 81], with N the spin degeneracy, in contrast to the Fermi liquid values, (3.62)–(3.64). As a consequence, the NCA fails to describe Fermi liquid behavior at temperatures $T \ll T_{\rm K}$, with spurious infrared singularities appearing in physical quantities at energies or temperatures $T \ll T_{\rm K}$ [70, 71, 79–81]. Since the NCA becomes formally exact for SU(N) symmetric models in the limit $N \to \infty$, with deviations appearing in $O(\frac{1}{N^2})$ [79, 80, 90], this low-T failure is less pronounced for $N \gg 1$. Note, however, that the deviation of the NCA infrared exponents α_f and α_e is of order $1/N$, not $1/N^2$ as one may have expected. In a magnetic field the NCA also fails even in the high-temperature regime, $T > T_{\rm K}$, producing a spurious resonance in the impurity spectrum at $\omega = 0$ in addition to the two Zeeman-split Kondo peaks. The low-T failure of the NCA can be traced back to its insufficient inclusion of coherent multiple spin–flip processes which are responsible for the formation of the Kondo singlet state. The origin of the failure in a magnetic field, on the other hand, lies in the fact that NCA neglects the exchange diagram to the conduction electron-impurity spin vertex at second order in the spin coupling J [68, 91]. As a consequence, logarithmic contributions in the potential scattering channel do not cancel even in leading logarithmic order, producing a spurious resonance which does not Zeeman-split in a magnetic field.

3.5.4.2 Anderson Impurity Model with Finite U: SUNCA

At finite Coulomb interaction U, the spin exchange interaction J acquires contributions from both, virtual excitations to the empty and to the doubly occupied impurity states via a Schrieffer–Wolff transformation [92],

$$J = \frac{|V|^2}{|\epsilon_f|} + \frac{|V|^2}{|\epsilon_f + U|}. \tag{3.65}$$

Neglecting either one of these contributions would lead to an exponentially wrong Kondo scale T_{K}, because of the exponential dependence of T_{K} on J. A simple generalization of NCA to this case, i.e., adding the second order self-consistent perturbation theory for the two processes, fails to capture the simultaneous contribution of both channels in each order of bare perturbation theory. For a correct treatment of both terms, there must be included, for each diagram with an empty boson line G_{e}, the corresponding diagram with G_{e} replaced by a doubly occupied boson line G_{d} (which amounts to the exchange diagram of the former), and vice versa, *on the level of bare perturbation theory* [93, 94]. The corresponding vertex corrections have first been evaluated in leading self-consistent order by Sakai et al. [95] and by Pruschke and Grewe [96]. The first conserving approximation for finite U, fully symmetric with respect to the empty and double occupied fluctuation channels, was formulated and evaluated by Haule et al. [93] and termed as the symmetrized finite-U NCA (SUNCA). On the level of renormalized perturbation theory (generating functionals), it means that for each dressed b-line there must be included a ladder vertex function with a-lines as rungs, and vice versa. The SUNCA is tractable with relatively moderate numerical effort, since it can be formulated in terms of no higher than three-point vertex functions. The results of a fully self-consistent evaluation of the impurity electron spectral function within SUNCA are shown in Fig. 3.5[1] in comparison with NRG results. It is seen that the correct Kondo scale (width of the Kondo peak) is reproduced. However, like the NCA, the SUNCA solution still develops a spurious low-T singularity.

There is evidence that both failures of NCA, at low temperature and in a magnetic field, can be cured by a systematic resummation of coherent spin–flip terms to infinite order, which will be discussed in the next section.

3.5.5 Conserving T-Matrix Approximation (CTMA)

3.5.5.1 Construction of the CTMA

As a minimal precondition to obtain a gauge symmetric description of the Fermi liquid fixed point, a conserving approximation must reproduce the correct

[1] We are grateful to T.A. Costi for providing the NRG data.

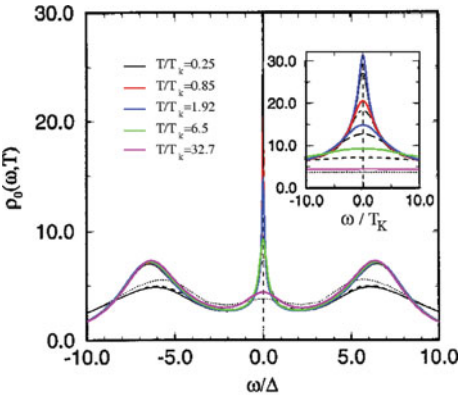

Fig. 3.5 Physical impurity electron spectral function of the Anderson impurity model for $U = -2\epsilon_f$. *Solid lines*: SUNCA results [93] and *dashed lines*: NRG results

pseudoparticle infrared exponents, (3.62)–(3.64), whose dependence on n_a is characteristic for a Fermi liquid ground state. It is easily seen by power counting arguments that any summation of a finite number of skeleton self-energy diagrams merely reproduces the incorrect NCA exponents [90], like, e.g., the post-NCA considered by Anders and Grewe [97] (diagrams up to $O(\Gamma^4)$ in Fig. 3.4). Hence, the generating functional must be comprised of an *infinite* class of skeleton diagrams in order to describe Fermi liquid behavior. Since the latter is a consequence of the singlet state formed at low T between the impurity and the conduction electron spins, one may expect that higher than two-particle correlation functions need not be considered in the single-channel case. The approximations to the total vertex functions between conduction electrons (c) and impurity DoF (pseudofermions f, slave bosons e) are then two-particle T-matrices, $T^{(cf)}$ and $T^{(ce)}$. As the irreducible parts of these T-matrices we select the single (renormalized) e or f particle lines, since (1) in the Kondo regime these terms are the leading contributions in the small parameter VN_0 and (2) in the spirit of principal diagrams, this choice gives rise to the maximum number of spin and charge fluctuation processes, respectively, in the T-matrices at any given order of (renormalized) perturbation theory. The Luttinger–Ward functional generating these ladder vertex terms (and others) for $U \to \infty$ is shown in Fig. 3.4. It is comprised of all closed pseudoparticle rings (skeletons) with each conduction electron line spanning at most two hybridization vertices and has been termed the conserving T-matrix approximation (CTMA) [67, 98]. Via the self-consistent inclusion of the self-energies, the vertex equations for T^{cf} and T^{ce} have parquet character. The CTMA integral equations, the analytical expressions for the self-energies, $\Sigma_{f\sigma}$ and Σ_e, and the impurity electron Green's function, $G_{a\sigma}$, are given explicitly in [98, 99]. Keiter and co-workers [100, 101] efficient decoupling scheme for the CTMA integral equations greatly facilitates the evaluation, and preserves the Fermi liquid values of the auxiliary particle infrared exponents.

3.5.5.2 Principal Results

Pseudoparticle Spectral Functions As a first indication for description of Fermi liquid behavior within CTMA it has been checked if the CTMA reproduces the correct Fermi liquid values of the pseudoparticle threshold exponents [67]. The exponents of the CTMA solution are displayed in Fig. 3.6 (left panel) and show, within the error bars, good agreement with the exact values, especially the dependence on the impurity occupation number n_d, characteristic for the Fermi liquid fixed point.

The static spin susceptibility of the impurity was calculated from the spin-dependent occupation numbers $n_{a\sigma}$ in a small magnetic field $\mathcal{B}$ as

$$\chi_i(T) = \left.\frac{\mathrm{d}M}{\mathrm{d}\mathcal{B}}\right|_{\mathcal{B}=0}, \tag{3.66}$$

where $M = g\mu_\mathrm{B}\sum_\sigma \sigma n_{a\sigma}$ is the impurity magnetization and

$$n_{a\sigma} = \lim_{\lambda\to\infty} \frac{\int \mathrm{d}\omega e^{-\beta\omega}\ \mathrm{Im}G_{f\sigma}(\omega - i0)}{\int \mathrm{d}\omega e^{-\beta\omega}\ \mathrm{Im}[\sum_\sigma G_{f\sigma}(\omega - i0) + G_e(\omega - i0)]}\,. \tag{3.67}$$

$\chi_i(T)$ is shown in Fig. 3.6 (left panel). It exhibits T-independent Pauli behavior for $T \lesssim 0.5T_\mathrm{K}$ with no singular behavior appearing down to the lowest T considered [99], indicative of the Fermi liquid ground state with a completely screened local moment. As expected, $\chi_i(T)$ obeys scaling for at least a range of T_K within a factor 10 [102], when plotted as a function of T/T_K. For details of the comparison with the Bethe ansatz results in Fig. 3.6 (left panel) see [99].

Fermi liquid behavior of the impurity electron Green's function

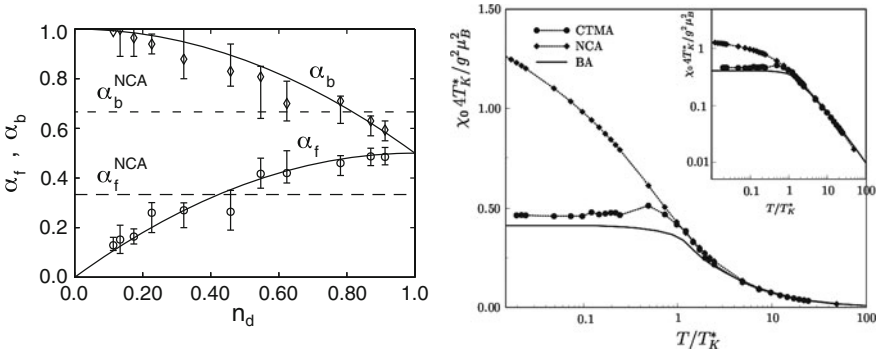

Fig. 3.6 *Left panel*: CTMA results (*symbols* with *error bars*) for the threshold exponents α_f and α_e for $U \to \infty$, $B = 0$. *Solid lines*: Exact values, (3.62) and (3.63), *dashed lines*: NCA results. *Right panel*: Static spin susceptibility as a function of temperature; Bethe ansatz, CTMA and NCA results (see text). Model parameters used: $\epsilon_d/D = -0.81$ and $\Gamma/D = 0.2$

$$G_{a\sigma}(\omega) = [\omega - \epsilon_f - i\Gamma - \Sigma_{a\sigma}(\omega)]^{-1} , \tag{3.68}$$

and the impurity electron interaction self-energy $\Sigma_{d\sigma}(\omega)$ is of prime interest especially for applications within DMFT. $\Sigma_{a\sigma}$ exhibits many features of Fermi liquid behavior [99]. It has quadratic dependence on both, ω and T, at low ω and T, with no sign of a spurious low-energy singularity down to the lowest T considered ($T \simeq 0.01\ T_{\rm K}$). As discussed in detail in [99], the curvature of the quadratic behavior in ω and T is found to be in good agreement with the exact Fermi liquid result, $\Sigma_{a\sigma}(\omega) = a[\omega^2 + (\pi T)^2]/T_{\rm K}^2$, where a is an exactly known prefactor. However, the position ω_0 of the minimum of $\mathrm{Im}\Sigma_{a\sigma}(\omega)$ is incorrectly shifted to $\omega_0 \approx -T_{\rm K}$, and $\mathrm{Im}\Sigma_{a\sigma}(\omega - i0)$ acquires negative values, thus violating the Friedel sum rule. When searching for the origin of this shortcoming, one must keep in mind that $\Sigma_{a\sigma}(\omega \approx 0)$ is determined via (3.68) by both $\mathrm{Im}G_{a\sigma}(\omega)$ and $\mathrm{Re}G_{a\sigma}(\omega)$, and thus, via the Kramers–Kronig relation, by high-energy (potential scattering) contributions to $G_{a\sigma}(\omega)$. Hence, the erroneous shift ω_0 may result from an unprecise calculation of $G_{a\sigma}(\omega)$ at high energies, either numerically or due to neglect of high-order potential scattering terms.[2] To correct this shortcoming, it has been suggested to add an appropriate, phenomenological *real constant* to $\Sigma_{a\sigma}(\omega)$. Through self-consistent it acts like a chemical potential and shifts the minimum of $\mathrm{Im}\Sigma_{a\sigma}(\omega)$ to $\omega = 0$. By inclusion of that single, real parameter, motivated by potential scattering contributions, the full Fermi liquid behavior of $\Sigma_{a\sigma}(\omega)$ is recovered, and $G_a(\omega)$ obeys the unitarity sum rule with good precision [99].

3.6 Renormalization Group Approaches

The renormalization group method is a powerful tool to calculate properties of interacting many-body systems. In the context of strongly correlated electron systems a first and seminal application of the method to the Kondo problem has been proposed by Anderson [103], who coined the term "poor man's scaling" for his method. While Anderson's treatment is perturbative in the exchange coupling constant of the Kondo model, and is, therefore, valid only at not too low energies, Wilson devised a numerical RG method for the Kondo model which he successfully implemented to cover the complete range of energies [104]. While these early RG formulations were phrased in terms of successive mappings of the Hamiltonian with continuously decreasing band width compensated by correspondingly adjusted coupling constants, the later functional renormalization group (FRG) schemes make use of a mapping of the complete set of Green's functions under a continuous change of an infrared cutoff. For the Kondo model, this is most conveniently formulated using the pseudofermion representation of the local spin 1/2. The "poor man's scaling" approach to the equilibrium Kondo model is essentially a simplified FRG formulation, in which the RG flow of the pseudofermion self-energy and the energy

[2] Note, however, that CTMA includes all potential scattering terms up to $O(\Gamma^4)$.

dependence of the exchange coupling function, as well as all higher-order couplings are neglected. It may be generalized to out of equilibrium situations, if the self-energy (the imaginary part of which describes the relaxation rate of the local spin) and the energy dependence of the coupling are kept. It is actually even possible to extend the RG approach into the strong coupling regime, provided the effect of the self-energy in controlling the flow of the coupling is treated correctly. Finally, the FRG has been successfully used to treat another strong coupling problem, that of frustrated quantum spin systems on a lattice in pseudofermion representation.

3.6.1 *"Poor Man'S Scaling" in the Equilibrium Kondo Model*

The interaction of a Fermi sea with a quantum impurity, e.g., a local spin exchange coupled to the local spin density of the Fermi system, gives rise to the Kondo effect. Initially introduced to describe magnetic impurities in metals, the Kondo problem is by now a ubiquitous phenomenon. The corresponding s–d exchange Hamiltonian, or Kondo Hamiltonian for short, reads

$$H = \sum_{k\sigma} \varepsilon_k c^\dagger_{k\sigma} c_{k\sigma} + J \sum_{kk'\sigma\sigma'} \frac{1}{2} c^\dagger_{k\sigma} \boldsymbol{\tau}_{\sigma\sigma'} c_{k'\sigma'} \cdot \boldsymbol{S} \,, \tag{3.69}$$

where $\boldsymbol{S} =$ is the local spin 1/2 operator in pseudofermion representation, (3.1), J is the exchange coupling constant and $\boldsymbol{\tau}_{\sigma\sigma'}$ is the vector of Pauli matrices. We take a flat conduction electron DOS, $N(\omega) = \frac{1}{2D}\theta(D_0 - |\omega|)$. If we now transform this Hamiltonian to an equivalent one with reduced band width $D = D_0 - dD$, the physical properties of the system remain unchanged, provided the exchange coupling is changed correspondingly. In order to calculate the required adjusted value of J at the scale D, denoted $J(D)$, one may determine the effective coupling in perturbation theory in J. The lowest (second order in J) correction terms δJ are shown in Fig. 3.7, where the dashed (solid) lines depict bare retarded pseudofermion (local Keldysh conduction electron) Green's functions. Differentiating δJ with respect to the running band width D one finds that in lowest order in the dimensionless coupling constant $g_0 = N(0)J$ the renormalized coupling $g(D)$ obeys the RG equation [103]

$$\frac{\mathrm{d}g(D)}{\mathrm{d}\ln D} = -2g^2 + O(g^3) \tag{3.70}$$

with the solution

$$g(D) = \frac{1}{2\ln(D/T_\mathrm{K})}, \tag{3.71}$$

where $T_\mathrm{K} = D_0 \exp(-1/2g_0)$ is the Kondo temperature. One observes that $g(D)$ diverges at $D = T_\mathrm{K}$. The divergence is removed by taking the pseudofermion self-energy into account, the imaginary part of which describes the relaxation rate for

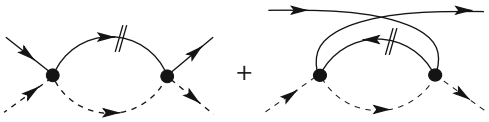

Fig. 3.7 Diagrammatic representation of the RG vertex in second order of J. The *strokes* indicate differentials with respect to the band cutoff D. *Dashed* and *solid lines* depict bare retarded pseudofermion and local Keldysh conduction electron Green's functions, respectively

spin–flip processes (see below). The above perturbative RG result provides a valid description as long as the running coupling is small, $g(D) << 1$. In the perturbative regime the Kondo effect, i.e., the screening of the local spin by the conduction electron spins, begins to build up. The screening is completed at the energy or temperature scale $T << T_{\rm K}$, i.e., in the strong coupling regime, not accessible by perturbative methods.

3.6.2 Functional RG for the Kondo Model Out of Equilibrium

The Kondo problem experienced a revival in the mid 1990s, when it was found that charge transport through nano-structures may be dominated by the Kondo effect, in the sense that the Kondo resonance at a quantum dot may lead to perfect conductance through the dot, although it is only weakly coupled to the leads [105,106]. The corresponding Kondo Hamiltonian of a local spin exchange coupled to the conduction electron spin densities in the two leads at the dot and the transfer operator through the dot reads

$$H = \sum_{\boldsymbol{k}\sigma\alpha}(\varepsilon_{\boldsymbol{k}} - \mu_\alpha)c^\dagger_{\boldsymbol{k}\sigma\alpha}c_{\boldsymbol{k}\sigma\alpha} + \sum_{\boldsymbol{k}\boldsymbol{k}'\sigma\sigma'} J_{\alpha\alpha'}\frac{1}{2}c^\dagger_{\boldsymbol{k}\sigma\alpha}\boldsymbol{\tau}_{\sigma\sigma'}c_{\boldsymbol{k}'\sigma'\alpha'}\cdot\boldsymbol{S}\,, \tag{3.72}$$

where $\alpha = L, R$ labels the leads and $\mu_\alpha = \pm eV/2$ denotes the chemical potential shifts in the leads induced by an applied d.c. bias voltage V. In the following the matrix of exchange couplings is assumed isotropic ($J_{\alpha\alpha'} = J$).

As first proposed in [107–110], and later derived within a full functional RG treatment [111], the "poor man's scaling" approach may be generalized to nonequilibrium by keeping the energy dependence of the coupling function $g(\omega) = J(\omega)N(0)$ and by observing that the RG flow is cutoff at the scale of the spin relaxation rate Γ. The role of Γ in suppressing the Kondo effect has been demonstrated within NCA in [112]. The generalized RG equation may be formulated as

$$\frac{{\rm d}g(D;\omega)}{{\rm d}\ln D} = -\sum_{\alpha=\pm} g^2(D;\alpha eV/2)\theta(D - \sqrt{(\omega + \alpha\, eV/2)^2 + \Gamma^2}). \tag{3.73}$$

This equation must be solved simultaneously with the relaxation rate at scale D

$$\Gamma = \pi \sum_{\alpha=L,R} \int \mathrm{d}\omega g^2(D;\omega) f(\omega - \mu_\alpha)[1 - f(\omega - \mu_{\alpha'})], \tag{3.74}$$

where $f(\omega)$ is the Fermi function of the conduction electrons. The charge current is given by the expression

$$I = e\frac{3\pi}{4} \int \mathrm{d}\omega g^2(D;\omega)\{f(\omega - \mu_L)[1 - f(\omega - \mu_R)] - (L \Leftrightarrow R)\} \tag{3.75}$$

Excellent agreement of the above theory with experimental data is found provided the bias voltage, the applied magnetic field (Zeeman splitting B) or temperature is sufficiently high, eV, $B, T >> T_\mathrm{K}$ so that the perturbative expression of the RG β-function is applicable [107, 108, 113–118].

3.6.3 RG Approach to the Kondo Model at Strong Coupling

Motivated by the success of the generalized RG in nonequilibrium and the insights gained from there, one may ask whether that formulation may be extended to the strong coupling domain. This has been attempted back in the 1970s by Larsen and Mattuck (LM) [119, 120], who discovered that the weak coupling RG equation, combined with the cutoff Γ provided by the relaxation rate, shifts the singularity in $g(D)$ from $D = T_\mathrm{K}$ down to $D = 0$. In the limit $T \to 0$ the relaxation rate Γ was found to approach the value T_K, the exchange coupling developed the singular behavior $g(T) \propto 1/T$ and the conductance assumed the exact unitarity limit. It is worth noting that LM employed the pseudofermion representation with particle–hole symmetric projection. The flaw in LM's derivation was that the leading low-temperature corrections turned out to be linear in T instead of quadratic, as required by Fermi liquid theory. This difficulty has been recently resolved by using the correct form of the RG β-function in the strong coupling regime [121]. In addition, it may be shown that while LM considered only the lowest (single particle–hole excitation) contribution to Γ, all the higher-order contributions may be subsummed to give the same structure, but with modified numerical coefficients. As shown in [121] the β-function in the strong coupling regime grows as g^3, and, therefore, faster than in the weak coupling limit. A useful interpolation expression is

$$\frac{\mathrm{d}g(D)}{\mathrm{d}\ln D} = -\frac{[g(2+g)]^2}{2+2g}\theta(D - \Gamma). \tag{3.76}$$

This equation may be solved analytically to give

$$g(D) = \sqrt{1 + 1/\ln(\sqrt{D^2 + \Gamma^2}/T_\mathrm{K})} - 1, \tag{3.77}$$

where T_K has been defined above. Here the relaxation rate is given by the self-consistent equation

$$\Gamma(T) = 3g^2(T) \int d\omega d\omega' f(\omega') f(\omega - \omega') f(-\omega) \frac{\Gamma}{\omega^2 + \Gamma^2}, \tag{3.78}$$

where $f(\omega)$ denotes the Fermi functions of both, conduction electrons and pseudofermions. At temperatures $T \ll \Gamma$ the integral may be evaluated to give $\Gamma^2(T) = c_\Sigma g^2(T) T^2$ ($c_\Sigma = 3\pi^2/4$), which when combined with (3.69) results in a finite value of the relaxation rate, $\Gamma = T_K + c_\Gamma T^2/T_K + O(T^3)$ ($c_\Gamma = c_\Sigma - 1/2$) and a diverging coupling $g(T) = c_g/T$. The conductance G through a Kondo dot in the linear response regime is given by

$$\frac{G}{G_0} = 3g^2 Y \int d\omega d\omega' f(\omega') f(\omega - \omega') f(-\omega) \frac{\Gamma}{\omega^2 + \Gamma^2} \frac{\Gamma}{(\omega')^2 + \Gamma^2}, \tag{3.79}$$

where G_0 is the conductance quantum and Y is the correction factor introduced in the section on particle–hole symmetric projection above ($Y \to 1$ as $T \to 0$). One observes that $G/G_0 \to 1$ as $T \to 0$ (unitarity).

3.6.4 Functional RG Approach to Frustrated Heisenberg Models

The pseudofermion representation of spin operators may also be used in lattice models. Here again the p–h symmetric projection is useful, particularly when one is interested in the low-temperature behavior ($T \ll J$). The more conventional approximation schemes (random phase approximation and FLEX approximation) when applied to the Heisenberg model on the square lattice have been shown to provide good results [72]. The effect of the exact projection via Popov–Fedotov has been found to be only important at higher temperatures, as expected [19]. Systems of recent interest are frustrated Heisenberg models, with competing interactions. Any approximate treatment should be carefully balanced as to not prefer one type of correlations over another one. A systematic treatment of all interaction channels may be achieved by employing the functional RG method [122–124]. A first study of a frustrated Heisenberg model at $T = 0$ using the pseudofermion representation and the FRG method provided excellent results on the phase diagram of the $J_1 - J_2$-model [22]. Similar to the treatment of the Kondo model by FRG mentioned above, one keeps only the two first RG-equations out of the infinite hierarchy of equations. It is necessary to keep the full energy dependence of self-energy and two-particle vertex functions. In addition, it turns out to be important to keep a three-particle correlation contribution in the form of self-energy insertions in the β-function of the RG equation for the couplings [125]. In this way the Ward identities may be approximately satisfied. While this method cannot yet be used to calculate properties of the ordered state, it allows us to identify phases without long-range order by

studying whether the RG flow runs smoothly all the way to $\Lambda = 0$ (Λ is the infrared cutoff parameter of FRG), which indicates a phase without LRO, or whether the flow becomes unstable at some finite Λ , pointing to the existence of LRO. A recent application to the $J_1 - J_2 - J_3$ model yielded again excellent agreement with complementary methods [126].

3.7 Conclusion

We have reviewed the most prominent auxiliary particle techniques and their applications to strongly correlated electron systems, using a variety of approximation schemes, ranging from SPA, possibly with Gaussian fluctuations, to conserving approximations to renormalization group methods. It was seen that the Kotliar–Ruckenstein representation, especially in its spin rotation invariant and spin–charge symmetric formulation, is particularly useful for identifying complex spin and/or charge ordered groundstates in mean-field like approximations, since it treats all spin and charge states on a lattice site on the same footing. Regarding the Hubbard model on the square lattice, unrestricted Hartree–Fock calculations point towards a huge number of solutions. An indication that this is also realized using slave bosons on the saddle-point level is provided by Fig. 3.3, but identifying the numerous competing phases remains a challenge. Conserving approximations provide a versatile tool for implementing the constraint on the auxiliary particle number and, hence, for a thorough treatment of fluctuations without resorting to a mean-field as a starting point. This appears necessary, in particular, for quantum impurity systems, where magnetic ordering does not occur. The NCA and the CTMA have been discussed as the most prominent examples of this approach. Finally, Larsen's particle–hole symmetric projection technique may provide a promising tool for describing the complete RG flow from the weak coupling regime to the strong coupling fixed point within a diagrammatic RG.

Acknowledgments We thank Jan Brinckmann, Theo Costi, Klaus Doll, Michael Dzierzawa, Peter Hirschfeld, Stefan Kirchner, Thilo Kopp, Gabi Kotliar, Burkhard Möller, Kandkher Muttalib, Andrzej Oleś, Henni Ouerdane, Jens Paaske, Marcin Raczkowski, Johannes Reuther, Achim Rosch, Holger Schmidt, and Walter Zimmermann for valuable collaboration. RF gratefully acknowledges financial support by the Région Basse-Normandie and the Ministère de la Recherche. JK is grateful for financial support from the Deutsche Forschungsgemeinschaft through SFB 608 and through grant No. KR 1727/1-2 and PW acknowledges support through the DFG Research Center for Functional Nanostructures (CFN) and the DFG Research Unit 960 "Quantum Phase Transitions".

References

1. M. Imada, A. Fujimori, Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
2. P.A. Lee, T.M. Rice, J. Serene, L.J. Sham, J.W. Wilkins, Comments on Condens. Matter Phys. **12,** 99 (1986)

3. P. W. Anderson, Science **235**, 1196 (1987); Experiment: J. G. Bednorz and K. A. Müller, Z. Physik B **64**, 189 (1986); B. Raveau, C. Michel, M. Hervieu, D. Groult, A. Maignan, J. Provost, J. Supercond. **5**, 203 (1992)
4. A.C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, London, 1993)
5. T. Holstein, H. Primakoff, Phys. Rev. **58**, 1094 (1940)
6. J. Schwinger, in *Quantum Theory of Angular Momentum*, ed. by H. Biedenharn, H. Van Dam (Academic, New York, 1965)
7. A.A. Abrikosov, Physics (Long Island City, N.Y.) **2**, 5 (1965)
8. B. Coqblin, J.R. Schrieffer, Phys. Rev. **185**, 847 (1969)
9. S.E. Barnes, J. Phys. F **6**, 1375 (1976); **7**, 2637 (1977)
10. P. Coleman, Phys. Rev. B **29**, 3035 (1984)
11. M. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
12. G. Kotliar, A.E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986)
13. T.C. Li, P. Wölfle, P.J. Hirschfeld, Phys. Rev. B **40**, 6817 (1989)
14. R. Frésard, P. Wölfle, Int. J. of Mod. Phys. B **6**, 685 (1992); ibid., p. 3087 (1992)
15. R. Frésard, G. Kotliar, Phys. Rev. B **56**, 12 909 (1997)
16. F. Lechermann, A. Georges, G. Kotliar, O. Parcollet, Phys. Rev. B **76**, 155102 (2007)
17. U. Larsen, Z. Physik **256**, 65 (1972)
18. V.N. Popov, S.A. Fedotov, Sov. Phys. JETP **67**, 535 (1988)
19. J. Brinckmann, P. Wölfle, Physica B **359–361**, 798 (2005)
20. M.N. Kiselev, R. Oppermann, JETP Lett. **71**, 250 (2000)
21. R. Dillenschneider, J. Richert, Phys. Rev. B **73**, 024409 (2006)
22. J. Reuther, P. Wölfle, Phys. Rev. B **81**, 144410 (2010)
23. J. Kroha, P. Hirschfeld, K.A. Muttalib, P. Wölfle Solid State Comm. 83 (12), 1003 (1992)
24. N. Read, D.M. Newns, J. Phys. C **16**, L1055 (1983)
25. N. Read, D.M. Newns, J. Phys. C **16**, 3273 (1983)
26. D.M. Newns, N. Read, Adv. Phys. **36,** 799 (1987)
27. R. Frésard, T. Kopp, Nucl. Phys. B **594**, 769 (2001)
28. R. Frésard, H. Ouerdane, T. Kopp, Nucl. Phys. B **785**, 286 (2007)
29. J.W. Rasul, T. Li, J. Phys. C **21**, 5119 (1988)
30. T.C. Li, J.W. Rasul, Phys. Rev. B **39**, 4630 (1989)
31. J.W. Rasul, T.C. Li, H. Beck, Phys. Rev. B **39**, 4191 (1989)
32. M. Lavagna, Phys. Rev. B **41**, 142 (1990); Helvetica Physica Acta **63**, 310 (1990); Int. J. Mod. Phys. B **5**, 885 (1991)
33. Th. Jolicœur and J.C. Le Guillou, Phys. Rev. B **44**, 2403 (1991)
34. Y. Bang, C. Castellani, M. Grilli, G. Kotliar, R. Raimondi, Z. Wang, Int. J. Mod. Phys. B **6**, 531 (1992)
35. R. Frésard, T. Kopp, Phys. Rev. B. **78**, 073108 (2008)
36. W. Zimmermann, R. Frésard, P. Wölfle, Phys. Rev. B **56**, 10 097 (1997)
37. B. Möller, K. Doll, R. Frésard, J. Phys. Condens. Matter **5** 4847 (1993)
38. R. Frésard, M. Dzierzawa, P. Wölfle, Europhys. Lett. **15**, 325 (1991)
39. E. Arrigoni, G.C. Strinati, Phys. Rev. B **44**, 7455 (1991)
40. R. Frésard, P. Wölfle, J. Phys. Condens. Matter **4** 3625 (1992)
41. G. Seibold, E. Sigmund, V. Hizhnyakov, Phys. Rev. B **57**, 6937 (1998)
42. M. Raczkowski, R. Frésard, A.M. Oleś, Phys. Rev. B **73**, 174525 (2006)
43. M. Raczkowski, R. Frésard, A.M. Oleś, Europhys. Lett. **76**, 128 (2006)
44. A. Isidori, M. Capone, Phys. Rev. B **80**, 115120 (2009)
45. A. Ruegg, S. Pilgram, M. Sigrist, Phys. Rev. B **75**, 195117 (2007)
46. P. Korbel, W. Wojcik, A. Klejnberg, J. Spałek, M. Acquarone, M. Lavagna, Eur. Phys. J. B **32**, 315 (2003)
47. G. Kotliar, E. Lange, M.J. Rozenberg, Phys. Rev. Lett. **84**, 5180 (2000)
48. G. Seibold, Phys. Rev. B **58**, 15520 (1998)
49. J. Lorenzana, G. Seibold, Low Temp. Phys. **32**, 320 (2006)

50. W.F. Brinkman, T.M. Rice, Phys. Rev. B. **2**, 4302 (1970)
51. D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984)
52. D. Vollhardt, P. Wölfle, P.W. Anderson, Phys. Rev. B. **35**, 6703 (1987)
53. S. Doniach, Physica B, C **91**, 231 (1977)
54. H.v. Löhneysen, A. Rosch, M. Vojta, P. Wölfle, Rev. Mod. Phys. **79**, 1015 (2007)
55. B. Möller, P. Wölfle, Phys. Rev. B **48**, 10320 (1993)
56. M. Guerrero, R.M. Noack, Phys. Rev. B **53**, 3707 (1996)
57. T.C. Li, Y.S. Sun, P. Wölfle, Z. Phys. B Condens. Matter **82**, 369 (1991)
58. T.C. Li, P. Bénard, Phys. Rev. B. **50**, 17 387 (1994)
59. P. Limelette, H. Muguerra, S. Hébert, Phys. Rev. B **82**, 035123 (2010)
60. K. Doll, M. Dzierzawa, R. Frésard, P. Wölfle, Z. Phys. B Condens. Matter **90**, 297 (1993)
61. R. Frésard, W. Zimmermann, Phys. Rev. B **58**, 15288 (1998)
62. M. Dzierzawa, unpublished
63. G. Baym, L.P. Kadanoff, Phys. Rev. **124** (1961) 287
64. G. Baym, Phys. Rev. **127**, 1391 (1962)
65. P.W. Anderson, Phys. Rev. Lett. **18**, 1049 (1967); Phys. Rev. **164**, 352 (1967)
66. B. Menge, E. Müller-Hartmann, Z. Phys. B **73**, 225 (1988)
67. J. Kroha, P. Wölfle, T.A. Costi, Phys. Rev. Lett. **79**, 261 (1997)
68. J. Kroha, P. Wölfle, J. Phys. Soc. Jpn. **74** (1), 16 (2005)
69. T.A. Costi, P. Schmitteckert, J. Kroha, P. Wölfle, Phys. Rev. Lett. **73**, 1275 (1994); Physica **C235–240**, 2287 (1994)
70. E. Müller-Hartmann, Z. Physik **B 57**, 281 (1984)
71. Y. Kuramoto, E. Müller-Hartmann, J. Mag. Mag. Mat. **52**, 122 (1985)
72. J. Brinckmann, P. Wölfle, Phys. Rev. B **70**, 174445 (2004)
73. H. Keiter, J.C. Kimball, Int. J. Magn. **1** 233 (1971)
74. N. Grewe, H. Keiter, Phys. Rev. B **24** 4420 (1981)
75. Y. Kuramoto, Z. Physik B **53**, 37 (1983)
76. H. Kojima, Y. Kuramoto, and M. Tachiki, Z. Physik **B 54**, 293 (1984)
77. Y. Kuramoto, H. Kojima, Z. Physik **57**, 95 (1984)
78. Y. Kuramoto, Z. Physik **B 65**, 29 (1986)
79. N.E. Bickers, D.L. Cox, J.W. Wilkins, Phys. Rev. **B 36**, 2036 (1987)
80. N.E. Bickers, Rev. Mod. Phys. **59**, 845 (1987)
81. T.A. Costi, J. Kroha, P. Wölfle, Phys. Rev. **B 53**, 1850 (1996)
82. P. Nozières, A. Blandin, J. Phys. (Paris) **41**, 193 (1980)
83. K. Haule, A. Rosch, J. Kroha, P. Wölfle, Phys. Rev. Lett. **89**, 236402 (2002)
84. K. Haule, A. Rosch, J. Kroha, P. Wölfle, Phys. Rev. B **68**, 155119 (2003)
85. F. B. Anders, M. Jarrel, D. Cox, Phys. Rev. Lett. **78**, 2000 (1997)
86. A. Schiller, F. B. Anders, D. Cox, Phys. Rev. Lett. **81**, 3235 (1998)
87. M. Arnold, J. Kroha, Phys. Rev. Lett. **100**, 046404 (2008)
88. F. Reinert et al., Phys. Rev. Lett. **87**, 106401 (2001)
89. D. Ehm et al., Phys. Rev. B **76**, 045117 (2007)
90. D.L. Cox, A.R. Ruckenstein, Phys. Rev. Lett. **71**, 1613 (1993)
91. S. Kirchner, J. Kroha, J. Low Temp. Phys. **126**, 1233 (2002)
92. J.R. Schrieffer, P.A. Wolff, Phys. Rev. **149**, 491 (1966)
93. K. Haule, S. Kirchner, J. Kroha, P. Wölfle: Phys. Rev. B **64**, 155111 (2001)
94. J. Holm, K. Schönhammer, Solid State Comm. **69**, 969 (1989)
95. O. Sakai, M. Motizuki, T. Kasuya, *Springer Series in Solid-State Sciences*, vol. 81 (Springer, Berlin, 1988), p. 45
96. Th. Pruschke, N. Grewe, Z. Phys. B **74**, 439 (1989)
97. F.B. Anders, N. Grewe, Europhys. Lett. **26**, 551 (1994)
98. J. Kroha, P. Wölfle, in *Theoretical Methods for Strongly Correlated Electrons*, CRM Series in Mathematical Physics (Springer, New York, 2003), p. 297
99. S. Kirchner, J. Kroha, P. Wölfle, Phys. Rev. B **70**, 165102 (2004)
100. H. Keiter, K. Baumgartner, Phys. Stat. Sol. (b) **242**(2), 377 (2005)

101. H. Keiter, K. Baumgartner, S. Duffe, Phys. Stat. Sol. (b) **244** (7), 2357 (2007)
102. J. Kroha, P. Wölfle, unpublished
103. P.W. Anderson, J. Phys. C **3**, 2439 (1970)
104. K.G. Wilson, Rev. Mod. Phys. **47**, 773 (1975)
105. D. Goldhaber-Gordon et al., Nature (London) **391**, 156 (1998)
106. S.M. Cronenwett, T.H. Oosterkamp, L.P. Kouwenhoven, Science **281**, 540 (1998)
107. A. Rosch, J. Paaske, J. Kroha, P. Wölfle, Phys. Rev. Lett. **90**, 076804 (2003)
108. A. Rosch, J. Paaske, J. Kroha, P. Wölfle, J. Phys. Soc. Jpn. **74**, 118 (2005)
109. J. Paaske, A. Rosch, P. Wölfle, Phys. Rev. B **69**, 155330 (2004)
110. J. Paaske, A. Rosch, J. Kroha, P. Wölfle, Phys. Rev. B **70**, 155301 (2004)
111. H. Schmidt, P. Wölfle, Ann. Phys. (Berlin) **19**, 60 (2010)
112. A. Rosch, J. Kroha, P. Wölfle, Phys. Rev. Lett. **57**, 156802 (2003)
113. J. Paaske, A. Rosch, P. Wölfle, C.M. Marcus, J. Nygård, Nat. Phys. **2**, 460 (2006)
114. V. Koerting, P. Wölfle, J. Paaske, Phys. Rev. Lett. **99**, 036807 (2007)
115. V. Koerting, J. Paaske, P. Wölfle, Phys. Rev. B **77**, 165122 (2008)
116. S. Schmaus, V. Koerting, J. Paaske, T.S. Jespersen, J. Nygård, P. Wölfle, Phys. Rev. B **79**, 045105 (2009)
117. C.-H. Chung, G. Zarand, P. Wölfle, Phys. Rev. **77**, 035120 (2008)
118. C.-H. Chung, K. LeHur, M. Vojta, P. Wölfle, Phys. Rev. Lett. **102**, 216803 (2009)
119. U. Larsen, R.D. Mattuck, *Proceedings of the International Conference on Magnetism*, Moscow, U.S.S.R., 1973, vol. 5 (Nauka, Moscow, 1974), p. 88
120. U. Larsen, Phys. Lett. A **40**, 39 (1972)
121. J. Paaske, H. Schmidt, P. Wölfle, to be published (2011)
122. M. Salmhofer, *Renormalization: An Introduction* (Springer, Berlin, 1998)
123. V. Meden, *Advances in Solid State Physics*, vol. 46, ed. by R. Haug (Springer, New York, 2007)
124. W. Metzner, in *Functional Renormalization Group Aapproach to Interacting Systems*. Lecture Notes in Physics (Springer, Berlin, 2007)
125. A.A. Katanin, Phys. Rev. **70**, 115109 (2004)
126. J. Reuther, P. Wölfle, R. Darradi, W. Brenig, M. Arlego, J. Richter, Phys. Rev. B **83**, 064416 (2011)

Chapter 4
The Composite Operator Method (COM)

Adolfo Avella and Ferdinando Mancini

Abstract The composite operator method (COM) is formulated, its internals illustrated in detail and some of its most successful applications reported. COM endorses the emergence, in strongly correlated systems (SCS), of composite operators, optimally deals with their unusual features and implements algebra constraints, and other relevant symmetries, in order to properly compute the unconventional properties of SCS.

4.1 Strong Correlations and Composite Operators

In the last decades, a large part of the research activity in solid state and condensed matter physics has been devoted to the study of electronic systems with very unconventional properties (i.e., with properties not consistent with the Fermi liquid theory). It is commonly believed that the origin of such anomalous behaviors should be traced back to the presence of strong electronic correlations in such systems.

A. Avella (✉)
Dipartimento di Fisica "E.R. Caianiello," Università degli Studi di Salerno, 84084 Fisciano (SA), Italy

CNR-SPIN, UoS di Salerno, 84084 Fisciano (SA), Italy

Unità CNISM di Salerno, Università degli Studi di Salerno, 84084 Fisciano (SA), Italy
e-mail: avella@physics.unisa.it

F. Mancini
Dipartimento di Fisica "E.R. Caianiello," Università degli Studi di Salerno, 84084 Fisciano (SA), Italy

Unità CNISM di Salerno, Università degli Studi di Salerno, 84084 Fisciano (SA), Italy
e-mail: mancini@physics.unisa.it

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_4,

Weakly correlated systems are described by Hamiltonians where the multibody terms, and the interactions they describe, feature so little coupling constants with respect to the average energies of the one-body terms (e.g., wide-band s and p metals) that we can successfully resort to perturbation theories. In particular, in such cases, mean-field-like approximations allow us to diagonalize the Hamiltonian in terms of new independent particles, described by *canonical operators*, that is, operators bounded to satisfy canonical commutation relations. On the contrary, in *strongly correlated* systems (e.g., narrow-band d and f metals) [1,2], the interactions are sufficiently intense to make completely useless any perturbation scheme that will be just bound to fail. In some relevant cases, even more than the strength of the interactions, it is their true nature, that is, their actual operatorial form, to bring all troubles (e.g., Kondo model [3]).

In order to tackle this problem, we need to change perspective and relax some of the constraints implied by the quest for new independent particles. In particular, we should not expect to be able to correctly describe strongly correlated systems only by means of canonical operators. Strong interactions modify dramatically the properties of the original particles. What are observed are new particles with new peculiar properties entirely determined by the dynamics and by the boundary conditions (i.e., all elements characterizing the physical situation under study). These new objects appear as the final result of the modifications imposed by the interactions on the original particles and embed, since the very beginning, the effects of correlations. Thus, the choice of new fundamental operators, whose properties will be self-consistently determined by dynamics, symmetries and boundary conditions, becomes relevant.

Let us consider a very simple, but very pedagogical, system in order to concretely illustrate this occurrence. A lattice of ions sited sufficiently far apart to avoid hopping of the electrons among the sites. For the sake of simplicity, we can also imagine that only one s-like orbital is active per ion and that the Coulomb repulsion is sufficiently intense only between same-site, that is, same-orbital, electrons:

$$H = -\mu \sum_{\mathbf{i}\sigma} c_\sigma^\dagger(i) c_\sigma(i) + U \sum_{\mathbf{i}} n_\uparrow(i) n_\downarrow(i), \tag{4.1}$$

where μ is the chemical potential, $\mathbf{i}$ is a vector of the lattice, $c_\sigma(i)$ is the destruction operator in the Heisenberg picture ($i = (\mathbf{i}, t)$) of an electron of spin σ in a Wannier state centered at the site $\mathbf{i}$, U is the strength of the on-site Coulomb repulsion and $n_\sigma(i) = c_\sigma^\dagger(i) c_\sigma(i)$ is the electronic number operator.

Although the Hamiltonian (4.1) looks very simple, there is neither known canonical transformation capable to diagonalize this Hamiltonian nor perturbation scheme, at any order, getting any closer to the exact solution or even just grasping its main features. This is simply due to the impossibility to describe this system in terms of independent particles represented by canonical operators.

If we analyze the equation of motion of the electronic operator $c_\sigma(i)$

$$\mathrm{i}\frac{\partial}{\partial t}c_\sigma(i) = [c_\sigma(i), H] = -\mu c_\sigma(i) + U n_{\bar{\sigma}}(i) c_\sigma(i), \tag{4.2}$$

we encounter a *fermionic operator*, that is, an operator constituted by an odd number of electronic operators, describing an electron of spin σ dressed by the charge fluctuations of the electrons of opposite spin ($\bar{\sigma}$). Let us name this new operator $\eta_\sigma(i) = n_{\bar{\sigma}}(i) c_\sigma(i)$. The equation of motion of η reads

$$\mathrm{i}\frac{\partial}{\partial t}\eta_\sigma(i) = [\eta_\sigma(i), H] = (U - \mu)\eta_\sigma(i). \tag{4.3}$$

It is immediate to see that there exists a companion of η, which we named $\xi_\sigma(i) = c_\sigma(i) - \eta_\sigma(i)$, whose equation of motion reads

$$\mathrm{i}\frac{\partial}{\partial t}\xi_\sigma(i) = [\xi_\sigma(i), H] = -\mu\xi_\sigma(i). \tag{4.4}$$

Equations (4.3) and (4.4) resemble that of independent particles as the *current* (i.e., the r.h.s of the equation of motion) is directly proportional to the operator itself. We define, as *eigenoperator* of a given Hamiltonian, an operator whose current is just proportional to the operator itself. We will call *eigenenergy* the proportionality constant between the eigenoperator and its current (e.g., $(U - \mu)$ for η).

ξ, together with η, constitutes a *fermionic closed basis* for this system. In our definition, a closed basis is the smallest set of operators necessary to build up the original operators as a linear combination (e.g., $c = \xi + \eta$) and which closes the equations of motion. That is, a set of operators $(O_1, O_2, \ldots, O_n)$ constitutes a closed basis for a system described by the Hamiltonian H and constituted, for instance, by c electrons iff $\mathrm{i}\frac{\partial}{\partial t}O_p = [O_p, H] = \sum_q a_{pq} O_q$ and $c = \sum_p b_p O_p$.

Often, by means of a fermionic closed basis, it is possible to *diagonalize* the related Hamiltonian, that is, to rewrite the Hamiltonian only in terms of number-like operators of elements of the operatorial basis. For instance, the Hamiltonian (4.1) can be rewritten as

$$H = -\mu \sum_{\mathbf{i}\sigma} \xi^\dagger_\sigma(i)\xi_\sigma(i) + \left(\frac{1}{2}U - \mu\right) \sum_{\mathbf{i}\sigma} \eta^\dagger_\sigma(i)\eta_\sigma(i). \tag{4.5}$$

This diagonalization, although exact, is formal as ξ and η are not canonical operators and they close commutation relations with their *number-like operators* $\xi^\dagger\xi$ and $\eta^\dagger\eta$ very different from that closed by the electronic operator c with $n = c^\dagger c$ (see Table 4.1; site indexes have been neglected: they would simply lead to Kronecker deltas).

Equation (4.5) shows that, owing to the presence of the interactions, the original electrons c are no more observable and new stable elementary excitations, described

Table 4.1 Anticommutation relations for ξ and η

$\{\xi_\sigma^\dagger, \xi_{\sigma'}\} = \delta_{\sigma\sigma'}(1 - c_{\bar{\sigma}}^\dagger c_{\bar{\sigma}}) + \delta_{\sigma\bar{\sigma}'} c_\sigma^\dagger c_{\bar{\sigma}}$	$[\xi_\sigma^\dagger \xi_\sigma, \xi_{\sigma'}] = -\delta_{\sigma\sigma'}\xi_\sigma$
$\{\eta_\sigma^\dagger, \eta_{\sigma'}\} = \delta_{\sigma\sigma'} c_{\bar{\sigma}}^\dagger c_{\bar{\sigma}} - \delta_{\sigma\bar{\sigma}'} c_\sigma^\dagger c_{\bar{\sigma}}$	$[\xi_\sigma^\dagger \xi_\sigma, \eta_{\sigma'}] = \delta_{\sigma\bar{\sigma}'}\eta_{\bar{\sigma}}$
$\{\xi_\sigma, \eta_{\sigma'}\} = \delta_{\sigma\bar{\sigma}'} c_\sigma c_{\bar{\sigma}}$	$[\eta_\sigma^\dagger \eta_\sigma, \eta_{\sigma'}] = -\eta_{\sigma'}$
$\{\xi_\sigma, \xi_{\sigma'}\} = \{\eta_\sigma, \eta_{\sigma'}\} = \{\xi_\sigma^\dagger, \eta_{\sigma'}\} = 0$	$[\eta_\sigma^\dagger \eta_\sigma, \xi_{\sigma'}] = 0$

by the operators ξ and η, appear. The electronic bare energy level $E = -\mu$ splits into two levels ($E_1 = -\mu$ and $E_2 = U - \mu$) and the original electrons turn out to be exactly the worst place to start: no perturbation scheme leads to the level splitting, which is the main and only feature of this very simple system. On the basis of this evidence, one is induced to move the attention from the original electronic operators to the new operators generated by the interaction. Such new operators can be written in terms of the original ones and are known as *composite operators*. A formulation which would treat composite operators as fundamental objects, instead of the original electronic operators, would be very promising. As a matter of fact, such a formulation would allow us to set up approximation schemes where some amount of the interaction is already exactly taken into account through the chosen operatorial basis, and would permit us to overcome the hoary problem of finding an appropriate expansion parameter.

However, one price must be paid. In general, composite operators are neither Fermi nor Bose operators, since they do not satisfy canonical (anti)commutation relations (see, for instance, Table 4.1). This very simple fact makes a tremendous difference with respect to the case in which one deals with the original electronic operators, which satisfy a canonical algebra. In developing perturbation schemes where the building blocks are propagators of composite operators, one cannot use the consolidated scheme any more: diagrammatic expansions, Wick's theorem and many other standard tools are no more valid or applicable. The formulation of the whole Green's function method must be revisited [4, 5] and new frameworks of calculations have to be formulated. In the following section, we will formulate the composite operator method (COM) [5] and illustrate its internals; COM systematically endorses the emergence, in strongly correlated systems, of composite operators and optimally deals with their unusual features.

4.2 The Composite Operator Method (COM)

4.2.1 Basis ψ

Given a Hamiltonian H describing a solid state system, first we have to choose a *basis* ψ constituted of a finite number of composite operators, either fermionic or bosonic. A fermionic basis will be used if the original interacting particles are electrons and we want to analyze thermodynamic and single-particle properties of the system. A bosonic basis can be either the *natural* basis for systems constituted

by bosons (e.g., spins, phonons, etc.) or a set of bosonic operators, constituted by an even number of electronic operators, representing the electronic number operator (charge), the electronic spin operator, ... and necessary to study the charge, spin, ... response of the system. In this latter case, the dynamics described by the bosonic basis is strongly coupled to that described by the related fermionic basis and vice versa. As a matter of fact, this is the case even if we do not explicitly open the bosonic *sector* as the self-consistent parameters appearing in the fermionic sector are just thermal averages of the related bosonic operators. A fully simultaneous self-consistent solution of both sectors is usually required. This consequence of the non-canonical algebra satisfied by composite operators is a manifestation of the relevance (actually, the dominance) of spin, charge, ... correlations in all properties exhibited by strongly correlated systems.

Obviously, a closed basis is the best basis one can adopt. It is possible to obtain a closed basis by just crossing the entire hierarchy of the equations of motion of the relevant canonical operator (e.g., the electronic operator, its number or spin operator, etc.). Namely, one has to: (a) compute high-order time derivatives of the relevant canonical operator by repeatedly commuting it with the Hamiltonian; (b) acquire in the basis all distinct composite operators appearing in the high-order currents; (c) terminate as soon as no new composite operator appears. This well-defined procedure, being equivalent to compute all spectral moments (i.e., the moments of the relevant canonical operator Green's function) [6–8], catches, in principle, all scales of energy featured by the system under analysis.

Following this procedure, one can see that there is a large class of systems (finite systems [4, 9], bulk systems with interacting localized electrons [10–12], Ising-like systems [13–16], etc.) where it is possible to obtain a closed basis with a finite number of elements and one can aim at the exact solution. In all other bulk systems, this procedure will lead to a closed basis with an infinite number of components. In such cases, one can: rewrite H as $H = H_0 + H_I$, where H_0 is the most relevant part whose finite closed basis is known; adopt the closed basis of H_0 as a *truncated basis* for H; and treat H_I as a *perturbation* to H_0. If the system under study features only *analytical scales of energy* (i.e., scales of energy whose mathematical expressions are expandable in power series of Hamiltonian model parameters), or if the range of temperatures/frequencies of interest does not involve *not-analytical scales of energy*, the adoption of a truncated basis can lead to an accurate description of the system under analysis. Such a description can be systematically improved adding to the truncated basis more and more elements according to the above described procedure. For instance, in the Hubbard model, we need at least a two-element basis to correctly describe the Hubbard on-site Coulomb repulsion U and at least a four-element basis to correctly describe the virtual exchange energy $J = 4t^2/U$ [9, 17].

Not-analytical scales of energy can be present in some models. For example, in the single-impurity Anderson model [3], the emergence of a not-analytical Kondo energy scale is directly connected to the virtual exchange processes triggered by the hybridization between conduction and impurity electrons and dominated by on-site Coulomb repulsion at the impurity site. This occurrence forces us to explore different routes to construct a proper operatorial basis. In fact, in such cases, the

adoption of any truncated basis would lead to a description of the system under analysis completely lacking any reference to not-analytical scales of energy. Such a description would turn out really poor at low enough temperatures/frequencies where neglected energy scales usually manifest themselves triggering quite unconventional behaviors (impurity spin screening, superconductivity, etc.). Actually, in such cases, one has to properly tweak the equations of motion of specific composite operators and smartly exploit self-consistency in order to get sensible results by means of a basis with a finite number of elements [18, 19].

Summarizing, many recipes can assure a correct and controlled description of relevant aspects of the dynamics. One can put in the basis:

- The higher-order composite operators emerging from the hierarchy of the equations of motion: The conservation of a definite number of spectral moments is guaranteed; polynomial and analytical scales of energy can be resolved with desired accuracy [5, 8].
- The eigenoperators of Hamiltonian terms describing specific interactions or virtual processes: the latter are correctly taken into account [9, 20].
- The eigenoperators of the problem reduced to a small cluster: all polynomial and analytical energy scales and their competitions/interplays are just exactly taken into account at short distances/large momenta [9, 21–23].
- The composite operator describing the Kondo-like singlet [18, 19, 24, 25].

4.2.2 Equations of Motion of ψ

Once the basis $\psi(i)$, constituted of n composite operators, has been chosen, we adopt a matricial notation and write

$$\psi(i) = \begin{pmatrix} \psi_1(i) \\ \vdots \\ \psi_n(i) \end{pmatrix}, \tag{4.6}$$

where we have not specified the nature, fermionic or bosonic, of the basis. In the case of fermionic operators, it is understood that we use the spinorial representation

$$\psi_m(i) = \begin{pmatrix} \psi_{m\uparrow}(i) \\ \psi_{m\downarrow}(i) \end{pmatrix}. \tag{4.7}$$

The current J of ψ can be exactly rewritten as

$$J(i) = \mathrm{i}\frac{\partial}{\partial t}\psi(i) = [\psi(i), H] = \sum_{\mathbf{j}} \varepsilon(\mathbf{i}, \mathbf{j})\psi(\mathbf{j}, t) + \delta J(i), \tag{4.8}$$

where the matrix ε is known as the *energy matrix*. δJ is named *residual current* and describes the component of the current J *orthogonal* to the basis ψ

$$\left\langle \left[\delta J(\mathbf{i},t), \psi^\dagger(\mathbf{j},t)\right]_\eta \right\rangle = 0 \Rightarrow \varepsilon(\mathbf{i},\mathbf{j}) = \sum_{\mathbf{l}} m(\mathbf{i},\mathbf{l}) I^{-1}(\mathbf{l},\mathbf{j}). \tag{4.9}$$

It is absolutely worth noticing that iff ψ is a closed basis, δJ is identically zero.

$m(\mathbf{i},\mathbf{j}) = \left\langle \left[J(\mathbf{i},t), \psi^\dagger(\mathbf{j},t)\right]_\eta \right\rangle$ is simply named m *matrix* and $I(\mathbf{i},\mathbf{j}) = \left\langle \left[\psi(\mathbf{i},t), \psi^\dagger(\mathbf{j},t)\right]_\eta \right\rangle$ is known as the *normalization matrix*. The matrix I is hermitian by definition, while the matrix m is hermitian because of the time independence of the matrix I: $0 = \mathrm{i}\frac{\partial}{\partial t} I = m - m^\dagger$. $\langle \cdots \rangle$ stands for the quantum mechanical average in the (grand-)canonical ensemble. $\eta = 1, -1$ marks anticommutators $\{,\}$ and commutators $[,]$, respectively. For fermionic (bosonic) composite operators, the choice $\eta = 1$ ($\eta = -1$) is much more convenient (i.e., it makes the calculations much simpler), although the opposite choice is also possible.

Once a basis has been chosen, either closed or truncated, we need to calculate the normalization I and the m matrices in order to construct the energy matrix ε, which will soon turn out to be a fundamental quantity. The calculation of the relevant (anti)commutators is straightforward, but it can be lengthy and cumbersome if the basis contains many elements. The thermal averaging procedure is instead much less mechanical as it involves choosing the various phases to be studied. All correlation functions present in the normalization I and in the energy ε matrices need to be computed self-consistently and form a first block of unknowns in the theory.

4.2.2.1 Weight and Orthogonality of Composite Operators

The entries of the normalization matrix I give a measure of both the *weight* of a composite operator (diagonal entries) and the degree of *orthogonality* between two of them (off-diagonal entries). These two concepts (weight and orthogonality) have obvious meanings for canonical electronic operators whose anticommutation relations are just $\mathbb{C}$-numbers ($\{c_{l\sigma}, c^\dagger_{l'\sigma'}\} = \delta_{\sigma\sigma'}\delta_{ll'}$): an electronic operator always *weights* just 1 and is *orthogonal* by definition to any other electronic operator, clearly describing a single-particle state with different quantum numbers (l, σ). As for composite operators, owing to the non-canonical commutation relations they obey, these concepts become more relevant and less obvious: both properties strongly depend on the actual operatorial form of the composite operators as well as, through the thermal averaging, on the Hamiltonian describing the system under analysis and the boundary conditions. Such a dependence manifests itself through the emergence, in the I matrix entries, of thermal averages of bosonic composite operators, whose determination is not always straightforward. Actually, it is just the presence of these parameters in the I matrix, to leave sometimes a problem unsolved

although a closed basis has been individuated. If the basis is not closed ($\delta J \neq 0$), the same kind of averages appears in the ε matrix too.

These features of composite operators (non-trivial weights and imperfect orthogonality) are distinctive landmarks of strong correlations. They are fundamental to the occurrence of well-known effects such as the transfer of spectral weight (directly connected to composite operator weights) between (sub-)bands on changing model or external (temperature, filling, pressure, external fields, etc.) parameters, and the interplay among scales of energies (connected to the *overlapping* of composite operators).

For instance, if we choose $\psi = (\xi, \eta)$ as (closed) fermionic basis for the toy model (4.1), we would find the following I and ε matrices:

$$
\begin{aligned}
I_{\sigma\sigma'}(\mathbf{i},\mathbf{j}) &= \delta_{\mathbf{ij}}\left[\delta_{\sigma\sigma'}\begin{pmatrix} 1-\langle n_{\bar{\sigma}}(\mathbf{i})\rangle & 0 \\ 0 & \langle n_{\bar{\sigma}}(\mathbf{i})\rangle \end{pmatrix} + \delta_{\sigma\bar{\sigma}'}\langle c^{\dagger}_{\sigma}(\mathbf{i}) c_{\bar{\sigma}}(\mathbf{i})\rangle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] \\
\varepsilon_{\sigma\sigma'}(\mathbf{i},\mathbf{j}) &= \delta_{\mathbf{ij}}\delta_{\sigma\sigma'}\begin{pmatrix} -\mu & 0 \\ 0 & U-\mu \end{pmatrix}.
\end{aligned}
\tag{4.10}
$$

As the chosen basis is closed, the energy matrix is free of unknown parameters, while the normalization matrix contains anyway some of them to be computed self-consistently according to the chosen phases to be investigated (e.g., paramagnetic, ferromagnetic, Neél, CDW, SDW, etc.).

4.2.3 Dyson Equation for G

We can now define a generalized matricial Green's function G for the basis ψ

$$
G^{\mathcal{Q}}(i,j) = \left\langle \mathcal{Q}[\psi(i)\psi^{\dagger}(j)]_{\eta}\right\rangle = \begin{cases} \theta(t_i - t_j)\left\langle [\psi(i), \psi^{\dagger}(j)]_{\eta}\right\rangle & \text{for} \quad \mathcal{Q} = \mathcal{R} \\ -\theta(t_j - t_i)\left\langle [\psi(i), \psi^{\dagger}(j)]_{\eta}\right\rangle & \text{for} \quad \mathcal{Q} = \mathcal{A} \\ \theta(t_i - t_j)\left\langle \psi(i)\psi^{\dagger}(j)\right\rangle & \text{for} \quad \mathcal{Q} = \mathcal{C} \\ -\eta\theta(t_j - t_i)\left\langle \psi^{\dagger}(j)\psi(i)\right\rangle & \end{cases},
\tag{4.11}
$$

where $\mathcal{R}$, $\mathcal{A}$ and $\mathcal{C}$ stand for the usual retarded, advanced and causal operators, respectively.

For the sake of simplicity, we pick up a phase with spatial homogeneity and move to momentum and frequency spaces through Fourier transform. Then, according to (4.8), G satisfies the following equation of motion

$$
G^{\mathcal{Q}}(\mathbf{k},\omega) = G_0^{\mathcal{Q}}(\mathbf{k},\omega) + G_0^{\mathcal{Q}}(\mathbf{k},\omega)T^{\mathcal{Q}}(\mathbf{k},\omega)G_0^{\mathcal{Q}}(\mathbf{k},\omega),
\tag{4.12}
$$

where $G_0^{\mathcal{Q}}(\mathbf{k},\omega)$, determined by the equation

$$[\omega - \varepsilon(\mathbf{k})]\, G_0^{\mathcal{Q}}(\mathbf{k},\omega) = I(\mathbf{k}), \tag{4.13}$$

will play the role of fundamental building block and the scattering matrix $T^{\mathcal{Q}}(\mathbf{k},\omega)$ is defined as

$$T^{\mathcal{Q}}(\mathbf{k},\omega) = I^{-1}(\mathbf{k})\, B^{\mathcal{Q}}(\mathbf{k},\omega)\, I^{-1}(\mathbf{k}), \tag{4.14}$$

with $B^{\mathcal{Q}}(i,j) = \left\langle \mathcal{Q}[\delta J(i)\delta J^{\dagger}(j)]_{\eta} \right\rangle$.

Now, we can introduce the irreducible self-energy $\Sigma^{\mathcal{Q}}(\mathbf{k},\omega)$, defined through the relation $T^{\mathcal{Q}}(\mathbf{k},\omega)G_0^{\mathcal{Q}}(\mathbf{k},\omega) = I^{-1}(\mathbf{k})\,\Sigma^{\mathcal{Q}}(\mathbf{k},\omega)G^{\mathcal{Q}}(\mathbf{k},\omega)$, and get

$$G^{\mathcal{Q}}(\mathbf{k},\omega) = \frac{1}{\omega - \varepsilon(\mathbf{k}) - \Sigma^{\mathcal{Q}}(\mathbf{k},\omega)} I(\mathbf{k}) = \frac{1}{[G_0^{\mathcal{Q}}(\mathbf{k},\omega)]^{-1} - I^{-1}(\mathbf{k})\,\Sigma^{\mathcal{Q}}(\mathbf{k},\omega)}. \tag{4.15}$$

Equation (4.15) generalizes the well-known Dyson equation to the case of Green's functions of composite operators [4, 5]. Together with the apparent similarities, it is absolutely worth noting the striking differences:

- All ingredients are matrices either $n \times n$ (bosonic) or $2n \times 2n$ (fermionic).
- The equivalent of the usual non-interacting Green's function, G_0, is, instead, fully interacting and, in principle, exact (if the basis ψ is closed then we have $\delta J = 0$, consequently $\Sigma = 0$, and finally $G = G_0$).
- The normalization matrix I evidently assumes the role of overall weight of the matricial Green's function G (I is its 0th moment: $\lim_{\omega\to\infty} G \rightsquigarrow I/\omega$).
- The eigenvalues of the energy matrix ε clearly play the fundamental role of poles, which will be more or less severely modified by the self-energy Σ.
- The self-energy Σ describes both the dynamics *orthogonal* to that described by the chosen basis ψ and the *residual* interactions among the elements of the latter. Smarter is the choice of ψ, up to a closed basis, less relevant is the knowledge of Σ to fully understand the system under study. Hence, we decided to rename the generalized irreducible self-energy Σ as *residual self-energy*: it is just the irreducible propagator of the residual current δJ.

4.2.4 Propagator G_0

Let us take a step back, the propagator G_0 satisfies the following equation

$$[\omega - \varepsilon(\mathbf{k})]\, G_0^{\mathcal{Q}}(\mathbf{k},\omega) = I(\mathbf{k}), \tag{4.16}$$

whose most general solution is [4, 5]

$$G_0^{\mathcal{Q}}(\mathbf{k},\omega) = \sum_{l=1}^{n} \left\{ \mathcal{P}\left[\frac{\sigma^{(l)}(\mathbf{k})}{\omega - \omega_l(\mathbf{k})}\right] - \mathrm{i}\pi\delta\left[\omega - \omega_l(\mathbf{k})\right] g^{(l)\mathcal{Q}}(\mathbf{k}) \right\}. \tag{4.17}$$

$\sigma^{(l)}(\mathbf{k})$ are the spectral density functions in matrix form, fully determined by the energy matrix ε and the normalization matrix I as

$$\sigma_{ab}^{(l)}(\mathbf{k}) = \Omega_{al}(\mathbf{k}) \sum_{c} \Omega_{lc}^{-1}(\mathbf{k}) I_{cb}(\mathbf{k}), \tag{4.18}$$

where $\Omega(\mathbf{k})$ is the matrix whose columns are the eigenvectors of the energy matrix ε. $\omega_l(\mathbf{k})$ are the eigenvalues of this latter. $g^{(l)\mathcal{Q}}(\mathbf{k})$ are unknown momentum functions, in matrix form, not fixed by the equations of motion (i.e., they correspond to constants in time), to be determined taking into account the boundary conditions specific of the type of propagator under study (retarded, advanced, or causal).

For both fermionic ($\eta = 1$) and bosonic ($\eta = -1$) basis, the boundary conditions turn out to be sufficient to fully determine $g^{(l)\mathcal{R},\mathcal{A}}(\mathbf{k})$ and we obtain a Lindhard-like representation for $G_0^{\mathcal{R},\mathcal{A}}(\mathbf{k},\omega)$

$$G_0^{\mathcal{R},\mathcal{A}}(\mathbf{k},\omega) = \sum_{l=1}^{n} \frac{\sigma^{(l)}(\mathbf{k})}{\omega - \omega_l(\mathbf{k}) \pm \mathrm{i}\delta}. \tag{4.19}$$

Contrarily, $g^{(l)\mathcal{C}}(\mathbf{k})$ and, consequently, the causal Green's function, $G_0^{\mathcal{C}}(\mathbf{k},\omega)$, can be fully determined by the boundary conditions only for the fermionic basis

$$G_0^{\mathcal{C}}(\mathbf{k},\omega) = \sum_{l=1}^{n} \sigma^{(l)}(\mathbf{k}) \left[\frac{1 - f_{\mathrm{F}}(\omega)}{\omega - \omega_l(\mathbf{k}) + \mathrm{i}\delta} + \frac{f_{\mathrm{F}}(\omega)}{\omega - \omega_l(\mathbf{k}) - \mathrm{i}\delta}\right], \tag{4.20}$$

where $f_{\mathrm{F}}(\omega)$ is the Fermi distribution function.

We conclude the discussion about fermionic basis reporting the expression for the correlation function $C(i,j) = \langle \psi(i)\psi^{\dagger}(j) \rangle$

$$C(\mathbf{k},\omega) = 2\pi \sum_{l=1}^{n} \left[1 - f_{\mathrm{F}}(\omega_l(\mathbf{k}))\right] \sigma^{(l)}(\mathbf{k}) \delta\left[\omega - \omega_l(\mathbf{k})\right]. \tag{4.21}$$

It is necessary noticing that the correlation function C can be computed by means of (4.21) only if the basis ψ is closed or if we neglect the residual self-energy. Otherwise, we should use the more general expression

$$C(\mathbf{k},\omega) = \mp 2\left[1 - f_{\mathrm{F}}(\omega)\right] \Im\left[G^{\mathcal{R},\mathcal{A}}(\mathbf{k},\omega)\right]. \tag{4.22}$$

For bosonic basis, instead, $g^{(l)\mathcal{C}}(\mathbf{k})$ cannot be fully determined by the boundary conditions, but we can still obtain a modified Lindhard-like representation for $G_0^{\mathcal{C}}(\mathbf{k},\omega)$ [4,5]

$$G_0^{\mathcal{C}}(\mathbf{k},\omega) = -2\pi \mathrm{i}\Gamma(\mathbf{k})\delta(\omega) + \sum_{l=1}^{n}\sigma^{(l)}(\mathbf{k})\left[\frac{1+f_{\mathrm{B}}(\omega)}{\omega-\omega_l(\mathbf{k})+\mathrm{i}\delta} - \frac{f_{\mathrm{B}}(\omega)}{\omega-\omega_l(\mathbf{k})-\mathrm{i}\delta}\right], \tag{4.23}$$

where $f_{\mathrm{B}}(\omega)$ is the Bose distribution function, provided that we explicitly keep in the final expression an unknown zero frequency momentum function in matrix form $\Gamma(\mathbf{k})$.

The actual value of $\Gamma(\mathbf{k})$ is directly related to the degree of ergodicity of the dynamics of the basis ψ driven by the Hamiltonian H

$$\Gamma(\mathbf{i}-\mathbf{j}) = \lim_{\mathsf{T}\to\infty}\frac{1}{T}\int_0^{\mathsf{T}} \mathrm{d}t\,\left\langle \psi(\mathbf{i},0)\psi^{\dagger}(\mathbf{j},t)\right\rangle. \tag{4.24}$$

If we would know, by any source of information, that the dynamics of ψ driven by H is fully ergodic, the last equation would simply give

$$\Gamma(\mathbf{i}-\mathbf{j}) = \langle\psi(\mathbf{i})\rangle\left\langle\psi^{\dagger}(\mathbf{j})\right\rangle, \tag{4.25}$$

leading to an effective removal of the unknown Γ function from the theory. Unfortunately, although many people just overlook this deliberately, there is absolutely no way to asses ergodicity in advance and $\Gamma(\mathbf{k})$ forms another block of unknowns in the theory. In particular, ergodicity cannot be supposed a priori for finite systems treated in statistical ensembles different from the microcanonical one and for bosonic operators commuting with the Hamiltonian (i.e., for constants of motion).

$\Gamma(\mathbf{k})$ also appears in the expression for the correlation function $C(i,j) = \left\langle\psi(i)\psi^{\dagger}(j)\right\rangle$

$$C(\mathbf{k},\omega) = 2\pi\Gamma(\mathbf{k})\delta(\omega) + 2\pi\sum_{l=1}^{n}\left[1+f_{\mathrm{B}}(\omega_l(\mathbf{k}))\right]\sigma^{(l)}(\mathbf{k})\delta\left[\omega-\omega_l(\mathbf{k})\right], \tag{4.26}$$

but this is absolutely not the only issue with the above expression (4.26).

If $\sigma^{(l)}(\mathbf{k})$ is not identically zero for all values of $\mathbf{k}$ where $\omega_l(\mathbf{k})$ vanishes, the correlation function $C(\mathbf{k},\omega)$ diverges as $1/\beta\omega_l(\mathbf{k})$ for the same values of $\mathbf{k}$ in the limit $\omega \to 0$ (i.e., at all times). Such a behavior of $C(\mathbf{k},\omega)$ usually manifests the establishing of long-range spatial correlations in the system, but it is admissible iff the corresponding correlation function in real space $C(\mathbf{r},t)$ stays always finite. Accordingly, the divergence should be integrable; this immediately excludes the possibility to have long-range order in finite systems and, at finite temperatures, in infinite systems (i.e., in the thermodynamic limit) with too low spatial dimension.

Actually, the lowest spatial dimension allowed to host long-range order will simply depend on the actual functional form of the vanishing $\omega_l(\mathbf{k})$ (e.g., if $\omega_l(\mathbf{k}) \propto |\mathbf{k}|^\alpha$, then the spatial dimension has to be strictly larger than α). No restriction applies to infinite systems of any dimension at zero temperature. This is just the content of Mermin–Wagner theorem [26].

For bosonic basis too, it is necessary noticing that the correlation function C can be computed by means of (4.26) only if the basis ψ is closed or if we neglect the residual self-energy. Otherwise, we should use the more general expression

$$C(\mathbf{k},\omega) = -2\frac{1+f_{\mathrm{B}}(\omega)}{1+2f_{\mathrm{B}}(\omega)}\Im\left[G^{\mathcal{C}}(\mathbf{k},\omega)\right]. \tag{4.27}$$

The use of the casual Green's function, for a bosonic basis, is strictly necessary in order to properly take into account the ergodicity issue that does not manifest in neither the advanced nor the retarded propagators, but only in the causal propagator and in the correlation function.

4.2.5 Residual Self-Energy $\boldsymbol{\Sigma}$

Obviously, we do not need to compute the residual self-energy at all (it is just identically zero) if we choose a closed basis; this is one of the main reasons why a closed basis is the best one we can choose. Accordingly, as much as a truncated basis is *large* enough less relevant will be the contribution of the residual self-energy to the description of the system under analysis. It is worth noticing that even if we would completely neglect the residual self-energy, the Green's function of the original interacting particles constituting the system, expressed in terms of the relevant entries of the propagator G_0, will anyway feature a fully momentum and frequency dependent irreducible self-energy with a $(n-1)$-polar structure, but sufficient to describe all scales of energy caught by the chosen basis.

Nevertheless, one cannot neglect the residual self-energy without being aware of the main drawback: one is actually promoting to the rank of *true* particle (i.e., with an infinite life-time) objects that, being still subject to all virtual processes not properly taken into account by a truncated basis, have finite life-times (i.e., they are quasiparticle) roughly inversely proportional to the largest neglected energy scale involving them (i.e., the transition described by a composite operator can or cannot be one of those necessary to construct the relevant virtual process). Clearly, this can be systematically controlled by enlarging the basis, although only on a quantitative level. Again, not-analytical energy scales can change so dramatically the properties of particles and quasiparticles to require a qualitative change of perspective also regarding the residual self-energy determination and role, as described in the Sect. 4.2.1.

As a matter of fact, it is sometimes convenient keeping the operatorial basis somewhat simpler than what is actually handleable in order to get simpler expressions for the residual currents too and effectively compute, starting from the latter, the residual self-energy. Such a procedure can lead to a description of the system under analysis featuring both some of the relevant energy scales and the decay effects inherent to the quasiparticle nature of composite operators belonging to a truncated basis.

In these years, we have been developing and testing different ways to compute residual self-energy; among others: the two-site resolvent approach [27, 28] and the non-crossing approximation (NCA) [29–36]. Any of them has its pros and cons and specific problems can be better tackled by one or the other (see Sect. 4.3).

4.2.6 Self-Consistency

The intrinsic complexity of the operatorial algebra obeyed by composite operators (for instance, see Table 4.1) hides a noteworthy possible exploitation of the same algebra. Composite operators, whose lattice sites overlap (any composite operator may span a certain number of lattice sites), obey algebra constraints directly coming from the Pauli exclusion principle. For instance, ξ and η satisfy the following exact relation: $\xi_\sigma(\mathbf{i})\eta^\dagger_{\sigma'}(\mathbf{i}) = 0$ that can be easily traced back to $c_\sigma(\mathbf{i})c_\sigma(\mathbf{i}) = 0$.

The relevance of such algebra constraints resides, on one side, in their capability to enforce the Pauli principle and its derivatives, and, on the other side, in the not-so-trivial request that they should be obeyed at the level of thermal averages too. In fact, the related thermal averages (e.g., $\langle \xi_\sigma(\mathbf{i})\eta^\dagger_{\sigma'}(\mathbf{i})\rangle = 0$), through the well-known relation existing between correlation functions and Green's functions (fluctuaction–dissipation theorem; see Sect. 4.2.4), depend on all unknown parameters appearing in the relevant Green's function (in I, ε, and Σ plus $\Gamma(\mathbf{k})$ for bosonic basis) and, in turn, can be used to fix them:

$$\langle \psi(\mathbf{i})\psi^\dagger(\mathbf{j})\rangle = C(\mathbf{i}-\mathbf{j}, t=0). \tag{4.28}$$

The l.h.s of (4.28), for $\mathbf{i}$ and $\mathbf{j}$ such that elements of the basis ψ span on, at least, one common site, would be fixed by the algebra, which imposes *contractions* (e.g., $\xi_\sigma(\mathbf{i})\eta^\dagger_{\sigma'}(\mathbf{i}) = 0$, $n(\mathbf{i})n(\mathbf{i}) = n(\mathbf{i}) + 2\eta^\dagger_\uparrow(\mathbf{i})\eta_\uparrow(\mathbf{i})$). The r.h.s. of (4.28) is given by the actual expression of the Green's function and contains all unknowns of the theory (for bosonic basis, it also contains $\Gamma(\mathbf{i}-\mathbf{j})$).

This deceptively simple conclusion has enormous implications on both the capability to solve, either exact or approximately, strongly correlated systems, and the *quality* of the solution. Not only algebra constraints allow us to find a solution, but they make this solution as closer as possible to the exact one because they embody the primary cause of electronic correlations: the Pauli principle.

It is also worth noticing that the symmetries enjoined by the Hamiltonian imply the existence of constants of motion and the possibility to formulate relations among matrix elements of the relevant Green's functions known as Ward–Takahashi identities [37, 38]. These identities can/should also be used to fix the unknowns and to constrain the theory.

4.2.7 Summary

Summarizing, COM framework envisages four main steps:

1. Choose a composite operator basis according to the system under analysis and all information we can gather from relevant numerical and exact solutions (see Sect. 4.2.1)
2. Compute I and m matrices and obtain ε matrix and propagator G_0 in terms of unknown correlators and unknown Γ function (see Sect. 4.2.1 and 4.2.4).
3. Choose a recipe to compute Σ or just neglect it (see Sect. 4.2.5).
4. Self-consistently compute the unknowns through algebra constraints and Ward–Takahashi identities (see Sect. 4.2.6).

In the last 15 years, COM has been applied to several models and materials: Hubbard [4, 5, 9, 17, 28, 32, 36, 39–42], p–d [43–45], t–J [9, 27], t–t'–U [46–48], extended Hubbard (t–U–V) [49, 50], Double-exchange [22], Kondo [18], Anderson [19], Kondo–Heisenberg [51], two-orbital Hubbard [52], Kondo lattice [53], Ising [10–14, 16], BEG [54], Heisenberg [55], Hubbard–Kondo [56], singlet-hole [57], cuprates [31, 58–60]. A comparison with the results of numerical simulations has been systematically carried on. The interested reader may refer to the works cited in [5] and, for the last years, at the web page: http://scs.physics.unisa.it.

In the second part of this chapter, as relevant application of the formalism, we will consider the Hubbard model and we will go through the different approximation schemes illustrated in the previous sections in a systematic way. A comprehensive comparison with the results of numerical simulations and with the experimental data for high-T_c cuprates will be also reported.

4.3 Case Study: The Hubbard Model

4.3.1 The Hamiltonian

The Hubbard model reads

$$H = \sum_{\mathbf{ij}} \left(-\mu\delta_{\mathbf{ij}} - 2\mathrm{d}t\alpha_{\mathbf{ij}}\right) c^{\dagger}(i)c(j) + U \sum_{\mathbf{i}} n_{\uparrow}(i)n_{\downarrow}(i). \tag{4.29}$$

The notation is the same as used in Sect. 4.1 with the following additions: t is the hopping and the energy unit; d is the dimensionality of the system; $\alpha_{\mathbf{ij}}$ is the projector on the nearest-neighbor sites, whose Fourier transform, for a d-dimensional cubic lattice with lattice constant a, is $\alpha(\mathbf{k}) = \frac{1}{d}\sum_{n=1}^{d}\cos(k_n a)$. The electronic operators $c(i)$ and $c^\dagger(i)$, as well as all other fermionic operators, are expressed in the spinorial notation [e.g., $c^\dagger(i) = \left(c^\dagger_\uparrow(i) \quad c^\dagger_\downarrow(i)\right)$. A detailed and comprehensive summary of the properties of Hamiltonian (4.29) are given in Sect. 1 of [5]. We here report a study of this model by means of the COM as formulated in Sect. 4.2. In order to proceed in a pedagogical way, we will go through different stages. In Sect. 4.3.1.1, we consider a truncated fermionic basis (two-pole approximation) given by the two Hubbard operators $\xi(i)$ and $\eta(i)$. In Sect. 4.3.1.2, we complement the analysis considering the related bosonic sector (charge and spin). In Sect. 4.3.2, we implement the study by considering the contribution of the residual self-energy. In Sect. 4.3.3, we enlarge the truncated basis by including higher-order composite fields (four-pole approximation). At all these stages, we will present COM results and compare them with experimental and/or simulation data. Due to the pedagogical nature of this manuscript, we will restrict the analysis to the paramagnetic state. Ordered phases (ferro and antiferromagnetic phases) have been also analyzed and the related results can be found in [5]. Finally, in Sect. 4.3.4, we will discuss the superconducting solution of the model in the d-wave channel and compare COM results with high-T_c cuprates experimental data.

4.3.1.1 Two-Pole Solution: Fermionic Sector

On the basis of the discussions reported in Sects. 4.1 and 4.2.1, we will adopt as fermionic basis

$$\psi(i) = \begin{pmatrix} \xi(i) \\ \eta(i) \end{pmatrix}, \tag{4.30}$$

where $\xi(i)$ and $\eta(i)$ are the Hubbard operators defined in Sect. 4.1. This field satisfies the equation of motion

$$\mathrm{i}\frac{\partial}{\partial t}\psi(i) = J(i) = \begin{pmatrix} -\mu\xi(i) - 2\mathrm{d}t\, c^\alpha(i) - 2\mathrm{d}t\,\pi(i) \\ (U-\mu)\eta(i) + 2\mathrm{d}t\,\pi(i) \end{pmatrix}, \tag{4.31}$$

with $n_\mu(i) = c^\dagger(i)\sigma_\mu c(i)$ and $\pi(i) = \frac{1}{2}\sigma^\mu n_\mu(i)c^\alpha(i) + c(i)c^{\dagger\alpha}(i)c(i)$. $\sigma^\mu = (-\mathbf{1}, \boldsymbol{\sigma})$ and $\sigma_\mu = (\mathbf{1}, \boldsymbol{\sigma})$, where $\boldsymbol{\sigma}$ are the Pauli matrices. Hereafter, given a generic operator $\Phi(i)$, we will use the notation $\Phi^\alpha(i) = \sum_{\mathbf{j}} \alpha_{\mathbf{ij}}\Phi(\mathbf{j}, t)$. The current $J(i)$ is projected on the basis (4.30) and the residual current $\delta J(i)$ is neglected. The normalization and energy matrices have the expressions

$$I(\mathbf{k}) = \begin{pmatrix} I_{11} & 0 \\ 0 & I_{22} \end{pmatrix} = \begin{pmatrix} 1-n/2 & 0 \\ 0 & n/2 \end{pmatrix} \quad \varepsilon(\mathbf{k}) = \begin{pmatrix} m_{11}(\mathbf{k}) I_{11}^{-1} & m_{12}(\mathbf{k}) I_{22}^{-1} \\ m_{12}(\mathbf{k}) I_{11}^{-1} & m_{22}(\mathbf{k}) I_{22}^{-1} \end{pmatrix}, \tag{4.32}$$

where the entries of the m matrix are given by

$$m_{11}(\mathbf{k}) = -\mu I_{11} - 2\mathrm{d}t\,[\Delta + \alpha(\mathbf{k})(1-n+p)] \tag{4.33}$$

$$m_{12}(\mathbf{k}) = 2\mathrm{d}t\,[\Delta + \alpha(\mathbf{k})(p - I_{22})] \tag{4.34}$$

$$m_{22}(\mathbf{k}) = (U-\mu) I_{22} - 2\mathrm{d}t\,[\Delta + \alpha(\mathbf{k})\, p]\,, \tag{4.35}$$

$n = 1/N \sum_{\mathbf{i}} \langle n(\mathbf{i}) \rangle$ is the particle number per site; the parameters Δ and p cause a constant shift of the bands and a bandwidth renormalization, respectively, and are defined as

$$\Delta = \langle \xi^{\alpha}(\mathbf{i}) \xi^{\dagger}(\mathbf{i}) \rangle - \langle \eta^{\alpha}(\mathbf{i}) \eta^{\dagger}(\mathbf{i}) \rangle \tag{4.36}$$

$$p = \frac{1}{4} \langle n_{\mu}^{\alpha}(\mathbf{i}) n_{\mu}(\mathbf{i}) \rangle - \langle \left[c_{\uparrow}(\mathbf{i}) c_{\downarrow}(\mathbf{i}) \right]^{\alpha} c_{\downarrow}^{\dagger}(\mathbf{i}) c_{\uparrow}^{\dagger}(\mathbf{i}) \rangle. \tag{4.37}$$

Δ is the difference between upper and lower intra-subband contributions to kinetic energy, while p is a combination of the nearest-neighbor charge–charge, spin–spin and pair–pair correlation functions.

The retarded $G^{\mathcal{R}}(i,j) = \left\langle \mathcal{R}\left[\psi(i) \psi^{\dagger}(j) \right] \right\rangle$ and the correlation $C(i,j) = \left\langle \psi(i) \psi^{\dagger}(j) \right\rangle$ functions are given by

$$G^{\mathcal{R}}(\mathbf{k}, \omega) = \sum_{n=1}^{2} \frac{\sigma^{(n)}(\mathbf{k})}{\omega - E_n(\mathbf{k}) + \mathrm{i}\delta} \tag{4.38}$$

$$C(\mathbf{k}, \omega) = \sum_{n=1}^{2} [1 - f_{\mathrm{F}}(\omega)]\, \sigma^{(n)}(\mathbf{k}) \delta[\omega - E_n(\mathbf{k})]. \tag{4.39}$$

The energy spectra $E_n(\mathbf{k})$ read

$$E_1(\mathbf{k}) = R(\mathbf{k}) + Q(\mathbf{k}) \qquad E_2(\mathbf{k}) = R(\mathbf{k}) - Q(\mathbf{k}), \tag{4.40}$$

where

$$R(\mathbf{k}) = \frac{1}{2}\,[U - 2\mu - 4\mathrm{d}t\,\alpha(\mathbf{k})] - \frac{m_{12}(\mathbf{k})}{2 I_{11} I_{22}} \qquad Q(\mathbf{k}) = \frac{1}{2} \sqrt{g^2(\mathbf{k}) + \frac{4 m_{12}^2(\mathbf{k})}{I_{11} I_{22}}}, \tag{4.41}$$

with $g(\mathbf{k}) = -U + \frac{1-n}{I_{11}I_{22}} m_{12}(\mathbf{k})$. The spectral functions $\sigma^{(n)}(\mathbf{k})$ have the following expressions

$$\begin{aligned} \sigma_{11}^{(1)}(\mathbf{k}) &= \frac{I_{11}}{2}\left[1 + \frac{g(\mathbf{k})}{2Q(\mathbf{k})}\right] & \sigma_{11}^{(2)}(\mathbf{k}) &= \frac{I_{11}}{2}\left[1 - \frac{g(\mathbf{k})}{2Q(\mathbf{k})}\right] \\ \sigma_{12}^{(1)}(\mathbf{k}) &= \frac{m_{12}(\mathbf{k})}{2Q(\mathbf{k})} & \sigma_{12}^{(2)}(\mathbf{k}) &= -\frac{m_{12}(\mathbf{k})}{2Q(\mathbf{k})} \\ \sigma_{22}^{(1)}(\mathbf{k}) &= \frac{I_{22}}{2}\left[1 - \frac{g(\mathbf{k})}{2Q(\mathbf{k})}\right] & \sigma_{22}^{(2)}(\mathbf{k}) &= \frac{I_{22}}{2}\left[1 + \frac{g(\mathbf{k})}{2Q(\mathbf{k})}\right] \end{aligned} \quad . \tag{4.42}$$

The determination of $G^{\mathcal{R}}(\mathbf{k}, \omega)$ and $C(\mathbf{k}, \omega)$ requires the knowledge of the chemical potential and of the two bosonic thermal averages Δ and p, a fermionic correlator and a bosonic one, respectively. These quantities are self-consistently determined by means of the following set of coupled equations

$$n = 2(1 - C_{11} - C_{22}) \tag{4.43}$$

$$\Delta = C_{11}^{\alpha} - C_{22}^{\alpha} \tag{4.44}$$

$$C_{12} = \langle \xi(\mathbf{i})\eta^{\dagger}(\mathbf{i})\rangle = 0, \tag{4.45}$$

where $C_{ab} = \langle \psi_a(\mathbf{i})\psi_b^{\dagger}(\mathbf{i})\rangle$ and $C_{ab}^{\alpha} = \langle \psi_a^{\alpha}(\mathbf{i})\psi_b^{\dagger}(\mathbf{i})\rangle$. The first equation fixes the chemical potential in terms of all other parameters; the second comes from the definition of Δ [cf. (4.37)]; the third comes from the constraint (4.28). Once this set of coupled self-consistent equations has been solved, we can calculate all relevant single-particle and thermodynamical properties of the model. The double occupancy per site: $D = 1/N \sum_{\mathbf{i}} \langle n_{\uparrow}(\mathbf{i}) n_{\downarrow}(\mathbf{i})\rangle = I_{22} - C_{22}$; the energy bands $E_n(\mathbf{k})$; the Fermi surface by means of the equation $E_n(\mathbf{k}) = 0$; the momentum distribution function $n(\mathbf{k})$ and the density of states $N(\omega)$:

$$n(\mathbf{k}) = 2\int_{-\infty}^{+\infty} \mathrm{d}\omega f_{\mathrm{F}}(\omega)\left[-\frac{1}{\pi}\Im[G_{cc}^{\mathcal{R}}(\mathbf{k}, \omega)]\right] \tag{4.46}$$

$$N(\omega) = \frac{a^d}{(2\pi)^d}\int_{\Omega_{\mathrm{B}}} \mathrm{d}^d k \left[-\frac{1}{\pi}\Im[G_{cc}^{\mathcal{R}}(\mathbf{k}, \omega)]\right], \tag{4.47}$$

where Ω_{B} is the volume of the first Brillouin zone, $G_{cc}^{\mathcal{R}}(\mathbf{k}, \omega) = \sum_{a,b=1}^{2} G_{ab}^{\mathcal{R}}(\mathbf{k}, \omega)$; the internal energy $E = \langle H\rangle = 4\mathrm{d}t \sum_{a,b=1}^{2} C_{ab}^{\alpha} + UD$; the specific heat $C = \mathrm{d}E/\mathrm{d}T$; the free energy $F(n, T) = \int_0^n \mu(n', T)\mathrm{d}n'$; the entropy $S = (E - F)/T$.

Once the fermionic propagator is known, there are several ways to compute response functions (i.e., the retarded propagators of the two-particle excitations: charge, spin, pair, ...). These techniques are usually based on diagrammatic expansions of the two-particle propagators in terms of the single-particle one. However, when operators with non-canonical commutations are involved, the complicated algebra invalidates the Wick theorem and, consequently, does not allow any simple

extension of decoupling schemes and more involved diagrammatic approximations [61, 62] are needed. Another technique [5, 63], the one-loop approximation for composite operators, has been developed by means of the equations of motion approach. By using this technique it is possible to calculate the causal function $\langle \mathcal{T}[n_\mu(i)n_\mu(j)]\rangle$ and then the charge and spin susceptibilities. We do not report the details of the calculations, which can be found in [5], and summarize the results. For the causal propagators we have

$$\langle \mathcal{T}[n(i)n(j)]\rangle = n^2 - \frac{n(2-n)}{n-2D}\sum_{a,b,c=1}^{2} I_{aa}^{-1} Q^{\mathcal{C}}_{abac}(i,j) \tag{4.48}$$

$$\langle \mathcal{T}[n_3(i)n_3(j)]\rangle = -\frac{n(2-n)}{n+2D-n^2}\sum_{a,b,c=1}^{2} I_{aa}^{-1} Q^{\mathcal{C}}_{abac}(i,j), \tag{4.49}$$

where $Q^{\mathcal{C}}_{abcd}(i,j) = G^{\mathcal{C}}_{ab}(i,j)G^{\mathcal{C}}_{cd}(j,i)$ is the fermionic loop constructed from the propagator $G^{\mathcal{C}}_{ab}(i,j) = \langle \mathcal{T}[\psi_a(i)\psi_b^\dagger(j)]\rangle$. For the charge and spin susceptibilities, we have

$$\chi_{\mathrm{c}}(\mathbf{k},\omega) = \frac{n(2-n)}{n-2D}\sum_{a,b,c=1}^{2} I_{aa}^{-1} Q^{\mathcal{R}}_{abac}(\mathbf{k},\omega) \tag{4.50}$$

$$\chi_{\mathrm{s}}(\mathbf{k},\omega) = \frac{n(2-n)}{n+2D-n^2}\sum_{a,b,c=1}^{2} I_{aa}^{-1} Q^{\mathcal{R}}_{abac}(\mathbf{k},\omega), \tag{4.51}$$

where $Q^{\mathcal{R}}_{abcd}(\mathbf{k},\omega)$ is the retarded part of $Q^{\mathcal{C}}_{abcd}(\mathbf{k},\omega)$ and has the expression

$$Q^{\mathcal{R}}_{abcd}(\mathbf{k},\omega) = \frac{a^d}{(2\pi)^d}\sum_{n,m=1}^{2}\int \mathrm{d}^d p \frac{\{f_{\mathrm{F}}[E_m(\mathbf{p})] - f_{\mathrm{F}}[E_n(\mathbf{k}+\mathbf{p})]\}\sigma^{(n)}_{ab}(\mathbf{k}+\mathbf{p})\sigma^{(m)}_{cd}(\mathbf{p})}{\omega + E_n(\mathbf{k}+\mathbf{p}) - E_m(\mathbf{p}) + \mathrm{i}\delta} \tag{4.52}$$

Results and Comparisons

Algebra Constraints

The violation of the algebra constraints in the Roth and Hubbard I approximations is analyzed in Fig. 4.1 (left panel), where the normalized Pauli amplitude $A_p = C_{12}/C_{22}$, which is identically zero in COM (as it should be due to the Pauli principle), is reported versus n at $T = 0$ for various values of U. We see that in the Roth [65, 66] and Hubbard I [1] schemes the Pauli principle is satisfied only in

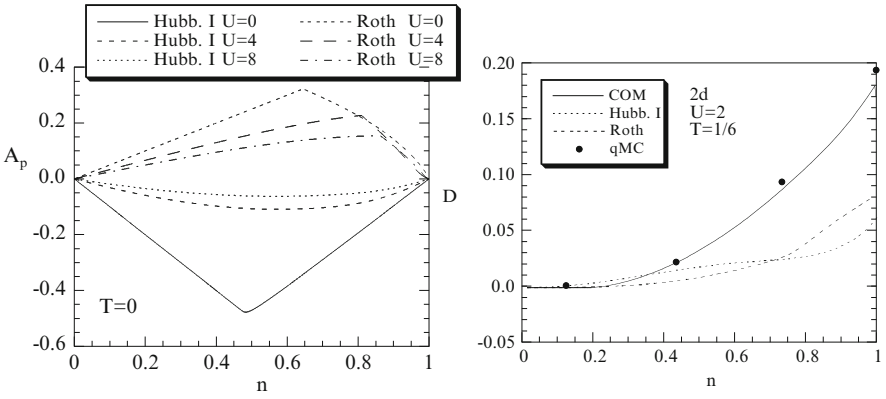

Fig. 4.1 (*Left*) Pauli amplitude $A_p = C_{12}/C_{22}$ vs. filling n at $T = 0$ for several values of U. (*Right*) Double occupancy D vs. filling n for $T = 1/6$ and $U = 2$. The *solid*, *dotted* and *dashed lines* stand for COM, Hubbard I and Roth, respectively. qMC data refer to a 12×12 cluster [64]

the cases $n = 0$ and $n = 1$. The deviation increases by decreasing U, reaching its maximum value in the non-interacting limit (i.e., $U = 0$). On the contrary, the Pauli principle is recovered in the limit $U \to \infty$ for any value of n. At any rate, the Hubbard I solution violates several other sum rules [67]. As a consequence of the fact that the Pauli principle is not satisfied, the particle–hole symmetry enjoined by the Hubbard Hamiltonian is also violated [5, 68]. In order to give a measure of how relevant is fulfilling algebra constraints, and in particular those directly coming from Pauli principle, in the right panel of the same Fig. 4.1, we show COM results for the double occupancy in comparison with the Roth and Hubbard I ones, and the data obtained on finite size clusters by quantum Monte Carlo method [64].

Chemical Potential

In Fig. 4.2 (left panel), we show the particle density versus chemical potential μ for various temperatures. COM results are compared with the Bethe ansatz (BA) ones [69]. The agreement is very good at low temperatures for densities smaller than 0.55. In the half-filled chain, $T \approx t$ is a relevant temperature as it signs the border between T-regions dominated by either spin or charge correlations [71, 72]. The agreement between COM and BA at $T \geq t$ is very good for the whole range of filling. Of course, at high temperatures, COM result reaches an excellent agreement with BA since the effect of correlations is completely suppressed. In Fig. 4.2 (right panel), we show COM results for the chemical potential as a function of the temperature T and of the filling for $U = 4$ in comparison with numerical results obtained by means of the Lanczos technique on a 4×4 cluster [70]. The agreement is rather good in particular for low filling and high temperatures where the numerical results are more reliable and the finite size of the cluster is less relevant.

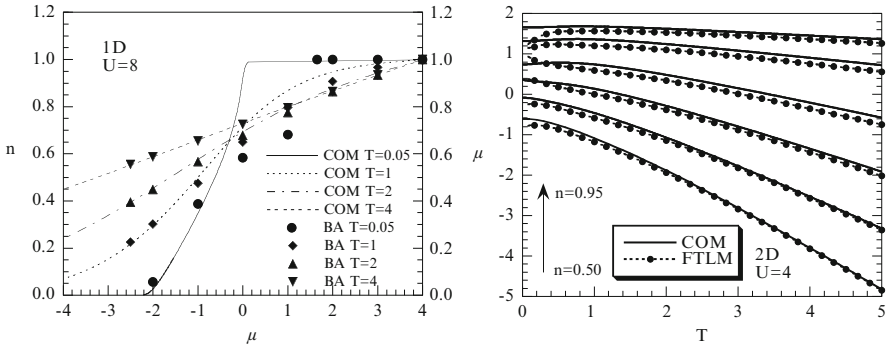

Fig. 4.2 (*Left*) Filling n vs. chemical potential μ for $T = 0.05, 1, 2$, and 4 and $U = 8$. BA result is taken from [69]. (*Right*) Chemical potential μ vs. temperature T for $U = 4$ and $n = 0.5 \rightarrow 0.95$. FTLM result is taken from [70]

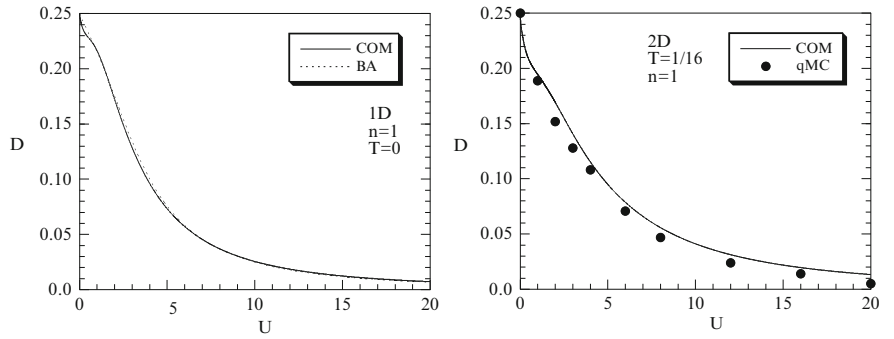

Fig. 4.3 (*Left*) Double occupancy D vs. Coulomb interaction U at $T = 0$ and $n = 1$. *Dotted lines* stand for BA results. (*Right*) Double occupancy D vs. Coulomb interaction U for $T = 1/6$ and $n = 1$, *solid symbols* stand for qMC data (4×4) [73]

Double Occupancy

The double occupancy for the one-dimensional case is shown in Fig. 4.3 (left panel), where COM is plotted versus Coulomb interaction U at $T = 0$ and for $n = 1$. The results are compared with the exact ones by Bethe ansatz. The agreement with BA is excellent. Such a good agreement is not reached by any other analytical approach, neither by the Gutzwiller or the ladder and the self-consistent ladder approximations (SCLA) [74–76]. In particular, these approaches fail to reproduce the correct asymptotic behavior. As shown in Fig. 4.3 (left panel), the double occupancy goes to zero as $U \rightarrow \infty$: the electrons localize only at infinite U, where the half-filled Hubbard chain is equivalent to the spin-$1/2$ AF Heisenberg chain. The double occupancy for the two-dimensional case is shown in Fig. 4.3 (right panel), where COM solution is plotted versus Coulomb interaction U at $T = 1/6$ and for

$n = 1$. The results are compared with the data of numerical simulation by qMC [73]. The agreement is excellent.

Internal energy

The internal energy at $n = 1$ and $T = 0$ is shown as a function of U in Fig. 4.4 (left panel). The results obtained by means of the BA [77] and other analytical approaches [74–76] are also reported. The agreement between COM and BA is excellent. The SCLA [75] shows also a very good agreement for all values of the coupling, but it does not have the correct behavior for an infinite value of the Coulomb interaction. Moreover, both the Ladder (LA) [75] and the Gutzwiller (GWA and GWF) [74] approximations go to zero at finite U, whereas the renormalization group (RG) [76] has the right asymptotic behavior for $U \to \infty$, but it does not reproduce the non-interacting limit. The doping dependence of the internal energy is shown in Fig. 4.4 (right panel) for $U = 4$. For comparison, we also report the BA results [77] and LA and SCLA approaches [75, 76]. COM agrees reasonably with the BA, reaching the best agreement at half-filling. The ladder approximation [75] deviates more and more from the BA as approaching half-filling; the SCLA [75] probes excellently at any doping. In Fig. 4.5 (left panel) we present the internal energy versus temperature in the range $0 \leq T \leq 5$ and $U = 4$ for several values of the filling. The results are compared with finite-temperature Lanczos method (FTLM) data [70]. The agreement is quite remarkable in the entire range of temperature and for all studied dopant concentrations. In Fig. 4.5 (right panel), we report the energy per site as a function of the particle density for different values of T and U and compare COM results with Lanczos data on a 4×4 lattice [70]. The agreement is very good on the whole range of filling and for all values of temperature. It is worth noticing that at $T = 0.1$ the numerical data report a strange kind of oscillation, not present in our results, that will be probably absent in the bulk system. It is worth noticing that many more comparisons, all showing a very good agreement, between COM results and numerical simulation data have

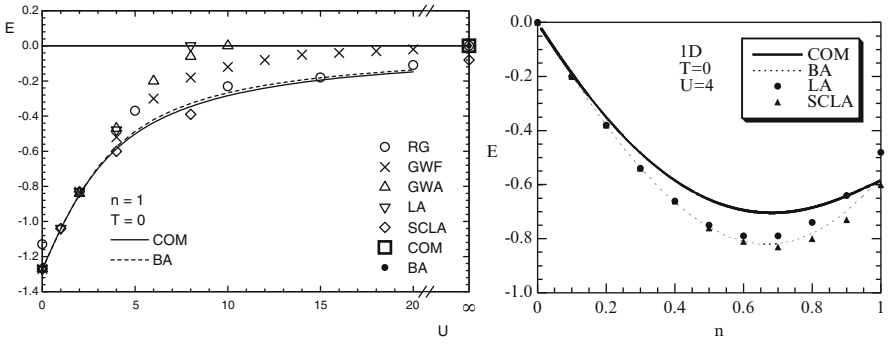

Fig. 4.4 Internal energy E vs.: (*left*) Coulomb interaction U for $T = 0$ and $n = 1$; (*right*) filling n for $T = 0$ and $U = 4$

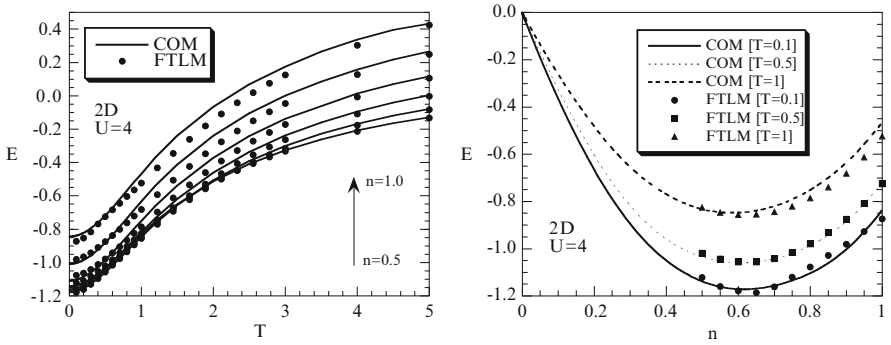

Fig. 4.5 Internal energy E vs.: (*left*) temperature T for $U = 4$ and $n = 0.5 \rightarrow 1$; (*right*) filling n for $U = 4$ and various values of temperature T. Lanczos data (4×4) are taken from [70]

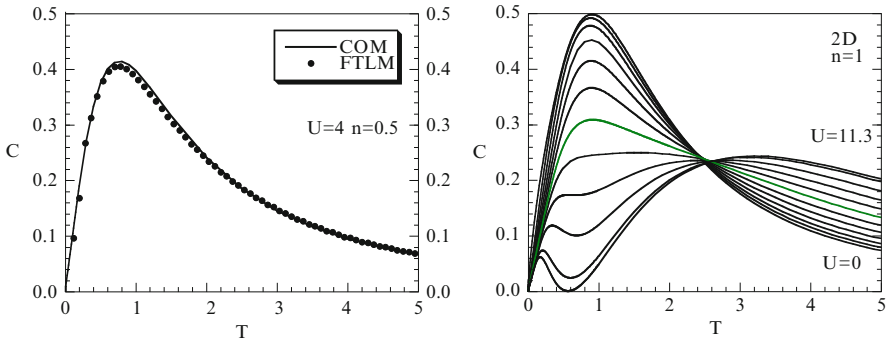

Fig. 4.6 Specific heat C vs. temperature T: (*left*) at $n = 0.5$ and $U = 4$; (*right*) at $n = 1$ and various values of Coulomb interaction U

been presented in [78, 79]. The overall picture emerging from such comparisons is that COM generally gives a very accurate description of the behavior of the internal energy as a function of the external parameters (T, U, and n) over a very wide range of their values.

Specific Heat

In Fig. 4.6 (left panel), we report the specific heat in 2D as a function of the temperature for $n = 0.5$ and $U = 4$. The agreement with FTLM data [P. Prelovšek, Private communication] is excellent in the whole range of temperatures. The presence of two peaks in the specific heat has been attributed to the spin and charge excitations. Our results show a double peak structure too. When U is weak the two peaks overlap and there is no resolution. By increasing U the position of the charge peak moves to higher temperatures and we distinguish the two contributions as shown in Fig. 4.6 (right panel), where C is plotted as a function of T at half-filling

for several values of U. A peak appears at low temperatures when U is rather large ($U \geq 8$). This behavior qualitatively reproduces the qMC results [80]. It is also relevant to observe the presence of a crossing point at a definite temperature. Such crossing points have been experimentally observed in several materials and obtained by means of analytical and numerical studies (see Sect. 3.4.2 of [5] for references and a detailed discussion).

Entropy

In Fig. 4.7 (left panel), the entropy S is reported as a function of the filling at $U = 4$ and various values of temperature T. The numerical data are taken from [70]. The agreement between COM and numerical results is rather good except for $T = 0.2$ at which quite anomalous oscillations appear in the Lanczos data. In Fig. 4.7 (right panel), we report the behavior of the entropy as a function of the temperature in comparison with some Lanczos data [70]. The agreement is very good over the whole range of temperature and for any reported value of the filling, except for low temperatures and low doping. The discrepancy can be related to the capability of our data to describe the effects related to the exchange interaction which is more and more important as low are the temperature and the doping.

4.3.1.2 Two-Pole Solution: Bosonic Sector

There is one more way to tackle the problem of computing the response functions: the two-particle excitations can be considered as a new sector in the dynamics of the system and we can choose also for them a suitable basis alike it has been done for fermions. This approach will be described in this section. We take as bosonic basis the following one [5, 42, 81]

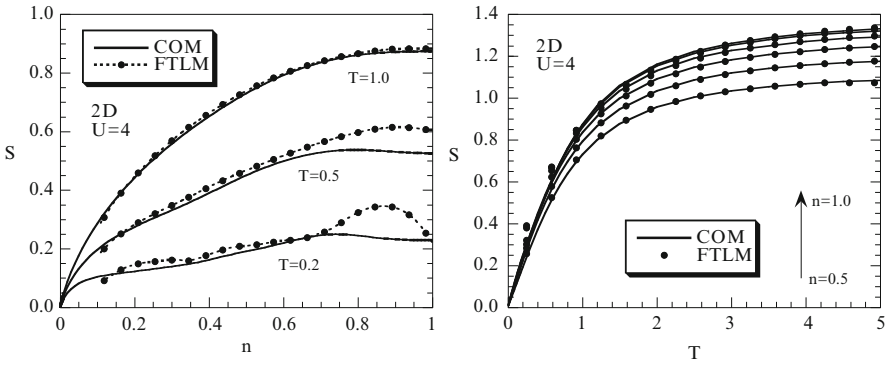

Fig. 4.7 Entropy S vs. filling n for: (*left*) $U = 4$ and various values of temperature T; (*right*) temperature T for $U = 4$ and $n = 0.5 \rightarrow 1$. Lanczos data are taken from [70]

$$N^{(\mu)}(i) = \begin{pmatrix} n_\mu(i) \\ \rho_\mu(i) \end{pmatrix} \qquad \begin{matrix} n_\mu(i) = c^\dagger(i)\sigma_\mu c(i) \\ \rho_\mu(i) = c^\dagger(i)\sigma_\mu c^\alpha(i) - c^{\alpha\dagger}(i)\sigma_\mu c(i) \end{matrix}. \tag{4.53}$$

This basis is directly generated by the hierarchy of the equations of motion; this will assure that the first four bosonic spectral moments have the correct functional form [5, 8]. The field $N^{(\mu)}(i)$ obeys the equation of motion

$$\mathrm{i}\frac{\partial}{\partial t}N^{(\mu)}(i) = J^{(\mu)}(i) = \begin{pmatrix} -2\mathrm{d}t\rho_\mu(i) \\ -2\mathrm{d}t l_\mu(i) + U\kappa_\mu(i) \end{pmatrix}, \tag{4.54}$$

where the higher-order bosonic operators are defined as $\kappa_\mu(i) = c^\dagger(i)\sigma_\mu\eta^\alpha(i) - \eta^\dagger(i)\sigma_\mu c^\alpha(i) + \eta^{\dagger\alpha}(i)\sigma_\mu c(i) - c^{\dagger\alpha}(i)\sigma_\mu\eta(i)$ and $l_\mu(i) = c^\dagger(i)\sigma_\mu c^{\alpha^2}(i) + c^{\dagger\alpha^2}(i)\sigma_\mu c(i) - 2c^{\dagger\alpha}(i)\sigma_\mu c^\alpha(i)$ [5, 42, 81]. We are using the notation $c^{\alpha^2}(\mathbf{i}, t) = \sum_{\mathbf{j}} \alpha^2_{\mathbf{ij}} c(\mathbf{j}, t) = \sum_{\mathbf{jl}} \alpha_{\mathbf{il}}\alpha_{\mathbf{lj}} c(\mathbf{j}, t)$. The current $J^{(\mu)}(i)$ is projected on the basis (4.53) and the residual current $\delta J^{(\mu)}$ is neglected (truncated basis). The normalization and energy matrices have the expressions

$$I^{(\mu)}(\mathbf{k}) = \begin{pmatrix} 0 & I^{(\mu)}_{12}(\mathbf{k}) \\ I^{(\mu)}_{12}(\mathbf{k}) & 0 \end{pmatrix} \quad \varepsilon^{(\mu)}(\mathbf{k}) = \begin{pmatrix} 0 & -2\mathrm{d}t \\ m^{(\mu)}_{22}(\mathbf{k})/I^{(\mu)}_{12}(\mathbf{k}) & 0 \end{pmatrix}, \tag{4.55}$$

where $I^{(\mu)}_{12}(\mathbf{k}) = 4[1 - \alpha(\mathbf{k})]C^\alpha$ and $m^{(\mu)}_{22}(\mathbf{k}) = -2\mathrm{d}t I_{l_\mu\rho_\mu}(\mathbf{k}) + U I_{\kappa_\mu\rho_\mu}(\mathbf{k})$. The parameter C^α is the electronic correlation function $C^\alpha = \langle c^\alpha(\mathbf{i})c^\dagger(\mathbf{i})\rangle$. The quantities $I_{l_\mu\rho_\mu}(\mathbf{k})$ and $I_{\kappa_\mu\rho_\mu}(\mathbf{k})$ are defined as $I_{l_\mu\rho_\mu}(\mathbf{k}) = \mathcal{F}\langle[l_\mu(\mathbf{i}, t), \rho^\dagger_\mu(\mathbf{j}, t)]\rangle$ and $I_{\kappa_\mu\rho_\mu}(\mathbf{k}) = \mathcal{F}\langle[\kappa_\mu(\mathbf{i}, t), \rho^\dagger_\mu(\mathbf{j}, t)]\rangle$. $\mathcal{F}$ stands for the Fourier transform operator. Because of the non-locality of the bosonic composite field (4.53), the analytical form of these quantities depends on the dimensionality of the system and it is necessary to separately discuss the different cases [5].

The causal $G^{C(\mu)}(i, j) = \langle\mathcal{T}[N^{(\mu)}(i)N^{(\mu)\dagger}(j)]\rangle$ and the correlation $C^{(\mu)}(i, j) = \langle N^{(\mu)}(i)N^{(\mu)\dagger}(j)\rangle$ functions are given by

$$G^{C(\mu)}(\mathbf{k}, \omega) = -2\mathrm{i}\pi\Gamma^{(\mu)}(\mathbf{k}) \tag{4.56}$$

$$+ \sum_{n=1}^{2} \sigma^{(n,\mu)}(\mathbf{k}) \left[\frac{1 + f_\mathrm{B}(\omega)}{\omega - \omega^{(\mu)}_n(\mathbf{k}) + \mathrm{i}\delta} - \frac{f_\mathrm{B}(\omega)}{\omega - \omega^{(\mu)}_n(\mathbf{k}) - \mathrm{i}\delta}\right] \tag{4.57}$$

$$C^{(\mu)}(\mathbf{k}, \omega) = 2\pi\Gamma^{(\mu)}(\mathbf{k})\delta(\omega) + 2\pi\sum_{n=1}^{2} \delta[\omega - \omega^{(\mu)}_n(\mathbf{k})][1 + f_\mathrm{B}(\omega)]\sigma^{(n,\mu)}(\mathbf{k}). \tag{4.58}$$

The energy spectra $\omega^{(\mu)}_n(\mathbf{k})$ are given by $\omega^{(\mu)}_n(\mathbf{k}) = (-)^n\sqrt{\varepsilon^{(\mu)}_{12}(\mathbf{k})\varepsilon^{(\mu)}_{21}(\mathbf{k})}$ and the spectral functions $\sigma^{(n,\mu)}(\mathbf{k})$ have the following expression

$$\sigma^{(n,\mu)}(\mathbf{k}) = \frac{I_{12}^{(\mu)}(\mathbf{k})}{2\omega_n^{(\mu)}(\mathbf{k})} \begin{pmatrix} \varepsilon_{12}^{(\mu)}(\mathbf{k}) & 1 \\ 1 & \varepsilon_{21}^{(\mu)}(\mathbf{k}) \end{pmatrix}. \tag{4.59}$$

The determination of $G^{C(\mu)}(\mathbf{k},\omega)$ and $C^{(\mu)}(\mathbf{k},\omega)$ requires the knowledge of a set of parameters: correlation functions of fermionic and bosonic operators and the unknown Γ function. According to the self-consistent scheme given in Sect. 4.2.6, these parameters are fixed by means of the following constraints: (1) all fermionic correlators are calculated within the fermionic sector (see previous subsection); (2) the conservation of the current; (3) the condition that the spin susceptibility be a single value finite function; (4) the local algebra constraint $\langle n_\mu(\mathbf{i})n_\mu(\mathbf{i})\rangle = n + 2(2\delta_{\mu,0} - 1)D$, where D is the double occupancy. The unknown $\Gamma^{(\mu)}(\mathbf{k})$ function can be calculated either by assuming the ergodicity of the system (e.g., $\Gamma_{11}^{(\mu)}(\mathbf{k}) = \delta_{\mu,0}(2\pi/a)^d\delta^{(d)}(\mathbf{k})n^2$) or by opening a new bosonic sector (the pair sector) [see [82]]. Once the self-consistent equations in the fermionic and in the bosonic sectors have been solved, we can calculate the Green's function and the charge and spin susceptibilities $\chi_\mu(\mathbf{k},\omega) = -\mathcal{F}\langle\mathcal{R}[n_\mu(i)n_\mu(j)]\rangle$ as well as the spin correlation functions $S(\mathbf{k}) = \mathcal{F}\langle n_3(\mathbf{i},t)n_3(\mathbf{j},t)\rangle$:

$$\chi_\mu(\mathbf{k},\omega) = -\sum_{n=1}^{2} \frac{\sigma^{(n,\mu)}(\mathbf{k})}{\omega - \omega_n^{(\mu)}(\mathbf{k}) + \mathrm{i}\delta} \tag{4.60}$$

$$S(\mathbf{k}) = -4\mathrm{d}tC^\alpha \frac{1-\alpha(\mathbf{k})}{\omega_2^{(3)}(\mathbf{k})} \coth\frac{\beta\omega_2^{(3)}(\mathbf{k})}{2} \tag{4.61}$$

Results and Comparisons

Spin Correlation Function

The behavior of $S(\mathbf{k})$, as a function of the momentum, is reported in Fig. 4.8 in comparison with some numerical data [83] for different values of filling, Coulomb repulsion, and temperature. We have a very good agreement with the numerical results for sufficiently high values of the doping for all shown values of the Coulomb repulsion. In the proximity of half-filling, the numerical data suffer from a saturation of the antiferromagnetic correlation length [73] that becomes comparable with the size of the cluster. For $U = 4$ and $n = 0.8$ (see Fig. 4.8 (left)), the correlation length is slightly smaller than the size of the cluster: our solution results are capable to describe this situation fairly well (the height of the peak at $\mathbf{Q}$ is exactly reproduced and the momentum dependence is qualitatively correct, again practically exact along the diagonal) except for the exact value of the numerical data along the main axes. This is probably due again to an overestimation of the correlations by the numerical analysis owing to the finite size of the cluster. For $U = 8$ and $n = 1$ (see Fig. 4.8 (right)), in order to reproduce the numerical data we need to increase the temperature so as to decrease our value of the correlation length and match that of the numerical

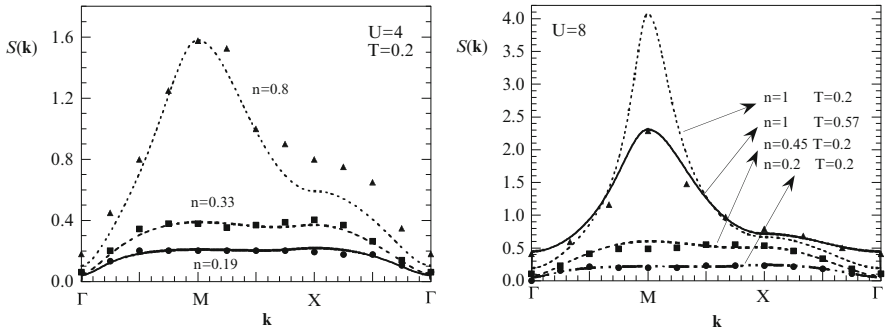

Fig. 4.8 Spin correlation function $S(\mathbf{k})$ vs. momentum $\mathbf{k}$ for: (*left*) $U = 4$, $T = 0.2$ and $n = 0.8$, 0.33, and 0.19; (*right*) $U = 8$, $T = 0.2$ and $T = 0.57$ and $n = 1$, 0.45, and 0.2. qMC data (8×8) are taken from [83]

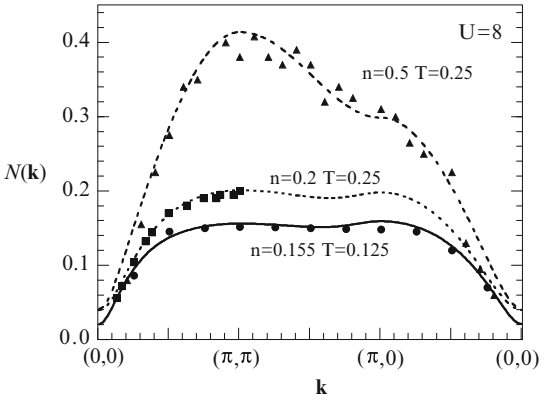

Fig. 4.9 Charge correlation function $N(\mathbf{k})$ vs. momentum $\mathbf{k}$ for $U = 8$, $T = 0.125$ and 0.25 and $n = 0.155$, 0.2, and 0.5; qMC data (8×8, 12×12, 16×16) are from [84]

analysis, which is stuck at the saturation value due to the finiteness of the clusters. The results of such a procedure are astonishing, we manage to exactly reproduce the numerical data for any value of the momentum, and not only at $\mathbf{Q}$, revealing the correctness and power of our approach and the limitations of the numerical analysis.

Charge Correlation Function

$N(\mathbf{k})$ is reported in Fig. 4.9, as a function of the momentum, for various fillings and temperatures and $U = 8$. We have again a very good agreement with quantum Monte Carlo results [84] for all shown values of the external parameters and of the momentum. The enhancement at $\mathbf{k} = \mathbf{Q} = M$ for $n = 0.5$ can be interpreted as the manifestation of a rather weak ordering of the charge with a checkerboard pattern.

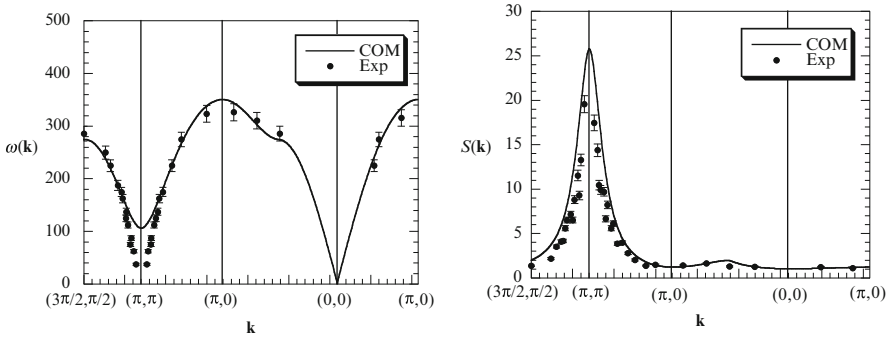

Fig. 4.10 (*Left*) Spin spectrum $\omega(\mathbf{k})$ along the PD for $n = 1$, $T = 10\,\mathrm{K} \approx 0.003t$ and $U = 8.8t$ ($t \approx 0.3\,\mathrm{eV}$). (*Right*) The spin-wave intensity along the PD for $n = 1$, $T = 295\,\mathrm{K} \approx 0.08t$ and $U = 8.8t$ ($t \approx 0.3\,\mathrm{eV}$). Experimental data are taken from [85]

COM results manage to reproduce such double peak structure showing a capability to quantitatively describe, also in a translational invariant phase, rather strong charge correlations.

Comparison with the Experimental Data for La_2CuO_4

In Fig. 4.10 (left), we report the energy spectrum of the spin–spin propagator along the principal directions and compare with the experimental data of [85] obtained for La_2CuO_4 by means of inelastic magnetic neutron scattering. We have fixed the temperature at $T = 10\,\mathrm{K}$ according to the experimental value; the value of the transfer integral ($t = 0.3\,\mathrm{eV}$) and of the Coulomb repulsion ($U = 8.8t$) have been chosen in order to fit the experimental points and they are within the ranges ($t = 0.3 \pm 0.02\,\mathrm{eV}$; $U = 2.2 \pm 0.4\,\mathrm{eV}$) suggested in [85]. The agreement with the experimental data is very good all over the momentum space except around (π, π). As a matter of fact, the experimental data refer to the antiferromagnetic phase of the material [the experimental spectrum gets completely soft at (π, π)] and our paramagnetic solution obviously cannot fully describe such behavior. However, it is worth noting that the Hubbard model at half-filling presents so strong antiferromagnetic correlations that they also show in the paramagnetic phase through a quite deep minimum. The spin-wave intensity is reported in Fig. 4.10 (right) and again compared with the experimental data. The agreement is again very good over all the momentum space and shows once more the capability of the Hubbard model within our formulation to catch the physics of such a strongly correlated system.

Incommensurability

In Fig. 4.11, we present the static spin susceptibility $\chi_s(\mathbf{k})$ for various temperatures. The choice for the parameters is $U = 4$, $t' = -0.1$ and the particle filling has

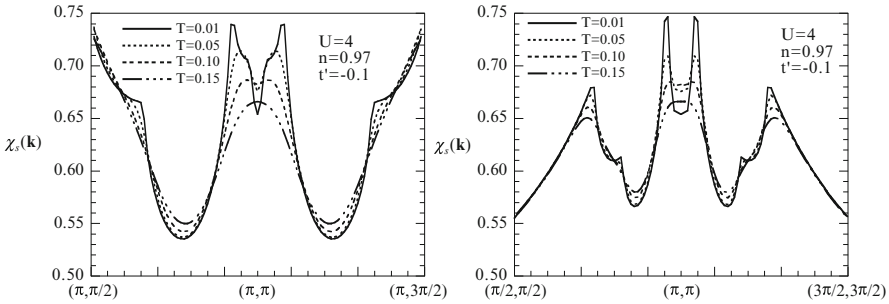

Fig. 4.11 Static spin susceptibility $\chi_s(\mathbf{k})$ for various values of temperature T, $U = 4$, $t' = -0.1$ and $n = 0.97$ along (*left*) $\mathbf{k} = (\pi, k)$ and (*right*) $\mathbf{k} = (k, k)$

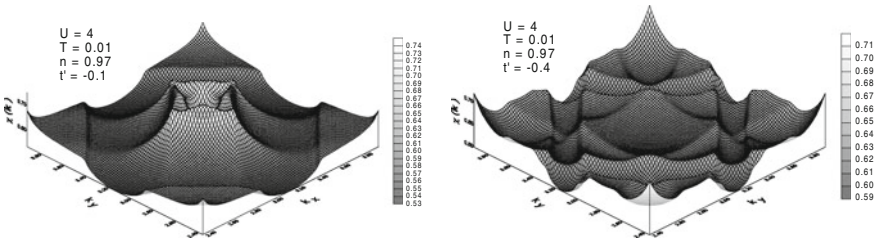

Fig. 4.12 Static spin susceptibility $\chi_s(\mathbf{k})$ for $U = 4$, $n = 0.97$ and $T = 0.01$ and (*left*) $t' = -0.1$ [(*right*) $t' = -0.4$]

been fixed as $n = 0.97$. t' is the next-nearest-hopping integral and the related COM formulation (i.e., for the t–t'–U Hubbard model) in the 2-pole approximation can be found in [40, 46–48, 86–88, 100]. In (left) and (right), $\chi_s(\mathbf{k})$ is reported along the lines $\mathbf{k} = (\pi, k)$ and $\mathbf{k} = (k, k)$, respectively. In both cases by increasing the temperature the incommensurate double-peak structure becomes a broad maximum centered at (π, π). The intensity over the whole Brillouin zone can be clearly observed in Fig. 4.12. The important features of the data are: the overall square symmetry of the scattering with the sides of the square parallel to the (k, k) and to the $(k, -k)$ lines and the accumulation of intensity near the corners of the square. These features reproduce the experimental situation for $La_{2-x}(Ba,Sr)_xCuO_4$ [89–95, 101]. In Fig. 4.13 (left panel), the incommensurability amplitude $\delta(x)$ is shown as a function of doping. For comparison we report the experimental data of [92, 93, 95, 101]. The linear behavior of $\delta(x)$, observed in the low-doping region, agrees exceptionally well with the experimental data. One of the most striking features of the results presented in Fig. 4.13 (left panel) is the resemblance between the incommensurability amplitude $\delta(x)$ and the critical temperature T_c. $\delta(x)$ is maximum in the region of optimal doping where T_c is maximum. It has already been experimentally observed in [95] that there is a linear relation between $\delta(x)$ and T_c up to the optimal doping level $x \approx 0.15$. Our theoretical results confirm this behavior and show that a close similarity between $\delta(x)$ and T_c exists in the entire region

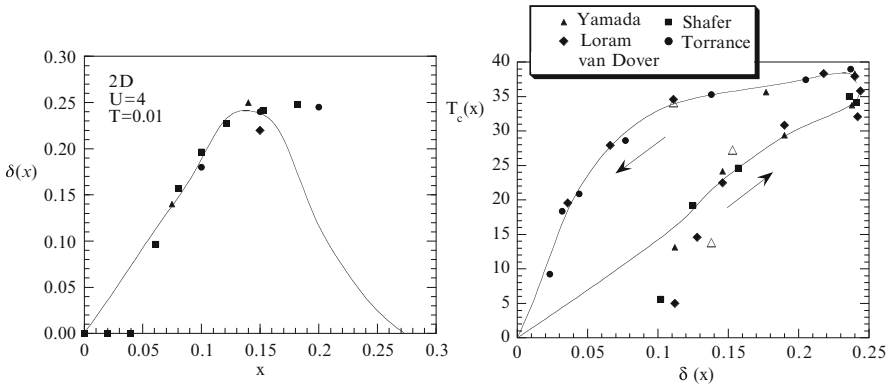

Fig. 4.13 (*Left*) Incommensurability amplitude $\delta(x)$ vs. doping x. COM results (solid line) refer to $U = 4$ and $T = 0.01$. (*Right*) Experimental values of critical temperature $T_c(x)$ for $La_{2-x}Sr_xCuO_4$, taken from Yamada et al. [95], Loram et al. [96], van Dover et al. [97], Shafer et al. [98], and Torrance et al. [99], vs. calculated incommensurability amplitude $\delta(x)$. The *solid line* is a guide to the eyes

of doping. This can be seen in Fig. 4.13 (right panel) where experimental values for T_c, taken from [95–99], are reported versus our calculated incommensurability amplitude $\delta(x)$.

4.3.2 The Residual Self-Energy $\Sigma(\mathbf{k}, \omega)$

We now come to the problem of taking into account the residual self-energy $\Sigma(\mathbf{k}, \omega)$ appearing in the full propagator $G(\mathbf{k}, \omega)$ once the truncated basis (4.30) has been adopted. In particular, we have computed [29–36] $\Sigma(\mathbf{k}, \omega)$ in the NCA by noting that it was possible to exactly rewrite the residual current $\delta J(i)$ as the sums of products of a bosonic operator, embodying either charge $n(i)$, spin $\boldsymbol{n}(i)$ or pair $p(i) = c_\uparrow(i)c_\downarrow(i)$ operator, and of a fermionic operator, which turned to be either $\xi(i)$ or $\eta(i)$. Then, following the prescription of NCA, we expressed the residual self-energy $\Sigma(\mathbf{k}, \omega)$ as sums of products of the charge-charge and spin-spin propagators (we neglected pair-pair propagator) and of the fermionic propagator. The bosonic propagators have been computed in the two-pole approximation as described in the previous section.

The whole framework closes the self-consistency cycle reported in Fig. 4.14, where G_0 and B_0 are the fermionic and bosonic (charge and spin) propagators in the two-pole approximation (in parentheses are the unknown parameters they depend on), Σ_n is the residual self-energy at the nth step (it depends on G_n and B_n, which are equal to G_0 and B_0 at the startup), G_n is the full fermionic propagator (it depends on the unknowns in round parentheses and on Σ_{n-1}) and, finally, B_n is the bosonic (charge and spin) propagators in the two-pole approximation at the nth

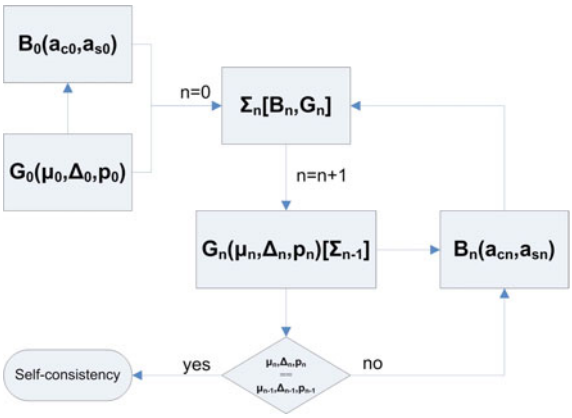

Fig. 4.14 Self-consistency scheme to compute the fermionic full propagator G in terms of the charge–charge and spin–spin two-pole propagator B and the residual self-energy Σ in the COM(2p+NCA(2p))

step (it depends on G_n, besides unknown parameters). The self-consistency cycle terminates when the fermionic parameters μ, Δ, and p do not change; to get a six-digit precision for the latter fermionic parameters, we usually need about ten cycles (it varies very much with doping, temperature, and interaction strength) on a 3D grid of 128×128 points in momentum space and 4096 Matsubara frequencies.

In Fig. 4.15 (top panel), the spectral function $A(\mathbf{k}, \omega)$ along the principal directions ($\Gamma = (0,0) \rightarrow S = (\pi/2, \pi/2) \rightarrow M = (\pi, \pi) \rightarrow X = (\pi, 0) \rightarrow S = (\pi/2, \pi/2) \rightarrow Y = (0, \pi) \rightarrow \Gamma = (0,0)$) is reported for Coulomb repulsion $U = 8$, filling $n = 0.92$, and temperature $T = 0.02$ in the frequency range in the proximity of the chemical potential ($\omega = 0$). We can clearly see many unconventional features: kinks in the dispersion along the main diagonal ($\Gamma \rightarrow M$) and the side of the Brillouin zone ($M \rightarrow X, Y$) at energies comparable with the effective exchange energy in the system ($J = 4t^2/U$); extended regions in momentum where the imaginary part of the self-energy is strong enough to selectively suppress the spectral weight; almost doubling of the Brillouin zone according to the rather high intensity of the not-so-short-range antiferromagnetic correlations; formation of a hole pocket centred along the main diagonal ($\Gamma \rightarrow M$); formation of almost-closed electron pockets centred at $X = (\pi, 0)$ and $Y = (0, \pi)$; high-intensity of the spectral weight at $X = (\pi, 0)$ and $Y = (0, \pi)$ (van Hove points) although well below the Fermi surface; no-flat dispersion along the antidiagonal ($X \rightarrow Y$) generated dynamically (t' is not present in the original Hamiltonian); strong suppression of the spectral weight in the proximity of $M = (\pi, \pi)$ (again, due to the rather high intensity of the not-so-short-range antiferromagnetic correlations) which will lead to the appearance of a pseudogap. This scenario is just the one claimed for high-T_c cuprates in the underdoped region by ARPES and many other experimental techniques.

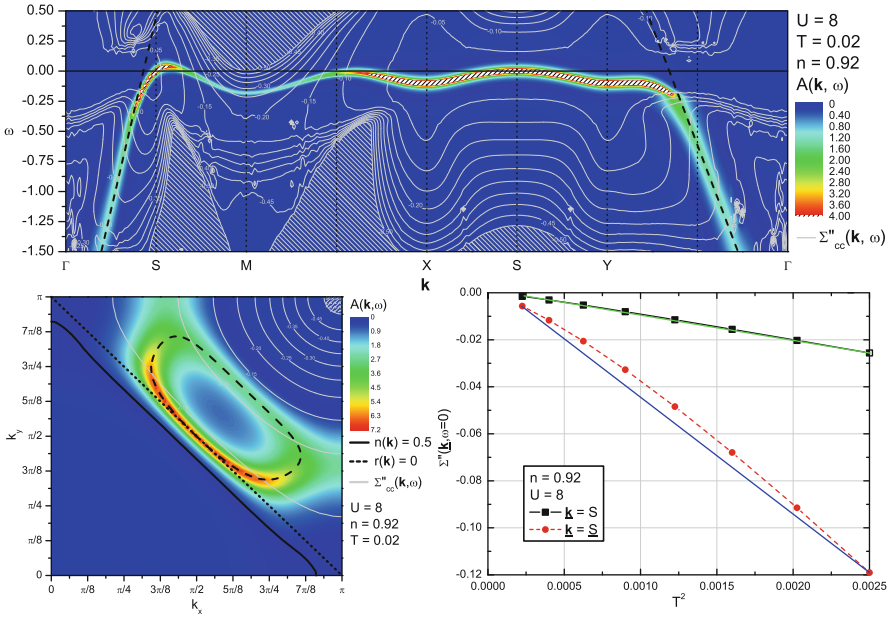

Fig. 4.15 $U = 8$, $n = 0.92$, and $T = 0.02$. (*Top*) Spectral function $A(\mathbf{k}, \omega)$ along the principal directions. (*Bottom left*) Spectral function at the chemical potential $A(\mathbf{k}, \omega = 0)$ as a function of momentum $\mathbf{k}$. (*Bottom right*) Imaginary part of the self-energy at the chemical potential $\Sigma''(\mathbf{k}, \omega = 0)$ as a function of temperature squared T^2 at two momenta: $S = (\pi/2, \pi/2)$ and $\underline{S}$, where the main diagonal of Brillouin zone touches the *phantom* portion of the hole pocket

In Fig. 4.15 (bottom left panel), the spectral function at the chemical potential $A(\mathbf{k}, \omega = 0)$ as a function of momentum $\mathbf{k}$ is reported for Coulomb repulsion $U = 8$, filling $n = 0.92$, and temperature $T = 0.02$. The hole pocket and the electron quasipocket are now clearly visible. The Fermi surface is deconstructed: there is coexistence of a small Fermi surface (the hole pocket) and of a large one partially solving the hoary dichotomy signalled by main experimental techniques; the actual Fermi surface, the locus formed by the momenta where the spectral weight reaches the highest intensity (this is also the only Fermi surface detectable by ARPES), is open (it is just an arc – half of the hole pocket) in contrast to what Fermi liquid picture would require; it is just the imaginary part of the residual self-energy to *eat up* a portion of the hole pocket and make this latter just a *phantom* arc. The residual self-energy results stronger and stronger close to $M = (\pi, \pi)$ as one should expect for high-intensity not-so-short-range antiferromagnetic correlations.

In Fig. 4.15 (bottom right panel), the imaginary part of the self-energy at the chemical potential $\Sigma''(\mathbf{k}, \omega = 0)$ as a function of the temperature squared T^2 at two momenta: $S = (\pi/2, \pi/2)$ and $\underline{S}$. The latter corresponds to the point where the main diagonal ($\Gamma \rightarrow M$) touches the *phantom* arc of the hole pocket; it is clearly visible in Fig. 4.15 both (top panel) and (bottom left panel). It is striking

the qualitative difference in the dependence of the imaginary part of the residual self-energy on T^2 in the two points: S *lives* in an ordinary Fermi liquid, where the dependence on temperature is exactly quadratic; $\underline{S}$ belongs to a portion of the system interested to a non-Fermi liquid description according to its dependence far from quadratic. Also the dependence of frequency (not shown) turns out to be almost perfectly quadratic in S and with strong linear components in $\underline{S}$. This unconventional behavior can be at the basis of the anomalous features shown by resistivity in these materials.

4.3.3 Four-Pole Solution

In the previous section, we have shown how the two-pole solution can be improved by taking into account residual dynamical corrections. However, as discussed in Sect. 4.2.5, the two-pole solution can be also improved by enlarging the basis through the addition of higher-order composite operators. This aspect is illustrated in this section.

As shown in Sect. 4.3.1.1, the higher order field $\pi(i)$ appears in the equations of motion of $\xi(i)$ and $\eta(i)$. The effect of this field was approximately taken into account by projecting it on the basis (4.30). Here we promote this field to the rank of new composite operator. Actually, we divide $\pi(i)$ into two operators $\pi(i) = \xi_s(i) + \eta_s(i)$, similarly to what we have done for the electronic field $[c(i) = \xi(i) + \eta(i)]$, defined as

$$\begin{aligned} \xi_s(i) &= \frac{1}{2}\sigma^{\mu} n_{\mu}(i)\xi^{\alpha}(i) + \eta^{\dagger\alpha}(i)c(i)c(i) \\ \eta_s(i) &= \frac{1}{2}\sigma^{\mu} n_{\mu}(i)\eta^{\alpha}(i) + \xi^{\dagger\alpha}(i)c(i)c(i) \end{aligned} \tag{4.62}$$

and take as a new basis

$$\psi(i) = \begin{pmatrix} \xi(i) \\ \eta(i) \\ \xi_s(i) \\ \eta_s(i) \end{pmatrix}. \tag{4.63}$$

The reason for such a choice is that this basis is the fermionic closed basis for the Hubbard model on two sites [9]. In such a case, $\xi_s(i)$ and $\eta_s(i)$ are eigenoperators of the Hamiltonian with eigenenergies given by $E_3(\mathbf{k}) = -\mu + t\alpha(\mathbf{k}) - J_U$ and $E_4(\mathbf{k}) = -\mu + t\alpha(\mathbf{k}) + U + J_U$. It is worth noting that the energy term $J_U = [\sqrt{U^2 + 16t^2} - U]/2$ returns, in the limit $U/t \gg 1$, the virtual exchange energy $J = 4t^2/U$.

COM, as formulated in Sect. 4.2, requires as a first step the calculation of the normalization I and energy ε matrices (4×4 matrices). The detailed expression

of the matrix I is reported in [17]. In order to effectively perform calculations, the elements of I which contain three-site correlation functions of the form

$$\langle A^{\alpha}(\mathbf{i})B^{\alpha}(\mathbf{i})C(\mathbf{i})D(\mathbf{i})\rangle = \frac{1}{4}\langle (A(\mathbf{i})B(\mathbf{i}))^{\alpha}C(\mathbf{i})D(\mathbf{i})\rangle + \sum_{\mathbf{j}\neq\mathbf{k}} \alpha_{\mathbf{ij}}\alpha_{\mathbf{ik}}\langle A(\mathbf{j})B(\mathbf{k})C(\mathbf{i})D(\mathbf{i})\rangle \tag{4.64}$$

have been computed through the following decoupling procedure: (1) the first term is reducible into two-site correlation functions, which can be directly evaluated in terms of the propagators under analysis; (2) the second one has been decoupled in terms of two-site correlation functions. This procedure preserves the particle–hole symmetry enjoined by the Hamiltonian.

The energy matrix can be straightforwardly calculated by means of the anticommuting algebra given in Table 4.1. However, due to the complexity of the equations of motion of the fields $\xi_s(i)$ and $\eta_s(i)$, a direct calculation will lead to the appearance of an enormous number of unknown correlation functions. In order to determine all correlation functions, we would then be forced to use some uncontrolled decoupling procedure and we would completely lose every control over the approximation procedure. Hence, we have opted for a controlled approximation at the level of equations of motion by neglecting irreducible three-site operators and paying attention to evaluating exactly all one- and two-site components. By means of this approximation procedure, the matrix ε can be easily computed. Then, we have all ingredients, following the procedure reported in Sects. 4.2.6 and 4.3.1, to calculate all single-particle and thermodynamic properties.

Results and Comparisons

In Fig. 4.16, we provide a detailed comparison of the band structure obtained by the present formulation with QMC results [102]. As can be easily seen, we have a good agreement with the QMC data, especially for the low-energy band around

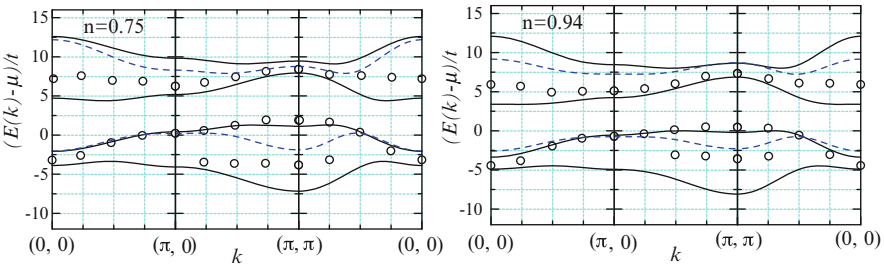

Fig. 4.16 The dispersion relation at $U/t = 8$, $T/t = 0.5$ and (*left*) $n = 0.75$ and (*right*) 0.94. The two-pole solution (*dashed line*) and QMC data (*circle*) of [102] are also reported

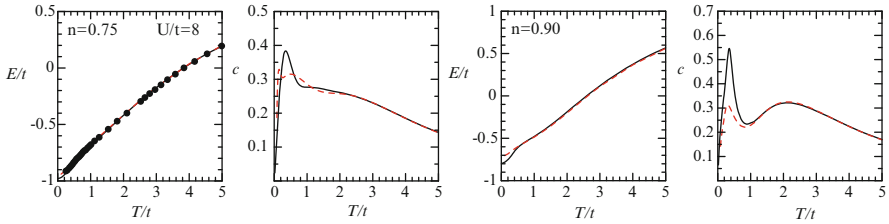

Fig. 4.17 Internal energy E and specific heat C per site at $U/t = 8$ and (*left*) $n = 0.75$ and (*right*) 0.90. Data from finite-temperature Lanczos [70] and QMC [80, 103] are provided for comparison

the Fermi level. Internal energy E and specific heat C per site at $U/t = 8$ and $n = 0.75$ and 0.90 are reported in Fig. 4.17. Data from finite-temperature Lanczos [70] and QMC [80, 103] are provided for comparison. As regards the internal energy, the agreement with the Lanczos data is excellent except for the low temperature region. As regards the specific heat, we observe a sharp peak around $T/t = 0.3$ and a fairly broad peak in the higher temperature region $T/t = 1$. The two-peak structure is more pronounced for $U/t = 8$, but not so much for $U/t = 4$ (not shown). This tendency is also observed in several numerical simulations [70, 80, 103, 104]. Usually, the sharp peak at lower temperatures and the broad peak at higher temperatures are interpreted as consequences of spin and charge fluctuations related to the energy scales of J and U, respectively. The main difference between our results and numerical ones regards the height of the peak in the specific heat around $T/t = 0.3$ that comes from the decrease in the internal energy. This is an indication of well-established spin ordering which cannot be correctly evaluated on a small cluster. Numerical simulation cannot describe spin and charge ordering in the case that the correlation lengths exceed the cluster size. On the other hand, in our formulation, there is a tendency to have too pronounced spin and charge correlations.

4.3.4 Superconducting Solution

The various solutions of the Hubbard model presented up to now refer to a paramagnetic homogeneous state. However, the same formulation can be easily applied to states with completely different symmetry properties. Ferromagnetic, antiferromagnetic, and superconducting solutions of the model have been reported in [5]. As an example, we here summarize the superconducting solution of the model in the d-wave channel.

We consider the Nambu representation and introduce the four-vector composite field [5, 105, 106]

$$\psi(i) = \begin{pmatrix} \xi_\uparrow(i) \\ \eta_\uparrow(i) \\ \xi^\dagger_\downarrow(i) \\ \eta^\dagger_\downarrow(i) \end{pmatrix}. \tag{4.65}$$

At a first level of approximation, the current of this operator is projected on the basis and the residual current is neglected. The effects of the residual current have also been analyzed, but they will not be presented here. Although the model admits various superconducting phases with different symmetries of the order parameter (see [5]), on the basis of experimental evidence for high T_c superconductors, we restrict the analysis to singlet pairing and d-wave symmetry for a two-dimensional lattice. Then, we impose as boundary conditions $\langle \psi_\uparrow(\mathbf{i})\psi_\downarrow(\mathbf{i})\rangle = 0$ and $\langle \psi_\uparrow(\mathbf{i})\psi^\alpha_\downarrow(\mathbf{i})\rangle = 0$. Furthermore, we assume translational and spin rotational invariance. Accordingly, the normalization and m matrices have the expressions

$$I(\mathbf{k}) = \begin{pmatrix} I_{11} & 0 & 0 & 0 \\ 0 & I_{22} & 0 & 0 \\ 0 & 0 & I_{11} & 0 \\ 0 & 0 & 0 & I_{22} \end{pmatrix} \qquad \begin{matrix} I_{11} = 1 - \frac{n}{2} \\ I_{22} = \frac{n}{2} \end{matrix} \tag{4.66}$$

$$m = \begin{pmatrix} m_1 & m_2 \\ m_2 & -m_1 \end{pmatrix} \qquad m_1 = \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix} \qquad m_2 = m_{13} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \tag{4.67}$$

with

$$m_{11}(\mathbf{k}) = -\mu I_{11} - 4t\Delta - 4t\alpha(\mathbf{k})(1 - n + p) \tag{4.68}$$

$$m_{12}(\mathbf{k}) = 4t\Delta + 4t\alpha(\mathbf{k})(p - I_{22}) \tag{4.69}$$

$$m_{13}(\mathbf{k}) = 4t\gamma(\mathbf{k})f \tag{4.70}$$

$$m_{22}(\mathbf{k}) = (U - \mu)I_{22} - 4t\Delta - 4t\alpha(\mathbf{k})p \tag{4.71}$$

and $\gamma(\mathbf{k}) = [\cos(k_x a) - \cos(k_y a)]/2$. The parameters Δ and p have been defined in Sect. 4.3.1.1, the parameter f plays the role of order parameter and is defined as $f = \langle c^\dagger_\downarrow(\mathbf{i})c_\downarrow(\mathbf{i})[c_\downarrow(\mathbf{i})c_\uparrow(\mathbf{i})]^\gamma\rangle + \langle [c^\dagger_\uparrow(\mathbf{i})c_\uparrow(\mathbf{i})]^\gamma c_\downarrow(\mathbf{i})c_\uparrow(\mathbf{i})\rangle$. The retarded and correlation functions, which are now 4×4 matrices containing the anomalous components, are then known functions of the internal parameters μ, Δ, p, and f. The self-consistent determination of these parameters is performed by means of the equations given in Sect. 4.3.1.1 and by using the equations of motion for the latter.

Results and Comparisons

In Fig. 4.18, we report the critical temperature T_c as a function of the doping δ for $U = 4t$ and $t \approx 0.17$ eV. The experimental data for $La_{2-\delta}Sr_\delta CuO_4$ are taken

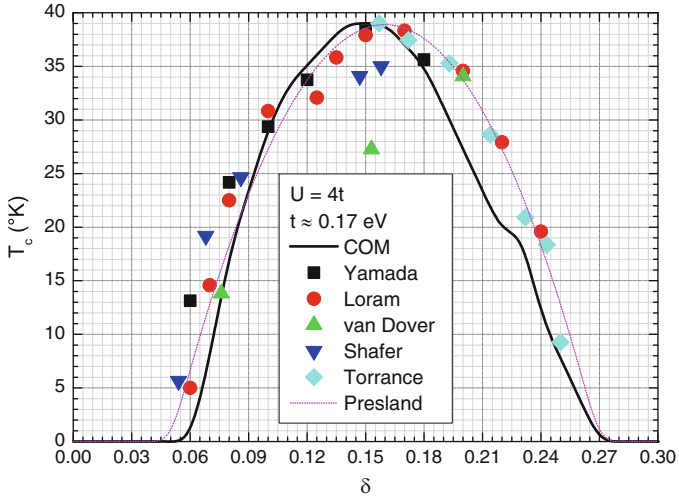

Fig. 4.18 Critical temperature T_c as a function of the doping δ for $U = 4t$ and $t \approx 0.17$ eV. Experimental data for $La_{2-\delta}Sr_{\delta}CuO_4$ from Yamada et al. [95], Loram et al. [96], van Dover et al. [97], Shafer et al. [98] and Torrance et al. [99]. Phenomenological curve from Presland et al. [107]

from Yamada et al. [95], Loram et al. [96], van Dover et al. [97], Shafer et al. [98] and Torrance et al. [99]. The phenomenological curve suggested by Presland et al. [107] is also reported. The agreement is really very good: the position and presence of *endpoints* (critical fillings at zero temperature), the position of the *peak* (critical filling at highest temperature) and the overall shape. This clearly shows that, together with the very remarkable capability to describe the very unconventional physics of the underdoped region (see Sect. 4.3.2), COM manages to catch many relevant aspects of cuprate physics.

4.4 Conclusions and Outlook

In the last decades, the discovery of many families of materials with unconventional properties and anomalous behaviors has been calling for the developing of new theoretical methods to properly take into account electronic correlations in many-body systems. Understanding and describing such systems, featuring both competition and interplay among all degrees of freedom involved (orbital, spin, lattice, etc.), is one of the most challenging open issues in modern condensed matter theory. Among the various methods reported in this volume, the COM, described in detail in this chapter, has the peculiar characteristic to endorse, since its foundations, the systematic emergence, in any strongly correlated system, of composite operators obeying non-canonical algebras. COM is designed to optimally deal with the unusual features of composite operators and to exploit

algebra constraints and other relevant symmetries in order to properly compute the unconventional properties of the system under analysis. The formalism, presented in a systematic way together with its internals in Sect. 4.2, is based on two main concepts: the use of the propagators of the emerging composite operators as proper building blocks of a more efficient perturbation scheme and the use of the constraints dictated by the non-canonical algebra closed by the emerging composite operators to fix the representation of the Green's functions and keep the algebraic and symmetry properties of the Hamiltonian under study. Although the method has been used to study a great variety of models, we have preferred just to present its application to the Hubbard model. The formulation is illustrated in detail in Sect. 4.3 by systematically going through a well-defined chain of increasing levels of approximations, and reporting, at each stage, comparisons with the results of exact treatments and numerical techniques. Comparisons with experimental data from high-T_c cuprates have been also reported. It clearly emerges that the COM is a quite powerful and general theoretical tool to study any kind of strongly correlated Hamiltonian system (lattice, impurity, spin, ...) and the actual materials described by them.

COM has already been applied to a variety of Hamiltonian models (see Sect. 4.2.7 for an exhaustive list). However, many interesting problems are still on the carpet, and many other still to come, and much work remains to be done. What are the advantages of COM? For the study of strongly correlated systems, it allows us to exploit all information, clues and hints one has about the specific targeted system, still retaining a clear-cut path of increasing levels of complexity and deepness, an intrinsic high tunability and relatively simple internals. For mechanical statistical problems, where exact solutions can be addressed, the use of COM is very promising, especially for dimensions higher than one, where transfer matrix methods are not easily applicable. Accordingly, we think that COM should be considered as a first attempt to construct a modern theory of strongly correlated systems, not heavily relying on numerics (as DMFT and SFA and derivatives are), beyond both traditional Fermi liquid theory and electronic perturbation schemes.

References

1. J. Hubbard, Proc. Roy. Soc. A **276**, 238 (1963)
2. P.W. Anderson, Science **177**, 393 (1972)
3. A.C. Hewson, *The kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, 1997)
4. F. Mancini, A. Avella, Euro. Phys. J. B **36**, 37 (2003)
5. F. Mancini, A. Avella, Adv. Phys. **53**, 537 (2004)
6. O.K. Kalashnikov, E.S. Fradkin, Sov. Phys. JETP **28**, 317 (1969)
7. W. Nolting, Z. Phys. **255**, 25 (1972)
8. F. Mancini, Phys. Lett. A **249**, 231 (1998)
9. A. Avella, F. Mancini, T. Saikawa, Euro. Phys. J. B **36**, 445 (2003)
10. F. Mancini, Europhys. Lett. **70**, 485 (2005)
11. F. Mancini, F.P. Mancini, Phys. Rev. E **77**, 061120 (2008)

12. F. Mancini, F.P. Mancini, Euro. Phys. J. B **73**, 581 (2010)
13. F. Mancini, Euro. Phys. J. B **45**, 497 (2005)
14. F. Mancini, A. Naddeo, Phys. Rev. E **74**, 061108 (2006)
15. A. Avella, F. Mancini, Euro. Phys. J. B **50**, 527 (2006)
16. F. Mancini, F.P. Mancini, Condens. Matter Phys. **11**, 543 (2008)
17. S. Odashima, A. Avella, F. Mancini, Phys. Rev. B **72**, 205121 (2005)
18. D. Villani, E. Lange, A. Avella, G. Kotliar, Phys. Rev. Lett. **85**, 804 (2000)
19. A. Avella, R. Hayn, F. Mancini, Euro. Phys. J. B **37**, 465 (2004)
20. A. Avella, F. Mancini, S. Odashima, Physica C **388**, 76 (2003)
21. A. Avella, S.S. Feng, F. Mancini, Physica B **312**, 537 (2002)
22. M. Bak, F. Mancini, Physica B **312**, 732 (2002)
23. A. Avella, F. Mancini, S. Odashima, J. Magn. Magn. Mat. **272**, E311 (2004)
24. A. Avella, F. Mancini, R. Hayn, Acta Phys. Pol. B **34**, 1345 (2003)
25. A. Avella, S. Krivenko, F. Mancini, Physica B **329**, 955 (2003)
26. N.D. Mermin, H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966)
27. H. Matsumoto, T. Saikawa, F. Mancini, Phys. Rev. B **54**, 14445 (1996)
28. H. Matsumoto, F. Mancini, Phys. Rev. B **55**, 2095 (1997)
29. A. Avella, S. Krivenko, F. Mancini, N.M. Plakida, J. Magn. Magn. Mat. **272**, 456 (2004)
30. S. Krivenko, A. Avella, F. Mancini, N. Plakida, Physica B **359**, 666 (2005)
31. A. Avella, F. Mancini, Phys. Rev. B **75**, 134518 (2007)
32. A. Avella, F. Mancini, J. Phys. Condens. Matter **19**, 255209 (2007)
33. A. Avella, F. Mancini, E. Plekhanov, J. Magn. Magn. Mat. **310**, 999 (2007)
34. A. Avella, F. Mancini, Physica C **460**, 1096 (2007)
35. A. Avella, F. Mancini, Acta Phys. Pol. A **113**, 395 (2008)
36. A. Avella, F. Mancini, J. Phys. Condens. Matter **21**, 254209 (2009)
37. J.C. Ward, Phys. Rev. **78**, 182 (1950)
38. Y. Takahashi, Nuovo Cimento **6**, 370 (1957)
39. M. Sánchez-Lopez, A. Avella, F. Mancini, Europhys. Lett. **44**, 328 (1998)
40. A. Avella, F. Mancini, R. Münzner, Phys. Rev. B **63**, 245117 (2001)
41. A. Avella, F. Mancini, M. Sànchez-Lopez, Euro. Phys. J. B **29**, 399 (2002)
42. A. Avella, F. Mancini, V. Turkowski, Phys. Rev. B **67**, 115123 (2003)
43. S. Ishihara, H. Matsumoto, S. Odashima, M. Tachiki, F. Mancini, Phys. Rev. B **49**, 1350 (1994)
44. V. Fiorentino, F. Mancini, E. Zasinas, A.F. Barabanov, Phys. Rev. B **63**, 214515 (2001)
45. A. Avella, F. Mancini, F.P. Mancini, E. Plekhanov, J. Phys. Conf. Series **273**, 012091 (2011)
46. A. Avella, F. Mancini, D. Villani, Phys. Lett. A **240**, 235 (1998)
47. A. Avella, F. Mancini, D. Villani, H. Matsumoto, Euro. Phys. J. B **20**, 303 (2001)
48. A. Avella, F. Mancini, Physica C **408**, 284 (2004)
49. A. Avella, F. Mancini, Euro. Phys. J. B **41**, 149 (2004)
50. A. Avella, F. Mancini, J. Phys. Chem. Sol. **67**, 142 (2006)
51. H. Matsumoto, A. Allega, S. Odashima, F. Mancini, Physica C **235**, 2227 (1994)
52. E. Plekhanov, A. Avella, F. Mancini, F.P. Mancini, J. Phys. Conf. Series **273**, 012147 (2011)
53. F. Mancini, N.B. Perkins, N.M. Plakida, Phys. Lett. A **284**, 286 (2001)
54. F.P. Mancini, J. Phys. Conf. Series **200**, 022030 (2010)
55. A.M. Allega, S. Odashima, H. Matsumoto, F. Mancini, Physica C **235**, 2229 (1994)
56. A. Avella, F. Mancini, Physica B **378**, 700 (2006)
57. F. Mancini, V. Oudovenko, D. Villani, Czecho. J. Phys. **46**, 1873 (1996)
58. A. Avella, F. Mancini, D. Villani, Sol. Stat. Comm. **108**, 723 (1998)
59. F. Mancini, D. Villani, H. Matsumoto, Phys. Rev. B **57**, 6145 (1998)
60. A. Avella, F. Mancini, Euro. Phys. J. B **32**, 27 (2003)
61. H. Matsumoto, H. Umezawa, Phys. Rev. B **31**, 4433 (1985)
62. Y.A. Izyumov, B.M. Letfulov, J. Phys. Condens. Matter **2**, 8905 (1990)
63. F. Mancini, S. Marra, H. Matsumoto, Physica C **252**, 361 (1995)
64. A. Moreo, D.J. Scalapino, R.L. Sugar, S.R. White, N.E. Bickers, Phys. Rev. B **41**, 2313 (1990)

65. L.M. Roth, Phys. Rev. **184**, 451 (1969)
66. L.M. Roth, Phys. Rev. **186**, 428 (1969)
67. F. Gebhard, *The Mott Metal-Insulator Transition: Models and Methods* (Springer, Berlin, 1997)
68. A. Avella, F. Mancini, D. Villani, L. Siurakshina, V.Y. Yushankhai, Int. J. Mod. Phys. B **12**, 81 (1998)
69. N. Kawakami, T. Usuki, A. Okiji, Phys. Lett. A **137**, 287 (1989)
70. J. Bonča, P. Prelovšek, Phys. Rev. B **67**, 85103 (2003) and private communications
71. H. Shiba, P.A. Pincus, Phys. Rev. B **5**, 1966 (1972)
72. J. Schulte, M.C. Böhm, Phys. Rev. B **53**, 15385 (1996)
73. S.R. White, D.J. Scalapino, R.L. Sugar, E.Y. Loh, J.E. Gubernatis, R.T. Scalettar, Phys. Rev. B **40**, 506 (1989)
74. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **59**, 121 (1987)
75. F.D. Buzatu, Mod. Phys. Lett. B **9**, 1149 (1995)
76. J.E. Hirsch, Phys. Rev. B **22**, 5259 (1980)
77. H. Shiba, Phys. Rev. B **6**, 930 (1972)
78. F. Mancini, S. Marra, H. Matsumoto, Physica C **250**, 184 (1995)
79. F. Mancini, H. Matsumoto, D. Villani, J. Phys. Stud. **3**, 474 (1999)
80. D. Duffy, A. Moreo, Phys. Rev. B **55**, 12918 (1997); **56**, 7022 (1997)
81. F. Mancini, S. Marra, A. Allega, H. Matsumoto, in *Superconductivity and Strongly Correlated Electron Systems*, ed. by C. Noce, A. Romano, G. Scarpetta (World Scientific, Singapore, 1994), p. 271
82. A. Avella, F. Mancini, Physica C **408**, 287 (2004)
83. Y.M. Vilk, L. Chen, A.M.S. Tremblay, Phys. Rev. B **49**, 13267 (1994)
84. Y.C. Chen, A. Moreo, F. Ortolani, E. Dagotto, T.K. Lee, Phys. Rev. B **50**, 655 (1994)
85. R. Coldea, S. Hayden, G. Aeppli, T. Perring, C. Frost, T. Mason, S.W. Cheong, Z. Fisk, Phys. Rev. Lett. **86**, 5377 (2001)
86. A. Avella, F. Mancini, D. Villani, H. Matsumoto, Physica C **282**, 1759 (1997)
87. R. Münzner, A. Avella, F. Mancini, Physica B **284**, 1577 (2000)
88. A. Avella, F. Mancini, *The t-t'-U Model and the Cuprate Materials*, vol. 241 (Nova Science Publishers, New York, 2003), p. 215
89. T.R. Thurston et al., Phys. Rev. B **40**, 4585 (1989)
90. G. Shirane et al., Phys. Rev. Lett. **63**, 330 (1989)
91. R. Birgenau, et al., Phys. Rev. B **38**, 6614 (1988)
92. S.W. Cheong et al., Phys. Rev. Lett. **67**, 1791 (1991)
93. G. Shirane, R.J. Birgeneau, Y. Endoh, M.A. Kastner, Physica B **197**, 158 (1994)
94. K. Yamada et al., Phys. Rev. Lett. **75**, 1626 (1995)
95. K. Yamada et al., Phys. Rev. B **57**, 6165 (1998)
96. J.W. Loram et al., Physica C **162**, 498 (1989)
97. R.B. van Dover et al., Phys. Rev. B **35**, 5337 (1987)
98. M.W. Shafer, T. Penny, B.L. Olson, Phys. Rev. B **36**, 4047 (1987)
99. J.B. Torrance et al., Phys. Rev. Lett. **61**, 1127 (1988)
100. A. Avella, F. Mancini, H. Matsumoto, D. Villani, Physica B **230**, 912 (1997)
101. S. Petit et al., *Spin Dynamics Study of LSCO by Inelastic Neutron Scattering*. Communication at ECNS'96 1st European Conference on Neutron Scattering Interlaken, Switzerland, October 8–11 1996
102. N. Bulut, D.J. Scalapino, S.R. White, Phys. Rev. Lett. **73**, 748 (1994)
103. D. Duffy, A. Nazarenko, S. Haas, A. Moreo, J. Riera, E. Dagotto, Phys. Rev. B **56**, 5597 (1997)
104. T. Paiva, R.T. Scalettar, C. Huscroft, A.K. McMahan, Phys. Rev. B **63**, 125116 (2001)
105. T. Di Matteo, F. Mancini, H. Matsumoto, V.S. Oudovenko, Physica B **230**, 915 (1997)
106. A. Avella, F. Mancini, D. Villani, H. Matsumoto, Physica C **282**, 1757 (1997)
107. M.R. Presland, J.L. Tallon, R.G. Buckley, R.S. Liu, N.E. Flower, Physica C **176**, 95 (1991)

Chapter 5
LDA+GTB Method for Band Structure Calculations in the Strongly Correlated Materials

Sergey G. Ovchinnikov, Vladimir A. Gavrichkov, Maxim M. Korshunov, and Elena I. Shneyder

Abstract We present the multielectron LDA+GTB (local density approximation plus generalized tight-binding) approach to the electronic structure calculations for the Mott insulators. This method is a straightforward generalization of the Hubbard perturbation theory which starts from the atomic limit. All local interactions within the unit cell are treated by the exact diagonalization of the multiband $p - d$ Hamiltonian with the parameters calculated within LDA. Intercell Hoppings and interactions between the unit cells are considered as perturbation within the Hubbard X-operators representation. We also discuss the application of the LDA+GTB method to cuprates, manganites, and cobaltites.

S.G. Ovchinnikov (✉)
Kirensky Institute of Physics, Siberian Branch of RAS, Krasnoyarsk 660036, Russia

Siberian Federal University, 79 Svobodny Prospect, Krasnoyarsk 660041, Russia

Reshetnev Siberian State Aerospace University, Krasnoyarsk 660014, Russia
e-mail: sgo@iph.krasn.ru

V.A. Gavrichkov
Kirensky Institute of Physics, Siberian Branch of RAS, Krasnoyarsk 660036, Russia

Siberian Federal University, 79 Svobodny Prospect, Krasnoyarsk 660041, Russia
e-mail: gav@iph.krasn.ru

M.M. Korshunov
Kirensky Institute of Physics, Siberian Branch of RAS, Krasnoyarsk 660036, Russia

Department of Physics, University of Florida, Gainesville, FL 32611, USA
e-mail: korshunov@phys.ufl.edu

E.I. Shneyder
Kirensky Institute of Physics, Siberian Branch of RAS, Krasnoyarsk 660036, Russia

Reshetnev Siberian State Aerospace University, Krasnoyarsk 660014, Russia
e-mail: shneyder@iph.krasn.ru

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_5,

5.1 Introduction

One of the most attractive feature of the condensed matter physics is the ability to explain and even predict properties of the materials which comprise our world. Significant leap in this direction was made in mid 1960 when Hohenberg et al. [1, 2] formulated a density functional theory (DFT). Because its starting point is the Shrödinger equation for the particular atomic arrangement and orbital and spin configurations, this theory is often referred to as the "first principles" or the ab initio calculations. Augmented with the local density approximation (LDA) or the generalized gradient approximation (GGA) for the initially unknown quantity, exchange-correlation energy, DFT provides quantitative description of the ground state energy and the band structure of various atoms, molecules, and crystalline solids (see e.g., [3]).

Despite its success for s and p atoms in solids, LDA failed to describe transition metal oxides with partially filled $3d$-orbitals. The most pronounced failure is that LDA predicts La_2CuO_4 to be a metal whereas experimentally it is an insulator. The root of the problem is the unscreened on-site Coulomb interaction (Hubbard repulsion) [4]. In a single-band system on the mean-field level if Hubbard repulsion U is larger than the bandwidth W, it splits the band into two Hubbard subbands with a gap $\propto U$. The spectral weight of a quasiparticle is redistributed between these subbands. At a half-filling the Fermi level is inside the gap and the system is an insulator. In a multiorbital system, along with the Hubbard repulsion other local interactions like the Hund's exchange J_{H} and the interorbital Hubbard repulsion U' are present and provide a rich set of physical properties. The opening of the Hubbard gap and moreover the major role played by the local interactions near the half-filling are beyond the scope of LDA and GGA.

There are several extensions to LDA which includes or simulates the effects of the on-site interactions. One of them is LDA+U [5] and another one is SIC–LSDA (self-interaction–corrected local spin density approximation) [6]. Both methods consider local interactions in the Hartree–Fock sense and result in the antiferromagnetic insulator as the ground state for La_2CuO_4 contrary to LDA, but the origin of the insulating gap is incorrect. In both LDA+U and SIC–LSDA, it is formed by the local single-electron states splitted by the spin or orbital polarization. Therefore, the paramagnetic phase above the Néel temperature T_{N} of the undoped La_2CuO_4 will be metallic in spite of strong correlation regime $U \gg W$. There is one more significant drawback in these approximations, namely, they disregard the redistribution of the spectral weight between the Hubbard subbands. The latter effect is incorporated in a different approach to ab initio calculations for strongly correlated systems – LDA+DMFT (LDA+dynamical mean-field theory) [7–10]. The method is based on the self-consistent procedure where the LDA band structure is used to calculate the electron self-energy in DMFT. DMFT utilizes the fact that in the infinite dimensional limit of the Hubbard model, $D \to \infty$, the self-energy is momentum independent, $\Sigma(\mathbf{k}, \omega) \to \Sigma(\omega)$ [11–13]. The remaining frequency dependence is exact in $D \to \infty$ limit and carries very important information

about dynamical correlations and Mott–Hubbard transition. On the other hand, the spatial correlations become crucial in low-dimensional systems like layered high-T_c cuprates. That is why the correct band dispersion and spectral intensities for these systems cannot be obtained within LDA+DMFT. The natural extension of this method, LDA+cluster or cellular DMFT [14–17], and SDFT (spectral density functional theory) [18–22] provides momentum dependent self-energy and thus allow for the nonlocal correlations.

The LDA+GTB method is not just "another one" approach to study the Hubbard model. From the very beginning, the generalized tight-binding (GTB) method has been suggested to extend the microscopic band structure calculations to take the strong electronic correlations (SECs) into account in the Mott–Hubbard insulators like the transition metal oxides [23]. Similar to conventional tight-binding (TB) approach we start with a particular local electron states (with all multiorbital effects, symmetry, and chemistry) and then by a Fourier transform move to the momentum space and obtain a band structure. Because of SEC we cannot use free electron local states, our local fermion in a d-orbital system is a quasiparticle given by the excitations between multielectron d^n and $d^{n\pm1}$ terms contrary to the conventional TB. In other words, GTB is the strongly correlated version of the TB method. The first computer codes and successful application of GTB have been developed for cuprates [24]. That version used the multiband $p-d$ model for La_2CuO_4 [25] with a lot of empirical parameters in the Hamiltonian. To create the ab initio approach, the hybrid LDA+GTB method has been developed in a combined effort of teams from Krasnoyarsk and Yekaterinburg [26]. Afterward, similar ideas have been used to study the GTB band structure of manganites $La_{1-x}Ca_xMnO_3$ [27] and cobaltites $LaCoO_3$ [28].

The LDA+GTB may be considered as the straightforward development of the Hubbard atomic representation approach [4] to real materials like $3d$ metal oxides. Indeed, the GTB is a specific version of cluster perturbation theory (CPT) in the Hubbard X-operators representation [29].

The rest of this chapter is organized as follows. In Sect. 5.2 we discuss the exact Lehmann representation for the electronic Green function that gives us the answer what is an electron in a SEC system. In Sect. 5.3 we provide the main ideas and technical steps of the LDA+GTB. In Sect. 5.4 we consider different approximations to solve the Dyson equation in the X-representation. In Sect. 5.5 we discuss the LDA+GTB band structure of $La_{2-x}Sr_xCuO_4$ in long-range antiferromagnetic (AFM) and short-range AFM spin-liquid regions. Sections 5.6 and 5.7 are devoted to manganites and cobaltates, respectively. Section 5.8 is the conclusion.

5.2 Quasiparticle Definition of an Electron From the Lehmann Representation

Before discussing how to calculate reliably the electronic structure in SEC materials the more fundamental question has to be answered: What is an electron in SEC system? Let us clarify this question first. Due to the strong interactions the free

electrons are so heavily renormalized, so that the new objects are unlikely to be called "the electrons." Still, we would like to generate the output of the theory in terms of the Green function that is the one-particle property, the "electron Green function." The goal is to find the electron Green function or, in other words, to define the electron.

There are a lot of speculations in the literature about spinless fermions and composite particles that are very often based on the exact solution of the one-dimensional Hubbard model. In two and higher dimensions different approaches result in the conclusion that the electron in SEC system is a sum of quasiparticles (QPs) with a charge e, spin $1/2$, and the renormalized spectral weight [30, 31]. We use the exact Lehmann representation to define the electron in SEC system [32]. At $T = 0$ the electron Green function can be written as

$$G_\sigma(\boldsymbol{k}, \omega) = \sum_m \left[\frac{A_m(\boldsymbol{k}, \omega)}{\omega - \Omega_m^+} + \frac{B_m(\boldsymbol{k}, \omega)}{\omega - \Omega_m^-} \right], \tag{5.1}$$

where $\Omega_m^+ = E_m(N+1) - E_0(N) - \mu$, $\Omega_m^- = E_0(N) - E_m(N-1) - \mu$, μ is the chemical potential, and numerators are equal to

$$\begin{aligned} A_m(\boldsymbol{k}, \omega) &= |\langle 0, N | a_{\boldsymbol{k}\sigma} | m, N+1 \rangle|^2, \\ B_m(\boldsymbol{k}, \omega) &= |\langle m, N-1 | a_{\boldsymbol{k}\sigma} | 0, N \rangle|^2. \end{aligned}$$

Here, $|m, N\rangle$ is the mth eigenstate of the N electron system, $H|m, N\rangle = E_m |m, N\rangle$.

Since each single-pole contribution on the right-hand side of (5.1) corresponds to some QP, we interpret the Lehmann representation in the following way: electron is the linear superposition of QPs with the energies Ω_m^+ (Ω_m^-) for electron addition (removal) and with the spectral weights A_m (B_m). At finite temperatures, the Lehmann representation determines the QP as the excitation between two arbitrary $|m, N+1\rangle$ and $|n, N\rangle$ terms with the energy $\Omega_{mn} = E_m - E_n$ and a temperature-dependent spectral weight [32]. This definition is very clear. Unfortunately it cannot be used straightforwardly because the exact eigenstates $|m, N\rangle$ and the eigenenergies E_m are unknown. Later we will show that the GTB method is the perturbative realization of the Lehmann representation.

We would like to emphasize two important points. First, it is clear that the Landau Fermi liquid QP is a specific particular case of (5.1) with only one QP close to the Fermi level. Second, for the free electron with energy ε_0, all QP energies are equal to $\Omega_m^+ = \Omega_m^- = \varepsilon_0 - \mu$.

5.3 The Main Steps of the LDA+GTB Method

As any other CPT approach, the GTB method starts with the exact diagonalization (ED) of the intracell part (H_c) of the multielectron Hamiltonian and treats the intercell part (H_{cc}) by a perturbation theory. Thus we make a realization of the

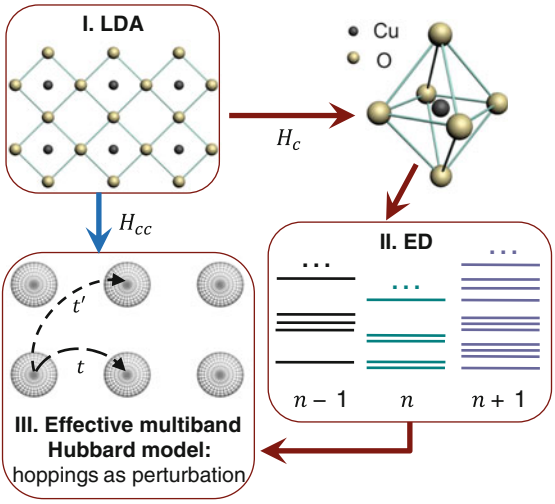

Fig. 5.1 Schematic representation of three LDA+GTB steps for the CuO_2 layer as an example. First step consists of the band structure calculation via LDA and extraction of parameters for the multiorbital Hamiltonian, then the subdivision of the Hamiltonian into intra- and inter cell parts, H_c and H_{cc}, respectively. Unit cell is the CuO_6 cluster = CuO_4 + two apical oxygens. Second step is the exact diagonalization of H_c which results in a set of energy levels. Eigenstates of H_c are used to construct a multiband Hubbard model in the X-representation with hoppings (t, t', etc.) between clusters from H_{cc}. After that, the perturbation theory for X-operators is used to obtain the electron Green function

Lehmann representation inside one unit cell with all local QP energies and spectral weights calculated via ED. The total procedure consists of the following steps [26], illustrated in Fig. 5.1:

Step I: LDA. Calculation of the LDA band structure, construction of Wannier functions with the given symmetry, and computation of the one- and two-electron matrix elements of the TB Hamiltonian with the local and nearest-neighbor Coulomb interactions.

Step II: ED. Separation of the total Hamiltonian H into the intra- and inter cell parts, $H = H_c + H_{cc}$, where H_c represents the sum of the orthogonal unit cells, $H_c = \sum_f H_f$. ED of a single unit cell term, H_f, and construction of the Hubbard X-operators $X_f^{pq} = |p\rangle\langle q|$ using the complete orthogonal set of eigenstates $\{|p\rangle\}$ of H_f.

Step III: Perturbation theory. Within the X-representation, local interactions are diagonal and all intercell hoppings and long-range Coulomb interaction terms have the bilinear form in the X-operators. Various perturbation approaches known for the Hubbard model in the X-representation can be used. The most general one includes treatment within the generalized Dyson equation obtained by the diagram technique [29].

Below we discuss each step in detail.

5.3.1 Step I: LDA

Let us assume that LDA provides us a set of Bloch functions $|\Psi_{\lambda \boldsymbol{k}}\rangle$ (λ is the band index) and band energies $\varepsilon_\lambda(\boldsymbol{k})$. For example, LDA band structure calculation for La_2CuO_4 and Nd_2CuO_4 was done within the TB-LMTO-ASA (linear muffin-tin orbitals using atomic sphere approximation in the tight-binding) method [33]. Using the Wannier functions (WFs) formalism [34], we obtain hopping integrals of the TB model via a projection procedure [26]. A number and a symmetry of chosen WFs are determined by the energy window that we are interested in. For example, it is widely believed that the low-energy physics of cuprates is captured by the three-band model [35,36] which includes bonding of Cu-$d_{x^2-y^2}$ and O-$p_{x,y}$ orbitals. Nevertheless if we are interested in any physics involving apical oxygen we have to add p_z oxygen and d_{z^2} copper orbital that results in the five band $p-d$ model [25]. If one wants to study UPS (ultraviolet photoelectron spectroscopy) or XPS (X-ray photoelectron spectroscopy), then all $3d$ orbitals have to be included. If we restrict ourself to the five band model, it allows us to consider ARPES (angle-resolved photoemission spectroscopy) with the binding energy up to 2 eV, RIXS (resonant inelastic X-ray spectroscopy), and other phenomena including polarization dependence. For other transition metal oxides, the set of essential Wannier functions is chosen separately. The projection procedure with the WFs (see details in [26, 34]) gives the single-electron parameters, ε_λ and $T_{fg}^{\lambda\lambda'}$, of the TB Hamiltonian:

$$H = \sum_{f,\lambda,\sigma} (\epsilon_\lambda - \mu) n_{f\lambda\sigma} + \sum_{f \neq g} \sum_{\lambda,\lambda',\sigma} T_{fg}^{\lambda\lambda'} c_{f\lambda\sigma}^\dagger c_{g\lambda'\sigma} + \frac{1}{2} \sum_{f,g,\lambda,\lambda'} \sum_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} V_{fg}^{\lambda\lambda'} c_{f\lambda\sigma_1}^\dagger c_{f\lambda\sigma_3} c_{g\lambda'\sigma_2}^\dagger c_{g\lambda'\sigma_4}, \tag{5.2}$$

where $c_{f\lambda\sigma}$ is the annihilation operator in the Wannier representation of the hole at the site f on the orbital λ and with the spin σ, $n_{f\lambda\sigma} = c_{f\lambda\sigma}^\dagger c_{f\lambda\sigma}$. The values of Coulomb parameters $V_{fg}^{\lambda\lambda'}$ are obtained by LDA supercell calculations [37]. For Cu in La_2CuO_4, the Hubbard parameter U and the Hund's exchange J are equal to 10 eV and 1 eV, respectively [38].

Another possible approach to calculation of the matrix elements in (5.2) from LDA is to use the NMTO method [39]. For example, the NMTO orbitals have been obtained for the $p-d$ model in $Nd_{2-x}Ce_xCuO_4$ [40]; the Fourier transform of the NMTO Hamiltonian results in the parameters ε_λ and $T_{fg}^{\lambda\lambda'}$ of the TB Hamiltonian.

5.3.2 Step II: Exact Diagonalization

In transition metal (Me) oxides, the unit cell may be chosen as the MeO_n ($n = 6, 5, 4$) cluster and usually there is a common oxygen shared by two adjacent

cells. All other ions provide the electroneutrality and contribute to the high-energy electronic structure. In the low-energy sector, they are inactive. Before ED calculations we should solve the problem of nonorthogonality of the oxygen molecular orbitals of adjacent cells. For the σ-bonding of the $3d$ metal e_g electrons and a_{1g}, b_{1g} oxygen orbitals this problem is solved explicitly using the diagonalization in k-space [41]. We have used the same procedure for the t_{2g} orbitals. Such orthogonalization results in the renormalization of the hopping and Coulomb matrix elements in (5.2).

The effect of renormalization is very important, so we would like to discuss it. Due to the common oxygen in the two unit cells, the MOLCAO (molecular orbital as a liner combination of atomic orbitals) constructed from atomic p_x, p_y, and p_z σ-bonding orbitals are nonorthogonal. The diagonalization of the TB Hamiltonian in k-space [41] results in the following linear combination

$$p_{\boldsymbol{k}\lambda} = \sum_{\mu} c_{\lambda\nu}(\boldsymbol{k})\tilde{p}_{\boldsymbol{k}\nu}, \quad (\lambda = x, y, z), \tag{5.3}$$

with the Fourier transforms of $\tilde{p}_{f\nu}$ (the so-called oxygen group orbitals) that are orthogonal in different cells. In the coordinate space, the group orbital $\tilde{p}_{f\nu}$ is a Wannier-like oxygen wave function centered at the $3d$-metal site $\boldsymbol{R}_f$; it spreads over several neighboring sites. Thus, the renormalizations of the $p-d$ and $p-p$ hoppings are given by

$$\begin{aligned}
\sum_{\boldsymbol{k}} T_{pd}^{\lambda\lambda'} p_{\boldsymbol{k}\lambda}^{\dagger} d_{\boldsymbol{k}\lambda'} &= \sum_{\boldsymbol{k},\nu} T_{pd}^{\lambda\lambda'} c_{\lambda\nu}^{*}(\boldsymbol{k})\tilde{p}_{\boldsymbol{k}\nu}^{\dagger} d_{\boldsymbol{k}\lambda'} = \sum_{f,g} T_{fg}^{\nu\lambda'} \tilde{p}_{f\nu}^{\dagger} d_{g\lambda'}, \\
\sum_{\boldsymbol{k}} T_{pp}^{\lambda\lambda} p_{\boldsymbol{k}\lambda}^{\dagger} p_{\boldsymbol{k}\lambda'} &= \sum_{\boldsymbol{k},\nu,\nu'} T_{pp}^{\lambda\lambda'} c_{\lambda\nu}^{*}(\boldsymbol{k}) c_{\lambda'\nu'}(\boldsymbol{k})\tilde{p}_{\boldsymbol{k}\nu}^{\dagger} p_{\boldsymbol{k}\nu'} = \sum_{f,g,\nu,\nu'} T_{fg}^{\nu\nu'} \tilde{p}_{f\nu}^{\dagger} \tilde{p}_{\boldsymbol{k}\nu'}.
\end{aligned} \tag{5.4}$$

The renormalized $p-d$ hopping parameters $T_{fg}^{\nu\lambda'}$ decrease with distance $\boldsymbol{R} = \boldsymbol{R}_f - \boldsymbol{R}_g$ and are non-zero for the next nearest neighbors even if the initial hopping parameters, T_{pd} and T_{pp}, are taken to be nonzero only for the nearest neighbors. For cuprates as an example, these parameters were calculated up to the sixth coordination sphere and are nonvanishing until the third coordination sphere [24]. It is the reason why the long-range hopping parameters t' and t'' appear and are necessary in the $t-t'-t''-J$ model for cuprates. Similarly, the intercell (long-range, but restricted to several nearest neighbors) Coulomb interaction with participation of oxygen hole is also strongly distance-dependent. The $p-d$ and $p-p$ Coulomb interactions are equal to

$$\begin{aligned}
H_{pd}^{\text{int}} &= \sum_{f,i,j} \sum_{\nu,\lambda,\sigma,\sigma'} V_{pd}\Phi_{fij} n_{df\lambda\sigma} \tilde{p}_{i\nu\sigma'}^{\dagger} \tilde{p}_{j\nu\sigma'}, \\
H_{pp}^{\text{int}} &= \sum_{f,g,i,j} \sum_{\nu} U_p \Psi_{fgij} p_{f\nu\uparrow}^{\dagger} p_{g\nu\uparrow} p_{i\nu\downarrow}^{\dagger} p_{j\nu\downarrow}.
\end{aligned} \tag{5.5}$$

The intracell $p-d$ coefficient is $\Phi_{000}=0.918$, while the intercell ones are $\Phi_{001}=-0.13$ and $\Phi_{002}=-0.02$ [41]. The $p-p$ coefficients are $\Phi_{0000}=0.21$ and $\Phi_{0001}=-0.03$. Thus, the Coulomb interaction is mainly localized inside the one unit cell in the Hamiltonian H_f.

Later we will work with the renormalized parameters. After the orthogonalization, the Hamiltonian (5.2), can be written as a sum of intracell and intercell contributions

$$H = H_c + H_{cc}, \quad H_c = \sum_f H_f, \quad H_{cc} = \sum_{f,g} H_{fg} \tag{5.6}$$

with orthogonal states in different cells described by H_f.

The ED of H_f gives us the set of eigenstates $|p\rangle = |n_p, i_p\rangle$ with the energy E_p. Here, n_p is the number of electrons per unit cell and i_p denotes all other quantum numbers like spin, orbital moment, etc. For cuprates, the hole language is more convenient, so $n=0$ means the hole vacuum ($d^{10}p^6$ configuration), $n=1$ is given by the hole molecular orbitals ($d^9p^6+d^{10}p^5$ configurations), $n=2$ represents the two-hole orbitals ($d^9p^5+d^{10}p^4+d^8p^6$); the Zhang–Rice singlet is the important but not exclusive contribution to the $n=2$ ground state. For manganites, $n=5,4,3$ electron configurations are important and are given by the high-spin terms d^5p^6 ($S=5/2$), $d^4p^6+d^5p^5$ ($S=2$), $d^3p^6+d^4p^5$ ($S=3/2$), correspondingly [27]. For cobaltates, $n=7,6,5$ electron configurations are involved [28]. We should emphasize that we perform the ED with all possible excited eigenstates, not the Lancoz procedure.

How to determine which configurations are relevant? They are found from the local electroneutrality. Let us consider $LaMeO_3$ with Me being a $3d$ element as an example. The ionic valency is $La^{3+}Me^{3+}O_3^{-2}$, thus the $3d$ cation is Me^{3+} (d^4 for Mn^{3+} and d^6 for Co^{3+}). Due to the covalency, the ground state of the unit cell is given by the hybridization of $d^np^6+d^{n+1}p^5$ configurations (in the spectroscopic notations $d^n+d^{n+1}\underline{L}$, where $\underline{L}$ means a ligand hole [42]. Electron addition process results in d^{n+1} subspace of the Hilbert space with a mixture of $d^{n+1}+d^{n+2}\underline{L}$ configuration. Similarly, electron removal results in d^{n-1} subspace with a $d^{n-1}+d^n\underline{L}+d^{n+1}\underline{L}^2$ mixture. Thus for stoichiometric compound, the three relevant subspaces, d^{n-1}, d^n, and d^{n+1}, of the Hilbert space are shown in Fig. 5.2. For each subspace, the ED provides a set of multielectron states $|n,i\rangle$ with an energy $E_i(n)$, $i=0,1,2,\ldots,N_n$. Within this set of multielectron terms, the charge-neutral Bose-type excitations with the energy $\omega_i=E_i(n)-E_0(n)$ are shown by the vertical wavy lines. Electron addition excitations (local Fermi-type QP) have energies $\Omega_{c,i}=E_i(n+1)-E_0(n)$. Here, the index "c" means that QPs form the empty conductivity band. Similarly, the valence band is formed by the electron removal Fermi-type QPs with energies $\Omega_{v,i}=E_0(n)-E_i(n-1)$ (a hole creation with an energy $E_i(n-1)-E_0(n)$ is shown in Fig. 5.2). This multielectron language has been used in the spectroscopy, see for example [42]. The proper mathematical tool to study both the local QP and their intercell hoppings is given by the Hubbard X-operators.

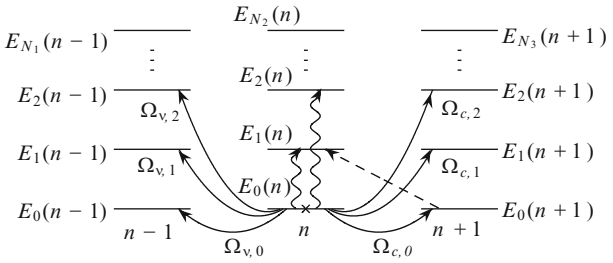

Fig. 5.2 The occupied ground term $E_0(n)$, marked by the *cross*, and excited terms $E_i(n)$ of the d^n configuration as well as electron removal [addition] d^{n-1} [d^{n+1}] sectors of the Hilbert space with energies $E_i(n-1)$ [$E_i(n+1)$]. *Vertical wavy lines* show local Bose-type excitons. *Solid lines* with *arrows* show electron removal (index v) and electron addition (index c) Fermi-type excitations. *Dashed line* shows the virtual Fermi-type excitation from $E_0(n+1)$ to $E_1(n)$

One more question to discuss is why we do not consider the La electrons. It is clear that due to the high stability of the La^{3+} closed shell, the excitations involving La electrons will have a very high energy and are irrelevant in the low-energy physics. Thus for $LaMeO_3$ we take the MeO_6 cluster as the unit cell. The situation may change if instead of La we will have another transition or rare-earth element. How to treat doped or nonstoichiometric compounds will be discussed later in Sect. 5.4.2.

The set of eigenstates $\{|p\rangle\}$ is the complete orthonormalized one and we use it to construct the Hubbard X-operators,

$$X_f^{pq} = |p\rangle \langle q| \tag{5.7}$$

with algebra given by the multiplication rule

$$X_f^{pq} X_f^{rs} = \delta_{qr} X_f^{ps}, \tag{5.8}$$

and by the completeness condition

$$\sum_p X_f^{pp} = 1. \tag{5.9}$$

The last two equations reflect the fact that X-operators are the projective operators; they project the state $|q\rangle$ onto $|p\rangle$. This is the Hubbard definition [4] except for one important difference: Hubbard used the intraatomic states $\{|p\rangle = |0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$, while we use the intracell states$\{|p\rangle\}$. In the Hubbard model, X^{pq} is the 4×4 matrix. In our approach, it is the $N \times N$ matrix, where N is given by the total number of the cell eigenstates; it depends on the number of initial Wannier functions included in the considered energy window. For example, for cuprates $N \approx 100$ and for manganites $N \approx 250$. Nevertheless, in the low-energy limit $N \leq 10$.

The operator X_f^{pq} describes the transition from the initial state $|q\rangle$ to the final state $|p\rangle$, $X^{pq}|q\rangle = |p\rangle$. The important property of X-operators is that any local operator is given by a linear combination of X-operators. Indeed,

$$\hat{O}_f = \hat{1} \cdot \hat{O} \cdot \hat{1} = \sum_{p,q} |p\rangle \langle p| \hat{O}_f |q\rangle \langle q| = \sum_{p,q} \langle p|\hat{O}_f |q\rangle X_f^{pq}. \tag{5.10}$$

The commutation rule for Hubbard operators follows from the (5.8) and is rather awkward. Nevertheless, if $n_p - n_q$ is odd (± 1, ± 3, etc.), then the X^{pq} is called the quasifermionic operator; if $n_p - n_q$ is even (0, ± 2, etc.), the X^{pq} is a quasibosonic one [4]. A single-electron (hole) creation operator is also given by a linear combination in the X-representation:

$$\begin{aligned} c_{f\lambda\sigma} &= \sum_{p,q} \gamma_{\lambda\sigma}(p,q) X_f^{pq}, \\ \gamma_{\lambda\sigma}(p,q) &= \langle p| c_{f\lambda\sigma} |q\rangle . \end{aligned} \tag{5.11}$$

In the 2D plane (p,q), each pair (p,q) corresponds to some vector $\boldsymbol{\alpha}(p,q)$ that is called "root vector" in the diagram technique [43]. The number of states in our Hilbert space is finite, thus we can numerate these vectors $\boldsymbol{\alpha}_m$, $m = 0, 1, 2, ..., N$. We will use a simplified notation

$$X_f^{pq} \to X_f^{\boldsymbol{\alpha}_m} \to X_f^m$$

with each X_f^{pq} being the excitation from $|q\rangle$ to $|p\rangle$. The number m just enumerate the excitation and plays the role of the QP band index. In this notation, (5.11) becomes

$$c_{f\lambda\sigma} = \sum_m \gamma_{\lambda\sigma}(m) X_f^m. \tag{5.12}$$

For practical calculations, a table of correspondence between (p,q) and m is very convenient. In the X-representation, the intracell part of the Hamiltonian is diagonal,

$$H_c = \sum_{f,p} \left(E_p - n_p \mu\right) X_f^{pp}. \tag{5.13}$$

The intercell hopping is given by

$$H_{cc} = \sum_{f \neq g} \sum_{m,m'} t_{fg}^{mm'} X_f^{m\dagger} X_g^{m'}, \tag{5.14}$$

where the matrix elements are

$$t_{fg}^{mm'} = \sum_{\sigma,\lambda,\lambda'} T_{fg}^{\lambda\lambda'} \gamma_{\lambda\sigma}^{*}(m)\gamma_{\lambda'\sigma}(m'). \tag{5.15}$$

All intraatomic $d-d$ Coulomb interactions are included in H_f and since H_f is diagonal in X-operators, they were treated exactly. The dominant part of the $p-d$ and $p-p$ Coulomb interactions, (5.5), is also included in H_f and gives contribution to energies E_p, while a small part of it ($\sim$ 10%, as was discussed after (5.5)) provides the intercell Coulomb interaction that is also bilinear in the X-operators,

$$H_{cc}^{\text{Coul}} = \sum_{f\neq g} \sum_{p,q,p',q'} V_{fg}^{pq,p'q'} X_f^{pq} X_g^{p'q'}. \tag{5.16}$$

Matrix elements $V_{fg}^{pq,p'q'}$ can be written straightforwardly using (5.5) and (5.11).

The single-electron Green function for a particle with a momenta $\boldsymbol{k}$, energy E, spin σ, and orbital indices λ and λ', $G_{\boldsymbol{k}\sigma}^{\lambda\lambda'}(E) \equiv \left\langle\!\left\langle c_{\boldsymbol{k}\lambda\sigma} \middle| c_{\boldsymbol{k}\lambda'\sigma}^{\dagger} \right\rangle\!\right\rangle_E$, is given by a linear combination of the Hubbard operator's Green functions $D_{\boldsymbol{k}\sigma}^{mn}(E) = \left\langle\!\left\langle X_{\boldsymbol{k}\sigma}^{m} \middle| X_{\boldsymbol{k}\sigma}^{n\dagger} \right\rangle\!\right\rangle_E$:

$$G_{\boldsymbol{k}\sigma}^{\lambda\lambda'}(E) = \sum_{mm'} \gamma_{\lambda\sigma}(m)\gamma_{\lambda'\sigma}^{*}(m') D_{\boldsymbol{k}\sigma}^{mm'}(E). \tag{5.17}$$

The intracell Green function is found exactly to be

$$G_{(0)\sigma}^{\lambda\lambda'}(E) = \sum_{m} |\gamma_{\lambda\sigma}(m)|^2 \delta_{\lambda\lambda'} D_{(0)\sigma}^{m}(E), \tag{5.18}$$

where

$$D_{(0)\sigma}^{m}(E) = \frac{F_m}{E - \Omega_m + i\delta}. \tag{5.19}$$

Here, $\Omega_m = E_p(n+1) - E_q(n)$ is the mth QP local energy, and the numerator is equal to $F_m = \langle X^{pp}\rangle + \langle X^{qq}\rangle$. We call F_m "the occupation factor"; it provides a nonzero spectral weight for the QP excitation between at least partially filling eigenstates and gives zero spectral weight for excitations between empty states.

Comparing the exact local Green function in (5.18) with the Lehmann representation in (5.1), we can say that the electron in (5.18) is a linear combination of local fermions (Hubbard fermions) with QP energy Ω_m and a spectral weight $W_m = |\gamma_{\lambda\sigma}(m)|^2 F_m$. It is exactly the same language as in the Lehmann representation. The difference is that it is realized locally and both QP energy and spectral weight are calculated explicitly.

5.4 Step III: Perturbation Theory

The characteristic local energy scale is given by the effective Hubbard parameter $U_{\text{eff}} = E_0(n+1) + E_0(n-1) - 2E_0(n)$ that is given by the difference of the initial $d^n + d^n$ and the excited $d^{n-1} + d^{n+1}$ configurations [42]. The same energy can be obtained as the local gap between the conductivity and valence bands, $U_{\text{eff}} = \Omega_{c,0} - \Omega_{v,0}$. Depending on the ratio of the bare Hubbard U and the charge excitation energy $\Delta_{pd} = \varepsilon_p - \varepsilon_d$, the U_{eff} may represent the Mott–Hubbard gap for $U < \Delta_{pd}$ or the charge transfer (CT) gap E_{CT} for $U > \Delta_{pd}$ [44]. The intercell hopping and interaction result in the dispersion and decrease the energy gap, $E_g < U_{\text{eff}}$. The intercell hoppings, (5.14), and the non-local Coulomb interactions, (5.16), can be treated by a perturbation theory. We would like to emphasize that in the X-representation the perturbation, (5.14), has exactly the same structure as the hopping Hamiltonian in the conventional Hubbard model. That is why the accumulated experience of the Hubbard model study in the X-representation can be used here.

5.4.1 Perturbation Theory in the X-Operators Representation

Before we start the discussion of the perturbation theory it is useful to make general estimations typical for a CPT. If we use the N single electron wave functions to construct the multielectron one, then the ED eigenstates $|p\rangle$ has a normalization factor $1/\sqrt{N}$, $|p\rangle \sim 1/\sqrt{N}$. The X-operator $X^{pq} = |p\rangle\langle q| \sim 1/N$. The electron creation operator $c^\dagger \sim X^{pq} \sim 1/N$. The kinetic energy $t_{fg} c_f^\dagger c_g \sim \frac{1}{N^2} t_{fg}$. The intercell Coulomb interaction $V_{fg} c_f^\dagger c_f c_g^\dagger c_g \sim \frac{1}{N^4} V_{fg}$. For MeO_6 cluster, $N \sim 10$ is a typical value. It means that the renormalized hopping for Hubbard fermions is $t \sim t_{fg}/N^2 \sim 0.1\,\text{eV}$ and intercell Coulomb interaction is $V_{fg}/N^4 \sim 10^{-2} \div 10^{-3}\,\text{eV}$. This estimation explains the values $\Phi_{001} = -0.13$ and $\Phi_{0001} = -0.03$ discussed above for (5.5). It is also clear from this analysis that the most important perturbation is the intercell hopping. For $N \to \infty$, it is the exact statement. A typical value of U_{eff} is $2 \div 4\,\text{eV}$ for transition metal oxides. Thus we have a good small parameter for the perturbation theory, $t/U_{\text{eff}} \sim 0.1$.

To calculate the X-operators' Green function $D^{mn}_{\boldsymbol{k}\sigma}(E)$, we use the equation of motion method. The exact equation of motion for the operator X_f^m is given by

$$i\dot{X}_f^m = \left[X_f^m, H_c + H_{cc}\right]_- = \Omega_m X_f^m + \left[X_f^m, H_{cc}\right]_- . \tag{5.20}$$

The intercell hopping parameters in (5.15) depend only on the distance $\boldsymbol{R} = \boldsymbol{R}_f - \boldsymbol{R}_g$. The exact commutator with H_{cc} is then equal to

$$\left[X_f^m, H_{cc}\right]_- = 2\sum_{R,n_1,n_2} t_{n_1n_2}(R)\left(E_f^{mn_1}X_{f+R}^{n_2} + D_f^{mn_2}X_{f+R}^{n_1\dagger}\right), \tag{5.21}$$

where the operators E_f^{mn} and D_f^{mn} describe the neutral and the doubly charged local bosons, correspondingly:

$$\begin{aligned}\left\{X_f^m, X_g^{n\dagger}\right\}_+ &= \delta_{fg}E_f^{m,n},\\ \left\{X_f^m, X_g^{n}\right\}_+ &= \delta_{fg}D_f^{m,n}.\end{aligned} \tag{5.22}$$

For example, in the $t-J$ model with only one Fermi-type QP $m=0$: $\alpha_0=(0,\sigma)$, $X_f^m = X_f^{0\sigma}$. We have diagonal Bose operators X_f^{00} and $X_f^{\sigma\sigma}$, and the spin operators $S_f^+ = X_f^{\uparrow\downarrow}$, $S_f^- = X_f^{\downarrow\uparrow}$, $S_f^z = \left(X_f^{\uparrow\uparrow} - X_f^{\downarrow\downarrow}\right)/2$. The doubly charged boson matrix $\hat{D}_f$ is zero and the neutral boson matrix $\hat{E}_f$ includes both charge and spin fluctuations,

$$E_f^{m,n} = \delta_{\sigma\sigma'}X_f^{00} + X_f^{\sigma'\sigma}. \tag{5.23}$$

For the single-band Hubbard model there are two Hubbard fermions, $\alpha_0=(0,\sigma)$ and $\alpha_1=(\bar{\sigma},2)$, where $\bar{\sigma}\equiv-\sigma$. The Bose excitations in (5.21) are given by 2×2 matrices,

$$E_f^{m,n} = \begin{pmatrix}\delta_{\sigma\sigma'}X_f^{00} + X_f^{\sigma'\sigma} & 0\\ 0 & \delta_{\sigma\sigma'}X_f^{22} + X_f^{\bar{\sigma},\bar{\sigma}'}\end{pmatrix},$$

$$D_f^{m,n} = \begin{pmatrix}0 & \delta_{\sigma\bar{\sigma}'}X_f^{02}\\ \delta_{\sigma\bar{\sigma}'}X_f^{02} & 0\end{pmatrix}.$$

In the Hubbard model the doubly charged excitations are additional ones to those appearing in the $t-J$ model. For a more complicated model, the explicit forms for the neutral and the charged bosons are obtained straightforwardly using the definition in (5.22).

Using the equation of motion (5.20), one can write down the exact equation for the X-operators' Green function

$$\begin{aligned}(E-\Omega_m)\left\langle\!\left\langle X_f^m \middle| X_g^{m'\dagger}\right\rangle\!\right\rangle_E &= \delta_{mm'}\delta_{fg}F_m\\ &+2\sum_{R,n_1,n_2} t^{n_1n_2}(R)\left[\left\langle\!\left\langle E_f^{mn_1}X_{f+R}^{n_2}\middle| X_g^{m'\dagger}\right\rangle\!\right\rangle_E\right.\\ &\left.+\left\langle\!\left\langle D_f^{mn_2}X_{f+R}^{n_1\dagger}\middle| X_g^{m'\dagger}\right\rangle\!\right\rangle_E\right]\end{aligned} \tag{5.24}$$

It is clear from (5.21) and (5.24) that due to the intercell hoppings, the quasiparticle moves from site f to $f+R$ and this process is accompanied by

the charge/spin fluctuation ($E_f^{mn_1}$) at the departure site f or another quasiparticle is created at the final site $f + R$ with the doublon fluctuation ($D_f^{mn_2}$) at the initial site f.

The simplest solution of (5.24) is obtained via the Hubbard I decoupling of the higher order Green functions [4]. It means that we neglect the local Boson fluctuations and substitute $\hat{E}_f$ and $\hat{D}_f$ with their average values. For the spatially uniform normal state, $E_f^{m,n} \to \left\langle E_f^{m,n} \right\rangle = \delta_{mn} F(m)$ and $D_f^{m,n} \to \left\langle D_f^{m,n} \right\rangle = 0$. This approximation results in the following equation for the Fourier transforms

$$D_{\boldsymbol{k}\sigma}^{mm'}(E) = D_{(0)\sigma}^{m}(E)\,\delta_{mm'} + D_{(0)\sigma}^{m}(E) \sum_{n} t^{mn}(\boldsymbol{k}) D_{\boldsymbol{k}\sigma}^{n,m'}(E) \tag{5.25}$$

with the Hubbard I solution in the matrix form:

$$\hat{D}_{\boldsymbol{k}\sigma}^{-1}(E) = \hat{D}_0^{-1}(E) - \hat{t}(\boldsymbol{k}). \tag{5.26}$$

The dispersion equation for the QP band structure of the Hubbard fermions in this case is given by

$$\det \left\| \delta_{mn} \left(E - \Omega_m\right) / F(m) - t^{mn}(\boldsymbol{k}) \right\| = 0. \tag{5.27}$$

The dispersion equation, (5.27), is similar to the conventional TB dispersion equation, but instead of a single-electron local energy ε_λ we have a local QP energy Ω_m. That is why we call this approach the "generalized TB method." One systematical way to go beyond the Hubbard I approximation is to use the projection operators technique that is considered in the other chapter of this book by Plakida. We will use it below in Sect. 5.5 to study the electronic structure of cuprates in the spin-liquid state.

The other regular way to treat perturbations is the diagram technique. For X-operators with their awkward algebra there is no standard Wick's theorem. Nevertheless, the generalized Wick's theorem has been proved quite some time ago [45]. The first convenient version of the diagram technique has been formulated in [43]. The general rules of the diagram technique for the X-operators are described in details in [29], where the generalized Dyson equation has been obtained

$$\hat{D}_{\boldsymbol{k}\sigma}(E) = \left[\left(E - \Omega_m\right) \delta_{mm'} - \hat{P}_{\boldsymbol{k}\sigma}(E) \hat{t}(\boldsymbol{k}) - \hat{\Sigma}_{\boldsymbol{k}\sigma}(E) \right]^{-1} \hat{P}_{\boldsymbol{k}\sigma}(E). \tag{5.28}$$

Besides the self-energy matrix $\hat{\Sigma}_{\boldsymbol{k}\sigma}(E)$, the unconventional term $\hat{P}_{\boldsymbol{k}\sigma}(E)$ called "the strength operator" appears due to the X-operators algebra. Similar term had been known for the spin-operator diagram technique [46]. This term determines the QP oscillation strength (spectral weight) as well as the renormalized bandwidth. The Hubbard I solution is obtained by setting $\hat{\Sigma}_{\boldsymbol{k}\sigma}(E) = 0$ and $\left[\hat{P}_{\boldsymbol{k}\sigma}(E)\right]_{mm'} = F_m \delta_{mm'}$.

It should be stressed that the LDA+GTB bands are not the single-electron conventional bands. There is no any single-particle Schrödinger equation with the effective potential that gives the LDA+GTB band structure. These QPs are excitations between different multielectron terms. The LDA+GTB bands depend on the multielectron term occupation numbers through $\hat{P}_{k\sigma}(E)$ and $\hat{\Sigma}_{k\sigma}(E)$ that should be calculated via the chemical potential equation. There is no rigid band behavior from the very beginning; the band structure depends on doping, temperature, pressure, and external fields.

5.4.2 Virtual and In-Gap States

The occupation numbers $\langle X^{pp}\rangle$ entering expressions for the X-operators' Green function $\hat{D}_{k\sigma}(E)$ are calculated self-consistently via the chemical potential equation,

$$n_e = \frac{\langle N_e\rangle}{N} = \frac{1}{N}\sum_{f,n,i} n\left\langle X_f^{n,i;\,n,i}\right\rangle. \tag{5.29}$$

The change of the concentration n_e redistributes the occupation numbers and due to the occupation factors, F_m changes the QP band structure.

The most unusual solutions of (5.27) are those with the zero spectral weight ($F_m = 0$). We know the QP energy and this QP has zero number of states. We call such QP the "virtual QP." One of them is shown in Fig. 5.2 by dashed line. To obtain the nonzero spectral weight $F_m \neq 0$ the nonzero occupation of the initial or final states for the excitation $\alpha_m = (p, q)$ is required. It may be achieved by doping, pressure, finite temperature, and external field. The virtual states have been found in GTB calculations for La_2CuO_4 [24] and for $FeBO_3$ [47]. Recently they have been experimentally observed in $FeBO_3$ by the IR (infrared) spectroscopy [48].

The doping dependence of the QP band structure is especially important for the physics of high-T_c superconductivity in cuprates and CMR (colossal magnetoresistance) in manganites. Let us consider a hole doped compound with the electron concentration $n_e = n - x$. At $T = 0$, the self-consistent solution of (5.29) is given by (all other terms have zero occupation)

$$\left\langle X^{n,0;\,n,0}\right\rangle = 1 - x, \quad \left\langle X^{n-1,0;\,n-1,0}\right\rangle = x. \tag{5.30}$$

The finite occupation of the d^{n-1} configuration results in the possibility to have the d^{n-2} final state after the electron removal process. The corresponding low-energy part of the Hilbert space is shown in Fig. 5.3. The QP band structure of the undoped compound will change in several ways. First, the spectral weight of the conductivity band $\Omega_{c,i}$ will decrease due to the occupation factor $F_{c,i} = \left\langle X^{n,0;\,n,0}\right\rangle + \left\langle X^{n+1,i;\,n+1,i}\right\rangle = 1 - x$. There is a redistribution of the spectral weight

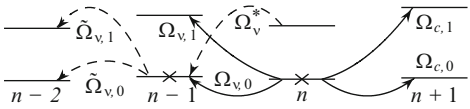

Fig. 5.3 The relevant sectors of the Hilbert space for hole doped compound. *Crosses* means the occupied ground states $E_0(n)$ and $E_0(n-1)$ according to (5.30). *Solid lines* correspond to the valence band and conductivity band QPs $\Omega_{v,i}$ and $\Omega_{c,i}$, respectively. *Dashed lines* show the in-gap QPs with the spectral weight proportional to the doping x

from the conductivity band to the valence band. Dashed lines in Fig. 5.3 show QPs with the occupation factors $F_m = x$; without doping they were the virtual states. Doping results in the finite spectral weight. Especially important is the Ω_v^* with the energy inside the gap, U_{eff}, close to the top of the valence band. Such in-gap states have been experimentally observed in doped cuprates and manganites.

In our general analysis, we neglect the degeneracy of the $E_i(n)$ terms due to spin and other quantum numbers. In practical calculations, this degeneracy is important; each spin sublevel has the occupation number $\sim 1/2S+1$. Simultaneously, the number of different excitations increases due to different spin projections. Details for cuprates can be found in [24, 26] and for manganites in [27].

The in-gap states look like deep impurity levels in the single particle density of states (DOS); they form a peak with intensity $\sim x$. Nevertheless, the physical nature is very different. While the impurity states stem from the fluctuations of the electronic potential, the in-gap states appear in a spatially uniform media with doping due to spectral weight redistribution induced by SEC. Of course, in reality there are impurity induced fluctuations in the doped or off-stoichiometric crystal and both mechanisms may form the in-gap states simultaneously.

5.4.3 *Summary of Step III*

We would like to summarize the perturbation theory approach. The hopping Hamiltonian H_{cc}, (5.14), of the multiband Hubbard model has the structure of the perturbation in the X-representation with a good small parameter $t/U_{\text{eff}} \ll 1$. Depending on the particular physical problem, we use one of the following approaches to treat the perturbation:

1. Hubbard I solution, (5.26), and the corresponding dispersion equation (5.27). This is the most simplest approximation though it provides such nontrivial solutions as the virtual and the in-gap states, and the nonrigid band behavior with the band structure depending on external parameters (doping, temperature, pressure, and external fields).
2. Go beyond the Hubbard I approximation by using the projection operators technique which allows further iterations of the equation of motions for the Green function, (5.24), and a more systematic decoupling in higher orders. One can do

a "maximal" mean-filed theory via this approach; it allows us to study the spin-liquid state of cuprates, as will be discussed in Sect. 5.5.

3. Use the diagram technique to calculate the self-energy, $\hat{\Sigma}_{k\sigma}(E)$, and the strength operator, $\hat{P}_{k\sigma}(E)$, to get the Green function $\hat{D}_{k\sigma}(E)$ via Dyson equation, (5.28). It is the most cumbersome but the most general approach. Virtually, one can obtain the exact solution, but practically it is a very complicated task even in the one-loop approximation.

5.5 LDA+GTB Band Structure of the Hole and Electron Doped Cuprates

In this section we present the results for cuprates with one CuO_2 layer in the unit cell: p-type (hole doped) cuprate $La_{2-x}Sr_xCuO_4$ and n-type (electron doped) cuprate $Sm_{2-x}Ce_xCuO_4$. Yet there is no self-consistent treatment of both electronic and magnetic structures; our LDA+GTB calculations have been carried out for a prescribed magnetic order. Thus, we consider separately the antiferromagnetic (AFM) phase with the long-range AFM order and the spin-liquid phase with the short-range AFM correlations.

5.5.1 LDA+GTB Band Structure of the Undoped La_2CuO_4

The ED of the multiband $p-d$ Hamiltonian (5.2) for the CuO_6 cluster, which includes the apical oxygens, results in the following local eigenstates (in the hole representation): 1) $n_h = 0$, the vacuum state $|0\rangle$ formed by $d^{10}p^6$ orbital configuration, 2) $n_h = 1$, the spin doublets $|\sigma, \lambda\rangle$ with different orbital symmetries. The lowest one is b_{1g}, $|\sigma\rangle$, and the first excited is a_{1g} molecular orbital. 3) $n_h = 2$. A set of two-hole singlets and triplets, spread in the energy region of about $U_d \sim 10\,\text{eV}$. The lowest one is the $^1A_1$ singlet $|S\rangle$ that includes the Zhang–Rice singlet among other two-hole singlets. The first excited triplet $|TM\rangle$ ($M = +1, 0, -1$) has the $^3B_{1g}$ symmetry. The total number of eigenstates is about 100.

The next practical step is the calculation of the matrix elements,

$$\langle 0|\, c_{f\lambda\sigma}\, \left|1\sigma'\lambda\right\rangle, \quad \left\langle 1\sigma'\lambda\right|\, c_{f\lambda\sigma}\, |2, i\rangle\,,$$

and construction of X-operators, (5.11), for all single-electron orbitals. Here, λ stands for Cu-$d_{x^2-y^2}$, Cu-d_{z^2}, O-b, O-a, or O-p_z orbital. For the AFM-ordered La_2CuO_4, we use the two-sublattice (A and B) version of the Hubbard I solution, (5.26), with two occupation factors, $F_{m,A}$ and $F_{m,B}$. Due to the effective molecular field, the local b_{1g} spin doublet is splitted so that at $T = 0$: $\left\langle X_A^{\uparrow\uparrow}\right\rangle = \left\langle X_B^{\downarrow\downarrow}\right\rangle = 1$, $\left\langle X_A^{\downarrow\downarrow}\right\rangle = \left\langle X_B^{\uparrow\uparrow}\right\rangle = 0$. The GTB band structure and the DOS in the wide energy

region with all excited two-hole states $|2,i\rangle$ have been calculated in [49]. The empty conductivity band is formed by only one Hubbard fermion, $X_f^{0,\sigma}$. It is separated by the CT gap $E_{\rm CT} \approx 2\,{\rm eV}$ from the filled valence band. The valence band is formed by a large number of Hubbard fermions $X_f^{\sigma,2i}$ and consists of a set of narrow bands with the total bandwidth about 6 eV. If we are interested in a smaller energy window around the $E_{\rm CT}$ (for example, to study ARPES), it is possible to simplify the calculation by neglecting the high-energy states from both $|2,i\rangle$ and $|1\sigma'\lambda\rangle$ sets. Then the minimal realistic basis is $\{|0\rangle, |\sigma\rangle, |S\rangle, |TM\rangle\}$.

The X-representation for the fermionic operators in this basis is

$$\begin{aligned}
c_{fd_{x^2-y^2}\sigma} &= uX_f^{0,\sigma} + 2\sigma\gamma_x X_f^{\bar\sigma,S},\\
c_{fp_b\sigma} &= vX_f^{0,\sigma} + 2\sigma\gamma_b X_f^{\bar\sigma,S},\\
c_{fp_a\sigma} &= \gamma_a(\sigma\sqrt{2}X_f^{\bar\sigma,T0} - X_f^{\sigma,T2\sigma}),\\
c_{fd_{z^2}\sigma} &= \gamma_z(\sigma\sqrt{2}X_f^{\bar\sigma,T0} - X_f^{\sigma,T2\sigma}),\\
c_{fp_z\sigma} &= \gamma_p(\sigma\sqrt{2}X_f^{\bar\sigma,T0} - X_f^{\sigma,T2\sigma}).
\end{aligned} \tag{5.31}$$

Here, $\bar\sigma \equiv -\sigma$ and $T2\sigma$ stands for $T(+1)$ or $T(-1)$ depending on the value of the spin label $\sigma = \pm 1/2$. The explicit form of the TB Hamiltonian (5.2) in this basis looks like the two-band singlet–triplet Hubbard model:

$$\begin{aligned}
H_{pd} = &\sum_f \left[\varepsilon_1 \sum_\sigma X_f^{\sigma,\sigma} + \varepsilon_{2S} X_f^{S,S} + \varepsilon_{2T}\sum_M X_f^{TM,TM}\right]\\
&+ \sum_{f\neq g,\sigma}\Big[t_{fg}^{00} X_f^{\sigma,0}X_g^{0,\sigma} + t_{fg}^{SS} X_f^{S,\bar\sigma}X_g^{\bar\sigma,S}\\
&+2\sigma t_{fg}^{0S}\left(X_f^{\sigma,0}X_g^{\bar\sigma,S} + h.c.\right)\\
&+t_{fg}^{ST}\left\{\left(\sigma\sqrt{2}X_f^{T0,\bar\sigma} - X_f^{T2\sigma,\sigma}\right)\left(vX_g^{0,\sigma} + 2\sigma\gamma_b X_g^{\bar\sigma,S}\right) + h.c.\right\}\\
&+ t_{fg}^{TT}\left(\sigma\sqrt{2}X_f^{T0,\bar\sigma} - X_f^{T2\sigma,\sigma}\right)\left(\sigma\sqrt{2}X_g^{\bar\sigma,T0} - X_g^{\sigma,T2\sigma}\right)\Big].
\end{aligned} \tag{5.32}$$

The hopping parameters, (5.15), of the effective Hubbard model are expressed through the microscopic ab initio parameters t_{pd} and t_{pp},

$$t_{fg}^{00} = -2t_{pd}\mu_{fg}2uv - 2t_{pp}\nu_{fg}v^2, \tag{5.33}$$

$$t_{fg}^{SS} = -2t_{pd}\mu_{fg}2\gamma_x\gamma_b - 2t_{pp}\nu_{fg}\gamma_b^2,$$

$$t_{fg}^{0S} = -2t_{pd}\mu_{fg}(v\gamma_x + u\gamma_b) - 2t_{pp}\nu_{fg}v\gamma_b, \tag{5.34}$$

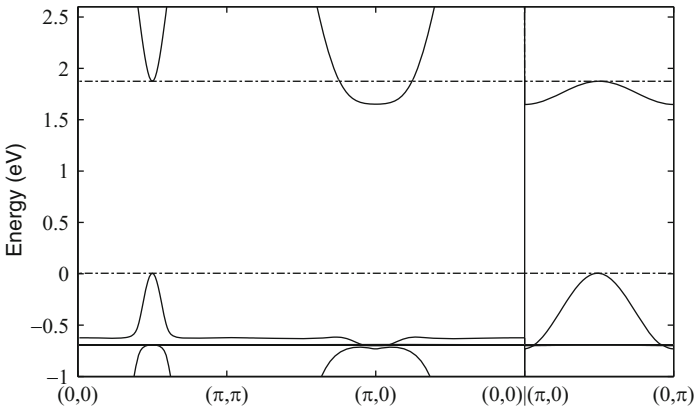

Fig. 5.4 The LDA+GTB band structure for the AFM La_2CuO_4 along the principal cuts of the Brillouin zone [26]

$$t_{fg}^{TT} = \frac{2}{\sqrt{3}} t_{pd} \lambda_{fg} 2\gamma_a \gamma_z + 2t_{pp} \nu_{fg} \gamma_a^2 - 2t'_{pp} \lambda_{fg} 2\gamma_p \gamma_a, \tag{5.35}$$

$$t_{fg}^{ST} = \frac{2}{\sqrt{3}} t_{pd} \xi_{fg} \gamma_z + 2t_{pp} \chi_{fg} \gamma_a - 2t'_{pp} \xi_{fg} \gamma_p. \tag{5.36}$$

Here, μ_{fg}, ν_{fg}, λ_{fg}, ξ_{fg}, and χ_{fg} are the coefficients of the oxygen group orbitals construction, and $u, v, \gamma_x, \gamma_b, \gamma_a, \gamma_p$, and γ_z are the matrix elements $\gamma_{\lambda\sigma}(p,q)$, (5.11) (see [24] for details). The QP band structure of La_2CuO_4 is shown in Fig. 5.4 for the $\Gamma(0,0) - M(\pi,\pi) - X(\pi,0) - \Gamma(0,0)$ and $X(\pi,0) - Y(0,\pi)$ cuts of the square Brillouin zone. Zero at the energy scale is not the Fermi level but rather fixed by the condition $\varepsilon_{d_{x^2-y^2}} = 0$.

The top of the valence band is at the $\bar{M} = (\pi/2, \pi/2)$ point, while the bottom of the conductivity band is at the X- point. The dispersion of the valence band was determined by the hybridization of the two bands, which formed by either $X_f^{\bar{\sigma}S}$ or $X_f^{\bar{\sigma}TM}$ Hubbard fermions. The hybridization between them is provided by the t^{ST} hopping matrix elements in (5.32). These are fermionic bands, but frequently in the literature terms "singlet band" and "triplet band" are used. These terms reflect the final two-hole states involved in the QP excitations. The dominant spectral weight in the singlet band stems from the oxygen b_{1g} states, while for the bottom of the conductivity band it is from the $d_{x^2-y^2}$ states of Cu.

5.5.2 *Low-Energy Effective $t - t' - t'' - J^*$ Model and the Fermi Surface of $Sm_{2-x}Ce_xCuO_4$*

In the doped cuprates the chemical potential shifts either to the valence band (the upper Hubbard band, UHB, for holes) or to the conductivity band (the LHB for

holes). These correspond to p- or n-type dopings, respectively. Because the other band (LHB or UHB) is not participating in the low-energy physics directly, it is possible to project it onto the band forming the Fermi surface. That way, it is possible to construct a more simple $t-J$-type model by the unitary transformation of the multiband Hubbard Hamiltonian (5.32) to fulfill the constraint due to SEC ("no-double occupancy" condition) [50, 51]. Thus the effective $t-t'-t''-J^*$ model is obtained with the Hamiltonian (for electron doping)

$$
\begin{aligned}
H &= H_{t-J} + H_{(3)} \\
H_{t-J} &= \sum_{f,\sigma} (\varepsilon - \mu)\, X_f^{\sigma\sigma} + \sum_{f,g,\sigma} t_{fg} X_f^{\sigma 0} X_g^{0\sigma} + \sum_{f,g} J_{fg} \left(\boldsymbol{S}_f \cdot \boldsymbol{S}_g - \frac{1}{4} n_f n_g \right) \\
H_{(3)} &= \sum_{f \neq g \neq m,\sigma} \frac{\tilde{t}_{fm} \tilde{t}_{mg}}{U_{\text{eff}}} \left(X_f^{\sigma 0} X_m^{\bar{\sigma}\bar{\sigma}} X_g^{0\sigma} - X_f^{\sigma 0} X_m^{\bar{\sigma}\sigma} X_g^{0\bar{\sigma}} \right).
\end{aligned}
\tag{5.37}
$$

Here, t_{fg} is the hopping inside the LHB and equal to t_{fg}^{00} from (5.36). The exchange interaction, $J_{fg} = 2\left(\tilde{t}_{fg}\right)^2 / U_{\text{eff}} = 2\left(t_{fg}^{0S}\right)^2 / U_{\text{eff}}$ is determined by the interband (between the UHB and LHB) hopping t_{fg}^{0S}. $H_{(3)}$ is the three-cite correlated hopping term with the same amplitude $\propto \left(t_{fg}^{0S}\right)^2 / U_{\text{eff}}$ as the exchange interaction. The fact that the latter terms are important was shown in AFM phase by comparing the results of the exact diagonalization for the Hubbard and the $t-J$ models [52]. As discussed above, $U_{\text{eff}} = 2\,\text{eV}$. Due to the Wannier functions partial delocalization, there is a gradual decrease of the hopping parameters t_{fg} with the distance. We have analyzed how the dispersion depends on the number of coordination spheres taken into account and observed the following [26]: there are large differences in dispersions when the first, second, and third nearest-neighbor hopping parameters are nonzero, and negligibly small differences between the dispersions when the third, fourth, and fifth nearest-neighbor hoppings are included. It is a microscopic proof of the necessity of t' and t'' in the $t-t'-t''-J$ model since all hopping and exchange parameters are calculated via LDA. For $Sm_{2-x}Ce_xCuO_4$ these parameters are

$$
\begin{aligned}
t &= -0.59\,\text{eV},\ t' = -0.08t,\ t'' = 0.15t, \\
J &= 0.92|t|, \quad J' = 0.01|t|,\ J'' = 0.02|t|.
\end{aligned}
\tag{5.38}
$$

Electron doped cuprates have the AFM long-range order until large doping values $x \sim 0.1$. $Sm_{2-x}Ce_xCuO_4$ with $x = 0.14$ is at the verge of the AFM order. We have calculated the band structure and the Fermi surface for paramagnetic and AFM phases [53]. A comparison with ARPES [54], see Fig. 5.5, shows that AFM calculations are closer to the experimental data.

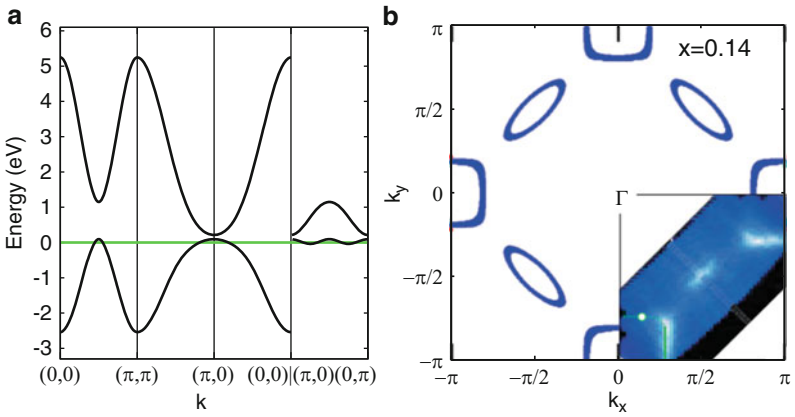

Fig. 5.5 The band structure (**a**) and the Fermi surface (**b**) in the AFM phase for $Sm_{2-x}Ce_xCuO_4$ within the LDA+GTB method for doping concentration $x = 0.14$. In (**a**), zero corresponds to the Fermi level. Inset in (**b**) represents the ARPES data for the Fermi surface [54]

5.5.3 *Doping-Dependent Evolution of the Fermi Surface and Lifshitz Quantum Phase Transitions in $La_{2-x}Sr_xCuO_4$*

The $t - t' - t'' - J^*$ Hamiltonian for p-type $La_{2-x}Sr_xCuO_4$ (LSCO) looks similar to (5.37). The main difference is that instead of the LHB we deal with the UHB fermion $X_f^{\bar{\sigma}2}$. Also, we have disregarded the triplet state $|TM\rangle$, which was present in the Hamiltonian (5.32). The reason is that the triplet itself and the singlet–triplet excitations do not contribute to the near-Fermi level physics.

The ab initio derived parameters for LSCO are (in eV)

$$\begin{aligned} t &= 0.93,\ t' = -0.12,\ t'' = 0.15, \\ \tilde{t} &= 0.77,\ \tilde{t}' = -0.08,\ \tilde{t}'' = 0.12, \\ J &= 0.29,\ J' = 0.003,\ J'' = 0.007. \end{aligned} \tag{5.39}$$

In the wide-range of doping a strong AFM short-range order (spin fluctuations) is known to persist in $La_{2-x}Sr_xCuO_4$. To study its effect on the electronic structure we should go beyond the Hubbard I approximation and include spin correlations at the different sites. It has been done using the Mori-type projection operators technique [55]. The electron Green function has the same structure as given by the Dyson equation (5.28),

$$G_\sigma(\boldsymbol{k}, E) = \frac{(1+x)/2}{E - \varepsilon_0 + \mu - \frac{1+x}{2} t_{\boldsymbol{k}} - \frac{1-x^2}{4} \frac{(\tilde{t}_{\boldsymbol{k}})^2}{U_{\text{eff}}} - \Sigma(\boldsymbol{k})} \tag{5.40}$$

with the strength operator $P_{k\sigma}(E) = F_{\sigma S} = (1+x)/2$ and the static self-energy

$$\Sigma(\boldsymbol{k}) = \frac{2}{1+x}\frac{1}{N}\sum_{\boldsymbol{q}}\left\{\left[t_{\boldsymbol{k}-\boldsymbol{q}} - \frac{1-x}{2}J_{\boldsymbol{q}} + \frac{1-x}{2}\frac{\tilde{t}^2_{\boldsymbol{k}-\boldsymbol{q}}}{U_{\text{eff}}} - \frac{1+x}{2}\frac{2\tilde{t}_{\boldsymbol{k}}\tilde{t}_{\boldsymbol{k}-\boldsymbol{q}}}{U_{\text{eff}}}\right]\right.$$
$$\left.\times\left[\frac{3}{2}C(\boldsymbol{q}) + K(\boldsymbol{k}-\boldsymbol{q})\right] - \frac{1+x}{2}\frac{\tilde{t}^2_{\boldsymbol{q}}}{U_{\text{eff}}}K(\boldsymbol{q})\right\}. \tag{5.41}$$

Here, $K(\boldsymbol{q})$ and $C(\boldsymbol{q})$ are the kinetic and the spin correlation functions,

$$\begin{aligned} K(\boldsymbol{q}) &= \sum_{f-g} \mathrm{e}^{-\mathrm{i}(f-g)\boldsymbol{q}}\left\langle X_f^{S\bar{\sigma}} X_g^{\bar{\sigma}S}\right\rangle, \\ C(\boldsymbol{q}) &= \sum_{f-g} \mathrm{e}^{-\mathrm{i}(f-g)\boldsymbol{q}}\left\langle X_f^{\sigma\bar{\sigma}} X_g^{\bar{\sigma}\sigma}\right\rangle = 2\sum_{f-g} \mathrm{e}^{-\mathrm{i}(f-g)\boldsymbol{q}}\left\langle S_f^z S_g^z\right\rangle. \end{aligned} \tag{5.42}$$

At low temperatures, the spin fluctuations are much slower than the electronic density variations that is why it is possible to neglect a frequency dependence of the spin correlation function and, consequently, the electronic self-energy (see discussion in [56]). We believe that the spatial dependence of the short-range AFM order can be described by the isotropic spin-liquid model with all local spin components being zero, $\left\langle S_f^{\alpha}\right\rangle = 0$, $\alpha = x, y, z$, and a nonzero correlation functions $\left\langle S_f^x S_g^x\right\rangle = \left\langle S_f^y S_g^y\right\rangle = \left\langle S_f^z S_g^z\right\rangle$ for $f \neq g$. Following the method of [57], these functions were calculated self-consistently with the chemical potential and the electronic structure. The doping-dependent nonrigid band structure and Fermi surface [55] are shown in Fig. 5.6.

The strongest doping dependence takes place around the (π, π) point. The Fermi surface for small dopings is a small closed pocket near $(\pi/2, \pi/2)$ point. Its area increases with doping and at $x_{c1} = 0.15 \approx x_{\text{opt}}$ (optimal doping) we found the Lifshitz quantum phase transition (QPT) with the change of the Fermi surface topology. Above x_{c1}, the two Fermi surfaces centered at (π, π) point coexist up to $x_{c2} \approx 0.23$. At the second Lifshitz QPT, the smaller electronic Fermi surface collapses. For $x > x_{c2}$, we have the large-hole Fermi surface centered at (π, π) point. The analysis of the low-temperature anomalies induced by QPT and comparison with experimental data (ARPES, quantum oscillations, and transport) are given in [56, 58]. The coincidence of the x_{c1} and x_{opt} is not occasional. The logarithmic divergence of the DOS at the Fermi level at x_{c1} results in the increasing critical temperature of the d-wave superconductivity. The second QPT is related to the crossover from the Fermi liquid behavior for $x > x_{c2}$ to the non-Fermi liquid behavior for $x < x_{c2}$ [58].

Similar Fermi surface reconstructions have been obtained by the Mori-type projection technique with $\mathrm{Im}\Sigma(\boldsymbol{k}, E) = 0$ [59], with $\mathrm{Im}\Sigma(\boldsymbol{k}, E) \neq 0$ [60], and by the multielectron quantum chemical approach [61]. Recently, the exact diagonalization version of cluster DMFT (CDMFT+ED) was used to study the electronic structure of the doped Mott–Hubbard insulator [62, 63]. The sequence

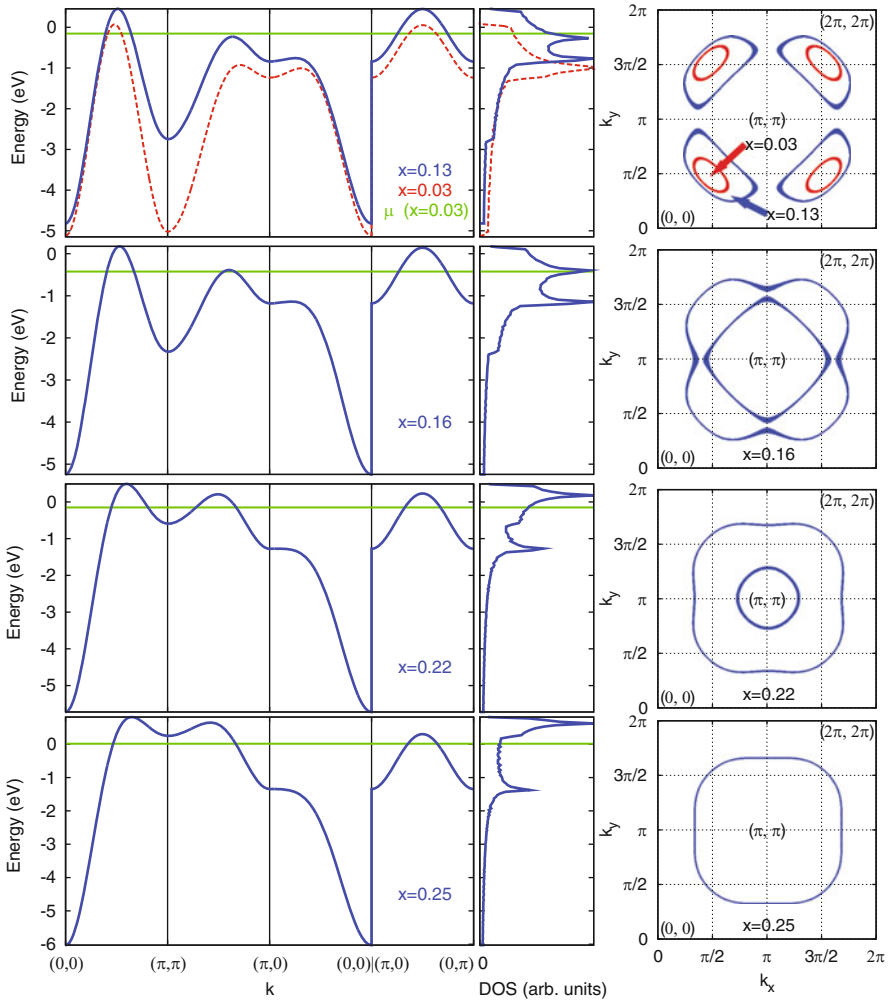

Fig. 5.6 Band structure (on the *left*), density of states (in the *middle*), and Fermi surface (on the *right*) evolution with doping concentration x within the $t-t'-t''-J^*$ model for p-type cuprates

of the FS transformations with doping in these references is very similar to ours in spite of some differences in details.

5.6 LDA+GTB Band Structure of Manganite $La_{1-x}Ca_xMnO_3$

Band structure calculations for manganites is more complicated for two reasons: (a) the orbital ordering double the unit cell and (b) the high spin values for d^4 ($S = 2$) and for $d^{4\pm1}$ ($S = 5/2$ for d^5 and $S = 3/2$ for d^3) increase the number

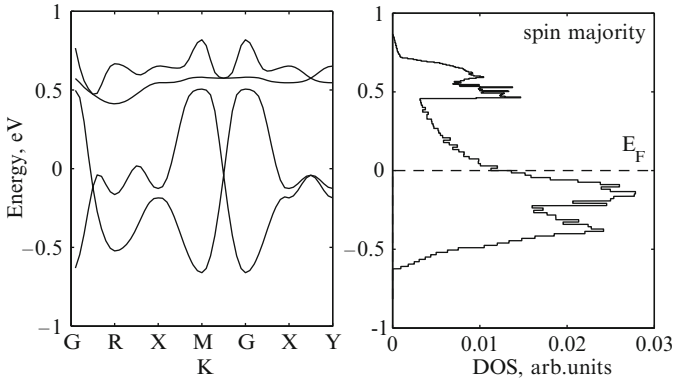

Fig. 5.7 LDA+GTB dispersion of the in-gap and valence bands (on the *left*) and spin majority density of states (on the *right*) for $La_{1-x}Sr_xMnO_3$ in the ferromagnetic phase ($x = 0.3$). The spin minority DOS in the same energy region is zero. Here, $G = (0, 0, 0)$, $R = (\pi, \pi, \pi)$, $X = (\pi, 0, 0)$, $M = (\pi, \pi, 0)$, and $Y = (0, \pi, 0)$ are the main symmetry points of the Brillouin zone

of states in the Hilbert space. The construction of the X-operators representation for $LaMnO_3$ within the high spin d^3, d^4, and d^5 configurations has been done in [27]. For the undoped $LaMnO_3$, the LDA+GTB calculations result in the QP band structure with the large CT gap $E_{CT} \approx 2\,\text{eV}$. Above the top of the occupied valence band there is a virtual state with the activation energy $\Delta\varepsilon \approx 0.4\,\text{eV}$. Doping or nonstoichiometry transforms this state into the in-gap narrow band. It is peculiar that the orbital ordering is necessary to get the insulating state in $LaMnO_3$. The similar LDA+GTB calculations for cubic $LaMnO_3$ without both the orbital ordering and the Jahn–Teller distortion result in the metallic ground state in spite of SEC. The LDA+GTB calculations for hole-doped $La_{1-x}Ca_xMnO_3$ for $x = 0.2 \div 0.3$ resulted in the half-metallic ground state with the 100% spin polarization at the Fermi level in the ferromagnetic phase, that is illustrated in Fig. 5.7.

Transition from the ferromagnetic to the paramagnetic phase results in the twofold band narrowing that is much stronger than the double exchange model provides [64].

5.7 Finite-Temperature Effect on the Electronic Structure of $LaCoO_3$

Perovskite $LaCoO_3$ is known for its unique magnetic properties and a related insulator–metal transition (IMT). All three spin states of Co^{3+} ion in the d^6 configuration [low-spin (LS) $S = 0$, intermediate-spin (IS) $S = 1$, and high-spin (HS) $S = 2$ states] are close in energy. The only general agreement in the literature is that the LS state is the ground state level at $T = 0$. The strong debates during the

last 15 years concern the order of the lowest excited states; their thermal population results in the paramagnetic susceptibility with a maximum around $T = 100$ K. The IMT is not a sharp phase transition but a smooth crossover in the region of 550–600 K. In this section we demonstrate that the LDA+GTB method allows us to treat the strong finite-temperature effect on the band structure of $LaCoO_3$ and to explain both the spin-state transition and the IMT.

Recently, it became clear that the first excited state is the HS one [65,66]. It has a spin $S = 2$ and the orbital moment $l = 1$. Due to the spin–orbit coupling, the ${}^5T_{2g}$ HS term is splitted into the $\tilde{J} = 1$, $\tilde{J} = 2$, and $\tilde{J} = 3$ sublevels [67]. The value of the spin gap Δ_{s-t} between the $\tilde{J} = 1$ HS sublevel and the ${}^1A_1$ LS is small, $\Delta_{s-t} = 140$ K. In the GTB scheme we have to include both d^5 and d^7 configurations to treat the electron removal and addition processes. The multielectron eigenstates for d^5, d^6, and d^7 configurations have been constructed recently in [68]. They include effects of SEC, covalence, and spin–orbital coupling and are presented in Fig. 5.8. The ground state term for the d^5 configuration is the HS ${}^6A_1$. The single electron removal from the ${}^1A_1 d^6$ LS state is prohibited by the spin selection rule. The solid lines with arrows in Fig. 5.8 show the Hubbard fermions that form the top of the valence band ($d^5 \rightarrow d^6$) and the bottom of the conductivity band ($d^6 \rightarrow d^7$) at $T = 0$. With the increasing temperature, the population of the d^6 excited HS states

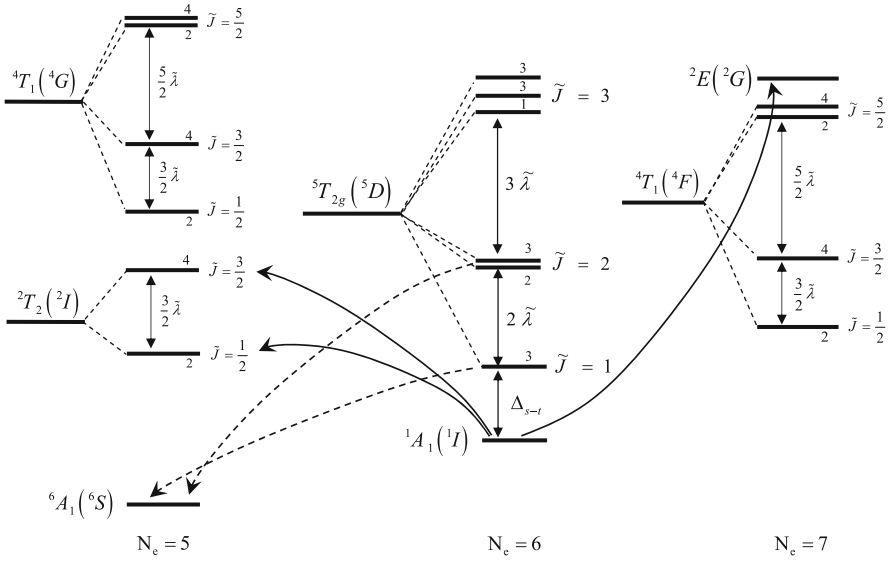

Fig. 5.8 The low-energy part of the Hilbert space for CuO_6 cluster with the electron numbers $N_e = 5, 6, 7$. Terms with a given N_e are the mixtures of d^{N_e}, $d^{N_e+1}\underline{L}$, and $d^{N_e+2}\underline{L}^2$ configurations. At $T = 0$, only the $N_e = 6$ low-spin term ${}^1A_1$ is occupied; the Fermi-type excitations from this term which form the top of the valence band ($d^6 \rightarrow d^5$) and the bottom of the conductivity band ($d^6 \rightarrow d^7$) are shown by the *solid lines* with *arrows*. The *dashed lines* denote the in-gap excitations with the spectral weight increasing with temperature due to the population of the HS excited d^6 terms

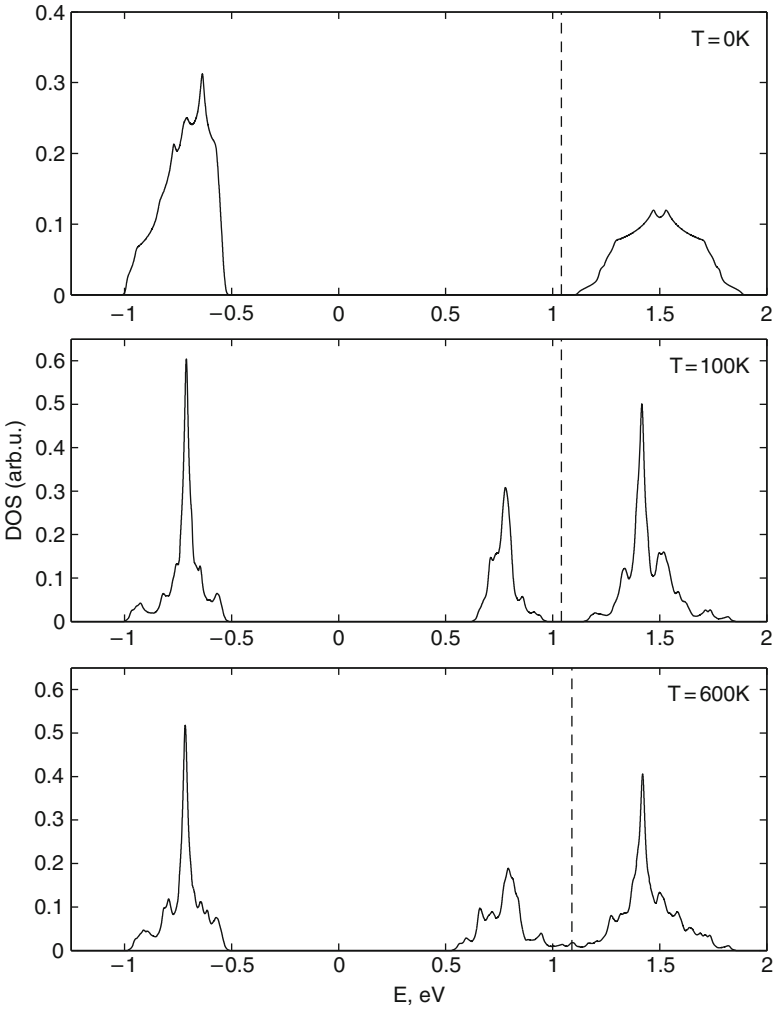

Fig. 5.9 Single-particle density of states at different temperatures. At $T = 0$, $LaCoO_3$ is the charge transfer insulator with the gap $E_g \approx 1.5$ eV. At finite temperatures, the in-gap band appears below the conductivity band with the temperature-dependent activation energy. At $T = 100$ K, $E_a \approx 0.1$ eV. At $T = T^* = 587$ K, $E_a = 0$, and above T^*, the band structure is of the metal type

with $\tilde{J} = 1$ and $\tilde{J} = 2$ results in the appearance of the new in-gap states which are shown by the dashed lines in Fig. 5.8. These states radically change the band structure (Fig. 5.9). At $T = 0$, we obtain the large charge gap $E_g = 1.5$ eV. The in-gap states are below the chemical potential; they acquire finite bandwidth and reduce the gap to $E_g = 0.2$ eV at $T = 100$ K. Their bandwidth is determined by the occupation factor that is equal to the thermal population of the HS states.

With further increase of the temperature, the gap at the Γ point tends to zero at some critical temperature $T = T_{\mathrm{IMT}}$. With the hopping parameters calculated via the LDA+GTB, we found $T_{\mathrm{IMT}} = 587\,\mathrm{K}$ [28]. It is necessary to emphasize that in spite of the large difference in the temperature of the spin-state transition (100 K) and the IMT (600 K), the underlying mechanism for both transitions is the same and is induced by the thermal population of the excited HS states.

5.8 Conclusions

We have presented the main ideas and some computational details of the LDA+GTB method. The space restriction did not allow us to discuss physical results, we have only mentioned them and provided the references. Being invented to study the high-T_{c} superconductivity in cuprates, LDA+GTB method appears to be a powerful approach to systems with SEC and useful for other Mott insulators. We have not discussed the electronic structure of boroxides like $FeBO_3$ and VBO_3 and its change at the spin crossover under the high pressure. One of the most unusual feature of the LDA+GTB band structure is its strong dependence on the external parameters, resulting in the in-gap formation in cuprates due to the doping and in cobaltates due to the temperature. Recently, we have demonstrated that the in-gap states in $FeBO_3$ are induced by the light irradiation [69].

The LDA+GTB method is not universal. Being a combination of the ab initio and model approaches it cannot go beyond the restriction of the model used in GTB. For example, the absence of the long-range Coulomb interaction which determines the Coulomb matrix elements in the large-wavelength limit prevents the correct description of the overdoped cuprates. Of course, it is the common deficiency of all Hubbard-type models. The modern version of the LDA+GTB cannot be used when the perturbation parameter t/U increases and the Mott transition is expected. Nevertheless, it works in the most difficult for conventional band theory region of strong electron correlations. Another variant of a combination of X-operators representation with the ab initio approach for the rare-earth compounds has been developed in [70].

Acknowledgments First of all, we would like to thank people with whom we collaborated to develop the LDA+GTB method: I.A. Nekrasov, Z.V. Pchelkina, Yu.S. Orlov, and V.I. Anisimov. We are also thankful to O.K. Andersen, A.F. Barabanov, P.J. Hirschfeld, K.I. Kikoin, N.M. Plakida, S. Sakai, A.-M.S. Tremblay, V.V. Val'kov, and R.O. Zaitsev for useful discussions.

This work was supported by the presidium of RAS program 18.7 "Quantum physics of condensed matter," the program OFN RAS "Strong electron correlations," RFFI Grants 09-02-00127, 09-02-00171, and 10-02-00251, the Siberian-Ural integration grant #40, Grant of President of Russia MK-1683.2010.2, Russian FTP SC-P891, and the Dynasty Foundation and International Centre for Fundamental Physics in Moscow.

References

1. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964)
2. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965)
3. R.O. Jones, O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
4. J.C. Hubbard, Proc. Roy. Soc. A **285**, 542 (1965)
5. V.I. Anisimov, J. Zaanen, O.K. Andersen, Phys. Rev. B **44**, 943 (1991)
6. A. Svane, O. Gunnarsson, Phys. Rev. Lett. **65**, 1148 (1990)
7. V.I. Anisimov, A.I. Poteryaev, M.A. Korotin, A.O. Anokhin, G. Kotliar, J. Phys. Condens. Matter **9**, 7359 (1997)
8. A.I. Lichtenstein, M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)
9. K. Held, I.A. Nekrasov, N. Blümer, V.I. Anisimov, D. Vollhardt, Int. J. Mod. Phys. B **15**, 2611 (2001)
10. G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
11. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
12. D. Vollhardt, in *Correlated Electron Systems*, ed. by V. J. Emery (World Scientific, Singapore, 1993), p. 57
13. A. Georges, G. Kotliar, W. Krauth, M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
14. M.H. Hettler, A.N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, H.R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998)
15. G. Kotliar, S.Y. Savrasov, G. Pálsson, G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
16. M. Potthoff, Eur. Phys. J. B **32**, 429 (2003)
17. T. Maier, M. Jarrell, T. Pruschke, M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
18. S.Yu. Savrasov, G. Kotliar, Phys. Rev. B **69**, 245101 (2004)
19. R. Del Sole, L. Reining, R.W. Godby, Phys. Rev. B **49**, 8024 (1994)
20. R.W. Godby, M. Schlter, L.J. Sham, Phys. Rev. Lett. **56**, 2415 (1986)
21. F. Aryasetiawan, O. Gunnarsson, Phys. Rev. Lett. **74**, 3221 (1995)
22. S.V. Faleev, M. van Schilfgaarde, T. Kotani, Phys. Rev. Lett. **93**, 126406 (2004)
23. S.G. Ovchinnikov, I.S. Sandalov, Physica C **161**, 607 (1989)
24. V.A. Gavrichkov, S.G. Ovchinnikov, A.A. Borisov, E.G. Goryachev, JETP **91**, 369 (2000)
25. Y.B. Gaididei, V.M. Loktev, Phys. Stat. Sol. B **147**, 307 (1988)
26. M.M. Korshunov, V.A. Gavrichkov, S.G. Ovchinnikov, I.A. Nekrasov, Z.V. Pchelkina, V.I. Anisimov, Phys. Rev. B. **72**, 165104 (2005)
27. V.A. Gavrichkov, S.G. Ovchinnikov, L.E. Yakimov, JETP **102**, 972 (2006)
28. S.G. Ovchinnikov, Yu.S. Orlov, I.A. Nekrasov, Z.V. Pchelkina, JETP **112**, 140 (2011)
29. S.G. Ovchinnikov, V.V. Val'kov, *Hubbard Operators in the Theory of Strongly Correlated Electrons* (Imperial College Press, London, 2004), p. 241
30. P. Fulde, *Electron Correlations in Molecular and Solids* (Springer, Berlin, 1991), p. 422
31. S. Maekawa, T. Tohyama, S.E. Barnes, S. Ishihara, W. Koshibae, G. Khaliullin, *Physics of Transition Metal Oxides* (Springer, Berlin, 2004), p. 377
32. D.N. Zubarev, Sov. Phys. Usp. **3**, 320 (1960)
33. O.K. Andersen, Z. Pawlowska, O. Jepsen, Phys. Rev. B **34**, 5253 (1986)
34. V.I. Anisimov, D.E. Kondakov, A.V. Kozhevnikov, I.A. Nekrasov, Z.V. Pchelkina, J.W. Allen, S.-K. Mo, H.-D. Kim, P. Metcalf, S. Suga, A. Sekiyama, G. Keller, I. Leonov, X. Ren, D. Vollhardt, Phys. Rev. B **71**, 125119 (2005)
35. V.J. Emery, Phys. Rev. Lett. **58**, 2794 (1987)
36. C.M. Varma, S. Smitt-Rink, E. Abrahams, Solid State Commun. **62**, 681 (1987)
37. O. Gunnarsson, O.K. Andersen, O. Jepsen, J. Zaanen, Phys. Rev. B **39**, 1708 (1989)
38. V.I. Anisimov, M.A. Korotin, I.A. Nekrasov, Z.V. Pchelkina, S. Sorella, Phys. Rev. B **66**, 100502(R) (2002)
39. O.K. Andersen, T. Saha-Dasgupta, Phys. Rev. B **62**, 16219 (R) (2000)

40. V.A. Gavrichkov, S.G. Ovchinnikov, I.A. Nekrasov, E.E. Kohorina, Z.V. Pchelkina, Phys. Solid State **49**, 2052 (2007)
41. R. Raimondi, J.H. Jefferson, L.F. Feiner, Phys. Rev. B **53**, 8774 (1996)
42. J. Zaanen, G.A. Sawatzky, J. Solid State Chem. **88**, 8 (1990)
43. R.O. Zaitsev, JETP **43**, 574 (1976)
44. J. Zaanen, G.A. Sawatzky, J.W. Allen, Phys. Rev. Lett. **55**, 418 (1985)
45. B. Westwanski, A. Pawlikovski, Phys. Lett. A **43**, 201 (1973)
46. V.G. Bar'yakhtar, V.N. Krivoruchko, D.A. Yablonskii, *Green's Functions in Magnetism Theory*, in Russian (Nauk. Dumka, Kiev, 1984), p. 336
47. S.G. Ovchinnikov, V.N. Zabluda, JETP **98**, 135 (2004)
48. S.G. Ovchinnikov, B.A. Gizhevskii, N.V. Kazak, V.V. Rudenko, A.V. Telegin, JETP Lett. **90**, 519 (2009)
49. S.G. Ovchinnikov, Phys. Rev. B **49**, 9891 (1994)
50. L.N. Bulaevskii, E.L. Nagaev, D.I. Khomskii, Sov. Phys. JETP **54**, 1562 (1968)
51. K.A. Chao, J. Spalek, A.M. Oles, Sol. Stat. Phys. **10**, 271 (1977)
52. K.J. von Szczepanski, P. Horsch, W. Stephan, M. Ziegler, Phys. Rev. B **41**, 2017 (1990)
53. M.M. Korshunov, E.V. Zakharova, I.A. Nekrasov, Z.V. Pchelkina, S.G. Ovchinnikov, J. Phys. Condens. Matter **22**, 015701 (2009)
54. S.R. Park, Y.S. Roh, Y.K. Yoon, C.S. Leem, J.H. Kim, H. Koh, H. Eisaki, N.P. Armitage, C. Kim, Phys. Rev. B **75**, 060501 (R) (2007)
55. M.M. Korshunov, S.G. Ovchinnikov, Euro. J. Phys. **57**, 271 (2007)
56. S.G. Ovchinnikov, M.M. Korshunov E.I. Shneyder, JETP **109**, 775 (2009)
57. V.V. Val'kov, D.M. Dzebisashvili, JETP **100**, 608 (2005)
58. S.G. Ovchinnikov, E.I. Shneyder, M.M. Korshunov, J. Phys. Condens. Matter **23**, 045701 (2011)
59. A.F. Barabanov, A.A. Kovalev, O.V. Urazaev, A.M. Belemuk, R. Hayn, JETP **92**, 667 (2001)
60. N.M. Plakida, V.S. Oudovenko, JETP **104**, 230 (2007)
61. L. Hozoi, M.S. Laad, P. Fulde, Phys. Rev. B **78**, 165107 (2008)
62. S. Sakai, Y. Motome, M. Imada, Phys. Rev. Lett. **102**, 056404 (2009)
63. S. Sakai, Y. Motome, M. Imada, Phys. Rev. B **82**, 134505 (2010)
64. V.A. Gavrichkov, S.G. Ovchinnikov, Z.V. Pchelkina, J. Phys. Conf. Series **200**, 012046 (2010)
65. S. Noguchi, S. Kawamata, K. Okuda, H. Najiri, M. Motokawa, Phys. Rev. B **66**, 094404 (2002)
66. M.W. Haverkort, Z. Hu, J.C. Cezar, T. Burnus, H. Hartmann, M. Reuther, C. Zobel, T. Lorenz, A. Tanaka, N.B. Brookes, H.H. Hsieh, H.-J. Lin, C.T. Chen, L.H. Tjeng, Phys. Rev. Lett. **97**, 176405 (2006)
67. Z. Ropka, R.J. Radwanski, Phys. Rev. B **67**, 172401 (2003)
68. Yu.S. Orlov, S.G. Ovchinnikov, JETP **109** 322 (2009)
69. S.G. Ovchinnikov, B.A. Gizhevskii, N.V. Kazak, V.V. Rudenko, A.V. Telegin, JETP Lett. **90**, 519 (2009)
70. U. Lundin, I. Sandalov, O. Eriksson, B. Johansson, Solid State Commun. **115**, 7 (2000)

Chapter 6
Projection Operator Method

Nikolay M. Plakida

Abstract A general projection operator method in the equation of motion method for the two-time Green functions is formulated. An exact Dyson equation for an arbitrary Green function is derived. The method is used to consider the single-particle electron Green functions for the Hubbard model within the non-crossing approximation for the self-energy. Strong-coupling superconductivity theory for the model is formulated and an equation for the superconducting T_c is analyzed. We argue that the d-wave pairing with high-T_c can occur in the repulsive Hubbard model, which is mediated by the antiferromagnetic exchange interaction and the spin-fluctuation scattering, both induced by the kinematic interaction for the Hubbard operators. A microscopic theory for the dynamic spin susceptibility within the projection operator method is formulated. The theory is applied to study spin-excitation spectrum for the t–J model within the mode-coupling approximation for the self-energy. A new approach to the theory of the magnetic resonance mode in cuprate superconductors is proposed.

6.1 Introduction

One of the basic models for studies of electronic spectra and superconductivity in strongly correlated electronic systems (SCES), as the cuprate high-temperature superconductors, is the Hubbard model [1]. In the simplest approximation, the model is specified by two parameters: a single-electron hopping matrix element t between the nearest neighbors and a single-site Coulomb energy U:

$$H = -t \sum_{i \neq j\sigma} a^{\dagger}_{i\sigma} a_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}, \tag{6.1}$$

N.M. Plakida (✉)
Joint Institute for Nuclear Researech, Dubna, Moscow region 141980, Russia
e-mail: plakida@theor.jinr.ru

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_6,

where $a^{\dagger}_{i\sigma}(a_{i\sigma})$ are the creation (annihilation) operators for an electron of spin σ on the lattice site i and $n_{i\sigma} = a^{\dagger}_{i\sigma}a_{i\sigma}$ is the electron occupation number. The model (6.1) permits one to consider both cases of weak correlations, $U \ll t$, and of strong correlations, $U \gg t$. In the weak correlation limit, a metallic state is observed, while in the strong correlation limit the model describes a Mott–Hubbard insulating state at half-filling (an average occupation electron number $n = 1$). For hole doping ($n < 1$) of the lower Hubbard subband (LHB), or for electron doping ($n > 1$) of the upper Hubbard subband (UHB) the model describes a strongly correlated metal.

Various methods have been used to investigate the Hubbard model, among which are numerical simulations for finite clusters (see, e.g., [2, 3]), dynamical mean field theory (DMFT) (see, e.g., [4, 5]), the dynamical cluster theory (see, e.g., [6, 7]), and other methods discussed in other chapters of this book. A rigorous analytical method in the limit of strong correlations is based on the Hubbard operator (HO) technique [8] since in this representation the local constraint of no double occupancy of any lattice site is rigorously implemented by the Hubbard operator algebra. However, since the HOs are composite operators with complicated commutation relations, the diagram technique for them is rather complex and only simplest set of diagrams can be taken into account in calculations (see, e.g., [9]).

A more convenient and straightforward technique is based on the equation of motion method for the thermodynamic two-time Green functions (GFs) introduced by Bogoliubov and Tyablikov [10]. By sequential differentiating GFs over time t or t', a chain of equations can be derived. To obtain a closed system of equations, an approximation should be used for higher order GFs, usually called as "truncating" or "decoupling" of GFs [11, 12]. We emphasize here that any type of a decoupling of higher order GFs for Fermi or Bose operators corresponds to a certain set of diagrams in the temperature diagram technique for the causal GFs (see, e.g., [13]). This enables to evaluate which set of diagrams are taken into account in the decoupling approximation and, therefore, to estimate the accuracy of the adopted approximation. It is possible to find such a correspondence also for the GFs for more complicated operators, as spin or Hubbard operators where the diagram technique is much more complex.

In recent years, a consistent way was developed for truncating a system of equation of motion for the GFs based on projection operator method (POM) similar to the Mori [14] memory function (for references, see [15]). The POM was used by various authors (see, e.g., [16–20]) and in the most general way was formulated by Tserkovnikov [21, 22]. In particular, it was shown in [17] that the POM for the two-time spin GF allows to reproduce results of the diagram technique for the spectrum of spin excitations in the ferromagnetic Heisenberg model. The method was also successfully used in studies of strongly anharmonic crystals close to the melting point [18]. An extensive study of the Hubbard model within the POM was performed by Mancini and Avella (for references see [19] and Chap. 3.7). Generally, it has been proved that the Mori-type projection technique for the time-dependent correlation functions provides reliable results in study spin dynamics, dense liquids, and phase

transitions where strong correlations are essential as in SCES (for references see, e.g., [23]).

In the next section, we formulate a general POM in the theory of GFs. The method is applied to study superconductivity within the Hubbard model in Sect. 2 and to consider spin excitations in Sect. 3 in the t–J model.

6.2 Equation of Motion Method for Green Functions

6.2.1 General Formulation

We introduce the thermodynamic retarded (r), advanced (a), and causal (c) two-time Green functions (GFs) as defined by Zubarev [11]:

$$
\begin{aligned}
G^r_{AB}(t-t') &\equiv \langle\langle A(t)|B(t')\rangle\rangle^r = -i\Theta(t-t')\langle A(t)B(t') - \eta B(t')A(t)\rangle,\\
G^a_{AB}(t-t') &\equiv \langle\langle A(t)|B(t')\rangle\rangle^a = i\Theta(t'-t)\langle A(t)B(t') - \eta B(t')A(t)\rangle,\\
G^c_{AB}(t-t') &\equiv \langle\langle A(t)|B(t')\rangle\rangle^c = -i\langle \hat{T}\, A(t)B(t')\rangle,
\end{aligned}
\tag{6.2}
$$

where the time-ordered product of operators is defined as $\hat{T}A(t)\,B(t') = \Theta(t-t')A(t)B(t') + \eta\Theta(t'-t)B(t')A(t)$. Here, the step-function $\Theta(t) = 1$, for $t > 0$ and $\Theta(t) = 0$ for $t < 0$, $A(t) = \exp(\mathrm{i}Ht)A\exp(-\mathrm{i}Ht)$ is a time-dependent operator in the Heisenberg representation and $\langle AB\rangle = (1/Z)\,Tr\{\exp(-\beta H)AB\}$, $Z = Tr\{\exp(-\beta H)\}$ is the statistical average (we set $\beta = 1/T$, $k_\mathrm{B} = 1$, $\hbar = 1$). The parameter $\eta = +1$ is taken for commutator GFs and $\eta = -1$ for anticommutator GFs: $[A, B]_\eta = AB - \eta BA$.

In this notation, equations of motion for all types of GFs are the same and can be written as

$$
i\frac{\mathrm{d}}{\mathrm{d}t}\langle\langle A(t)|B(t')\rangle\rangle = \delta(t-t')\,\langle[A(0), B(0)]_\eta\rangle + \langle\langle[A(t), H]|B(t')\rangle\rangle. \tag{6.3}
$$

The Fourier transformation for the GFs (6.2) and the respective time correlation functions read,

$$
G_{AB}(t-t') = \frac{1}{2\pi}\int_{-\infty}^{+\infty} G_{AB}(\omega)\,\mathrm{e}^{-\mathrm{i}\omega(t-t')}\mathrm{d}\omega, \tag{6.4}
$$

$$
\langle B(t')A(t)\rangle = \frac{1}{2\pi}\int_{-\infty}^{+\infty} J_{BA}(\omega)\,\mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{d}\omega, \tag{6.5}
$$

where Fourier transformation for the correlation function $\langle A(t)B(t')\rangle$ is similar to (6.5) with the spectral function $J_{AB}(\omega) = \mathrm{e}^{-\beta\omega}J_{BA}(-\omega)$. The spectral representation for the retarded (advanced) GF in (6.2) can be written as

$$G^{r,a}_{AB}(E) \equiv \langle\langle A|B\rangle\rangle^{r,a}_{E} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\mathrm{e}^{\beta\omega}-\eta}{E-\omega\pm i\epsilon}\, J_{BA}(\omega)\,\mathrm{d}\omega, \quad \epsilon\to 0^{+}. \quad (6.6)$$

The retarded GF $G^{r}_{AB}(E)$ is an analytic function in the upper half-plane of the complex variable E, $\mathrm{Im}E > 0$, while the advanced GF $G^{a}_{AB}(E)$ is an analytic function in the lower half-plane of the complex variable E, $\mathrm{Im}E < 0$ and therefore, a single analytic function can be introduced, $G_{AB}(E) = G^{r}_{AB}(E)$ for $\mathrm{Im}E > 0$ and $G_{AB}(E) = G^{a}_{AB}(E)$ for $\mathrm{Im}E < 0$ [10]. The analytical GF $G_{AB}(E)$ obeys the dispersion relation

$$\mathrm{Re}\, G^{r,a}_{AB}(\omega) = \pm\frac{1}{\pi}\, P\int_{-\infty}^{+\infty} \frac{\mathrm{d}z}{z-\omega}\,\mathrm{Im}\, G^{r,a}_{AB}(z)\,, \quad (6.7)$$

where P denotes the principal value of the integral. At the same time, the causal GF $G^{c}_{AB}(E)$ in (6.2) has a more complicated than (6.6) spectral representation, which shows that it cannot be analytically continued into the complex E plane at nonzero temperature, $T > 0$ (see [13]). In this respect, $G^{c}_{AB}(E)$ is less convenient for application than the analytical function $G_{AB}(E)$. Moreover, the retarded commutator GFs have a simple physical meaning. They are directly related to the complex admittance, or the generalized susceptibility, $\chi_{AB}(\omega) = -\langle\langle A|B\rangle\rangle^{r}_{\omega,\eta=+1}$, which describes an influence of an external perturbation determined by the operator B on the average value of the dynamical variable A.

From the spectral representation (6.6) we obtain the relation between the spectral density of the correlation function and the GF:

$$\begin{aligned} J_{BA}(\omega) &= i\,\{G_{AB}(\omega+i\epsilon) - G_{AB}(\omega-i\epsilon)\}\,(\mathrm{e}^{\beta\omega}-\eta)^{-1} \\ &= -2\mathrm{Im}\, G_{AB}(\omega+i\epsilon)\,(\mathrm{e}^{\beta\omega}-\eta)^{-1}, \end{aligned} \quad (6.8)$$

where the last relation holds for a real function $J_{BA}(\omega)$ for operators A, B of the same parity with respect to time inversion. The spectral density (6.8) is related to the dynamical structure factor of a system which is measured in scattering experiments and can be used in calculation of the time-dependent correlation functions (6.5).

In study of collective excitations, such as spin or charge fluctuations, it is more convenient to consider the Kubo-Mori relaxation function (cf. [15, 22]),

$$\begin{aligned} \Phi_{AB}(t-t') &\equiv ((A(t)|B(t'))) = -i\Theta(t-t')(A(t)|B(t')) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} ((A|B))_{\omega}\mathrm{e}^{-\mathrm{i}\omega(t-t')}\mathrm{d}\omega. \end{aligned} \quad (6.9)$$

Here, the Kubo–Mori scalar product is defined as,

$$(A(t)|B) = \int_{0}^{\beta} \mathrm{d}\lambda\langle A(t-i\lambda)B\rangle. \quad (6.10)$$

The retarded commutator GF in (6.2) and the relaxation function (6.9) are coupled by the equation

$$\omega((A|B))_\omega = (A|B) + \langle\langle A|B\rangle\rangle_\omega. \tag{6.11}$$

Useful relations follow from (6.2) and (6.9)–(6.11):

$$((i\dot{A}|B))_\omega = ((A|-i\dot{B}))_\omega = \langle\langle A|B\rangle\rangle_\omega, \tag{6.12}$$

$$(i\dot{A}|B) = (A|-i\dot{B}) = \langle[A,B]\rangle, \quad (A|B) = -\langle\langle A|B\rangle\rangle_{\omega=0}, \tag{6.13}$$

where $i\dot{A} = i\,\mathrm{d}A/\mathrm{d}t = [A,H]$.

However, it should be pointed out that the relation (6.8) and (6.12), (6.13) can be used for the commutator GF for ergodic systems only. In those systems, the time-dependent correlation function (6.5) decays with time, $\lim_{t\to\infty}\langle BA(t)\rangle \to 0$ (assuming that $\langle B\rangle = \langle A\rangle = 0$), and its spectral density is a regular function. For a nonergodic system, the spectral density can be written as,

$$J_{AB} = J'_{AB}(\omega) + 2\pi C_{AB}\delta(\omega), \tag{6.14}$$

where the first term is a regular function, while the second term is an irregular part. The latter can be calculated from the pole of the anticommutator GFs (6.2):

$$2C_{AB} = \lim_{\omega\to 0}\omega\,\mathrm{Re}\;G_{AB}(\omega)|_{\eta=-1}. \tag{6.15}$$

The nonergodic constant C_{AB} determines the difference between the isothermal thermodynamic susceptibility $\chi^T_{AB} = (A|B)$ and the Kubo (isolated) static susceptibility $\chi^K_{AB} = -G^r_{AB}(\omega = 0)$:

$$\beta C_{AB} = \chi^T_{AB} - \chi^K_{AB} = \lim_{\omega\to 0}\omega\,\Phi_{AB}(\omega). \tag{6.16}$$

The nonergodic behavior occurs for a dynamical variable A coupled to integrals of motion, $\langle AK_n\rangle \neq 0$, where $[K_n, H] = 0$ [24].

6.2.2 Projection Technique for Green Functions

In this section, we formulate a general projection technique method in the equations of motion for the GFs (6.2) or the relaxation function (6.9). The method is based on the differentiation the GFs over two times, t and t', that enables to derive a Dyson-type equation with a self-energy similar to the memory function in the Mori projection technique [14].

We consider the GF (6.2) for arbitrary dynamical operators A_k, A^+_k,

$$G_{k,k'}(t-t') = \langle\langle A_k(t)|A^+_{k'}(t')\rangle\rangle. \tag{6.17}$$

In a general case, the operators A_k and A_k^+ are vectors and the GF is a matrix. Using (6.3), an equation of motion for the GF (6.17) reads,

$$i\frac{\mathrm{d}}{\mathrm{d}t}\langle\langle A_k(t)|A_{k'}^+(t')\rangle\rangle = \delta(t-t')\,\langle[A_k, A_{k'}]_\eta\rangle + \langle\langle[A_k(t), H]|A_{k'}^+(t')\rangle\rangle. \quad (6.18)$$

It is convenient to extract a linear term in the equation of motion for the original operator A_k in (6.18) by using the Mori-type projection technique:

$$i\,\dot{A}_k = [A_k,\, H] = \sum_q E_{k,q}A_q + Z_k^{(\mathrm{ir})}. \quad (6.19)$$

The *irreducible* $Z_k^{(\mathrm{ir})}$-operator is determined by the orthogonality condition,

$$\langle[Z_k^{(\mathrm{ir})},\, A_{k'}^+]_\eta\,\rangle = \langle Z_k^{(\mathrm{ir})}A_{k'}^+ - \eta A_{k'}^+ Z_k^{(\mathrm{ir})}\rangle = 0\,. \quad (6.20)$$

This defines the *frequency matrix*,

$$E_{kq} = \sum_{k'}\langle[\,[A_k,\, H],\, A_{k'}]_\eta\rangle\; I_{k'q}^{-1}\,, \quad I_{kq} = \langle[A_k, A_q]_\eta\rangle. \quad (6.21)$$

Using the Fourier transformation (6.4), the equation for the GF (6.17) can be written as,

$$G_{k,k'}(\omega) = G_{k,k'}^{(0)}(\omega) + \sum_{q,q'} G_{k,q}^{(0)}(\omega)I_{q,q'}^{-1}\langle\langle Z_{q'}^{(\mathrm{ir})}|A_{k'}^+\rangle\rangle_\omega, \quad (6.22)$$

where we introduced the zero-order GF,

$$G_{k,k'}^{(0)}(\omega) = \sum_q\{\omega\delta_{k,q} - E_{kq}\}^{-1}\,I_{q,k'}\,, \quad (6.23)$$

which defines the excitation spectrum in the generalized mean-field approximation (GMFA).

Differentiating the many-particle GF $\langle\langle Z_q^{(\mathrm{ir})}(t)|A_{k'}^+(t')\rangle\rangle$ over the second time t' and using the same projection procedure as above, the equation (6.22) can be written in the form,

$$G_{k,k'}(\omega) = G_{k,k'}^{(0)}(\omega) + \sum_{q,q'} G_{k,q}^{(0)}(\omega)\,T_{q,q'}(\omega)\,G_{q',k'}^{(0)}(\omega). \quad (6.24)$$

We introduced here the scattering matrix,

$$T_{k,k'}(\omega) = \sum_{q,q'} I_{k,q}^{-1}\langle\langle Z_q^{(\mathrm{ir})}|(Z_{q'}^{(\mathrm{ir})})^+\rangle\rangle_\omega I_{q',k'}^{-1}. \quad (6.25)$$

To obtain the Dyson equation, we define the self-energy matrix $\Sigma_{k,k'}(\omega)$ by the equation,

$$T_{k,k'}(\omega) = \Sigma_{k,k'}(\omega) + \sum_{q,q'} \Sigma_{k,q}(\omega) G^{(0)}_{q,q'}(\omega)\, T_{q',k'}(\omega)\,, \tag{6.26}$$

which shows that the self-energy matrix is a *proper part* of the scattering matrix which has no parts coupled by a single zero-order GF. An exclusion of the free (GMFA) GF in the definition of the self-energy (6.26) is equivalent to an introduction of the projected Liouvillian superoperator for the memory function in the original Mori technique [14]. Thus, we obtain an exact Dyson equation for the GF,

$$G_{k,k'}(\omega) = G^{(0)}_{k,k'}(\omega) + \sum_{q,q'} G^{(0)}_{k,q}(\omega)\, \Sigma_{q,q'}(\omega)\, G_{q',k'}(\omega), \tag{6.27}$$

where the self-energy matrix, defined for the irreducible operators $Z^{(\mathrm{ir})}_k = [A_k,\, H] - \sum_q E_{k,q} A_q$, reads,

$$\Sigma_{k,k'}(\omega) = \sum_{q,q'} I^{-1}_{k,q} \langle\langle Z^{(\mathrm{ir})}_q | (Z^{(\mathrm{ir})}_{q'})^{+} \rangle\rangle^{(\mathrm{proper})}_{\omega} I^{-1}_{q',k'}. \tag{6.28}$$

In comparison with a standard diagram technique, where the self-energy operator in the Dyson equation is written as a full vertex and a product of full GFs for excitations under study, the self-energy operator in the projection technique method is defined by the zeroth-order vertices and the full multiparticle GF which describes inelastic scattering of quasiparticles.

To consider higher order contributions to the self-energy, a general theory proposed by Tserkovnikov[21, 22] can be used. In the theory, formally an exact system of equations is derived for the GFs (6.2) or (6.9) by differentiating them over time t and t'. For a sequence of operators $A_1, A_2, \ldots A_n$ where A_1 is the original dynamic operator for a physical quantity under study and the higher orders operators are given by time derivatives: $i\dot{A}_n = [A_n,\, H]$, an infinite system of coupled equations (in fact, identities) can be written:

$$\begin{aligned} \omega \langle\langle A_n | A^{+}_n \rangle\rangle_{\omega,n-1} &= \langle A_n | A^{+}_n \rangle + \{ \langle\langle i\dot{A}_n | A^{+}_n \rangle + \langle\langle i\dot{A}_n | - i\dot{A}^{+}_n \rangle\rangle_{\omega,n} \} \\ &\quad \times \langle A_n | A^{+}_n \rangle^{-1} \langle\langle A_n | A^{+}_n \rangle\rangle_{\omega,n-1}, \quad n = 1, 2, \ldots, \end{aligned} \tag{6.29}$$

Here, the *irreducible* GFs of the nth order are coupled by the recurrence relation,

$$\begin{aligned} \langle\langle A | B^{+} \rangle\rangle_{\omega,n} &= \langle\langle A | B^{+} \rangle\rangle_{\omega,n-1} \\ &\quad - \langle\langle A | B^{+}_n \rangle\rangle_{\omega,n-1} \langle\langle A_n | B^{+}_n \rangle\rangle^{-1}_{\omega,n-1} \langle\langle A_n | B^{+} \rangle\rangle_{\omega,n-1}, \end{aligned} \tag{6.30}$$

where $\langle\langle A | A^{+}_m \rangle\rangle_{\omega,n} = \langle\langle A_m | B^{+} \rangle\rangle_{\omega,n} = 0$, for $m \le n$. The higher order operators A_n are also irreducible operators being orthogonal in respect to the scalar products $\langle A_n | A^{+}_n \rangle$:

$$\langle A|B^+\rangle_n = \langle A|B^+\rangle_{n-1} - \langle A|B_n^+\rangle_{n-1}\langle A_n|B_n^+\rangle_{n-1}^{-1}\langle A_n|B^+\rangle_{n-1}, \qquad (6.31)$$

where $\langle A|B_m^+\rangle_n = \langle A_m|B^+\rangle_n = 0$, for $m \le n$. In principle, within this approach, higher order corrections to the self-energy can be found. However, in many cases, the so-called mode-coupling approximation (MCA) for the lowest order self-energy (6.28) (equivalent to the noncrossing approximation in the diagram technique) can be used to obtain physically reasonable results. We discuss this approximation in the next sections for particular models.

6.3 Superconducting Pairing in the Hubbard Model

In this section, we apply the POM to study superconducting pairing in the Hubbard model (6.1). There are still hot debates whether the repulsion interaction in the Hubbard model could provide superconducting pairing and explain high-temperature superconductivity observed in cuprates. Our numerical calculations have definitely confirmed such a possibility [25–27] (see also [28–31]).

6.3.1 Hubbard Model

We consider the Hubbard model on a square lattice in a hole representation usually used in describing cuprate superconductors,

$$H = E_1 \sum_{i,\sigma} X_i^{\sigma\sigma} + E_2 \sum_i X_i^{22} + \sum_{i\neq j,\sigma} t_{ij}\{X_i^{\sigma 0}X_j^{0\sigma} + X_i^{2\sigma}X_j^{\sigma 2} + \sigma(X_i^{2\bar{\sigma}}X_j^{0\sigma} + \text{H.c.})\}, \qquad (6.32)$$

where $X_i^{\alpha\beta} = |i\alpha\rangle\langle i\beta|$ are the Hubbard operators (HOs) for the four states: an empty state $|0\rangle$, a singly occupied state $|\sigma\rangle$ with spin $\sigma/2 = (\uparrow,\downarrow)$, $\sigma = \pm 1$, $\bar{\sigma} = -\sigma$, and a doubly occupied state $|2\rangle = |\uparrow\downarrow\rangle$. We denote the single-site repulsion energy by U and introduce $E_1 = \varepsilon_1 - \mu$ and $E_2 = 2E_1 + U$ as the energy levels for the one-hole (with a reference energy ε_1) and the two-hole states, respectively. The model (6.32) can be used to study cuprate superconductors where, within the cell-cluster perturbation theory, the repulsion energy is defined by the charge-transfer gap Δ_{pd} in cuprates, $U = \Delta_{pd}$ (see, e.g., [32, 33]). In this case, the one-hole band is the d-like copper band, $\varepsilon_1 = \varepsilon_d$, and the two-hole band is the Zhang–Rice (ZR) p-d singlet band, $\varepsilon_2 = \varepsilon_d + \varepsilon_p$, [34]. We consider a strongly correlated limit of the Hubbard model, $U \gg t$, with an insulating state at half-filling ($n = 1$) when the atomic representation in terms of HOs is appropriate.

The bare electron dispersion is defined by the hopping parameter t_{ij}, which $\mathbf{k}$-dependence is specified by the equation,

$$t(\mathbf{k}) = 4t\,\gamma(\mathbf{k}) + 4t'\,\gamma'(\mathbf{k}) + 4t''\,\gamma''(\mathbf{k}), \qquad (6.33)$$

where the hopping parameters are equal to t for the nearest neighbors and t', t'' for the second neighbors, which determine the bare (band) dispersion by the functions: $\gamma(\mathbf{k}) = (1/2)(\cos k_x + \cos k_y)$, $\gamma'(\mathbf{k}) = \cos k_x \cos k_y$, and $\gamma''(\mathbf{k}) = (1/2)(\cos 2k_x + \cos 2k_y)$. In the cell-cluster perturbation theory, one can take $t \simeq \Delta_{pd}/8 \simeq 0.4\,\mathrm{eV}$ [33]. The chemical potential μ depends on the average *hole* occupation number,

$$n = \langle N_i \rangle, \quad N_i = \sum_{\sigma} X_i^{\sigma\sigma} + 2X_i^{22}. \tag{6.34}$$

The spin operators in terms of HOs are defined as,

$$S_i^{\sigma} = X_i^{\sigma\bar{\sigma}}, \quad S_i^z = \sum_{\sigma=\pm 1} (\sigma/2)\, X_i^{\sigma\sigma}. \tag{6.35}$$

The HOs satisfy the completeness relation $X_i^{00} + X_i^{\uparrow\uparrow} + X_i^{\downarrow\downarrow} + X_i^{22} = 1$ and the multiplication rules $X_i^{\alpha\beta} X_i^{\gamma\delta} = \delta_{\beta\gamma} X_i^{\alpha\delta}$. From the latter follow, the commutation relations,

$$\left[X_i^{\alpha\beta}, X_j^{\gamma\delta} \right]_{\pm} = \delta_{ij} \left(\delta_{\beta\gamma} X_i^{\alpha\delta} \pm \delta_{\delta\alpha} X_i^{\gamma\beta} \right). \tag{6.36}$$

The upper sign pertains to Fermi-type operators such as $X_i^{0\sigma}$ which change the number of particles and the lower sign pertains to Bose-type operators, for example, the particle number operator N_i in (6.34) or the spin operators S_i^{α} (6.35).

We emphasize that the Hubbard model (6.32) does not involve a dynamical coupling of electrons (holes) to fluctuations of spins or charges. Its role is played by the *kinematic* interaction caused by the non-Fermi nature of commutation relations (6.36), as already has been noted by Hubbard [8]. For example, the equation of motion for the HO $X_i^{\sigma 2}$ has the form,

$$\begin{aligned} i\,\mathrm{d}X_i^{\sigma 2}/\,\mathrm{d}t &= [X_i^{\sigma 2}, H] = (E_1 + U) X_i^{\sigma 2} \\ &+ \sum_{l \neq i, \sigma'} t_{il} \left(B_{i\sigma\sigma'}^{22} X_l^{\sigma' 2} - \sigma B_{i\sigma\sigma'}^{21} X_l^{0\bar{\sigma}'} \right) - \sum_{l \neq i} t_{il}\, X_i^{02} \left(X_l^{\sigma 0} + \sigma X_l^{2\bar{\sigma}} \right), \end{aligned} \tag{6.37}$$

$$\begin{aligned} B_{i\sigma\sigma'}^{22} &= (X_i^{22} + X_i^{\sigma\sigma})\, \delta_{\sigma'\sigma} + X_i^{\sigma\bar{\sigma}}\, \delta_{\sigma'\bar{\sigma}} = (N_i/2 + S_i^z)\, \delta_{\sigma'\sigma} + S_i^{\sigma}\, \delta_{\sigma'\bar{\sigma}}, \\ B_{i\sigma\sigma'}^{21} &= (N_i/2 + S_i^z)\, \delta_{\sigma'\sigma} - S_i^{\sigma}\, \delta_{\sigma'\bar{\sigma}}. \end{aligned} \tag{6.38}$$

Here, $B_{i\sigma\sigma'}^{\alpha\beta}$ are Bose-like operators related to the particle number operator N_i and the spin operators S_i^{α} (6.35).

6.3.2 Dyson Equation

To consider the superconducting pairing in the Hubbard model (6.32), we introduce the 4×4 matrix anticommutator GFs (6.2) [35]

$$\mathsf{G}_{ij\sigma}(\omega) = \langle\langle \hat{X}_{i\sigma} \mid \hat{X}^{\dagger}_{j\sigma}\rangle\rangle_{\omega} = \begin{pmatrix} \hat{G}_{ij\sigma}(\omega) & \hat{F}_{ij\sigma}(\omega) \\ \hat{F}^{\dagger}_{ij\sigma}(\omega) & -\hat{G}_{ji\bar{\sigma}}(-\omega) \end{pmatrix}, \tag{6.39}$$

where $\hat{X}_{i\sigma}$ and $\hat{X}^{\dagger}_{i\sigma} = (X_i^{2\sigma}\ X_i^{\bar{\sigma}0}\ X_i^{\bar{\sigma}2}\ X_i^{0\sigma})$ are the four-component Nambu operators. Because of the two-subband nature of the model (6.32), the normal $\hat{G}_{ij\sigma}$ and anomalous $\hat{F}_{ij\sigma}$ components of the GF are 2×2 matrices which are coupled by the symmetry relations for the GFs [11].

To calculate the GF (6.39), we use the POM introduced above. As a result, we derive the following equation for the frequency matrix (6.21):

$$\mathsf{E}_{ij\sigma} = \langle\{[\hat{X}_{i\sigma}, H], \hat{X}^{\dagger}_{j\sigma}\}\rangle \mathsf{Q}^{-1}, \quad \{A, B\} = AB + BA, \tag{6.40}$$

$$\mathsf{Q} = \langle\{\hat{X}_{i\sigma}, \hat{X}^{\dagger}_{i\sigma}\}\rangle = \hat{\tau}_0 \times \hat{Q}\,, \quad \hat{Q} = \begin{pmatrix} Q_2 & 0 \\ 0 & Q_1 \end{pmatrix}. \tag{6.41}$$

Here, $\hat{\tau}_0$ is the 2×2 unit matrix and in a paramagnetic state the coefficients $Q_2 = \langle X_i^{22} + X_i^{\sigma\sigma}\rangle = n/2$ and $Q_1 = \langle X_i^{00} + X_i^{\bar{\sigma}\bar{\sigma}}\rangle = 1 - Q_2$ depend on the occupation number of holes (6.34) only. In the Q matrix, we neglected anomalous averages of the type $\langle X_i^{02}\rangle$ which give no contribution to the d-wave pairing.

Frequency matrix (6.40) determines quasiparticle (QP) spectrum in the GMFA given by poles of the zeroth-order GF:

$$\mathsf{G}^{(0)}_{\sigma}(\mathbf{q}, \omega) = \left(\omega\tilde{\tau}_0 - \mathsf{E}_{\sigma}(\mathbf{q})\right)^{-1} \mathsf{Q},$$

$$\mathsf{E}_{\sigma}(\mathbf{q}) = (1/N) \sum_{\mathbf{R_{ij}}} \mathsf{E}_{ij\sigma} \exp\{\mathrm{i}(\mathbf{q} \cdot \mathbf{R}_{ij})\}, \tag{6.42}$$

where $\tilde{\tau}_0$ is 4×4 unit matrix. According to (6.27), the Dyson equation in the $(\mathbf{q}, \omega)$-representation can be written as,

$$(\mathsf{G}_{\sigma}(\mathbf{q}, \omega))^{-1} = \left(\mathsf{G}^{0}_{\sigma}(\mathbf{q}, \omega)\right)^{-1} - \Sigma_{\sigma}(\mathbf{q}, \omega), \tag{6.43}$$

where the self-energy matrix (6.28) reads,

$$\Sigma_{\sigma}(\mathbf{q}, \omega) = \mathsf{Q}^{-1} \langle\langle \hat{Z}^{(\mathrm{ir})}_{\mathbf{q}\sigma} \mid \hat{Z}^{(\mathrm{ir})\dagger}_{\mathbf{q}\sigma}\rangle\rangle^{(\mathrm{prop})}_{\omega}\ \mathsf{Q}^{-1}. \tag{6.44}$$

The system of equations (6.40)–(6.44) give an exact representation for the GF (6.39). To obtain a closed system of equations, we have to evaluate the multiparticle GF in the self-energy operator (6.44), which describes processes of inelastic scattering of electrons (holes) on charge and spin fluctuations due to kinematic interaction.

6.3.3 Mean-Field Approximation

The superconducting pairing in the Hubbard model already occurs in the GMFA and is caused by the kinetic exchange interaction as proposed by Anderson [36,37]. It is therefore reasonable to consider the GMFA described by the GF (6.42) separately. Using commutation relations for the HOs (6.36), we evaluate the frequency matrix (6.40),

$$\mathsf{E}_{ij\sigma} = \begin{pmatrix} \hat{\varepsilon}_{ij\sigma} & \hat{\Delta}_{ij\sigma} \\ \hat{\Delta}^{\dagger}_{ij\sigma} & -\hat{\varepsilon}_{ji\bar{\sigma}} \end{pmatrix}, \quad \text{or} \quad \mathsf{E}_{\sigma}(\mathbf{k}) = \begin{pmatrix} \hat{\varepsilon}_{\sigma}(\mathbf{k}) & \hat{\Delta}_{\sigma}(\mathbf{k}) \\ \hat{\Delta}^{\dagger}_{\sigma}(\mathbf{k}) & -\hat{\varepsilon}_{\bar{\sigma}}(\mathbf{k}) \end{pmatrix}. \tag{6.45}$$

The matrix $\hat{\varepsilon}_{\sigma}(\mathbf{k})$ determines the QP spectrum in the two Hubbard subbands in the normal state (for details, see [32,38]),

$$\begin{aligned} \varepsilon_{1,2}(\mathbf{k}) &= (1/2)[\omega_2(\mathbf{k}) + \omega_1(\mathbf{k})] \mp (1/2)\Lambda(\mathbf{k}), \\ \Lambda(\mathbf{k}) &= \{[\omega_2(\mathbf{k}) - \omega_1(\mathbf{k})]^2 + 4W(\mathbf{k})^2\}^{1/2}, \end{aligned} \tag{6.46}$$

where $\omega_1(\mathbf{k}) = 4t\,\alpha_1\gamma(\mathbf{k}) + 4t'\,\beta_1\gamma'(\mathbf{k}) - \mu$, $\omega_2(\mathbf{k}) = 4t\,\alpha_2\gamma(\mathbf{k}) + 4t'\,\beta_2\gamma'(\mathbf{k}) + U - \mu$, and $W(\mathbf{k}) = 4t\,\alpha_{12}\gamma(\mathbf{k}) + 4t'\,\beta_{12}\gamma'(\mathbf{k})$. The kinematic interaction renormalizes the spectrum: $\alpha_{1(2)} = Q_{1(2)}[1 + C_1/Q^2_{1(2)}]$, $\beta_{1(2)} = Q_{1(2)}[1 + C_2/Q^2_{1(2)}]$, $\alpha_{12} = \sqrt{Q_1Q_2}[1 - C_1/Q_1Q_2]$, $\beta_{12} = \sqrt{Q_1Q_2}[1 - C_2/Q_1Q_2]$. Here, we go beyond the Hubbard I renormalization of hopping parameters given by the factors $Q_{1(2)}$ and take into consideration the renormalization caused by spin correlation functions for the nearest and the second nearest neighbors, respectively:

$$C_1 = \langle \mathbf{S}_i \mathbf{S}_{i\pm a_x/a_y} \rangle, \quad C_2 = \langle \mathbf{S}_i \mathbf{S}_{i\pm a_x \pm a_y} \rangle. \tag{6.47}$$

They considerably suppress the hopping parameters for the nearest neighbors: $\alpha_{1(2)} \ll 1$ due to short-range antiferromagnetic (AF) correlations which are large in the underdoped state, $C_1 < 0, |C_1| = 0.1 - 0.2$, and $Q_{1(2)} \sim 1/2$. It should be pointed out that the DMFT fails to take into account a rather strong renormalization induced by AF correlations. For large $U \gg 8t$, we can neglect charge correlations in the renormalization of the frequency matrix $\hat{\varepsilon}_{\sigma}(\mathbf{k})$ by using the mean-field approximation: $\langle N_i N_j \rangle = \langle N_i \rangle \langle N_j \rangle$.

Now we evaluate the anomalous component $\hat{\Delta}_{ij\sigma}$ of matrix (6.45), which determines the superconducting gap. In what follows, we consider only the singlet d-type pairing, which is determined by the anomalous averages at noncoincident sites, $(i \neq j)$. The diagonal matrix components have the forms,

$$\Delta^{22}_{ij\sigma} = -\sigma t_{ij} \langle X^{02}_i N_j \rangle / Q_2, \quad \Delta^{11}_{ij\sigma} = \sigma t_{ij} \langle N_j X^{02}_i \rangle / Q_1. \tag{6.48}$$

Expressing the Fermi operators in terms of the Hubbard operators as $a_{i\sigma} = X^{0\sigma}_i + \sigma X^{\bar{\sigma}2}_i$, we can write the anomalous averages in (6.48) as $\langle a_{i\downarrow} a_{i\uparrow} N_j \rangle = \langle X^{0\downarrow}_i X^{\downarrow 2}_i N_j \rangle = \langle X^{02}_i N_j \rangle$, because the other products of the Hubbard operators

do not contribute, in accordance with the multiplication rules, $X_i^{\alpha\gamma} X_i^{\lambda\beta} = \delta_{\gamma,\lambda} X_i^{\alpha\beta}$. This representation of the anomalous averages in terms of the Fermi operators shows that the pairing occurs at a single site but in different Hubbard subbands.

The anomalous averages $\langle X_i^{02} N_j \rangle$ can be calculated directly by using the equation for the pair commutator GF $L_{ij}(t-t') = \langle\langle X_i^{02}(t) \mid N_j(t')\rangle\rangle$ without *any decoupling* approximations [26]. Below we consider the pair correlation function of the two-hole subband, in which the pairing occurs under the hole doping, $n = 1 + \delta > 1$:

$$\langle X_i^{02} N_j \rangle = -\frac{1}{U} \sum_{m \neq i,\sigma} \sigma t_{im} \langle X_i^{\sigma 2} X_m^{\bar{\sigma} 2} N_j \rangle \simeq -\frac{4t_{ij}}{U} \sigma \langle X_i^{\sigma 2} X_j^{\bar{\sigma} 2} \rangle. \tag{6.49}$$

The last equation is obtained in the two-site approximation, $m = j$, which is typically used in the t–J model. As a result, the equation for the superconducting gap (6.48) in the case of hole doping can be written as,

$$\Delta_{ij\sigma}^{22} = -\sigma\, t_{ij} \langle X_i^{02} N_j \rangle / Q_2 = J_{ij} \langle X_i^{\sigma 2} X_j^{\bar{\sigma} 2} \rangle / Q_2. \tag{6.50}$$

This is equivalent to the gap equation in the t–J model which describes a pairing due to the exchange interaction $J_{ij} = 4\,(t_{ij})^2/U$ [25]. A similar gap equation can be obtained in the one-hole Hubbard subband in the case of electron doping: $\Delta_{ij\sigma}^{11} = J_{ij} \langle X_i^{0\bar{\sigma}} X_j^{0\sigma} \rangle / Q_1$. We thus conclude that anomalous averages (6.48) in the Hubbard model correspond to the anomalous averages in one of the Hubbard subbands (depending on the position of the chemical potential) in the t–J model. Therefore, this is just conventional pairing mediated by the exchange interaction which has been extensively studied within the t–J model (see, e.g., [25] and references therein).

The same anomalous pair correlation functions (6.48) were obtained in the GMFA for the original Hubbard model (6.1) in [28–31]. To calculate the anomalous correlation function $\langle c_{i\downarrow} c_{i\uparrow} N_j \rangle$ in [28, 31], the Roth procedure, based on a decoupling of the operators on the same lattice site in the time-dependent correlation function, $\langle c_{i\downarrow}(t) | c_{i\uparrow}(t') N_j(t') \rangle$, was used. However, the decoupling of the Hubbard operators on the same lattice site is not unique (as it has been really observed in [28, 31]) and unreliable. To avoid the uncontrollable decoupling, kinematical restrictions were imposed on the correlation functions for the Hubbard operators in [29,30]. However, it also has not produced a unique solution for the superconducting equations.

6.3.4 Self-Energy Operator

Self-energy operator (6.44) can be conveniently written in the same form as the GF (6.39),

$$\Sigma_{ij\sigma}(\omega) = \mathsf{Q}^{-1} \begin{pmatrix} \hat{M}_{ij\sigma}(\omega) & \hat{\Phi}_{ij\sigma}(\omega) \\ \hat{\Phi}^{\dagger}_{ij\sigma}(\omega) & -\hat{M}_{ji\bar{\sigma}}(-\omega) \end{pmatrix} \mathsf{Q}^{-1}, \tag{6.51}$$

where the matrices $\hat{M}$ and $\hat{\Phi}$ denote the respective normal and anomalous components of the self-energy operator.

The system of equations for the (4×4) matrix GF (6.39) and the self-energy (6.51) can be reduced to a system of equations for the normal $\hat{G}_\sigma(\mathbf{k}, \omega)$ and anomalous $\hat{F}_\sigma(\mathbf{k}, \omega)$ (2×2) matrix components. By using representations for the frequency matrix (6.45) and the self-energy (6.51), we derive for these components the following system of matrix equations,

$$\hat{G}_\sigma(\mathbf{k}, \omega) = \left(\hat{G}_N(\mathbf{k}, \omega)^{-1} + \hat{\varphi}_\sigma(\mathbf{k}, \omega)\, \hat{G}_N(\mathbf{k}, -\omega)\, \hat{\varphi}^\dagger_\sigma(\mathbf{k}, \omega) \right)^{-1} \hat{Q}\,, \quad (6.52)$$

$$\hat{F}^\dagger_\sigma(\mathbf{k}, \omega) = -\hat{G}_N(\mathbf{k}, -\omega)\, \hat{\varphi}^\dagger_\sigma(\mathbf{k}, \omega)\, \hat{G}_\sigma(\mathbf{k}, \omega). \quad (6.53)$$

In (6.52), we introduced the normal state matrix GF and the matrix superconducting gap function,

$$\hat{G}_N(\mathbf{k}, \omega) = \left(\omega \hat{\tau}_0 - \hat{\varepsilon}(\mathbf{k}) - \hat{M}(\mathbf{k}, \omega)/\hat{Q} \right)^{-1}, \quad (6.54)$$

$$\hat{\varphi}_\sigma(\mathbf{k}, \omega) = \hat{\Delta}_\sigma(\mathbf{k}) + \hat{\Phi}_\sigma(\mathbf{k}, \omega)/\hat{Q}. \quad (6.55)$$

To calculate the self-energy matrix (6.44), we use the noncrossing (NCA) or the self-consistent Born approximation (SCBA), which can be also called as the MCA. In this approximation, Fermi-like excitations described by the operators $X_j = X_j^{0\sigma}$ $(X_j^{\bar{\sigma}2})$ and Bose-like excitations described by the operators B_i (6.38) in the multiparticle GF (6.44) are assumed to propagate independently, and, therefore, their correlation functions at noncoincident lattice sites $(i \neq j, l \neq m)$ can be factored into a product of the corresponding functions:

$$\langle B_i(t) X_j(t) B_l(t') X_m(t') \rangle \simeq \langle X_j(t) X_m(t') \rangle \langle B_i(t) B_l(t') \rangle. \quad (6.56)$$

Using the spectral representation for these correlation functions, we get in the NCA the following expressions for the normal $M^{\alpha\alpha}_\sigma(\mathbf{q}, \omega)$ and anomalous $\Phi^{\alpha\alpha}_\sigma(\mathbf{q}, \omega)$ diagonal components of the self-energy:

$$M^{22}_\sigma(\mathbf{k}, \omega) = \frac{1}{N} \sum_{\mathbf{q}} \int_{-\infty}^{+\infty} \mathrm{d}z\, K^{(+)}(\omega, z|\mathbf{q}, \mathbf{k}-\mathbf{q}) \times \left\{ -(1/\pi)\, \mathrm{Im} \left[G^{22}_\sigma(\mathbf{q}, z) + G^{11}_\sigma(\mathbf{q}, z) \right] \right\}, \quad (6.57)$$

$$\Phi^{22}_\sigma(\mathbf{k}, \omega) = \frac{1}{N} \sum_{\mathbf{q}} \int_{-\infty}^{+\infty} \mathrm{d}z\, K^{(-)}(\omega, z|\mathbf{q}, \mathbf{k}-\mathbf{q}) \times \left\{ -(1/\pi)\, \mathrm{Im} \left[F^{22}_\sigma(\mathbf{q}, z) - F^{11}_\sigma(\mathbf{q}, z) \right] \right\}, \quad (6.58)$$

where $G^{\alpha\alpha}_\sigma(\mathbf{q}, z)$ and $F^{\alpha\alpha}_\sigma(\mathbf{q}, z)$ are given by the diagonal components of the matrices (6.52) and (6.53). Analogous expressions hold for $M^{11}_\sigma(\mathbf{k}, \omega)$ and $\Phi^{11}_\sigma(\mathbf{k}, \omega)$ [35].

The kernel of the integral equations (6.57), (6.58) has a form, similar to the strong-coupling Eliashberg theory [39],

$$K^{(\pm)}(\omega, z|\mathbf{q}, \mathbf{k}-\mathbf{q}) = |t(\mathbf{q})|^2 \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{d\nu}{\omega - z - \nu} \frac{1}{2} \left[\tanh \frac{z}{2T} + \coth \frac{\nu}{2T} \right] \times \mathrm{Im}\, \chi_{\mathrm{sc}}^{(\pm)}(\mathbf{k}-\mathbf{q}, \nu), \tag{6.59}$$

where the interaction is defined by the hopping parameter $t(\mathbf{q})$ (6.33). The spectral density of bosonic excitations in (6.59) is determined by the dynamic susceptibility for the Bose-like operators $B_i(t)$ – spin and number (charge) fluctuations,

$$\chi_{\mathrm{sc}}^{(\pm)}(\mathbf{q}, \nu) = \chi_{\mathrm{s}}(\mathbf{q}, \nu) \pm (1/4)\, \chi_{\mathrm{c}}(\mathbf{q}, \nu)\,, \tag{6.60}$$

where we introduced the dynamical spin and number susceptibility determined by the commutator GF for fluctuations of spins $\chi_{\mathrm{s}}(\mathbf{q}, \nu) = -\langle\langle \mathbf{S}_{\mathbf{q}} | \mathbf{S}_{-\mathbf{q}} \rangle\rangle_{\nu}$, and numbers $\chi_{\mathrm{c}}(\mathbf{q}, \nu) = \langle\langle N_{\mathbf{q}} | N_{-\mathbf{q}} \rangle\rangle_{\nu}$. The renormalized QP spectrum in the two-hole subband in the normal state is determined by the equation (6.54): $\tilde{\varepsilon}_2(\mathbf{k}) \simeq \varepsilon_2(\mathbf{k}) + \mathrm{Re} M_{\sigma}^{22}(\mathbf{q}, \omega = \tilde{\varepsilon}_2(\mathbf{k}))/Q_2$, and the gap function (6.55) is defined as $\varphi_{2,\sigma}(\mathbf{k}, \omega) = \Delta_{\sigma}^{22}(\mathbf{k}) + \Phi_{\sigma}^{22}(\mathbf{k}, \omega)/Q_2$.

In the NCA, the vertex corrections are neglected as in the Migdal–Eliashberg theory. For the electron–phonon system the vertex corrections are small in the adiabatic approximation, as shown by Migdal [40]. The kinematic interaction induced by the intraband hopping is of the same order as the bandwidth and vertex corrections may be important in obtaining quantitative results. However, in the NCA the self-energy is calculated self-consistently allowing to consider a strong coupling limit which plays an essential role both in the renormalization of QP spectra and in superconducting pairing. Thus, this approach can be considered as a first reasonable approximation. Concerning the spin fluctuation contribution, it should be pointed out that the NCA is quite reliable in this case since a certain set of diagrams, in particular the first crossing diagram [41], vanish due to kinematic restrictions for spin scattering processes.

6.3.5 *Equation for Superconducting Gap and* $\mathbf{T_c}$

In this section, we discuss equations for the superconducting gap (6.55) and T_{c} at the case of hole doping. The linearized gap equation in the Matsubara frequency representation, $\omega_n = i\pi T(2n+1)$, can be written as,

$$\varphi_{2,\sigma}(\mathbf{k}, i\omega_n) = \frac{T}{N} \sum_{\mathbf{q}} \sum_{m} \{ J(\mathbf{k}-\mathbf{q}) + \lambda^{(-)}(\mathbf{q}, \mathbf{k}-\mathbf{q} \mid i\omega_n - i\omega_m) \} \times G_N^{22}(\mathbf{q}, -i\omega_m)\, G_N^{22}(\mathbf{q}, i\omega_m)\, \varphi_{2,\sigma}(\mathbf{q}, i\omega_m). \tag{6.61}$$

The interaction function determined by the spin susceptibility is given by the equation,

$$\lambda^{(-)}(\mathbf{q}, \mathbf{k}-\mathbf{q} \mid i\omega_\nu) = -\frac{|t(\mathbf{q})|^2}{\pi} \int_0^\infty \frac{2\nu \, d\nu}{\nu^2 + (\omega_\nu)^2} [\mathrm{Im}\ \chi_{sc}^{(-)}(\mathbf{k}-\mathbf{q}, \nu)]. \quad (6.62)$$

To calculate T_c and to find the gap function, one should solve the linear equation (6.61) in $(\mathbf{k}, \omega_n)$-space which determines the eigenvalue and the eigenfunction. In the strong-coupling limit within the Eliashberg-type theory, the full normal-state GFs (6.54) calculated self-consistently with the self-energy (6.57) should be used as has been implemented for the single-band t–J model in [25].

To discuss the pairing in the Hubbard model, we consider at first a weak-coupling approximation (WCA) for the gap function equation (6.55). In the WCA, the kernel (6.59) of the integral equation (6.58) is approximated by its value near the Fermi energy for $|\omega, z| \ll \mu$ as follows:

$$K(\omega, z|\mathbf{q}, \mathbf{k}-\mathbf{q}) = -|t(\mathbf{q})|^2 \ \chi_{sc}^{(-)}(\mathbf{k}-\mathbf{q}) \ (1/2) \tanh(z/2T), \quad (6.63)$$

where $\chi_{sc}^{(-)}(\mathbf{q}) = \mathrm{Re}\, \chi_{sc}^{(-)}(\mathbf{q}, \nu = 0)$ is the static susceptibility. In the WCA, the self-energy contributions in the normal-state GFs in equation (6.61) are neglected and the equation for the gap function at the Fermi energy $\Delta(\mathbf{k}) = \varphi_{2,\sigma}(\mathbf{k}, \omega = 0)$ can be written approximately as:

$$\Delta(\mathbf{k}) = \frac{1}{N} \sum_{\mathbf{q}} \left[J(\mathbf{k}-\mathbf{q}) - |t(\mathbf{q})|^2 \, \chi_{sc}^{(-)}(\mathbf{k}-\mathbf{q}) \right] \frac{Z^2(\mathbf{q})}{2\varepsilon_2(\mathbf{q})} \tanh \frac{\varepsilon_2(\mathbf{q})}{2T}. \quad (6.64)$$

In this equation, we take into account only the contribution from the two-hole subband on the FS where $Z(\mathbf{q}) = 1 - b(\mathbf{q})$ is the QP weight reduced by the subband hybridization $b(\mathbf{q})$ (see [38]). A numerical solution of this reduced equation has been obtained in [26] in the limit of weak hybridization, $b(\mathbf{q}) \ll 1$, by taking into account both the exchange $J(\mathbf{k}-\mathbf{q})$ and the spin-fluctuation contribution for a model spin-susceptibility $\chi_s(\mathbf{k}-\mathbf{q})$. The d-wave pairing with high $T_c^{\max} \sim 200\,\mathrm{K}$ was found.

Concerning the mechanism of pairing in the Hubbard model, we can draw the following general conclusion from the gap equations (6.61) or (6.64). These equations show that there are two channels of pairing. The first one is mediated by interband hopping and is determined by the AF exchange interaction $J(\mathbf{k}-\mathbf{q})$, which is usually considered in the RVB-type theories [36, 37]. There are no retardation effects for the exchange pairing because of a large hopping energy $U \gg t$ that results in pairing of all electrons (holes) in the doped subband as shown in Fig. 6.1. The second contribution comes from the spin and charge-fluctuation pairing $\propto \chi_{sc}^{(-)}(\mathbf{k}-\mathbf{q})$ induced by intraband hopping, which is only possible in a range of energies $\pm\omega_s$ near the FS, as in the BCS theory. This type of pairing is usually considered in phenomenological spin-fermion models where only the

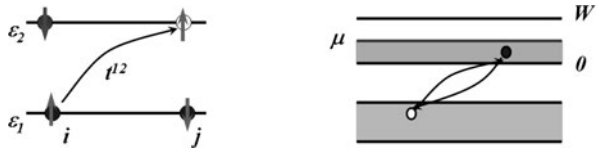

Fig. 6.1 Antiferromagnetic exchange pairing mediated by interband hopping t^{12}

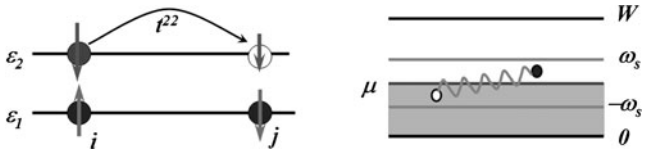

Fig. 6.2 Spin-fluctuation pairing mediated by intraband hopping t^{22}

spin-fluctuation contribution $\propto \chi_s(\mathbf{k}-\mathbf{q})$ is considered as sketched in Fig. 6.2 (for a review, see [42]). The spin-fluctuation interaction is repulsive and can produce pairing only for an alternating-sign gap on the FS, as the d-wave gap.

To estimate both contributions, we consider the equation for the d-wave gap $\Delta(\mathbf{k}) = \Delta_0\,(\cos k_x - \cos k_y)/2 \equiv \Delta_0\,\eta(\mathbf{k})$. By multiplying the equation (6.64) by $\eta(\mathbf{k})$ and integrating over $\mathbf{k}$, we obtain a conventional BCS equation for T_c:

$$1 = \int_{-\mu}^{\widetilde{W}-\mu} \frac{\mathrm{d}\epsilon}{2\epsilon}\,\tanh\frac{\epsilon}{2T_c}[J\,N_d(\epsilon) + \theta(\omega_s - |\epsilon|)\,\lambda_{sf}\,N_{sf}(\epsilon)]\,, \tag{6.65}$$

where $N_d(\epsilon)$ and $N_{sf}(\epsilon)$ are the density of electronic states for the exchange and spin-fluctuation interactions, respectively, weighted with the d-wave factor $[\eta(\mathbf{k})]^2$. In the first term, the integration is extended over the whole effective subband width $\widetilde{W}$, while in the second spin-fluctuation contribution $\lambda_{sf} \sim t_{eff}^2/\omega_s$ the integration is restricted to the characteristic spin-fluctuation energy $\omega_s \sim J$ [26]. By solving equation (6.65) in the conventional logarithmic approximation, we obtain for the exchange pairing $T_c^{ex} = [\mu(\widetilde{W}-\mu)]^{1/2}\exp(-1/V_{ex})$, and for the spin-fluctuation pairing $T_c^{sf} = \omega_s\,\exp(-1/V_{sf})$ where the effective coupling constants $V_{ex} \sim J\,N_d(0)$ and $V_{sf} \sim \lambda_{sf}\,N_{sf}(0)$. It is remarkable that $T_{c,ex}$ is proportional to Fermi energy and even for a weak coupling can be large for $\mu \sim (1/2)\widetilde{W}$. By taking into account both contributions, the superconducting temperature in WCA can be written as

$$T_c^{WCA} = \omega_s\,\exp(-1/\widetilde{V}_{sf}), \quad \widetilde{V}_{sf} = V_{sf} + \frac{V_{ex}}{1 - V_{ex}\ln(\mu/\omega_s)}\,. \tag{6.66}$$

In the conventional superconductors, the nonretarded Coulomb repulsion suppresses the pairing, while in (6.66) the nonretarded exchange attraction V_{ex} enhances the

pairing due to the large ratio μ/ω_s. Therefore, even for originally weak coupling, $V_{sf} \simeq V_{ex} \simeq 0.2$ (e.g., by assuming $\omega_s \simeq J \simeq 0.13\,\text{eV}$, $\lambda_s \simeq 0.2\,\text{eV}$, $N(\mu) \sim 1\,\text{eV}^{-1}$), and for a narrow bandwidth $\widetilde{W} \simeq 8t_{\text{eff}} = 1.6\,\text{eV}$ we obtain $T_c^{\text{ex}} \simeq 60\,\text{K}$ for a half-filled band $\mu = \widetilde{W}/2$ and $T_c^{\text{sf}} \simeq 10\,\text{K}$, while according to (6.66) $T_c^{\text{WCA}} \simeq 200\,\text{K}$ for $\widetilde{V}_{sf} \simeq 0.5$. However, a strong reduction of the WCA value for T_c is observed as we have found in solving of the full gap equation (6.61) in the strong-coupling theory.

To conclude, we have unambiguously proved that the repulsive Hubbard model provides superconducting d-wave pairing mediated by the AFM exchange interaction and spin-fluctuation scattering induced by the kinematic interaction characteristic to systems with strong correlations. These mechanisms of superconducting pairing are absent in the fermionic models (for a discussion, see [43]) and are generic for cuprates. Therefore, we believe that the proposed superconducting pairing is a relevant mechanism of high-temperature superconductivity in copper-oxide materials. For a quantitative comparison of the theoretical results with experiments in cuprates, one should also to consider contributions from the long-range part of the Coulomb interaction (known as the pseudopotential μ^*) and from the electron–phonon interaction. The latter may be important in explaining the isotope shift both for T_c and the magnetic penetration depth λ due to polaron effects in the QP spectra [44] but seems to be not efficient in the superconducting d-wave pairing (for a discussion see [27]).

6.4 Spin-Excitation Spectrum

In the previous section, we consider an application of the POM to study fermionic-type QP excitations in the Hubbard model. At the same time, the Mori projection technique was extensively used in considering bosonic-type excitations as spin waves or density fluctuations (see [15]). In this section, we demonstrate that the POM can be also effectively used in study of the dynamical spin susceptibility (DSS) in SCES within the t–J model, both in the normal and in the superconducting states [45–47]). A similar approach based on a simplified version of the POM and the original Mori memory function has been used in [48–51]. An application of the POM to study the magnetic resonance mode observed in the superconducting cuprates is discussed at the end of the section.

6.4.1 Dynamical Spin Susceptibility

Spin excitation spectrum is described by the DSS related to the retarded commutator GF defined in Sect. 6.2.1:

$$\chi(\mathbf{q},\omega) = -\langle\langle S_{\mathbf{q}}^{+}|S_{-\mathbf{q}}^{-}\rangle\rangle_{\omega} = -\sum_{\mathbf{R}_{ij}} \exp\{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_{ij}\}\,\langle\langle S_i^{+}|S_j^{-}\rangle\rangle_{\omega}. \tag{6.67}$$

For a spin-ordered state, ferromagnetic or antiferromagnetic, the POM can be directly applied to the two-time GF as discussed in Sect. 6.2.2. In particular, the method was used in [17] to show that the POM allows to reproduce results of the diagram technique for the spectrum of spin excitations in the ferromagnetic Heisenberg model [9]. In the paramagnetic state with zero sublattice magnetization, an average of the commutator $\langle [S_i^+, S_j^-] \rangle = 2\delta_{ij} \langle S_i^z \rangle$ vanishes. Since just this quantity enters as an initial condition ($t = 0$) in the equation of motion for the GF (6.3), it is more convenient to construct the self-consistent equation for the Kubo–Mori relaxation function (6.9) $\Phi(\mathbf{q}, \omega) = ((S_{\mathbf{q}}^+ | S_{-\mathbf{q}}^-))_\omega$ coupled to the DSS (6.67) by the equation (6.11),

$$\omega \, \Phi(\mathbf{q}, \omega) = \chi_{\mathbf{q}} - \chi(\mathbf{q}, \omega), \tag{6.68}$$

where $\chi_{\mathbf{q}} \equiv \chi(\mathbf{q}, 0) = (S_{\mathbf{q}}^+ | S_{-\mathbf{q}}^-)$ is the static spin susceptibility.

By differentiating the relaxation function $\Phi(\mathbf{q}, t - t') = ((S_{\mathbf{q}}^+(t) | S_{-\mathbf{q}}^-(t')))$ over two times, t and t', and using the POM we derive the equation,

$$\Phi(\mathbf{q}, \omega) = ((S_{\mathbf{q}}^+ | S_{-\mathbf{q}}^-))_\omega = \frac{\chi_{\mathbf{q}}}{\omega - M(\mathbf{q}, \omega)/\chi_{\mathbf{q}}} \, . \tag{6.69}$$

Here, the memory function (MF) $M(\mathbf{q}, \omega)$ is defined by the *proper* part of the scattering matrix for the spin currents $J_{\mathbf{q}}^{\pm} = i \dot{S}_{\mathbf{q}}^{\pm}$,

$$M(\mathbf{q}, \omega) = ((i \dot{S}_{\mathbf{q}}^+ | - i \dot{S}_{-\mathbf{q}}^-))_\omega^{(\text{proper})}. \tag{6.70}$$

The irreducible spin operator here $\{i \dot{S}_{\mathbf{q}}^{\pm}\}^{\text{irr}} = i \dot{S}_{\mathbf{q}}^{\pm}$ due to the orthogonality condition: $(i \dot{S}_{\mathbf{q}}^{\pm} | \, S_{-\mathbf{q}}^{\mp}) = \langle [S_{\mathbf{q}}^{\pm}, S_{-\mathbf{q}}^{\mp}] \rangle = 0$.

In the study of the single-particle spectrum in Sect. 6.3.4, it was sufficient to use the NCA (6.56) for the self-energy in terms of the first-order time derivatives of the Fermi-like HOs. However, the spectrum of spin-density fluctuations $S_{\mathbf{q}}^{\pm}$ should be defined by the forces $F_{\mathbf{q}}^{\pm} = i^2 \ddot{S}_{\mathbf{q}}^{\pm}$, i.e., the second time derivatives of the spin operators. Therefore, we should consider equations for the current-current relaxation function (6.70) in terms of the forces, $F_{\mathbf{q}}^{\pm} = i \dot{J}_{\mathbf{q}}^{\pm} = i^2 \ddot{S}_{\mathbf{q}}^{\pm}$. By analogy with the derivation of the equations for relaxation function (6.69), we write equations of motion for the MF $((i \dot{S}_{\mathbf{q}}^+(t) | - i \dot{S}_{-\mathbf{q}}^-(t')))$ by differentiating it over two times, t and t' and using the POM. Taking into account that the irreducible spin operators here $\{i \ddot{S}_{\mathbf{q}}^{\pm}\}^{\text{irr}} = i \ddot{S}_{\mathbf{q}}^{\pm}$ due to the orthogonality condition $(-\ddot{S}_{\mathbf{q}}^{\pm} | - i \dot{S}_{-\mathbf{q}}^{\mp}) = \langle [i \dot{S}_{\mathbf{q}}^+, -i \dot{S}_{-\mathbf{q}}^-] \rangle = 0$ the following equation for the MF is derived,

$$M(\mathbf{q}, \omega) = \frac{m(\mathbf{q})}{\omega - \Sigma(\mathbf{q}, \omega)}, \tag{6.71}$$

where the static correlation function

$$m(\mathbf{q}) = (i\,\dot{S}^{+}_{\mathbf{q}}| - i\,\dot{S}^{-}_{-\mathbf{q}}) = \langle [i\,\dot{S}^{+}_{\mathbf{q}}, S^{-}_{-\mathbf{q}}]\rangle. \tag{6.72}$$

The self-energy in (6.71) is defined by the *proper* part of the scattering matrix of the second order,

$$\Sigma(\mathbf{q},\omega) = \frac{1}{m(\mathbf{q})}((-\ddot{S}^{+}_{\mathbf{q}}| - \ddot{S}^{-}_{-\mathbf{q}}))^{(\mathrm{proper})}_{\omega}\,. \tag{6.73}$$

As was pointed out in Sect. 6.2.2, an introduction of the memory function (6.70) and the self-energy (6.73) as the proper parts in the respective equations for the scattering matrices is equivalent to consideration only the projected part of the temporal dependence of correlation functions in the original Mori technique [14].

Using representations (6.69) for the relaxation function and (6.71) for the MF, we obtain the final expression for the relaxation function,

$$\Phi(\mathbf{q},\omega) = \chi_{\mathbf{q}}\frac{\omega - \Sigma(\mathbf{q},\omega)}{\omega^2 - \omega^2_{\mathbf{q}} - \omega\,\Sigma(\mathbf{q},\omega)}, \tag{6.74}$$

where we introduced the frequency of spin excitations in a GMFA,

$$\omega^2_{\mathbf{q}} = m(\mathbf{q})/\chi_{\mathbf{q}}\,. \tag{6.75}$$

Therefore, using (6.68) we obtain an exact representation for the DSS,

$$\chi(\mathbf{q},\omega) = \chi_{\mathbf{q}} - \omega\,\Phi(\mathbf{q},\omega) = -\frac{m(\mathbf{q})}{\omega^2 - \omega^2_{\mathbf{q}} - \omega\,\Sigma(\mathbf{q},\omega)}. \tag{6.76}$$

The spectrum of spin excitations is defined by the imaginary part of the DSS

$$\mathrm{Im}\chi(\mathbf{q},\omega) = \chi''(\mathbf{q},\omega) = \frac{-\omega\,\Sigma''(\mathbf{q},\omega)\,m(\mathbf{q})}{[\omega^2 - \omega^2_{\mathbf{q}} - \omega\,\Sigma'(\mathbf{q},\omega)]^2 + [\omega\,\Sigma''(\mathbf{q},\omega)]^2}\,, \tag{6.77}$$

where we introduced the real part $\Sigma'(\mathbf{q},\omega) = -\Sigma'(\mathbf{q},-\omega)$ and the imaginary part $\Sigma''(\mathbf{q},\omega) = \Sigma''(\mathbf{q},-\omega)$ of the self-energy operator, $\Sigma(\mathbf{q},\omega) = \Sigma'(\mathbf{q},\omega) + i\,\Sigma''(\mathbf{q},\omega)$. They are coupled by the dispersion relation,

$$\Sigma'(\mathbf{q},\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega'}{\omega' - \omega}\,\Sigma''(\mathbf{q},\omega'). \tag{6.78}$$

Coherent propagation of spin excitations is described by the static frequency $\omega_{\mathbf{q}}$, while the self-energy operator $\Sigma(\mathbf{q},\omega)$ defines the incoherent part of the spectrum, determined by inelastic scattering processes. To study the spin-excitation spectrum for a particular system, a certain approximations should be adopted in calculations of the self-energy (6.73), while the static susceptibility $\chi_{\mathbf{q}}$ in (6.75) can be estimated from the dispersion relation for the DSS,

$$\chi_{\mathbf{q}} = \chi(\mathbf{q}, 0) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{\omega} \chi''(\mathbf{q}, \omega). \tag{6.79}$$

However, this nonlinear equation is difficult to treat self-consistently and usually a certain approximate calculation of the static spin-excitation spectrum (6.75) is performed which gives an approximate expression for the static susceptibility.

6.4.2 Spin Susceptibility in the $t-J$ Model

In this section, we derive an expression for the DSS (6.76) for the t–J model. The model follows from the Hubbard model (6.1) in the limit of strong electron correlations, $U \gg t$, by exclusion of virtual hopping between the Hubbard subbands. In terms of the Hubbard operators (HOs), the model reads

$$H = H_t + H_J = -\sum_{i \neq j, \sigma} t_{ij} X_i^{\sigma 0} X_j^{0\sigma} - \mu \sum_{i\sigma} X_i^{\sigma\sigma} + \frac{1}{4} \sum_{i \neq j, \sigma} J_{ij} \left(X_i^{\sigma\bar{\sigma}} X_j^{\bar{\sigma}\sigma} - X_i^{\sigma\sigma} X_j^{\bar{\sigma}\bar{\sigma}} \right), \tag{6.80}$$

where t_{ij} is the hopping integral and J_{ij} is the kinematic exchange interaction. Here, the HOs describe transitions between three possible states at a site i on a square lattice: an empty state $|i, \alpha\rangle = |i, 0\rangle$ and a singly occupied state $|i, \alpha\rangle = |i, \sigma\rangle$ with spin $\sigma/2 = \pm(1/2)$, $(\bar{\sigma} = -\sigma)$. In terms of the HOs, the spin and the number operators read (cf. (6.34) and (6.35)),

$$S_i^{\sigma} = X_i^{\sigma\bar{\sigma}}, \quad S_i^z = \sum_{\sigma} (\sigma/2)\, X_i^{\sigma\sigma}, \quad N_i = \sum_{\sigma} X_i^{\sigma\sigma}. \tag{6.81}$$

The completeness relation in the t–J model for the HOs takes the form: $X_i^{00} + \sum_{\sigma} X_i^{\sigma\sigma} = 1$. The chemical potential μ is determined from the equation for the average electron density $n = \langle N_i \rangle = 1 - \delta$, where $\delta = \langle X_i^{00} \rangle$ is the hole concentration.

6.4.2.1 Static Magnetic Properties

The DSS (6.76) is determined by the static spin-excitation spectrum $\omega_{\mathbf{q}}$ (6.75) and the self-energy (6.73). To calculate the static spin-excitation spectrum $\omega_{\mathbf{q}}$, the equality

$$m(\mathbf{q}) = (-\ddot{S}_{\mathbf{q}}^{+}, S_{-\mathbf{q}}^{-}) = \omega_{\mathbf{q}}^2 (S_{\mathbf{q}}^{+}, S_{-\mathbf{q}}^{-}), \tag{6.82}$$

is used, where the correlation function $(-\ddot{S}_{\mathbf{q}}^{+}, S_{-\mathbf{q}}^{-})$ is evaluated in the GMFA by a decoupling procedure in the site representation as described in [46]. This procedure is equivalent to the MCA for the equal-time correlation functions which results in the spin-excitation spectrum

$$\omega_{\mathbf{q}}^2 = 4J^2(1-\gamma_{\mathbf{q}})\big[\,\lambda_2\frac{n}{2} - \alpha_1 C_{1,0}(4\gamma_{\mathbf{q}}+1) + \alpha_2(2C_{1,1}+C_{2,0})\,\big] + 8t^2\lambda_1(1-\gamma_{\mathbf{q}})[1-n-F_{2,0}-2F_{1,1}]. \tag{6.83}$$

Here t and J are the hopping integral and the exchange interaction for the nearest neighbors, respectively, and $\gamma_{\mathbf{q}} = (1/2)\,(\cos q_x + \cos q_y)$ (we take the lattice spacing $a_x = a_y$ to be unity). The static electron and spin correlation functions are defined as

$$F_{n,m} \equiv F_{\mathbf{R}} = \langle X_{\mathbf{0}}^{\sigma 0}\, X_{\mathbf{R}}^{0\sigma}\rangle = \frac{1}{N}\sum_{\mathbf{q}} F_{\mathbf{q}} e^{i\mathbf{q}\mathbf{R}}, \tag{6.84}$$

$$C_{n,m} \equiv C_{\mathbf{R}} = \langle S_{\mathbf{0}}^{+}\, S_{\mathbf{R}}^{-}\rangle = \frac{1}{N}\sum_{\mathbf{q}} C_{\mathbf{q}} e^{i\mathbf{q}\mathbf{R}}, \tag{6.85}$$

where $\mathbf{R} = n\mathbf{a}_x + m\mathbf{a}_y$. The GMFA spectrum (6.83) is calculated self-consistently by using the MFA approximation for the static correlation function (6.85),

$$C_{\mathbf{q}} = \langle S_{\mathbf{q}}^{+} S_{-\mathbf{q}}^{-}\rangle = \frac{m(\mathbf{q})}{2\,\omega_{\mathbf{q}}}\,\coth\frac{\beta\,\omega_{\mathbf{q}}}{2}\,. \tag{6.86}$$

The decoupling parameters α_1, α_2 and λ_1, λ_2 in (6.83) take into account the vertex renormalization for the spin–spin and electron–spin interaction, respectively, as explained in [46]. In particular, the parameters α_1, α_2 are evaluated from the results for the Heisenberg model at $\delta = 0$ and are kept fixed for $\delta \neq 0$. The parameters λ_1, λ_2 are calculated from the sum rule $C_{0,0} = \langle S_0^+ S_0^-\rangle = (1/2)(1-\delta)$ with a fixed ratio $\lambda_1/\lambda_2 = 0.378$. The variation of the parameters λ_1, λ_2 below the superconducting transition is negligibly small and practically has no influence on the spectrum $\omega_{\mathbf{q}}$. The direct calculation of $m(\mathbf{q})$ yields

$$m(\mathbf{q}) = -8\,(1-\gamma_{\mathbf{q}})\,[J\,C_{1,0} + t\,F_{1,0}]\,. \tag{6.87}$$

Thus, the static susceptibility $\chi_{\mathbf{q}} = m(\mathbf{q})/\omega_{\mathbf{q}}^2$ is explicitly determined by (6.83) and (6.87).

The AF correlation length ξ is calculated by expanding the static susceptibility in the neighborhood of the AF wave vector $\mathbf{Q} = (\pi, \pi)$, $\chi_{\mathbf{Q}+\mathbf{k}} = \chi_{\mathbf{Q}}/(1+\xi^2 k^2)$ which results in the expression,

$$\xi^2 = 8J^2\alpha_1|C_{1,0}|/\omega_{\mathbf{Q}}^2\,. \tag{6.88}$$

The critical behavior of the model (6.80) is reflected by the divergence of $\chi_{\mathbf{Q}}$ and ξ as $T \to 0$, i.e., by $\omega_{\mathbf{Q}}(T=0) = 0$. In the phase with AF long-range order which, in two dimensions, may occur at $T = 0$ only, the correlation function $C_{\mathbf{R}}$ (6.85) should be written as

$$C_{\mathbf{R}} = \frac{1}{N} \sum_{\mathbf{q} \neq \mathbf{Q}} C_{\mathbf{q}} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}} + C\, \mathrm{e}^{\mathrm{i}\mathbf{Q}\mathbf{R}}. \tag{6.89}$$

The condensation part C determines the staggered magnetization which is defined in the spin-rotation-invariant form as,

$$m^2 = \frac{3}{2N} \sum_{\mathbf{R}} C_{\mathbf{R}} \mathrm{e}^{-\mathrm{i}\mathbf{Q}\mathbf{R}} = \frac{3}{2} C. \tag{6.90}$$

The static susceptibility, the correlation functions, the correlation length, and the magnetization were calculated in the GMFA for arbitrary temperatures and doping in [46]). A good agreement with numerical simulations for finite clusters and experiments in cuprates was found.

6.4.2.2 Spin Dynamics

To study spin dynamics and the spin-excitation spectrum determined by the DSS (6.77), we should evaluate the self-energy operator (6.73). For this, the MCA is used, as in Sect. 6.3.4, for the time-dependent multi-particle correlation function which appears in the spectral representation of the self-energy,

$$\Sigma''(\mathbf{q}, \omega) = -\frac{1}{2\omega\, m(\mathbf{q})} [I(\mathbf{q}, -\omega) - I(\mathbf{q}, \omega)],$$
$$I(\mathbf{q}, \omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} \langle \ddot{S}^{-}_{\mathbf{q}} | \ddot{S}^{+}_{-\mathbf{q}}(t) \rangle^{\mathrm{proper}}. \tag{6.91}$$

The second order time derivative of the spin operator,

$$-\ddot{S}^{+}_i = [[S^{+}_i, (H_t + H_J)], (H_t + H_J)] \equiv \sum_{\alpha} F^{\alpha}_i, \tag{6.92}$$

determines the force operators F^{α}_i denoted by the index $\alpha = tt, tJ, Jt, JJ$ respective to the commutation with H_t or H_J terms in the Hamiltonian (6.80). Therefore, the two-time correlation function in (6.91) yields 16 terms for the self-energy $\Sigma_{\alpha,\beta}(\mathbf{q}, \omega)$. In the Heisenberg limit at zero doping, $\delta = 0$, only the exchange interaction gives a contribution to the self-energy, $\Sigma_J(\mathbf{q}, \omega) = ((F^{JJ}_{\mathbf{q}} | (F^{JJ}_{\mathbf{q}})^{+}))_{\omega} / m(\mathbf{q})$. At a finite doping $\delta \sim 0.1$, the term F^{tt}_i gives the leading contribution. Using the commutation relations for the Hubbard operators for the spin operator $S^{+}_i = X^{+-}_i$ we get for this term,

$$F^{tt}_i = [[S^{+}_i, H_t], H_t] = \sum_{j,n} t_{ij} \Big\{ t_{jn} \Big[H^{-}_{ijn} + H^{+}_{nji} \Big] - t_{in} \Big[H^{-}_{jin} + H^{+}_{nij} \Big] \Big\},$$
$$H^{\sigma}_{ijn} = X^{\sigma 0}_i X^{+-}_j X^{0\sigma}_n + X^{+0}_i (X^{00}_j + X^{\sigma\sigma}_j) X^{0-}_n. \tag{6.93}$$

After the Fourier transformation to the $\mathbf{q}$-space, we obtain the force–force correlation function

$$\begin{aligned}\langle [F^{tt}_{\mathbf{q}}]^{\dagger} \mid F^{tt}_{\mathbf{q}}(t)\rangle &= (2t)^4 \sum_{\mathbf{q}_1,\mathbf{q}_2} \sum_{\mathbf{q}'_1,\mathbf{q}'_2} \\ &\langle [\Lambda_{\mathbf{q}'_1,\mathbf{q}'_2,\mathbf{q}'_3}\, H^{-}_{\mathbf{q}'_1,\mathbf{q}'_2,\mathbf{q}'_3} + \Lambda_{\mathbf{q}'_3,\mathbf{q}'_2,\mathbf{q}'_1}\, H^{+}_{\mathbf{q}'_1,\mathbf{q}'_2,\mathbf{q}'_3}]^{\dagger} \mid \\ &[\Lambda_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\, H^{-}_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}(t) + \Lambda_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1}\, H^{+}_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}(t)]\rangle,\end{aligned} \tag{6.94}$$

where $\mathbf{q}_3 = \mathbf{q} - \mathbf{q}_1 - \mathbf{q}_2$ and $\mathbf{q}'_3 = \mathbf{q} - \mathbf{q}'_1 - \mathbf{q}'_2$. Here we introduce the vertex function

$$\Lambda_{\mathbf{q}_1\mathbf{q}_2\mathbf{q}_3} = 4(\gamma_{\mathbf{q}_3+\mathbf{q}_2} - \gamma_{\mathbf{q}_1})\,\gamma_{q_3} + \gamma_{\mathbf{q}_2} - \gamma_{\mathbf{q}_1+\mathbf{q}_3}, \tag{6.95}$$

where the terms linear in $\gamma_{\mathbf{q}}$ reflect the exclusion of terms in F^{tt}_i with coinciding sites.

In the MCA, we assume that the propagation of electronic- and bosonic-type excitations at different lattice sites in (6.93) occurs independently, which results in the decoupling of the correlation function (6.94) into a product of the corresponding time-dependent correlation functions. In this approximation, the spin-diagonal correlation functions in the normal state reads,

$$\begin{aligned}\langle [H^{\sigma}_{\mathbf{q}'_1,\mathbf{q}'_2,\mathbf{q}'_3}]^{\dagger} \mid H^{\sigma}_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}(t)\rangle &= \langle X^{\sigma 0}_{\mathbf{q}'_3} S^{-}_{-\mathbf{q}'_2} X^{0\sigma}_{\mathbf{q}'_1} \mid X^{\sigma 0}_{\mathbf{q}_1}(t) S^{+}_{\mathbf{q}_2}(t) X^{0\sigma}_{\mathbf{q}_3}(t)\rangle \\ &= \langle X^{0\sigma}_{\mathbf{q}_1} X^{\sigma 0}_{\mathbf{q}_1}(t)\rangle \langle S^{-}_{-\mathbf{q}_2} S^{+}_{\mathbf{q}_2}(t)\rangle \langle X^{\sigma 0}_{\mathbf{q}_3} X^{0\sigma}_{\mathbf{q}_3}(t)\rangle \delta_{\mathbf{q}_1,\mathbf{q}'_1}\,\delta_{\mathbf{q}_2,\mathbf{q}'_2}\delta_{\mathbf{q}_3,\mathbf{q}'_3}.\end{aligned} \tag{6.96}$$

Here, the contribution from the charge excitations given by $(X^{00}_j + X^{\sigma\sigma}_j)$ in (6.93) is neglected in comparison with the spin-excitation contribution given by $X^{+-}_j = S^{+}_j$. In the superconducting state, the pair correlation functions should be additionally taken into account in the spin off-diagonal terms

$$\begin{aligned}\langle [H^{\bar{\sigma}}_{\mathbf{q}'_1,\mathbf{q}'_2,\mathbf{q}'_3}]^{\dagger} \mid H^{\sigma}_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}(t)\rangle &= \langle X^{\bar{\sigma}0}_{\mathbf{q}'_3} S^{-}_{-\mathbf{q}'_2} X^{0\bar{\sigma}}_{\mathbf{q}'_1} \mid X^{\sigma 0}_{\mathbf{q}_1}(t) S^{+}_{\mathbf{q}_2}(t) X^{0\sigma}_{\mathbf{q}_3}(t)\rangle \\ &= -\langle X^{0\bar{\sigma}}_{-\mathbf{q}_3} X^{0\sigma}_{\mathbf{q}_3}(t)\rangle \langle S^{-}_{-\mathbf{q}_2} S^{+}_{\mathbf{q}_2}(t)\rangle \langle X^{\bar{\sigma}0}_{-\mathbf{q}_1} X^{\sigma 0}_{\mathbf{q}_1}(t)\rangle \delta_{\mathbf{q}_1,-\mathbf{q}'_3}\,\delta_{\mathbf{q}_2,\mathbf{q}'_2}\delta_{\mathbf{q}_3,-\mathbf{q}'_1}.\end{aligned} \tag{6.97}$$

Substituting the MCA correlation functions (6.96) and (6.97) into (6.91) we write the multi-particle correlation function in the form,

$$\begin{aligned}I(\mathbf{q},\omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} \frac{1}{N^2} \sum_{\mathbf{q}_1,\mathbf{q}_2} &\langle S^{-}_{-\mathbf{q}_2} S^{+}_{\mathbf{q}_2}(t)\rangle \\ &\times\big[(\Lambda^2_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3} + \Lambda^2_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1})\langle X^{0\sigma}_{\mathbf{q}_1} X^{\sigma 0}_{\mathbf{q}_1}(t)\rangle \langle X^{\sigma 0}_{\mathbf{q}_3} X^{0\sigma}_{\mathbf{q}_3}(t)\rangle \\ &-\Lambda_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\Lambda_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1} \sum_{\sigma} \langle X^{\bar{\sigma}0}_{-\mathbf{q}_1} X^{\sigma 0}_{\mathbf{q}_1}(t)\rangle \langle X^{0\bar{\sigma}}_{-\mathbf{q}_3} X^{0\sigma}_{\mathbf{q}_3}(t)\rangle\big].\end{aligned} \tag{6.98}$$

Using the spectral representation for the time-dependent correlation functions (6.5) and (6.8) after integration over time t, the self-energy takes the form,

$$\Sigma_t''(\mathbf{q},\omega) = -\frac{\pi(2t)^4(e^{\beta\omega}-1)}{m(\mathbf{q})\,\omega}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}\omega_1 \mathrm{d}\omega_2 \mathrm{d}\omega_3$$
$$\frac{1}{N^2}\sum_{\mathbf{q}_1,\mathbf{q}_2} N(\omega_2)[1-n(\omega_1)]n(\omega_3)\delta(\omega+\omega_1-\omega_2-\omega_3)$$
$$B_{\mathbf{q}_2}(\omega_2)\big[(\Lambda^2_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}+\Lambda^2_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1})\,A^N_{\mathbf{q}_1}(\omega_1)\,A^N_{\mathbf{q}_3}(\omega_3)$$
$$-2\Lambda_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\Lambda_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1}\,A^S_{\mathbf{q}_1\sigma}(\omega_1)\,A^S_{\mathbf{q}_3\sigma}(\omega_3)\big], \tag{6.99}$$

where $\mathbf{q}_3 = \mathbf{q}-\mathbf{q}_1-\mathbf{q}_2$. The Fermi and Bose functions are denoted by $n(\omega) = (e^{\beta\omega}+1)^{-1}$ and $N(\omega) = (e^{\beta\omega}-1)^{-1}$, respectively. Here, we introduced the spectral functions: $A^N_{\mathbf{q}}(\omega) = -(1/\pi)\mathrm{Im}\langle\langle X^{0\sigma}_{\mathbf{q}}|X^{\sigma 0}_{\mathbf{q}}\rangle\rangle_\omega$, $A^S_{\mathbf{q}\sigma}(\omega) = -(1/\pi)\mathrm{Im}\langle\langle X^{0\sigma}_{\mathbf{q}}|X^{0\bar{\sigma}}_{-\mathbf{q}}\rangle\rangle_\omega$, and $B_{\mathbf{q}}(\omega) = (1/\pi)\,\chi''(\mathbf{q},\omega)$, where $A^{N,S}_{\mathbf{q}}(\omega)$ are determined by the GFs for electrons in the normal and superconducting states. The real part of the self-energy $\Sigma_t'(\mathbf{q},\omega)$ can be found from the dispersion relation (6.78).

The spin-excitation spectrum defined by the DSS (6.77) with the self-energy operator $\Sigma(\mathbf{q},\omega) = \Sigma_J(\mathbf{q},\omega)+\Sigma_t(\mathbf{q},\omega)$ was studied in a broad range of doping and temperatures in [46]. At low doping, a spin-wave-type behavior was found as in the Heisenberg model, while at higher doping a strong damping caused by hole hopping occurs, and a relaxation-type spin dynamics was observed in agreement with numerical simulations and experiments.

6.4.3 Magnetic Resonance Mode

In the superconducting state, the spin-excitation spectrum of high-T_c cuprates is dominated by a sharp magnetic peak at the planar AF wave vector $\mathbf{Q} = \pi(1,1)$, which is called *the resonance mode* (RM). The spin-excitation dispersion close to the RM exhibits a peculiar "hour-glass"-like shape with upward and downward dispersions. Whereas the RM energy changes with doping, no essential temperature dependence of it was found. In the optimal doping region, the RM and both dispersion branches vanish above T_c. In the strongly underdoped YBCO crystal, only the downward dispersion vanishes above T_c, whereas the upward dispersion and the RM are observed in the normal pseudogap state (for reviews see, e.g., [27, 52]). Since the energy of the RM scales with the superconducting temperature, $E_r \approx 5.3k_BT_c$, [52] it has been argued that it might constitute the bosonic excitation mediating superconducting pairing in cuprates which has motivated an extensive study of the RM phenomenon (see, e.g. [53]).

To explain the RM in superconducting cuprates, two basic approaches can be singled out. In the first one, the RM is considered as a particle-hole bound state,

usually referred to as a spin-1 exciton. In this approach, a Fermi-liquid model of itinerant electrons is assumed and the DSS is calculated within the random phase approximation (RPA) (for references, see, e.g., [53–56]). The state is formed below the continuum of particle-hole excitations which is gapped at a threshold energy $\omega_c \leq 2\Delta(\mathbf{q}^*)$ determined by the superconducting d-wave gap $2\Delta(\mathbf{q}^*)$ at a particular wave vector $\mathbf{q}^*$ on the Fermi surface (FS). In the Fermi-liquid approach, the d-wave symmetry of the gap and a shape of the Fermi surface which should cross the AF Brillouin zone (BZ) are essential in explaining the $\mathbf{q}$- and ω-dependence of the RM. In particular, a difference of the magnetic neutron scattering in the LSCO and YBCO compounds was explained by distinctions in the FS of these compounds [54].

In the second approach, an importance of strong electron correlations is stressed which are usually treated within the t–J model suitable for consideration of low-energy spin dynamics. To take into account a projected character of electron operators, the slave-boson technique (see, e.g., [57] and references therein) or a more rigorous diagram technique for HOs [58] were used. In particular, calculations of the DSS within the RPA revealed a sharp RM caused by opening of the superconducting gap below T_c and low-energy collective excitations similar to the Fermi-liquid models [59].

A more general Mori projection technique in the equation of motion method for the relaxation function was used by several groups (see, e.g., [60–63]. To clarify some of the open problems in describing the RM phenomenon (e.g., an appearance of the RM above T_c and its temperature dependence), we calculated the DSS in the superconducting state as described in the previous section and analyzed the RM phenomenon [47]. Our general expression for the DSS (6.76) is similar to that one obtained within the original Mori projection technique in previous studies of the t–J model. However, in those calculations the DSS was approximated by the bubble-type diagrams similar to the RPA which ignores an important role of spin excitations in the decay process. We have found out that an energy gap in the spin-excitation spectrum at the AF wave vector $\mathbf{Q}$ strongly reduces the damping at low temperatures, $T \ll E_r \simeq 5k_BT_c$ and bring about a sharp peak in the spectral function, the RM excitation. In the low doping region, where the damping is extremely small, the RM is found even above T_c. In the overdoped region at hole concentration $\delta \sim 0.2$ and $T \sim T_c$, the spin-excitation damping is large and opening of the superconducting gap enhances the intensity of the RM, so that it becomes observable only below T_c. Thus, as compared with the spin-exciton scenario, we propose an alternative explanation of the RM which is driven by the spin gap at the AF wave vector $\mathbf{Q}$ instead of the superconducting gap 2Δ.

6.4.3.1 Numerical Results

As a first step, the self-energy (6.99) is calculated by perturbation using the mean-field approximation for the electron spectral functions and the spin-excitation spectral function, respectively,

$$A^N_{\mathbf{q}}(\omega) = Q \sum_{\omega_1 = \pm E_{\mathbf{q}}} [(\omega_1 + \varepsilon_{\mathbf{q}})/2\omega_1]\delta(\omega - \omega_1) \tag{6.100}$$

$$A^S_{\mathbf{q}\sigma}(\omega) = Q \sum_{\omega_1 = \pm E_{\mathbf{q}}} [\Delta_{\mathbf{q}\sigma}/2\omega_1]\delta(\omega - \omega_1) \tag{6.101}$$

$$B_{\mathbf{q}}(\omega) = m(\mathbf{q}) \sum_{\omega_1 = \pm \tilde{\omega}_{\mathbf{q}}} [1/\omega_1]\, \delta(\omega - \omega_1). \tag{6.102}$$

Here, $Q = 1 - n/2$ is the Hubbard weighting factor, $E_{\mathbf{q}} = \sqrt{\varepsilon^2_{\mathbf{q}} + \Delta^2_{\mathbf{q}}}$, and the superconducting gap function $\Delta_{\mathbf{q}\sigma} = (\mathrm{sgn}\ \sigma)\, \Delta_{\mathbf{q}}$. For the electron spectrum we take the simplest approximation, $\varepsilon_{\mathbf{q}} = -4t\, Q\, \gamma_{\mathbf{q}} - \mu\varepsilon_{\mathbf{q}}$, where t is the nearest-neighbor hopping (cf. (6.46)). The spectrum of spin excitations $\widetilde{\omega}_{\mathbf{q}}$ in (6.102) is determined by the pole of the DSS, $\widetilde{\omega}_{\mathbf{q}} = [\omega^2_{\mathbf{q}} + \widetilde{\omega}_{\mathbf{q}}\, \Sigma'(\mathbf{q}, \widetilde{\omega}_{\mathbf{q}})]^{1/2}$. After integration over energies in (6.99) using these spectral functions, the self-energy can be written in the form convenient for calculations in the limit $T \to 0$:

$$\begin{aligned}\Sigma''_t(\mathbf{q}, \omega) = \frac{\pi(2t)^4}{\omega m(\mathbf{q})} \frac{Q^2}{N^2} \sum_{\mathbf{q}_1, \mathbf{q}_2} \sum_{\omega_1 = \pm E_{\mathbf{q}_1}} \sum_{\omega_2 = \pm \tilde{\omega}_{\mathbf{q}_2}} \sum_{\omega_3 = \pm E_{\mathbf{q}_3}} \\ m(\mathbf{q}_2) \frac{N(\omega_2)n(-\omega_1)n(\omega_3) + N(-\omega_2)n(\omega_1)n(-\omega_3)}{8\omega_1\omega_2\omega_3} \\ \big[(\Lambda^2_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3} + \Lambda^2_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1})(\omega_1 + \varepsilon_{\mathbf{q}_1})(\omega_3 + \varepsilon_{\mathbf{q}_3}) \\ -2\Lambda_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\Lambda_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1}\Delta_{\mathbf{q}_1}\Delta_{\mathbf{q}_3}\big]\delta(\omega + \omega_1 - \omega_2 - \omega_3).\end{aligned} \tag{6.103}$$

It should be emphasized that the self-energy (6.99) or (6.103) is determined by the decay of a spin excitation with the energy ω and wave vector $\mathbf{q}$ into three excitations: a particle-hole pair and a spin excitation. This process is controlled by the energy and momentum conservation laws, $\omega = (\omega_3 - \omega_1) + \omega_2$ and $\mathbf{q} = \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$, respectively. In previous studies of the t–J model, the contribution of the additional spin excitation has been neglected (see, e.g., [59]) or approximated by static or mean-field-type expressions (see, e.g., [60] and [61]) which results in the fermion-bubble approximation for the self-energy as in the RPA. The fermion-bubble approximation follows from (6.99), if we ignore the time dependence of the spin correlation function in the MCA (6.96), (6.97) and exclude the spin-excitation wave vector $\mathbf{q}_2$ by averaging over it in (6.99). Using the spectral densities (6.100)–(6.102), we obtain the fermion-bubble approximation for the self-energy,

$$\begin{aligned}\widetilde{\Sigma}''_t(\mathbf{q}, \omega) = -\frac{\pi(2t)^4}{m(\mathbf{q})\omega} \frac{Q^2}{N^2} \sum_{\mathbf{q}_1, \mathbf{q}_2} C_{\mathbf{q}_2} \sum_{\omega_1 = \pm E_{\mathbf{q}_1}} \sum_{\omega_2 = \pm E_{\mathbf{q}-\mathbf{q}_1}} \frac{n(\omega_1) - n(\omega_2)}{4\omega_1\omega_2} \\ \big[\big(\Lambda^2_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3} + \Lambda^2_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1}\big)(\omega_1 + \varepsilon_{\mathbf{q}_1})(\omega_2 + \varepsilon_{\mathbf{q}-\mathbf{q}_1}) \\ -\big(\Lambda_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\Lambda_{\mathbf{q}_3,\mathbf{q}_2,\mathbf{q}_1}\big)\Delta_{\mathbf{q}_1}\Delta_{\mathbf{q}-\mathbf{q}_1}\big]\delta(\omega + \omega_1 - \omega_2)\,.\end{aligned} \tag{6.104}$$

In the numerical calculations in [47], both the d-wave and s-wave symmetry of the superconducting gap $\Delta_{\mathbf{q}}^{(d)} = (\Delta/2)(\cos q_x - \cos q_y)$ and $\Delta^{(s)} = \Delta$ were considered with the temperature-dependent amplitude $\Delta(T)$. To compare the numerical results with experiments, two doping values $\delta = 0.2$, close to the optimal doping, and $\delta = 0.09$ for the underdoped case were considered. Below we present only few results referring to a more interesting underdoped case when the RM is observed even above T_c, while a detailed discussion can be found in [47].

To elucidate the role of spin excitations in the damping and their relevance to the shape of the spectral function, (6.77), the temperature dependence of the spin-excitation damping at the AF wave vector, $\Gamma(\mathbf{Q},\omega) = -(1/2)\,\Sigma_t''(\mathbf{Q},\omega)$, was calculated. Figure 6.3 shows the damping for $\delta = 0.09$ at various temperatures in the case of the d-wave gap. A weak and smooth variation of the damping with energy is revealed below T_c contrary to a sharp suppression of the damping in the fermion-bubble approximation (6.104). As numerical calculations show, a sharp increase of $\Gamma_{\mathbf{q}}$ away from the AF wave-vector $\mathbf{Q}$ explains the resonance character of spin excitations at $\mathbf{Q}$. In Fig. 6.4, the temperature dependence of the spectral function $\chi''(\mathbf{Q},\omega)$ for $\delta = 0.09$ is plotted. The RM energy weakly depends on temperature and is still quite visible at $T = T_c$ and even at $T = 1.4T_c$. In comparison with the results for the overdoped case, the resonance energy decreases with underdoping, while the intensity of the RM greatly increases. A general good agreement with neutron-scattering experiments in cuprates was obtained concerning the doping

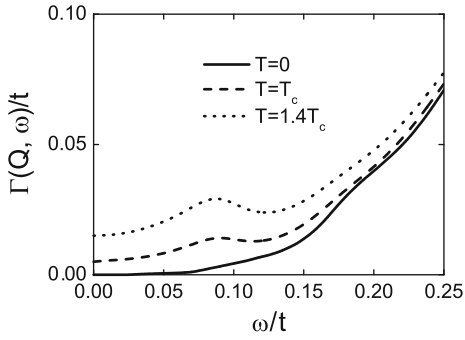

Fig. 6.3 Spin-excitation damping $\Gamma(\mathbf{Q},\omega)$ for $\delta = 0.09$ at $T \le T_c$ for the d-wave pairing

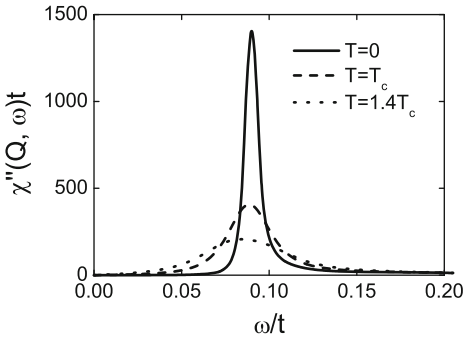

Fig. 6.4 Temperature dependence of the spectral function $\chi''(\mathbf{Q},\omega)$ at $\delta = 0.09$

and temperature dependence of the RM energy E_r. For instance, no temperature dependence was found for E_r in the underdoped ortho-II $YBCO_{6.5}$ single crystal with $E_r = 33\,\text{meV} = 6.5\,k_B T_c$ [64] and for the local spin susceptibility in the $YBCO_{6.6}$ single crystal with $T_c = 61$ K and $E_r = 38$ meV [65, 66].

Quite a different behavior of the damping and the spectral function is obtained for the self-energy (6.104). In the fermion-bubble approximation, a step-like dependence of the damping is observed for $\omega \leq 2\Delta$ at $T \ll T_c$ for a large FS in the overdoped case as was also found in [60]. Moreover, the RM energy E_r noticeably decreases with increasing temperature caused by decreasing the superconducting gap $2\Delta(T)$, contrary to a negligible shift of the RM shown in Fig. 6.4 for $T = 0.4T_c$. This comparison demonstrates that in the fermion-bubble approximation (6.104) the superconducting gap plays a crucial role in the occurrence of the RM with $E_r(T) < 2\Delta(T)$, while in the full self-energy (6.103) the superconducting gap and details of the electron dispersion and the FS are less important.

This difference can be explained as follows. Whereas in the fermion-bubble approximation (6.104) the spin excitation with the energy ω at the wave vector $\mathbf{Q}$ can decay only into a particle-hole pair with the energy $\omega(\mathbf{Q}) = E_{\mathbf{Q}+\mathbf{q}} + E_{\mathbf{q}}$, in a more general process described by (6.103) an additional spin excitation participates in the scattering. In the limit $T \to 0$, the decay process in this case is governed by another energy-conservation law, $\omega(\mathbf{Q}) = E_{\mathbf{q}_3} + E_{\mathbf{q}_1} + \widetilde{\omega}_{\mathbf{q}_2}$ where the largest contribution from the spin excitation comes from $\widetilde{\omega}_{\mathbf{q}_2} \simeq \widetilde{\omega}_{\mathbf{Q}}$ due to the factor $m(\mathbf{q}_2)$ (6.87) in (6.103). This energy-momentum conservation law strongly reduces the phase space for the decay and suppresses the damping of the initial spin excitation with the energy $\omega(\mathbf{Q})$. In fact, the occurrence of an additional spin excitation in the scattering process with the finite energy $\widetilde{\omega}_{\mathbf{Q}}$ plays a role similar to the superconducting gap in the excitation of the particle-hole pair in (6.104). Therefore, the damping at low temperatures ($k_B T \ll \widetilde{\omega}_{\mathbf{Q}} \sim E_r$) appears to be small even in the normal state as numerical calculations show [47]. In the case of the particle-hole relaxation, the condition for the occurrence of the RM, $\omega(\mathbf{Q}) = E_{\mathbf{q}+\mathbf{Q}} + E_{\mathbf{q}} \leq 2\Delta(\mathbf{q}^*)$, imposes a strong restriction on the shape of the FS which should cross the AF Brillouin zone to accommodate the scattering vector $\mathbf{Q}$ and the vector $\mathbf{q}^*$ on the FS (see, e.g. [56]). In particular, in the underdoped case with a small FS the energy-momentum conservation law cannot be fulfilled and the RM does not appear.

To conclude, a detailed study of the DSS in the superconducting state has revealed an important role of the spin-excitation damping in the RM phenomenon. We have found that the low-temperature damping essentially depends on the gap $\widetilde{\omega}_{\mathbf{Q}} \simeq E_r$ in the spin-excitation spectrum at the AF wave vector $\mathbf{Q}$, while the opening of a superconducting gap 2Δ below T_c is less important contrary to the fermion-bubble approximation and the RPA in the Fermion-liquid models. A good agreement of our results for the temperature and doping dependence of the spin-excitation spectrum and the RM with inelastic neutron-scattering experiments provides a strong support for the proposed theory.

6.5 Conclusion

The projection operator method, as was demonstrated in previous sections, is a powerful tool in calculation of the excitation spectrum within the equation of motion method for the two-time Green functions. The method is especially useful in study of the SCES where a conventional diagram technique cannot be used, while, in the absence of a small parameter, the perturbation theory is unreliable. The method is equally effective both in study of the single-particle Fermion-like excitations considered in Sect. 6.3, and the collective spin excitations in Sect. 6.4. A completely self-consistent theory of electronic spectrum and superconducting pairing within the Hubbard model can be developed by using the projection operator method for the electronic Green functions and the dynamical spin (and charge) susceptibility.

References

1. J. Hubbard, Proc. Roy. Soc. (London) A **276**, 238 (1963)
2. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)
3. N. Bulut, Adv. Phys. **51**, 1587 (2002)
4. A. Georges, G. Kotliar, W. Krauth, M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
5. G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
6. Th. Maier, M. Jarrel, Th. Pruschke, M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
7. A-M.S. Tremblay, B. Kyung, D. Sénéchal, Fiz. Nizk. Temp. (Low Temp. Phys., Ukraine) **32**, 561 (2006)
8. J. Hubbard, Proc. Roy. Soc. A (London) **285**, 542 (1965)
9. Yu.A. Izyumov, Yu.N. Scryabin, *Statistical Mechanics of Magnetically Ordered Systems* (Consultant Bureau, New York, 1989)
10. N.N. Bogoliubov, S.V. Tyablikov, Sov. Phys. Doklady **4**, 604 (1959)
11. D.N. Zubarev, Sov. Phys. Uspekhi, **3**, 320 (1960)
12. D.N. Zubarev, V. Morozov, G. Röpke, *Statistical Mechanics of Non-Equilibrium Processes* (Akademie, Berlin, 1996)
13. A.A. Abrikosov, L.P. Gorkov, I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, NJ, 1963)
14. H. Mori, Prog. Theor. Phys. **33**, 423 (1965)
15. D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry and Correlation Functions* (Benjamin, New York, 1975)
16. M. Ichianagi, J. Phys. Soc. Jpn. **32**, 604 (1972)
17. N.M. Plakida, Phys. Lett. A **43**, 481 (1973)
18. N.M. Plakida, in *Statistical Physics and Quantum Field Theory*, ed. by N.N. Bogolubov. Green Function Method in the Theory of Anharmonic Crystals (Nauka, Moscow, 1973), pp. 205–240
19. F. Mancini, A. Avella, Adv. Phys. **53**, 537 (2004)
20. A.L. Kuzemsky, Nuovo Cimento B **109**, 829 (1994)
21. Yu.A. Tserkovnikov, Theor. Math. Phys. **49**, 993 (1981)
22. Yu.A. Tserkovnikov, Theor. Math. Phys. **52**, 712 (1982)
23. U. Balucani, M.H. Lee, V. Tognetti, Phys. Rep. **373**, 409 (2003)
24. M. Suzuki, Physica **51**, 277 (1971)
25. N.M. Plakida, V.S. Oudovenko, Phys. Rev. B **59**, 11949 (1999)
26. N.M. Plakida, L. Anton, S. Adam, Gh. Adam, JETP **97**, 331 (2003)

27. N.M. Plakida, *High Temperature Cuprate Superconductors* (Springer, Berlin, 2010), pp. 428–454
28. J. Beenen, D.M. Edwards, Phys. Rev. B **52**, 13636 (1995)
29. A. Avella, F. Mancini, D. Villani, H. Matsumoto, Physica C **282–287**, 1757 (1997)
30. T. Di Matteo, F. Mancini, H. Matsumoto, V.S. Oudovenko, Physica B, **230–232**, 915 (1997)
31. T.D. Stanescu, I. Martin, Ph. Phillips, Phys. Rev. B **62**, 4300 (2000)
32. N.M. Plakida, R. Hayn, J.-L. Richard, Phys. Rev. B **51**, 16599 (1995)
33. L.F. Feiner, J.H. Jefferson, R. Raimondi, Phys. Rev. B **53**, 8751 (1996)
34. F.C. Zhang, T.M. Rice, Phys. Rev. B **37**, 3759 (1988)
35. N.M. Plakida, Physica C, **282–287**, 1737 (1997)
36. P.W. Anderson, Science **235**, 1196 (1987)
37. P.W. Anderson, *The Theory of Superconductivity in the High-T_c Cuprates* (Princeton University Press, Princeton, 1997)
38. N.M. Plakida, V.S. Oudovenko, JETP **104**, 230 (2007)
39. G.M. Eliashberg, Soviet Phys. JETP **11**, 696 (1960)
40. A.B. Migdal, Soviet Phys. JETP **7**, 996 (1958)
41. Z. Liu, E. Manousakis, Phys. Rev. B **45**, 2425 (1992)
42. A. Chubukov, D. Pines, J. Schmalian, in *Conventional, High-Transition Temperature, and Novel Superconductors*, vol. I, ed. by K.H. Benemann, J.B. Ketterson, (Springer, Heidelberg, 2003), p. 495
43. P.W.Anderson, Adv. Phys. **46**, 3 (1997)
44. R. Khasanov, A. Shengelaya, E. Morenzoni, K. Conder, I.M. Savič, H. Keller, J. Phys. Condens. Matter **16**, S4439 (2004)
45. A.A. Vladimirov, D. Ihle, N.M. Plakida, Theor. Math. Phys. **145**, 1576 (2005)
46. A.A. Vladimirov, D. Ihle, N.M. Plakida, Phys. Rev. B **80**, 104425 (2009)
47. A.A. Vladimirov, D. Ihle, N.M. Plakida, Phys. Rev. B **83**, 024411 (2011)
48. A. Sherman, M. Schreiber, Eur. Phys. J. B **32**, 203 (2003)
49. A. Sherman, Phys. Rev. B **70**, 184512 (2004)
50. A. Sherman and M. Schreiber, Fiz. Nizk. Temp. (Low Temp. Phys. Ukraine) **32**, 499 (2006)
51. P. Prelovšek, I. Sega, J. Bonča, Phys. Rev. Lett. **92**, 027002 (2004)
52. Y. Sidis, S. Pailhès, B. Keimer, Ph. Bourges, C. Ulrich, L.P. Regnault, Phys. Stat. Sol. (B) **241**, 1204 (2004)
53. M. Eschrig, Adv. Phys. **55**, 47 (2006)
54. Y. Zha, K. Levin, Phys. Rev. B **47**, 9124 (1993)
55. D.Z. Liu, Y. Zha, K. Levin, Phys. Rev. Lett. **75**, 4130 (1995)
56. I. Eremin, D.K. Morr, A.V. Chubukov, K. Bennemann, M.R. Norman, Phys. Rev. Lett. **94**, 147001 (2005)
57. J. Brinckmann, P.A. Lee, Phys. Rev. Lett. **82**, 2915 (1999)
58. Yu.A. Izyumov, B.M. Letfulov, J. Phys. Condens. Matter **3** 5373 (1991)
59. F. Onufrieva, P. Pfeuty, Phys. Rev. B **65**, 054515 (2002)
60. I. Sega, P. Prelovšek, J. Bonča, Phys. Rev. B **68**, 054524 (2003)
61. A. Sherman, M. Schreiber, Phys. Rev. B **68**, 094519 (2003)
62. I. Sega, P. Prelovšek Phys. Rev. B **73**, 092516 (2006)
63. P. Prelovšek, I. Sega, Phys. Rev. B **74**, 214501 (2006)
64. C. Stock, W.J.L. Buyers, R. Liang, D. Peets, Z. Tun, D. Bonn, W.N. Hardy, R.J. Birgeneau, Phys. Rev. B **69**, 014502 (2004)
65. V. Hinkov, Ph. Bourges, S. Pailhèes, Y. Sidis, A. Ivanov, C.D. Frost, T.G. Perring, C.T. Lin, D.P. Chen, B. Keimer, Nat. Phys. **3**, 780 (2007)
66. V. Hinkov, B. Keimer, A. Ivanov, Ph. Bourges, Y. Sidis, C.D. Frost, Neutron scattering study and analytical decription of the spin excitation spectrum of twin-free $YBa_2Cu_3O_{6.6}$. http://xxx.lanl.gov/arXiv:cond-mat/1006.3278 (2010)

Chapter 7
Dynamical Mean-Field Theory

Dieter Vollhardt, Krzysztof Byczuk, and Marcus Kollar

Abstract The dynamical mean-field theory (DMFT) is a widely applicable approximation scheme for the investigation of correlated quantum many-particle systems on a lattice, e.g., electrons in solids and cold atoms in optical lattices. In particular, the combination of the DMFT with conventional methods for the calculation of electronic band structures has led to a powerful numerical approach which allows one to explore the properties of correlated materials. In this introductory article we discuss the foundations of the DMFT, derive the underlying self-consistency equations, and present several applications which have provided important insights into the properties of correlated matter.

7.1 Motivation

7.1.1 Electronic Correlations

Earlier in 1937, at the outset of modern solid state physics, de Boer and Verwey [1] drew attention to the surprising properties of materials with incompletely filled $3d$-bands. This observation prompted Mott and Peierls [2] to discuss the interaction between the electrons. Ever since transition metal oxides (TMOs) were investigated intensively [3]. It is now well-known that in many materials with partially filled electron shells, such as the $3d$ transition metals V and Ni and their oxides, or

D. Vollhardt (✉) · M. Kollar
Theoretical Physics III, Center for Electronic Correlations and Magnetism, Institute of Physics, University of Augsburg, 86135 Augsburg, Germany
e-mail: Dieter.Vollhardt@physik.uni-augsburg.de; Marcus.Kollar@physik.uni-augsburg.de

K. Byczuk
Faculty of Physics, Institute of Theoretical Physics, University of Warsaw, ul. Hoża 69, 00-681 Warsaw, Poland
e-mail: Krzysztof.Byczuk@fuw.edu.pl

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_7,

$4f$ rare-earth metals such as Ce, electrons occupy narrow orbitals. The spatial confinement enhances the effect of the Coulomb interaction between the electrons, making them "strongly correlated." Correlation effects can lead to profound quantitative and qualitative changes of the physical properties of electronic systems as compared to non-interacting particles. In particular, they often respond very strongly to changes in external parameters. This is expressed by large renormalizations of the response functions of the system, e.g., of the spin susceptibility and the charge compressibility. In particular, the interplay between the spin, charge and orbital degrees of freedom of the correlated d and f electrons and with the lattice degrees of freedom leads to an amazing multitude of ordering phenomena and other fascinating properties, including high temperature superconductivity, colossal magnetoresistance and Mott metal–insulator transitions (MIT) [3].

7.1.2 The Hubbard Model

The simplest microscopic model describing the interaction with electrons in a solid is the one-band, spin-1/2 Hubbard model [4–6] where the interaction between the electrons is assumed to be so strongly screened that it is purely local. More generally the Hubbard model applies to lattice fermions with a point interaction, such as ultracold fermionic atoms in optical lattices where the interaction is indeed extremely short ranged. The Hamiltonian consists of two terms, the kinetic energy $\hat{H}_0$ and the interaction energy $\hat{H}_I$ (the operators here and in the following equations are denoted by a hat):

$$\hat{H} = \hat{H}_0 + \hat{H}_I \tag{7.1a}$$

$$\hat{H}_0 = \sum_{i,j} \sum_{\sigma} t_{ij} \hat{c}^+_{i\sigma} \hat{c}_{j\sigma} = \sum_{\boldsymbol{k},\sigma} \epsilon_{\boldsymbol{k}} \hat{n}_{\boldsymbol{k}\sigma} \tag{7.1b}$$

$$\hat{H}_{\mathrm{I}} = U \sum_{i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \tag{7.1c}$$

where $\hat{c}^+_{i\sigma}(\hat{c}_{i\sigma})$ are creation (annihilation) operators of fermions with spin σ at site $\boldsymbol{R}_i$ (for simplicity denoted by i), and $\hat{n}_{i\sigma} = \hat{c}^+_{i\sigma}\hat{c}_{i\sigma}$. The Fourier transform of the kinetic energy in (7.1b), where t_{ij} is the amplitude for hopping between sites i and j, involves the dispersion $\epsilon_{\boldsymbol{k}}$ and the momentum distribution operator $\hat{n}_{\boldsymbol{k}\sigma}$. This model provides the basis for most of the theoretical research on correlated electrons during the last decades.

The Hubbard model describes an interacting many-body system which cannot be solved analytically, except for dimension $d = 1$, for nearest-neighbor hopping [7]. In the absence of exact solutions there is clearly a great need for reliable, controlled approximation schemes for this model. However, such approximations are not

easy to construct as the following observation shows. When viewed as a function of time a given site of the lattice will sometimes be empty, singly occupied or doubly occupied. For strong repulsion U double occupations are energetically very unfavorable and are therefore strongly suppressed, implying $\langle \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\rangle \neq \langle \hat{n}_{i\uparrow}\rangle\langle \hat{n}_{i\downarrow}\rangle$. Therefore approximation schemes based on the factorization of the interaction term, e.g., Hartree–Fock-type mean-field theories, are generally insufficient to explain the physics of electrons in their paramagnetic phase beyond the limit of weak interactions. This is due to the fact that in such approximations the interaction is described only as an average, static potential, whereby correlations, i.e., dynamical many-body effects due to the interaction of individual electrons, are excluded from the beginning. Hence correlation phenomena such as the Mott–Hubbard MIT cannot be described by such approximations. This clearly shows the need for comprehensive approximation schemes, which are applicable for all values of the input parameters, e.g., coupling parameters and temperature, diagrammatically controlled, and thermodynamically consistent [8].

7.1.3 Construction of Comprehensive Mean-Field Theories for Many-Particle Models

There exists a well-established branch of approximation techniques which makes use of the simplifications that occur when some parameter is taken to be large (in fact, infinite), e.g., the length of the spins S, the spin degeneracy N, the spatial dimension d, or the coordination number Z, i.e., the number of nearest neighbors of a lattice site.[1] Investigations in this limit, supplemented if possible by an expansion in the inverse of the large parameter, often provide valuable insights into the fundamental properties of a system even when the inverse parameter is not very small.

One of the best-known mean-field theories in many-body physics is the Weiss molecular-field theory for the Ising model [10]. It is a prototypical *single-site mean-field theory* which becomes exact for infinite-range interaction, as well as in the limit of the coordination number $Z \to \infty$ or the dimension $d \to \infty$. In the latter case $1/Z$ or $1/d$ is a small parameter which can be used to improve the mean-field theory systematically. This mean-field theory is comprehensive in the sense discussed above. Namely, it contains no unphysical singularities, is applicable for

[1] The coordination number Z is determined by the dimension d and the lattice structure. Already in $d = 3$ the coordination number can be quite large, e.g., $Z = 6$ for a simple cubic lattice, $Z = 8$ for a bcc-lattice and $Z = 12$ for an fcc-lattice, making its inverse, $1/Z$, rather small. It is then natural to consider the limit $Z \to \infty$ to simplify the problem. For a hypercubic lattice, obtained by generalizing the simple cubic lattice in $d = 3$ to arbitrary dimensions, one has $Z = 2d$. The limit $d \to \infty$ is then equivalent to $Z \to \infty$. Several standard approximation schemes which are commonly used to explain experimental results in dimension $d = 3$ are exact only in $d, Z = \infty$ [9].

all values of the input parameters, i.e., the coupling parameter, magnetic field, and temperature, and is diagrammatically controlled [11].

Itinerant quantum mechanical models such as the Hubbard model and its generalizations are much more complicated than classical, Ising-type models. Generally there is no semiclassical approximations for such models that might serve as a starting point for further investigations. Under such circumstances the construction of a mean-field theory with the comprehensive properties of the Weiss molecular field theory for the Ising model will necessarily be much more complicated, too. Here the limit of high spatial dimensions d or coordination number Z has again been extremely useful since it provides the basis for the construction of a comprehensive *dynamical* mean-field theory (DMFT) for lattice fermions.

In this article we will first discuss (Sect. 7.2) the limit of high spatial dimensions d for lattice fermions, the scaling of the hopping amplitude which is necessary to obtain a meaningful limit $d \to \infty$ as well as the simplifications of the many-body perturbation theory occurring in this limit. In Sect. 7.3 the self-consistency equations obtained in the limit $d \to \infty$ are derived which provide the basis for the DMFT for correlated lattice fermions. An example for the many insights gained by the DMFT is the Mott–Hubbard MIT discussed in Sect. 7.4. The application of the DMFT to real correlated materials is described in Sect. 7.5. Brief introductions to the DMFT for correlated systems in the presence of disorder (Sect. 7.6), correlated bosons in optical lattices (Sect. 7.7), and systems in non-equilibrium (Sect. 7.8) are also given. In Sect. 7.9 a summary and outlook is presented.

7.2 Lattice Fermions in the Limit of High Dimensions

7.2.1 Scaling of the Hopping Amplitude

We consider the kinetic energy term (7.1a) since the interaction term is purely local and is thereby completely independent of the lattice structure and the dimension. For hopping between nearest-neighbor (NN) sites i and j with amplitude $t_{ij} \equiv -t$ on a d-dimensional hypercubic lattice with lattice spacing a, the dispersion $\epsilon_{\boldsymbol{k}}$ is given by

$$\epsilon_{\boldsymbol{k}} = -2t \sum_{n=1}^{d} \cos(k_n a). \tag{7.2}$$

The density of states (DOS) corresponding to $\epsilon_{\boldsymbol{k}}$ is

$$N_d(\omega) = \sum_{\boldsymbol{k}} \delta(\hbar\omega - \epsilon_{\boldsymbol{k}}), \tag{7.3}$$

which is the probability density for finding[2] $\omega = \epsilon_{\boldsymbol{k}}$ for a random choice of $\boldsymbol{k} = (k_1, \ldots, k_d)$. If the k_i are chosen randomly, $\epsilon_{\boldsymbol{k}}$ in (7.2) is the sum of (independent) random numbers $-2t \cos k_i$. The central limit theorem then implies [12] that in the limit $d \to \infty$ the DOS is given by a Gaussian

$$N_d(\omega) \overset{d\to\infty}{\longrightarrow} \frac{1}{2t\sqrt{\pi d}} \exp\left[-\left(\frac{\omega}{2t\sqrt{d}}\right)^2 \right]. \tag{7.4}$$

Unless t is scaled properly with d this DOS will become arbitrarily broad and featureless for $d \to \infty$. Clearly only the scaling

$$t \to \frac{t^*}{\sqrt{d}}, \; t^* = \text{const.}, \tag{7.5}$$

("quantum scaling") yields a non-trivial limit $d \to \infty$ for the DOS [12, 13].

The interaction term in (7.1) is seen to be purely local and independent of the surrounding; hence it is independent of the spatial dimension of the system. Consequently, the on-site interaction U need not be scaled. So we see that the scaled Hubbard Hamiltonian

$$\hat{H} = -\frac{t^*}{\sqrt{Z}} \sum_{\langle i,j \rangle} \sum_{\sigma} \hat{c}^+_{i\sigma} \hat{c}_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \tag{7.6}$$

has a nontrivial $Z \to \infty$ limit, where both terms, the kinetic energy and the interaction, are of the same order of magnitude and are thereby able to compete; here $\langle i, j \rangle$ denotes NN sites i and j. It is this competition between the two terms which leads to interesting many-body physics. Mathematically this is expressed by the fact that the generic matrix elements of the commutator between the kinetic and the interaction part of the Hamiltonian do not vanish in the $d \to \infty$ limit.

The quantum scaling (7.5) was determined within a $\boldsymbol{k}$-space formulation, but it can also be derived within a position-space formulation as will be discussed next.

7.2.2 Simplifications of the Many-Body Perturbation Theory

The most important consequence of the scaling (7.5) is the fact that it leads to significant simplifications in the investigation of Hubbard-type lattice models [12, 14–19]. To understand this point better we take a look at the perturbation theory in terms of U. At $T = 0$ and $U = 0$ the kinetic energy of the electrons is given by

[2] In the following we set the Planck constant $\hbar$, the Boltzmann constant k_B, and the lattice constant a equal to unity.

$$E^0_{\rm kin} = -t \sum_{\langle i,j \rangle} \sum_{\sigma} g^0_{ij,\sigma}. \tag{7.7}$$

Here $g^0_{ij,\sigma} = \langle \hat{c}^+_{i\sigma} \hat{c}_{j\sigma} \rangle_0$ is the one-particle density matrix which can be interpreted as the probability amplitude for hopping from site j to site i. The square of its absolute value is proportional to the probability for an electron to hop from j to i, i.e., $|g^0_{ij,\sigma}|^2 \sim 1/Z \sim 1/d$, since site j has $\mathcal{O}(d)$ NN sites i. The sum of $|g^0_{ij,\sigma}|^2$ over all NN sites i of j must then yield a Z or d independent constant. In the limit $d \to \infty$ we therefore find

$$g^0_{ij,\sigma} \sim \mathcal{O}\Big(\frac{1}{\sqrt{d}}\Big). \tag{7.8}$$

Since the sum over NN sites in (7.7) is of $\mathcal{O}(d)$, the NN hopping amplitude t must obviously be scaled according to (7.5) for $E^0_{\rm kin}$ to remain finite in the limit $d, Z \to \infty$. Hence, as expected, a real-space formulation yields the same results for the required scaling of the hopping amplitude.

The one-particle Green function $G^0_{ij,\sigma}(\omega)$ of the non-interacting system obeys the same scaling as $g^0_{ij,\sigma}$. This follows directly from its definition

$$G^0_{ij,\sigma}(t) \equiv -\langle T \hat{c}_{i\sigma}(t) \hat{c}^+_{j\sigma}(0) \rangle_0, \tag{7.9}$$

where T is the time ordering operator, and the time evolution of the operators is given by the Heisenberg representation. The one-particle density matrix is obtained as $g^0_{ij,\sigma} = \lim_{t \to 0-} G^0_{ij,\sigma}(t)$. If $g^0_{ij,\sigma}$ obeys (7.8) the one-particle Green function must follow the same scaling at all times since this property does not dependent on the time evolution and the quantum mechanical representation. The Fourier transform $G^0_{ij,\sigma}(\omega)$ also preserves this property.

It is important to realize that, although $G^0_{ij,\sigma} \sim 1/\sqrt{d}$ vanishes for $d \to \infty$, the particles are *not* localized, but are still mobile. Indeed, even in the limit $d \to \infty$ the off-diagonal elements of $G^0_{ij,\sigma}$ contribute, since a particle may hop to d the nearest neighbors with reduced amplitude $t^*/\sqrt{2d}$. For general i, j one finds [15, 20]

$$G^0_{ij,\sigma} \sim \mathcal{O}\Big(1/d^{||\boldsymbol{R}_i - \boldsymbol{R}_j||/2}\Big), \tag{7.10}$$

where $||\boldsymbol{R}|| = \sum_{n=1}^{d} |R_n|$ is the length of $\boldsymbol{R}$ in the so-called "New York metric" (also called "taxi cab metric," since particles only hop along horizontal or vertical lines, never along a diagonal).

It is the property (7.10) which is the origin of all simplifications arising in the limit $d \to \infty$. In particular, it implies the collapse of all connected, irreducible perturbation theory diagrams in the position space [12, 14, 15]. In general, any

two vertices which are connected by more than two separate paths[3] will collapse onto the same site. In particular, the external vertices of any irreducible self-energy diagram are always connected by three separate paths and hence always collapse. As a consequence the full irreducible self-energy becomes a purely local quantity [12], but retains its dynamics [14]

$$\Sigma_{ij,\sigma}(\omega) \overset{d\to\infty}{=} \Sigma_{ii,\sigma}(\omega)\delta_{ij}\,. \tag{7.11a}$$

In the paramagnetic phase we may write $\Sigma_{ii,\sigma}(\omega) \equiv \Sigma(\omega)$. The Fourier transform of $\Sigma_{ij,\sigma}$ is then momentum-independent:

$$\Sigma_{\sigma}(\boldsymbol{k},\omega) \overset{d\to\infty}{\equiv} \Sigma_{\sigma}(\omega). \tag{7.11b}$$

This leads to tremendous simplifications in all many-body calculations for the Hubbard model and related models. It should be noted that a $\boldsymbol{k}$-independence of Σ was sometimes *assumed* as a convenient approximation ("local approximation") [21–23]. Here we identified the limit where this is indeed exact.

7.2.3 Interactions Beyond the On-Site Interaction

In the case of more general interactions than the Hubbard interaction, e.g., the nearest-neighbor interactions such as

$$\hat{H}_{nn} = \sum_{\langle i,j\rangle} \sum_{\sigma\sigma'} V_{\sigma\sigma'} \hat{n}_{i\sigma} \hat{n}_{j\sigma'} \tag{7.12}$$

the interaction constant has to be scaled, too, in the limit $d \to \infty$. In the case of (7.12), which has the form of a classical interaction, the "classical" scaling

$$V_{\sigma\sigma'} \to \frac{V^*_{\sigma\sigma'}}{Z} \tag{7.13}$$

is required. Of course, the propagator still has the dependence (7.10).

Due to (7.13) all contributions, except for the Hartree term, are found to vanish in $d = \infty$ [14]. Hence nonlocal interactions only contribute via their Hartree contribution, which is purely static. This gives the Hubbard interaction a unique role: of all interactions for fermionic lattice models only the purely local Hubbard interaction remains dynamical in the limit $d \to \infty$ [14].

[3] Here a "path" is any sequence of lines in a diagram; they are "separate" when they have no lines in common.

7.2.4 Single-Particle Propagator

Due to the $\boldsymbol{k}$-independence of the irreducible self-energy, (7.11b), the one-particle propagator of an interacting lattice fermion system is given by

$$G_{\boldsymbol{k},\sigma}(\omega) = \frac{1}{\omega - \epsilon_{\boldsymbol{k}} + \mu - \Sigma_\sigma(\omega)}. \tag{7.14}$$

Most importantly, the $\boldsymbol{k}$ dependence of $G_{\boldsymbol{k}}(\omega)$ comes entirely from the energy dispersion $\epsilon_{\boldsymbol{k}}$ of the *non*-interacting particles. This means that for a homogeneous system with the propagator

$$G_{ij,\sigma}(\omega) = L^{-1} \sum_{\boldsymbol{k}} G_{\boldsymbol{k},\sigma}(\omega) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j)} \tag{7.15}$$

its local part, $G_{ii,\sigma}$, has the form [16]

$$G_{ii,\sigma}(\omega) = L^{-1} \sum_{\boldsymbol{k}} G_{\boldsymbol{k},\sigma}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}E \, \frac{N_\infty(E)}{\omega - E + \mu - \Sigma_\sigma(\omega)} \tag{7.16a}$$

$$\equiv G_\sigma(\omega). \tag{7.16b}$$

In the following we will limit our discussion to the paramagnetic phase and omit the spin index. The spectral function of the interacting system (often referred to as the DOS as in the non-interacting case) is given by

$$A(\omega) = -\frac{1}{\pi} \mathrm{Im} G(\omega + i0^+); \tag{7.16c}$$

for $U = 0$ one has $A(\omega) \equiv N(\omega)$. In the limit $d \to \infty$ two quantities then play the most important role: the local propagator $G(\omega)$ and the self-energy $\Sigma(\omega)$.

7.2.5 Consequences of the Momentum Independence of the Self-Energy

We now discuss some more consequences of the $\boldsymbol{k}$-independence of the self-energy as derived by Müller–Hartmann [16]. Let us consider the Hubbard model, or any one of its generalizations, in the paramagnetic phase, i.e., without a broken symmetry. At $T = 0$ the one-particle Green function (7.14) then reads

$$G_{\boldsymbol{k}}(\omega) = \frac{1}{\omega - \epsilon_{\boldsymbol{k}} + E_\mathrm{F} - \Sigma(\omega)}. \tag{7.17}$$

In general, even when $\Sigma(\omega)$ is $\boldsymbol{k}$-dependent, the Fermi surface is defined by the $\omega = 0$ limit of the denominator of (7.17) as

$$\epsilon_{\boldsymbol{k}} + \Sigma_{\boldsymbol{k}}(0) = E_{\mathrm{F}}. \tag{7.18a}$$

According to Luttinger and Ward [24] the volume within the Fermi surface is not changed by interactions, provided the effect of the latter can be treated in infinite-order perturbation theory (i.e., no broken symmetry). This is expressed by

$$n = \sum_{\boldsymbol{k}\sigma} \theta[E_{\mathrm{F}} - \epsilon_{\boldsymbol{k}} - \Sigma_{\boldsymbol{k}}(0)], \tag{7.18b}$$

where n is the particle density and $\theta(x)$ is the step function. In general, the $\boldsymbol{k}$-dependence of $\Sigma_{\boldsymbol{k}}(0)$ in (7.18a) implies that, in spite of (7.18b), the shape of the Fermi surface of the interacting system will be quite different from that of the non-interacting system (except for the fully rotation invariant case $\epsilon_{\boldsymbol{k}} \sim k^2$). For lattice fermion models in $d < \infty$, with $\Sigma_{\boldsymbol{k}}(\omega) \equiv \Sigma(\omega)$, (7.18a) implies that the Fermi surface itself (and hence the volume enclosed) is not changed by interactions.[4] The Fermi energy is simply shifted uniformly from its non-interacting value E_{F}^0, i.e., $E_{\mathrm{F}} = E_{\mathrm{F}}^0 + \Sigma(0)$, to keep n in (7.18b) as a constant. From (7.16a) we thus conclude that the $\omega = 0$ value of the local propagator, $G(0)$, and hence of the spectral function, $A(0) = -\frac{1}{\pi}\mathrm{Im}G(i0^+)$, is not changed by interactions. Renormalizations of $A(0)$ can only come from a $\boldsymbol{k}$-dependence of Σ, i.e., if $\partial\Sigma/\partial\boldsymbol{k} \neq 0$.

For $\omega \to 0$ the self-energy has the property

$$\mathrm{Im}\ \Sigma(\omega) \propto \omega^2 \tag{7.18c}$$

which implies quasiparticle (Fermi liquid) behavior. The effective mass

$$\frac{m^*}{m} = 1 - \left.\frac{\mathrm{d}\Sigma}{\mathrm{d}\omega}\right|_{\omega=0} = 1 + \frac{1}{\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega\ \frac{\mathrm{Im}\Sigma(\omega + i0^-)}{\omega^2} \geq 1 \tag{7.18d}$$

is seen to be enhanced. In particular, the momentum distribution

$$n_{\boldsymbol{k}} = \frac{1}{\pi}\int_{-\infty}^{0} \mathrm{d}\omega\ \mathrm{Im}G_{\boldsymbol{k}}(\omega) \tag{7.19}$$

has a discontinuity at the Fermi surface, given by $n_{k_{\mathrm{F}}^-} - n_{k_{\mathrm{F}}^+} = (m^*/m)^{-1}$, where $k_{\mathrm{F}}^{\pm} = k_{\mathrm{F}} \pm 0^+$.

[4] In $d = \infty$ limit the notion of a Fermi surface of a lattice system is complicated by the fact that the dispersion $\epsilon_{\boldsymbol{k}}$ is not a simple smooth function.

7.3 Dynamical Mean-Field Theory for Correlated Lattice Fermions

The limit of high spatial dimensions d or coordination number Z provides the basis for the construction of a comprehensive mean-field theory for lattice fermions which is diagrammatically controlled and whose free energy has no unphysical singularities. It starts from the scaled Hamiltonian (7.6) and makes use of the simplifications in the many-body perturbation theory discussed in Sect. 7.2.2. There we found that the local propagator $G(\omega)$, i.e., the probability amplitude for an electron to return to a lattice site, and the local, but fully dynamical self-energy $\Sigma(\omega)$ are the most important quantities in such a theory. Since the self-energy is a dynamical variable (in contrast to the Hartree–Fock theory where it is merely an average, static potential) the resulting mean-field theory is also dynamical and can thus describe genuine correlation effects such as the Mott–Hubbard MIT.

The self-consistency equations of this *dynamical mean-field theory* (DMFT) for correlated lattice fermions can be derived in different ways. All derivations make use of the fact that in the limit of high spatial dimensions Hubbard-type models, i.e., lattice models with a local interaction, reduce to a "dynamical single-site problem," where the d-dimensional lattice model is effectively described by the dynamics of the correlated fermions on a single site embedded in a "bath" provided by the other particles. In particular, the derivation by Janiš [25, 26] is a generalization of the coherent potential approximation (CPA) for disordered systems[5] to the Hubbard model. In the following we will present today's standard derivation by Georges and Kotliar [31] which is based on the mapping of the lattice problem onto a self-consistent single-impurity Anderson model; this approach was also employed by Jarrell [32]. Although the DMFT equations derived within the CPA approach and the self-consistent single-impurity approach, respectively, are identical it is the latter formulation which was immediately adopted by the community since it makes contact with the theory of quantum impurities and Kondo problems; for a review, see [33]. This is a well-understood branch of many-body physics [34] for whose solution efficient numerical codes had been developed already in 1980s, in particular by use of the quantum Monte-Carlo (QMC) method [35].

7.3.1 Construction of the DMFT as a Self-Consistent Single-Impurity Anderson Model

Following the presentation of Georges et al. [33] the DMFT equations will now be derived using the so-called cavity method.[6] This derivation starts by removing

[5] The CPA is the best single-site approximation for disordered, non-interacting lattice electrons [27–29]; it becomes exact in the limit $d, Z \to \infty$ [8, 30].

[6] We note that the sign of the hopping amplitude t_{ij} used here (see the definition in (7.1b)) is opposite to that in [33].

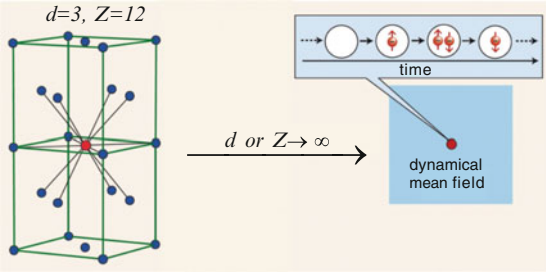

Fig. 7.1 Already in a $d = 3$ the coordination number Z can be quite large, as in the case of a face-centered cubic lattice with $Z = 12$. In the limit $d \to \infty$, i.e., $Z \to \infty$, the many-body lattice problem reduces to that of a single lattice site embedded in a dynamical mean field. As shown in the inset electrons can hop onto and off that site and interact as in the finite-dimensional Hubbard model. Therefore the DMFT describes the dynamics of the interacting electrons correctly

one lattice site together with its bonds from the rest of the lattice. The remaining lattice, which now contains a cavity, is replaced by a particle bath which plays the role of the dynamical mean field (see Fig. 7.1). Now comes a physically motivated step: the bath is coupled, via a hybridization, to the cavity. The resulting problem then amounts to the solution of an effective single-impurity Anderson model where the degrees of freedom of the bath, represented by an appropriate hybridization function, have to be determined self-consistently.

To be specific, we consider the partition function in the grand canonical ensemble

$$\mathcal{Z} = \int \prod_{i\,\sigma} Dc^*_{i\sigma} Dc_{i\sigma} \exp[-S]. \tag{7.20}$$

The action S for the Hubbard model is given by

$$S = \int_0^\beta \mathrm{d}\tau \left[\sum_{i\sigma} c^*_{i\sigma}(\tau)(\frac{\partial}{\partial\tau} - \mu)c_{i\sigma}(\tau) + \sum_{ij\,\sigma} t_{ij} c^*_{i\sigma}(\tau)c_{j\sigma}(\tau) \right.$$
$$\left. + \sum_i U c^*_{i\uparrow}(\tau)c_{i\uparrow}(\tau)c^*_{i\downarrow}(\tau)c_{i\downarrow}(\tau) \right], \tag{7.21}$$

where we use Grassmann variables $c^*_{i\sigma}$, $c_{i\sigma}$. We split the action S into three parts

$$S = S_0 + \Delta S + S^{(0)}, \tag{7.22}$$

where S_0 is the part containing only variables on site 0

$$S_0 = \int_0^\beta \mathrm{d}\tau \left[\sum_\sigma c^*_{0\sigma}(\tau)(\frac{\partial}{\partial\tau} - \mu)c_{0\sigma}(\tau) + U c^*_{0\uparrow}(\tau)c_{0\uparrow}(\tau)c^*_{0\downarrow}(\tau)c_{0\downarrow}(\tau) \right], \tag{7.23}$$

ΔS contains the hoppings between site 0 and other sites of the lattice $i \neq 0$

$$\Delta S = \int_0^\beta \mathrm{d}\tau \sum_{i\,\sigma} \left[t_{i0} c^*_{i\sigma}(\tau)c_{0\sigma}(\tau) + t_{0i} c^*_{0\sigma}(\tau)c_{i\sigma}(\tau) \right], \tag{7.24}$$

and the rest, denoted by $S^{(0)}$, is the part of the action where the site 0 and its bonds are removed, i.e., for $i, j \neq 0$ one has

$$S^{(0)} = \int_0^\beta d\tau \left[\sum_{i\neq 0\,\sigma} c^*_{i\sigma}(\tau)(\frac{\partial}{\partial\tau} - \mu)c_{i\sigma}(\tau) + \sum_{ij\neq 0\,\sigma} t_{ij} c^*_{i\sigma}(\tau)c_{j\sigma}(\tau) \right.$$
$$\left. + U \sum_{i\neq 0} c^*_{i\uparrow}(\tau)c_{i\uparrow}(\tau)c^*_{i\downarrow}(\tau)c_{i\downarrow}(\tau) \right]. \tag{7.25}$$

We now rewrite the partition function $\mathcal{Z}$ as

$$\mathcal{Z} = \int \prod_\sigma Dc^*_{0\sigma} Dc_{0\sigma} \exp[-S_0]$$
$$\times \int \prod_{i\neq 0\,\sigma} Dc^*_{i\sigma} Dc_{i\sigma} \exp[-S^{(0)}] \exp[-\Delta S] \tag{7.26}$$

and use the ensemble average

$$\langle X \rangle_{(0)} \equiv \frac{1}{\mathcal{Z}^{(0)}} \int \prod_{i\neq 0\,\sigma} Dc^*_{i\sigma} Dc_{i\sigma} X \exp[-S^{(0)}] \tag{7.27}$$

taken with respect to $S^{(0)}$ (the action where the site $i = 0$ is excluded), with $\mathcal{Z}^{(0)}$ being the corresponding partition function. Then the partition function reads

$$\mathcal{Z} = \mathcal{Z}^{(0)} \int \prod_\sigma Dc^*_{0\sigma} Dc_{0\sigma} \exp[-S_0] \langle \exp[-\Delta S] \rangle_{(0)}. \tag{7.28}$$

In the next step we expand the second exponent with respect to the action ΔS. A non-trivial limit $d \to \infty$ is obtained by scaling the hopping amplitudes t_{ij} as described in Sect. 7.2.2. Consequently, in the $Z \to \infty$ limit only the contribution $G^{(0)}_{jk\sigma}$, where

$$G^{(0)}_{jk\sigma}(\tau_1-\tau_2) = -\langle T_\tau c_{j\sigma}(\tau_1)c^*_{k\sigma}(\tau_2)\rangle_{(0)}, \tag{7.29}$$

or disconnected contributions made of products of $G^{(0)}_{jk\sigma}$'s remain. Applying the linked-cluster theorem and collecting only connected contributions in the exponential function one obtains the local action

$$\begin{aligned} S_{\rm loc} = \Bigg[& \int_0^\beta \mathrm{d}\tau \sum_\sigma c^*_{0\sigma}(\tau)(\frac{\partial}{\partial\tau}-\mu)c_{0\sigma}(\tau) \\ & + U\int_0^\beta \mathrm{d}\tau c^*_{0\uparrow}(\tau)c_{0\uparrow}(\tau)c^*_{0\downarrow}(\tau)c_{0\downarrow}(\tau) \\ & + \int_0^\beta \mathrm{d}\tau_1 \int_0^\beta \mathrm{d}\tau_2 \sum_\sigma \sum_{j,k\neq 0} t^*_{j0}t^*_{k0}G^{(0)}_{jk\sigma}(\tau_1-\tau_2)c^*_{0\sigma}(\tau_1)c_{0\sigma}(\tau_2) \Bigg], \end{aligned} \tag{7.30}$$

where the rescaled hoppings are denoted with a star. Introducing the hybridization function

$$\Delta_\sigma(\tau_1-\tau_2) = -\sum_{i,j\neq 0} t^*_{i0}t^*_{j0}G^{(0)}_{ij\sigma}(\tau_1-\tau_2), \tag{7.31}$$

and employing the free ("Weiss") mean-field propagator $\mathcal{G}_\sigma$ one can express the DMFT local action in the following form (here the site index $i = 0$ is omitted for simplicity)

$$\begin{aligned} S_{\rm loc} = & -\int_0^\beta \mathrm{d}\tau_1 \int_0^\beta \mathrm{d}\tau_2 \sum_\sigma c^*_\sigma(\tau_1)\mathcal{G}^{-1}_\sigma(\tau_1-\tau_2)c_\sigma(\tau_2) \\ & + U\int_0^\beta \mathrm{d}\tau c^*_\uparrow(\tau)c_\uparrow(\tau)c^*_\downarrow(\tau)c_\downarrow(\tau), \end{aligned} \tag{7.32}$$

where

$$\mathcal{G}^{-1}_\sigma(\tau_1-\tau_2) = -\left(\frac{\partial}{\partial\tau_1}-\mu\right)\delta_{\tau_1\tau_2} - \Delta_\sigma(\tau_1-\tau_2). \tag{7.33}$$

Finally, we need the relation between the Green function $G^{(0)}_{ij\sigma}(\tau-\tau')$ where the site $i = 0$ is removed and the full lattice Green function, i.e.,

$$G^{(0)}_{ij\sigma} = G_{ij\sigma} - G_{i0\sigma}G^{-1}_{00\sigma}G_{0j\sigma}, \tag{7.34}$$

which holds for a general lattice.

In order to obtain the full solution of the lattice problem it is convenient to express the relation between the local Green function $G_{00\sigma} \equiv G_\sigma$ and the "Weiss" mean

field[7] $\mathcal{G}_\sigma^{-1}$ in the form of a Dyson equation

$$[G_\sigma(i\omega_n)]^{-1} = \mathcal{G}_\sigma^{-1}(i\omega_n) - \Sigma_\sigma(i\omega_n) \tag{7.35}$$

$$= i\omega_n + \mu - \Delta_\sigma(i\omega_n) - \Sigma_\sigma(i\omega_n). \tag{7.36}$$

Then the lattice Green function (in $\boldsymbol{k}$-space) $G_{\boldsymbol{k}\sigma}(i\omega_n)$ is given by

$$G_{\boldsymbol{k}\sigma}(i\omega_n) = \frac{1}{i\omega_n - \epsilon_{\boldsymbol{k}} + \mu - \Sigma_\sigma(i\omega_n)}. \tag{7.37}$$

After performing the so-called lattice Hilbert transform we recover the local Green function

$$G_\sigma(i\omega_n) = \sum_{\boldsymbol{k}} G_{\boldsymbol{k}\sigma}(i\omega_n) = \sum_{\boldsymbol{k}} \frac{1}{i\omega_n - \epsilon_{\boldsymbol{k}} + \mu - \Sigma_\sigma(i\omega_n)} \tag{7.38}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}\epsilon \frac{N(\omega)}{i\omega_n - \epsilon + \mu - \Sigma_\sigma(i\omega_n)}. \tag{7.39}$$

The ionic lattice on which the electrons move, and its structure, are seen to enter only via the DOS of the non-interacting electrons. After analytic continuation to real frequencies the local ("$\boldsymbol{k}$ averaged") propagator reads

$$G_{\boldsymbol{k}\sigma}(\omega) = \frac{1}{\omega - \epsilon_{\boldsymbol{k}} + \mu - \Sigma_\sigma(\omega)}. \tag{7.40}$$

This completes the derivation of the self-consistent DMFT equations. Namely, the functional integral determining the local propagator

$$G_\sigma(i\omega_n) = -\frac{1}{\mathcal{Z}} \int \prod_\sigma Dc_\sigma^* Dc_\sigma [c_\sigma(i\omega_n) c_\sigma^*(i\omega_n)] \exp[-S_{\mathrm{loc}}], \tag{7.41}$$

where the partition function $\mathcal{Z}$ and the local action S_{loc} are given by

$$\mathcal{Z} = \int \prod_\sigma Dc_\sigma^* Dc_\sigma \exp[-S_{\mathrm{loc}}] \tag{7.42}$$

and (7.32), respectively, together with the expression (7.39) for the lattice Green function provide a closed set of equations for the local propagator $G_\sigma(i\omega_n)$ and the

[7] In principle, any one of the local functions $\mathcal{G}_\sigma(i\omega_n)$, $\Sigma_\sigma(i\omega_n)$, or $\Delta_\sigma(i\omega_n)$ can be viewed as a "dynamical mean field" acting on particles on a site, since they all appear in the bilinear term of the local action (7.32).

self-energy $\Sigma_\sigma(i\omega_n)$. These equations can be solved iteratively: Starting with an initial value for the self-energy one obtains the local propagator from (7.39) and thereby the bath Green function $\mathcal{G}_\sigma(i\omega_n)$ from (7.35). This determines the local action (7.32) which is used to solve the single-impurity problem (7.41), leading to a new value for the local propagator and, by employing the old self-energy, a new bath Green function, etc. The single-impurity problem is still a complicated many-body interacting problem which cannot, in general, be solved exactly.

7.3.2 *Solution of the Self-Consistency Equations of the DMFT*

Due to the purely on-site interaction in the local action (7.32) the dynamics of the Hubbard model, (7.1), remains complicated even in the limit $d \to \infty$. Exact evaluations are only feasible when there is no coupling between the frequencies as, for example, in the Falicov–Kimball model [36]. This model was solved analytically by Brandt and Mielsch [19] soon after the introduction of the $d \to \infty$ limit [12].

In general, the local action (7.32) is the most complicated part of the DMFT equations. To solve the self-consistency equations different techniques ("impurity solvers") have been developed which are either fully numerical and "numerically exact," or semi-analytic and approximate. The numerical solvers can be divided into renormalization group techniques such as the numerical renormalization group (NRG) [37, 38] and the density-matrix renormalization group (DMRG) [39], exact diagonalization (ED) [40–42], and methods based on the stochastic sampling of quantum and thermal averages, i.e., QMC techniques such as the Hirsch-Fye QMC algorithm [32, 33, 43, 44] and continuous-time (CT) QMC [45–47].

Semi-analytic approximations such as the iterated perturbation theory (IPT) [31, 33, 48], the non-crossing approximation (NCA) [33, 49], the fluctuation exchange approximation (FLEX) [50–53], the local moment approach (LMA) [54, 55], and the parquet approximation [56] can also provide valuable insight.

It quickly turned out that the DMFT is a powerful tool for the investigation of electronic systems with strong correlations. It provides a non-perturbative and thermodynamically consistent approximation scheme for finite-dimensional systems which is particularly valuable for the study of intermediate-coupling problems where perturbative techniques fail [8, 33, 57–59].

7.4 The Mott–Hubbard Metal–Insulator Transition

The correlation induced transition between a paramagnetic metal and a paramagnetic insulator, referred to as "Mott–Hubbard MIT", is one of the most intriguing phenomena in condensed matter physics [60–62]. This transition is a consequence of the competition between the kinetic energy of the electrons and their local interaction U. Namely, the kinetic energy prefers the electrons to move (a wave

effect) which leads to doubly occupied sites and thereby to interact between the electrons (a particle effect). For large values of U the doubly occupied sites become energetically very costly. The system may reduce its total energy by localizing the electrons. Hence the Mott transition is a localization–delocalization transition, demonstrating the particle–wave duality of electrons [59].

Mott–Hubbard MITs are found, for example, in transition metal oxides with partially filled bands. For such systems band theory typically predicts metallic behavior. The most famous example is V_2O_3 doped with Cr [63–65]. In particular, below $T = 380\,\mathrm{K}$ the MIT in paramagnetic $(V_{0.96}Cr_{0.04})_2O_3$ is of first order [64], with discontinuities in the lattice parameters and in the conductivity. However, the two phases remain isostructural.

Making use of the half-filled, single-band Hubbard model (7.1) the Mott–Hubbard MIT was studied intensively in the past [5, 60–62, 65]. Important early results were obtained by Hubbard [66] within a Green function decoupling scheme, and by Brinkman and Rice [67] within the Gutzwiller variational method [4], both at $T = 0$. Hubbard's approach yields a continuous splitting of the band into a lower and upper Hubbard band, but cannot describe quasiparticle features. By contrast, the Gutzwiller–Brinkman–Rice approach gives a good description of the low-energy, quasiparticle behavior, but cannot reproduce the upper and lower Hubbard bands. In the latter approach the MIT is signalled by the disappearance of the quasiparticle peak.

To solve this problem the DMFT has been extremely valuable since it provided detailed insights into the nature of the Mott–Hubbard MIT for all values of the interaction U and temperature T [33, 59].

7.4.1 DMFT and the Three-Peak Structure of the Spectral Function

The Mott–Hubbard MIT is monitored by the spectral function $A(\omega) = -\frac{1}{\pi}\mathrm{Im}G(\omega + i0^+)$ of the correlated electrons[8] [33]. While at small U the system can be described by coherent quasiparticles whose DOS still resembles that of the free electrons, the spectrum in the Mott insulator state consists of two separate incoherent "Hubbard bands" whose centers are separated approximately by the energy U. The latter originate from atomic-like excitations at the energies $\pm U/2$ broadened by the hopping of electrons away from the atom. At intermediate values of U the spectrum then has a characteristic three-peak structure as in the single-impurity Anderson model, which includes both the atomic features (i.e., Hubbard bands) and the narrow quasiparticle peak at low excitation energies, near $\omega = 0$. This corresponds to a strongly correlated metal. The general structure of the spectrum (lower Hubbard band, quasiparticle peak, and upper Hubbard band) is

[8] In the following we only consider the paramagnetic phase, whereas magnetic order is assumed to be suppressed ("frustrated").

rather insensitive to the specific form of the DOS of the non-interacting electrons. The width of the quasiparticle peak vanishes for $U \to U_{c2}(T)$. On decreasing U, the transition from the insulator to the metal occurs at a lower critical value $U_{c1}(T)$, where the gap vanishes.

It is important to note that the three-peak spectrum originates from a lattice model with only *one* type of electrons. This is in contrast to the single-impurity Anderson model whose spectrum shows very similar features, but is due to *two* types of electrons – the localized orbital at the impurity site and the free conduction band. Therefore the screening of the magnetic moment which gives rise to the Kondo effect in impurity systems has somewhat a different origin in interacting lattice systems. Namely, as explained by the DMFT the electrons provide both the local moments and the electrons which screen these moments [33].

Interestingly, for any typical spectral function $A(\omega)$ with three peaks, Kramers–Kronig relations and the DMFT self-consistency equations imply that the self-energy $\Sigma(\omega)$ abruptly changes slope *inside* the central peak at some frequency $\omega_\star$ [68], once at positive and once at negative frequency. While this behavior is not visible in $A(\omega)$ itself, it leads to "kinks" in the effective dispersion relation $E_{\boldsymbol{k}}$ of one-particle excitations, which is defined as the frequency for which the momentum-resolved spectral function $A(\boldsymbol{k},\omega) = -\mathrm{Im}G(\boldsymbol{k},\omega)/\pi = -(1/\pi)\,\mathrm{Im}[1/(\omega+\mu-\epsilon_{\mathbf{k}}-\Sigma(\omega))]$ is maximal. For frequencies below $\omega_\star$ the dispersion is given by Fermi-liquid (FL) theory, $E_{\boldsymbol{k}} = Z_{\mathrm{FL}}\epsilon_{\boldsymbol{k}}$, where $Z_{\mathrm{FL}} = (\partial \mathrm{Re}\Sigma(\omega)/\partial\omega)_{\omega=0}$ is the FL renormalization parameter. The FL regime terminates at the kink energy scale $\omega_\star$. This energy cannot be obtained within FL theory itself. Namely, it is determined by Z_{FL} and the non-interacting DOS, e.g., $\omega_\star = 0.41 Z_{\mathrm{FL}} D$, where D is an energy scale of the non-interacting system such as half the bandwidth [68]. Above $\omega_\star$ the dispersion is given by a different renormalization with a small offset, $E_{\boldsymbol{k}} = Z_{\mathrm{CP}}\epsilon_{\boldsymbol{k}} + \mathrm{const}$, where Z_{CP} is the weight of the central peak of $A(\omega)$. This theory explains kinks in the slope of the dispersion as a direct consequence of the electronic interaction [68]. The same mechanism may also lead to kinks in the low-temperature electronic specific heat [69]. These kinks have also been linked to maxima in the spin susceptibility [70]. Of course, *additional* kinks in the electronic dispersion may also arise from the coupling of electrons to bosonic degrees of freedom, such as phonons [71, 72] or spin fluctuations [73, 74]. Interestingly, recent experiments [75] have found evidence for kinks in Ni(1 1 0), which may be due to the electronic mechanism discussed here.

The evolution of the spectral function of the half-filled frustrated Hubbard model at a finite temperature is shown in Fig. 7.2. This temperature is above the temperature of the critical point so that there is no real transition but only a crossover from a metallic-like to an insulating-like solution. The height of the quasiparticle peak at the Fermi energy is no longer fixed at its zero temperature value. This is due to the temperature dependent imaginary part of the self-energy. The spectral weight of the quasiparticle peak is seen to be gradually redistributed and shifted to the upper (lower) edge of the lower (upper) Hubbard band. The inset of Fig. 7.2 shows the U-dependence of the spectral function at zero frequency $A(\omega = 0)$. For higher values of U the spectral density at the Fermi level is still finite and vanishes

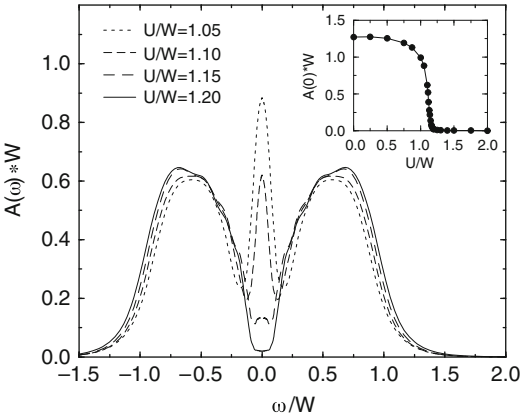

Fig. 7.2 Spectral function for the half-filled Hubbard model for various values of U at $T = 0.0276\,W$ (W: bandwidth) in the crossover region. The crossover from the metal to the insulator occurs via a gradual suppression of the quasiparticle peak at $\omega = 0$. The inset shows the U dependence of $A(\omega = 0)$, in particular the rapid decrease for $U \approx 1.1\,W$; from [76]

only in the limit $U \to \infty$. For the insulating phase DMFT predicts the filling of the Mott–Hubbard gap with increasing temperature. This is due to the fact that the insulator and the metal are not distinct phases in the crossover regime, implying that the insulator has a finite spectral weight at the Fermi level. This behavior has been detected experimentally by photoemission experiments [77].

Altogether, the thermodynamic transition line $U_c(T)$ corresponding to the Mott–Hubbard MIT is found to be of first order at finite temperatures, and is associated with a hysteresis region in the interaction range $U_{c1} < U < U_{c2}$, where U_{c1} and U_{c2} are the interaction values at which the insulating and metallic solution, respectively, vanish [33, 37, 76, 78–80]. As shown in Fig. 7.3 the hysteresis region terminates at a critical end point. At higher temperatures the transition changes into a smooth crossover from a bad metal to a bad insulator. At half-filling and for bipartite lattices in dimensions $d > 2$ (in $d = 2$ only at $T = 0$) the paramagnetic phase is, in fact, unstable against antiferromagnetic long-range order. The MIT is then completely hidden by the antiferromagnetic insulating phase [81].

In Fig. 7.3 it is seen that the slope of the phase transition line U_c is negative down to $T = 0$, which implies that for constant interaction U the metallic phase can be reached from the insulator by decreasing the temperature T, i.e., by cooling. This anomalous behavior (which corresponds to the Pomeranchuk effect [82] in ^{3}He, if we associate solid ^{3}He with the insulator and liquid ^{3}He with the metal) can be understood from the Clausius–Clapeyron equation $\mathrm{d}U/\mathrm{d}T = \Delta S/\Delta D$. Here ΔS is the difference between the entropy in the metal and in the insulator, and ΔD is the difference between the number of doubly occupied sites in the two phases. Within the DMFT there is no exchange coupling J between the spins of the electrons in the insulator, since the scaling (7.5) implies $J \propto -t^2/U \propto 1/d \to 0$ for

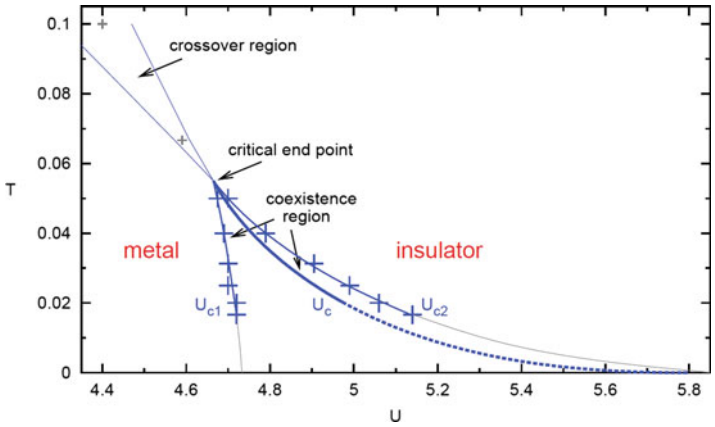

Fig. 7.3 Phase diagram of the Mott–Hubbard MIT showing the metallic phase and the insulating phase, respectively, at temperatures below the critical end point, as well as a coexistence region; from [80]

$d \to \infty$. Hence the insulating state is macroscopically degenerate, with entropy $S_{\text{ins}} = k_{\text{B}} \ln 2$ per electron down to $T = 0$. This is larger than the entropy $S_{\text{met}} \propto T$ per electron in the Landau Fermi liquid describing the metal, i.e., $\Delta S = S_{\text{met}} - S_{\text{ins}} < 0$. At the same time the number of doubly occupied sites is lower in the insulator than in the metal, i.e., $\Delta D = D_{\text{met}} - D_{\text{ins}} > 0$. The Clausius–Clapeyron equation then implies that the phase-transition line T versus U has indeed a negative slope down to $T = 0$. However, this is an artifact of the DMFT. Namely, there will usually exist an exchange coupling between the electrons which leads to a vanishing entropy of the insulator for $T \to 0$. Since the entropy of the insulator vanishes faster than linearly with the temperature, the difference $\Delta S = S_{\text{met}} - S_{\text{ins}}$ eventually becomes positive, whereby the slope also becomes positive[9]; this is indeed observed in cluster DMFT calculations [83]. However, since $\Delta S \to 0$ for $T \to 0$ the phase boundary must eventually terminate at $T = 0$ with infinite slope.

7.5 Theory of Electronic Correlations in Materials

7.5.1 The LDA + DMFT Approach

Although the Hubbard model is able to explain basic features of the phase diagram of correlated electrons it cannot describe the physics of real materials in any detail.

[9]Here we assume for simplicity that the metal remains a Fermi liquid and the insulator stays paramagnetic down to the lowest temperatures.

Clearly, realistic theories must take into account the explicit electronic and lattice structure of the systems.

Until recently the electronic properties of solids were investigated by two essentially separate communities, one using model Hamiltonians in conjunction with many-body techniques, the other employing density functional theory (DFT) [84,85]. DFT and its local density approximation (LDA) have the advantage of being ab initio approaches which do not require empirical parameters as input. Indeed, they are highly successful techniques for the calculation of the electronic structure of real materials [86]. However, in practice DFT/LDA is seriously restricted in its ability to describe strongly correlated materials where the on-site Coulomb interaction is comparable with the band width. Here, the model Hamiltonian approach is more powerful since there exist systematic theoretical techniques to investigate the many-electron problem with increasing accuracy. Nevertheless, the uncertainty in the choice of the model parameters and the technical complexity of the correlation problem itself prevent the model Hamiltonian approach from being a suitable tool for studying real materials. The two approaches are therefore complementary. In view of the individual power of DFT/LDA and the model Hamiltonian approach, respectively, a combination of these techniques for ab initio investigations of correlated materials including, for example, f-electron systems and Mott insulators, would be highly desirable. One of the first successful attempts in this direction was the LDA + U method [87, 88], which combines LDA with a static, i.e., Hartree–Fock-like, mean-field approximation for a multi-band Anderson lattice model with interacting and non-interacting orbitals. This method proved to be a very useful tool in the study of long-range ordered, insulating states of transition metals and rare-earth compounds. However, the paramagnetic metallic phase of correlated electron systems such as high-temperature superconductors and heavy-fermion systems clearly requires a treatment that goes beyond a static mean-field approximation and includes dynamical effects, i.e., the frequency dependence of the self-energy.

Here the recently developed LDA + DMFT method, a new computational scheme which merges electronic band structure calculations and the DMFT, has proved to be a breakthrough [49, 52, 59, 89–98]. Starting from conventional band structure calculations in the LDA the correlations are taken into account by the Hubbard interaction and a Hund's rule coupling term. The resulting DMFT equations are solved numerically, e.g., with a QMC algorithm. By construction, LDA + DMFT includes the correct quasiparticle physics and the corresponding energetics. It also reproduces the LDA results in the limit of weak Coulomb interaction U. More importantly, LDA + DMFT correctly describes the correlation induced dynamics near a Mott–Hubbard MIT and beyond. Thus, LDA + DMFT is able to account for the physics at all values of the Coulomb interaction and doping level.

In the LDA + DMFT approach the LDA band structure is expressed by a one-particle Hamiltonian $\hat{H}^0_{\mathrm{LDA}}$, and is then supplemented by the local Coulomb repulsion U and Hund's rule exchange J. This leads to a material specific generalization of the one-band model Hamiltonian

$$\hat{H} = \hat{H}^0_{\mathrm{LDA}} + U \sum_m \sum_i \hat{n}_{im\uparrow}\hat{n}_{im\downarrow} + \sum_{i,m\neq m',\sigma,\sigma'} (V - \delta_{\sigma\sigma'} J)\, \hat{n}_{im\sigma}\hat{n}_{im'\sigma'}. \quad (7.43)$$

Here m and m' enumerate those orbitals for which the interaction between the electrons is explicitly included, e.g., the three t_{2g} orbitals of the $3d$ electrons of transition metal ions or the $4f$ orbitals in the case of rare earth elements. The interaction parameters are related by $V = U - 2J$ which holds exactly for degenerate orbitals and is a good approximation for t_{2g} electrons. The actual values for U and V can be calculated by constrained LDA [49].

In the one-particle part of the Hamiltonian

$$\hat{H}^0_{\mathrm{LDA}} = \hat{H}_{\mathrm{LDA}} - \sum_i \sum_{m\sigma} \Delta\epsilon_d\, \hat{n}_{im\sigma}. \quad (7.44)$$

the energy term containing $\Delta\epsilon_d$ is a shift of the one-particle potential of the interacting orbitals. It cancels the Coulomb contribution to the LDA results, and can also be calculated by constrained LDA [49].

Within the LDA + DMFT scheme the self-consistency condition connecting the self-energy Σ and the Green function G at frequency ω reads:

$$G_{qm,q'm'}(\omega) = \frac{1}{V_{\mathrm{B}}} \int \mathrm{d}^3k \left(\left[\omega\mathbf{1} + \mu\mathbf{1} - H^0_{\mathrm{LDA}}(\boldsymbol{k}) - \boldsymbol{\Sigma}(\omega) \right]^{-1} \right)_{qm,q'm'}. \quad (7.45)$$

Here, $\mathbf{1}$ is the unit matrix, μ the chemical potential, $H^0_{\mathrm{LDA}}(\boldsymbol{k})$ is the orbital matrix of the LDA Hamiltonian derived, for example, in a linearized muffin-tin orbital (LMTO) basis, $\boldsymbol{\Sigma}(\omega)$ denotes the self-energy matrix which is nonzero only between the interacting orbitals, and $[...]^{-1}$ implies the inversion of the matrix with elements $n\ (=qm)$, $n'(=q'm')$, where q and m are the indices of the atom in the primitive cell and of the orbital, respectively.[10] The integration extends over the Brillouin zone with volume V_{B}.

For cubic transition metal oxides (7.45) can be simplified to

$$G(\omega) = G^0(\omega - \Sigma(\omega)) = \int \mathrm{d}\epsilon \frac{N^0(\epsilon)}{\omega - \Sigma(\omega) - \epsilon} \quad (7.46)$$

provided the degenerate t_{2g} orbitals crossing the Fermi level are well separated from the other orbitals [49]. For non-cubic systems the degeneracy is lifted. In this case (7.46) is an approximation where different $\Sigma_m(\omega)$, $N^0_m(\epsilon)$ and $G_m(\omega)$ have to be used for the three non-degenerate t_{2g} orbitals.

The Hamiltonian (7.43) is diagonalized within the DMFT where, for example, QMC techniques [35] are used to solve the self-consistency equations. From the imaginary time QMC Green function we calculate the physical (real frequency) spectral function with the maximum entropy method [99].

[10] We note that $\hat{H}^0_{\mathrm{LDA}}$ may include additional non-interacting orbitals.

During the last few years the LDA + DMFT and other DMFT based computational schemes have provided great progress in the understanding of the electronic, magnetic and structural properties of many correlated electron materials. These materials range from $3d$ transition metals and their oxides, and f electron systems, all the way to Heusler alloys, ferromagnetic half-metals, fullerenes, and zeolites [52, 95–98]. Nevertheless, this framework still needs to be considerably improved before it becomes a truly comprehensive ab initio approach for complex correlated matter with predictive power. In particular, the interface between the band-structure and the many-body components of the approach needs to be optimized. This includes, for example, a solution of the double counting correction problem and a fully self-consistent treatment of the spin, orbital, and charge densities. Another important goal are realistic computations of free energies and forces, and the development of efficient methods to treat non-local correlations with the quantum cluster methods [100–102].

7.5.2 Single-Particle Spectrum of Correlated Electrons in Materials

Transition metal oxides are an ideal laboratory for the study of electronic correlations in solids. Among these materials, cubic perovskites have the simplest crystal structure and thus may be viewed as a starting point for understanding the electronic properties of more complex systems. Typically, the $3d$ states in those materials form comparatively narrow bands with width $W \sim 2$–3 eV, which leads to strong Coulomb correlations between the electrons. Particularly simple are transition metal oxides with a $3d^1$ configuration since, among others, they do not show a complicated multiplet structure.

Photoemission spectra provide a direct experimental tool to study the electronic structure and spectral properties of electronically correlated materials. Intensive experimental investigations of spectral and transport properties of strongly correlated $3d^1$ transition metal oxides started with investigations by Fujimori et al. [103]. These authors observed a pronounced lower Hubbard band in the photoemission spectra which cannot be explained by conventional band structure theory. In photoemission spectroscopy (PES) a photon of a given energy is used to emit an electron whose properties (energy, angular distribution) are measured by a detector. Angular resolved PES is referred to as ARPES. These techniques measure the *occupied* electronic states, i.e., those states which are described by the full spectral function multiplied by the Fermi function $f(\omega, T)$. By contrast, inverse photoemission spectroscopy (IPES) measures the *unoccupied* electronic states, i.e., the states described by the full spectral function of a material multiplied by $1 - f(\omega, T)$. IPES is harder to perform and not as accurate as PES. But in many situations information about the unoccupied states is also available by X-ray absorption spectroscopy (XAS).

Spectroscopic techniques provide very valuable information about correlated electronic systems since they can directly measure the spectral function of a material, a quantity which can also be directly calculated as discussed in Sect. 7.4. In particular, photoemission techniques allow one to detect the correlation induced shift of spectral weight. In the following we will illustrate the computation of the $\boldsymbol{k}$-integrated electronic spectra of correlated materials within the LDA + DMFT scheme by investigating the two simple transition metal oxides $SrVO_3$ and $CaVO_3$.

7.5.2.1 $Sr_xCa_{1-x}VO_3$

$SrVO_3$ and $CaVO_3$ are simple transition metal compounds with a $3d^1$ configuration. The main effect of the substitution of Sr ions by the isovalent, but smaller, Ca ions is to decrease the V–O–V angle from $\theta = 180°$ in $SrVO_3$ to $\theta \approx 162°$ in the orthorhombically distorted structure of $CaVO_3$. Remarkably this rather strong bond bending results only in a 4% decrease of the one-particle bandwidth W and thus in a correspondingly small increase of the ratio U/W as one moves from $SrVO_3$ to $CaVO_3$ [104, 105].

LDA + DMFT(QMC) spectral functions of $SrVO_3$ and $CaVO_3$ were calculated by Sekiyama et al. [104] by starting from the respective LDA DOS of the two materials; they are shown in Fig. 7.4. These spectra show genuine correlation effects, i.e., the formation of lower Hubbard bands at about 1.5 eV and upper Hubbard bands at about 2.5 eV, with well-pronounced quasiparticle peaks at the Fermi energy. Therefore both $SrVO_3$ and $CaVO_3$ are strongly correlated metals. The small difference of the LDA bandwidth of $SrVO_3$ and $CaVO_3$ is only reflected in some additional transfer of spectral weight from the quasiparticle peak to the Hubbard bands, and minor differences in the positions of the Hubbard bands. The DOS of the two systems shown in Fig. 7.4 are quite similar. In fact, $SrVO_3$ is slightly less correlated than $CaVO_3$, in accord with their different LDA bandwidths. The inset of Fig. 7.4 shows that the effect of temperature on the spectrum is weak for $T \lesssim 700$ K. Detailed spectra of $SrVO_3$ and $CaVO_3$ were also computed by Pavarini et al. [106].

Since the three t_{2g} orbitals of this simple $3d^1$ material are almost degenerate the spectral function has the same three-peak structure as that of the one-band Hubbard

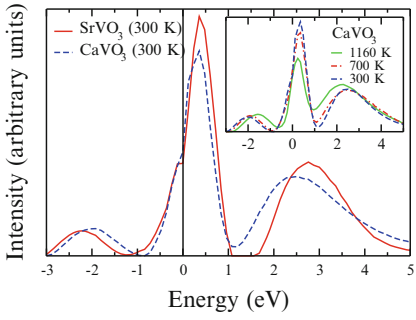

Fig. 7.4
LDA + DMFT(QMC) spectral functions of $SrVO_3$ (*solid line*) and $CaVO_3$ (*dashed line*) calculated at T = 300 K; inset: effect of temperature in the case of $CaVO_3$; after [104]

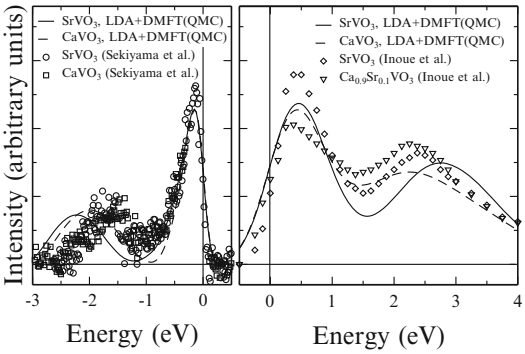

Fig. 7.5 Comparison of the calculated, parameter-free LDA + DMFT(QMC) spectral functions for $SrVO_3$ (*solid line*) and $CaVO_3$ (*dashed line*) with experiment. *Left*: Bulk-sensitive high-resolution PES ($SrVO_3$: *circles*; $CaVO_3$: *rectangles*). *Right*: 1s XAS for $SrVO_3$ (*diamonds*) and $Ca_{0.9}Sr_{0.1}VO_3$ (*triangles*) [107]. *Horizontal line*: experimental subtraction of the background intensity; after [105]

model shown in Fig. 7.2. The temperature induced decrease the quasiparticle peak height that is also clearly seen. We note that the actual form of the spectrum no longer resembles the LDA DOS used as input, i.e., it essentially depends only on the first three energy moments of the LDA DOS (electron density, average energy, and band width).

In the left panel of Fig. 7.5 the LDA + DMFT(QMC) spectral functions at 300 K are compared with experimental high-resolution bulk PES. For this purpose the full theoretical spectra were multiplied with the Fermi function at the experimental temperature (20 K) and were Gauss broadened with the experimental resolution of 0.1 eV [104]. The quasiparticle peaks in theory and experiment are seen to be in very good agreement. In particular, their height and width are almost identical for both $SrVO_3$ and $CaVO_3$. The difference in the positions of the lower Hubbard bands may be partly due to (1) the subtraction of the estimated oxygen contribution (which may also remove some $3d$ spectral weight below -2 eV), and (2) uncertainties in the ab initio calculation of the local Coulomb interaction strength. In the right panel of Fig. 7.5 comparison is made with XAS data of Inoue et al. [107]. To this end the full LDA + DMFT spectrum was multiplied with the inverse Fermi function at 80 K and was then Gauss broadened with the experimental resolution of 0.36 eV [I.H. Inoue, private communication (2003)]. The overall agreement of the weights and positions of the quasiparticle and upper t_{2g} Hubbard band is good, including the tendencies when going from $SrVO_3$ to $CaVO_3$ (in fact, $Ca_{0.9}Sr_{0.1}VO_3$ in the experiment). For $CaVO_3$ the weight of the quasiparticle peak is somewhat lower than in the experiment. In contrast to one-band Hubbard model calculations, the material specific results reproduce the strong asymmetry around the Fermi energy w.r.t. weights and bandwidths.

The experimentally determined spectra of $SrVO_3$ and $CaVO_3$ and the good agreement with parameter-free LDA + DMFT calculations confirm the existence of a pronounced three-peak structure in a correlated bulk material. Although the

DMFT had predicted such a behavior for the Hubbard model [33] it was not clear whether the DMFT would really be able to describe real materials. Now it is clear that the three-peak structure not only occurs in single-impurity Anderson models or the DMFT for the Hubbard model, but also is a characteristic feature of correlated *bulk* matter in $d = 3$.

7.6 Electronic Correlations and Disorder

The properties of real materials are influenced not only by the interaction between the electrons in the periodic crystal lattice, but also by the presence of randomness, e.g., impurities and lattice defects [108]. In particular, Coulomb correlations and disorder are separately the driving forces behind MITs (MITs) connected with the localization and delocalization of particles. The Mott–Hubbard MIT is caused by the electronic repulsion [5, 61, 109] and is characterized by the opening of a gap in the DOS at the Fermi level. By contrast, the Anderson localization transition is due to coherent backscattering of non-interacting particles from randomly distributed impurities [108, 110]. At the Anderson transition the character of the spectrum at the Fermi level changes from a continuous to a dense point spectrum. Both MITs can be characterized by a single quantity, the local density of states (LDOS). Although the LDOS is not an order parameter associated with a symmetry-breaking phase transition, it can distinguish between a metal and an insulator.

The DMFT can easily be extended to study correlated lattice electrons with local disorder [8, 111–116]. For this purpose a single-particle term with random local energies ϵ_i is added to the Hubbard model, leading to the Anderson–Hubbard model

$$\hat{H} = -t \sum_{ij,\sigma} \hat{c}^{+}_{i\sigma} \hat{c}_{j\sigma} + \sum_{i\sigma} \epsilon_i n_{i\sigma} + U \sum_{i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}. \tag{7.47}$$

The ionic energies ϵ_i describe the local, quenched disorder acting on the motion of the electrons. They are drawn from a probability distribution function $\mathcal{P}(\epsilon_i)$, which can be a continuous or a multi-modal function.

The DMFT provides a valuable, non-perturbative theoretical framework also for the investigation of correlated electrons in the presence of disorder. If in the DMFT the effect of local disorder is taken into account through the arithmetic mean of the LDOS one obtains, in the absence of interactions ($U = 0$), the coherent potential approximation (CPA) [30, 117], which does not describe the physics of Anderson localization. To overcome this deficiency Dobrosavljević and collaborators formulated a variant of the DMFT where the *geometrically* averaged LDOS is computed from the solutions of the self-consistent stochastic DMFT equations [118] which is then incorporated into the self-consistency cycle [119]. Thereby a mean-field theory of Anderson localization can be derived which reproduces many of the expected features of the disorder-driven MIT for non-interacting electrons [119]. This scheme uses only one-particle quantities and can therefore easily be

included in the DMFT to treat disordered electrons in the presence of phonons [120] or Coulomb correlations [121,122]. In particular, the DMFT with geometrical averaging allows one to compute the phase diagram for the Anderson–Hubbard model [121, 122]. For a continuous disorder distribution function and a half-filled band one finds that the metallic phase is *enhanced* at small and intermediate values of the interaction and disorder, but that the metallicity is eventually destroyed when the disorder is strong enough [121]. Surprisingly, the Mott and Anderson insulators are continuously connected.

New interesting phenomena are also expected in correlated electron systems when the disorder distribution function has a binary-alloy form. Namely, it was predicted that a disorder induced splitting of the band ("alloy-band splitting") can enhance the critical temperature for the onset of itinerant ferromagnetism [123,124]. Another direct consequence of the alloy-band splitting is the fact that, if the alloy subband is effectively half-filled, a Mott–Hubbard MIT can occur at non-integer filling [125]. The spectral-weight transfer in correlated systems with binary-alloy disorder was also investigated away from the MIT regime [126, 127]. For the periodic Anderson model with the binary-alloy disorder analogous behavior was predicted [128].

7.7 DMFT for Correlated Bosons in Optical Lattices

The observation of Bose–Einstein condensation (BEC) in ultra-cold atomic gases has greatly stimulated research into the properties of this fascinating quantum state of matter [129]. In particular, experiments with alkali atoms confined in optical lattices [130–132] have renewed the theoretical interest [133–135] in the physics of strongly correlated bosons on lattices, which promises significant new insights and even applications in fields such as quantum computing [136].

A lattice model of interacting bosons with purely local interaction, the bosonic Hubbard model, has the form

$$H_{\mathrm{BH}} = \sum_{\langle i,j\rangle} t_{ij}\hat{b}_i^\dagger \hat{b}_j + \frac{U}{2}\sum_{\mathbf{R}_i} \hat{n}_i(\hat{n}_1 - 1), \tag{7.48}$$

where $\hat{b}_i$ and $\hat{b}_i^\dagger$ are bosonic operators. In the ground state two characteristically different phases are expected to occur: a bosonic incompressible Mott phase with a correlation gap and a compressible superfluid phase characterized by a non-vanishing expectation value $\langle \hat{b}_i \rangle$ which serves as an order parameter [137].

The construction of a DMFT for the bosonic Hubbard model which – as in the case of the fermionic DMFT – becomes exact in the limit d or $Z \to \infty$ and is valid at all temperatures is made complicated by the fact that the system can Bose condense below a BEC temperature T_{BEC}. This has immediate consequences for the expectation value of the kinetic energy in (7.48). Namely, while in the normal

state only the product $\hat{b}_i^\dagger \hat{b}_j$ has a finite expectation value, in the Bose condensed phase even $\langle \hat{b}_j \rangle$ (the order parameter) is non-zero. For the expectation value of the kinetic energy in (7.48) to remain finite in the limit $Z \to \infty$ the hopping amplitude then has to be scaled differently in the two phases. In the normal phase the hopping amplitude needs to be scaled as in the fermionic case ("quantum scaling"), i.e., as $t_{ij} = t_{ij}^*/\sqrt{Z}$, while in the condensed phase a classical scaling $t_{ij} = t_{ij}^*/Z$ is required. Such a scaling of the hopping amplitudes cannot be performed on the level of the Hamiltonian, but is possible in the effective action [138]. The bosonic DMFT ("B-DMFT") derived thereby treats normal and condensed bosons on equal footing and thus includes the effects caused by their dynamic coupling. The self-consistency equations of the B-DMFT are those of a single bosonic impurity coupled to two baths, one corresponding to bosons in the normal state and one to bosons in the condensate. The B-DMFT derived in the limit d or $Z \to \infty$ is again a comprehensive mean-field theory, i.e., it is valid for all input parameters and all temperatures. It not only reproduces all exactly solvable limits, such as the non-interacting ($U = 0$) and the atomic ($t_{ij} = 0$) limit, but also well-known approximation schemes for interacting bosons. For example, by neglecting all terms containing the hybridization function in the local action one obtains the mean-field theory of Fisher et al. [137]; for a detailed discussion see [138].

The solution of the self-consistent bosonic impurity problem defined by the B-DMFT equations requires new theoretical/computational methods. So far exact diagonalization [139, 140] and continuous-time QMC [141] were employed. Thereby the bosonic Hubbard model was solved on the Bethe lattice for finite [139] and infinite [140] coordination number Z as well as on a simple-cubic lattice [141]. The phase diagram of correlated bosons on a simple-cubic lattice computed by the bosonic DMFT was found to agree with that obtained by numerically exact QMC to within 2% [141].

The B-DMFT is expected to be a valuable approximation scheme for the investigation of lattice bosons in situations where exact numerical computations are difficult to perform or inefficient, as in the case of bosons with disorder or many internal degrees of freedom, and for Bose–Fermi mixtures [142].

7.8 DMFT for Nonequilibrium

Recently the study of strongly correlated many-body systems out of equilibrium has received much attention [143, 144]. This is motivated in particular by experimental progress in the investigation of ultra-cold atomic gases [132] and time-resolved pump-probe spectroscopy on strongly correlated materials [145, 146]. In general the real-time dynamics of correlated systems can be described by the extension of DMFT to nonequilibrium, provided that they are dominated by local temporal fluctuations and spatial correlations are not crucial. In nonequilibrium DMFT an effective impurity problem is formulated using the Keldysh formalism [147, 148],

and as in equilibrium DMFT this mapping onto a single site becomes exact in the limit of infinite dimensions. For the Hubbard model in nonequilibrium the single-site DMFT action reads

$$S = -i \int_{\mathcal{C}} \mathrm{d}t \; H_{\mathrm{loc}}(t) - i \sum_{\sigma} \int_{\mathcal{C}} \mathrm{d}t \int_{\mathcal{C}} \mathrm{d}t' \, c_{\sigma}^{\dagger}(t) \Lambda(t,t') c_{\sigma}(t') \,, \tag{7.49}$$

where the Keldysh contour $\mathcal{C}$ runs from t_{min} to t_{max} on the real time axis, back to t_{min}, and finally to $-i\beta$ along the imaginary time axis [149, 150]. The first term contains the local part of the Hamiltonian, e.g., $H_{\mathrm{loc}}(t) = U(t) n_{\uparrow} n_{\downarrow} - \mu(n_{\uparrow} + n_{\downarrow})$ for a time-dependent interaction. The second term involves the hybridization function $\Lambda(t,t')$ which couples the impurity to a time-dependent bath which must be determined self-consistently. Local contour-ordered correlation functions such as the Green function $G(t,t')$ are obtained from the action (7.49) as expectation values $\langle A(t)B(t')\cdots\rangle = \mathrm{Tr}[\mathrm{T}_{\mathcal{C}} \exp(S) A(t) B(t') \cdots]/\mathcal{Z}$ at the appropriate times, where $\mathrm{T}_{\mathcal{C}}$ is the contour-ordering operator. For the Hubbard model this evaluation is the most demanding part of nonequilibrium DMFT and can so far be done with real-time quantum Monte Carlo methods [150, 151] for not too long times, and for sufficiently large U using a self-consistent perturbation expansion around the atomic limit [152]. For the Falicov–Kimball model, on the other hand, closed equations of motion govern the impurity Green function [19], which can be solved on the real time axis. The hybridization function $\Lambda(t,t')$ in (7.49) must be determined self-consistently by computing the local self-energy $\Sigma(t,t')$ from the Dyson equation of the impurity model, calculating the momentum-dependent Green function of the lattice model from the lattice Dyson equation, integrating over momentum to obtain the local lattice Green function, and finally equating it with the impurity Green function. While this procedure is necessary for a Gaussian or other general DOS, for a semielliptical DOS the self-energy can be eliminated and Λ expressed directly in terms of G [153].

Nonequilibrium DMFT was used to obtain the response of time-resolved photoemission [154–156] and optical spectroscopy [157] in correlated systems in terms of Green functions of the electronic system. The Falicov–Kimball and Hubbard models were studied in the presence of dc and ac electric fields [148, 156, 158–165], as well as for abrupt [150–152, 155, 157, 166] or slow changes [167, 168] of the interaction parameter as a function of time.

As an example, we consider a sudden change ("quench") in the interaction parameter of the Hubbard model from $U = 0$ (i.e., with the non-interacting ground state as initial state) to finite values of U for times $t > 0$. For this case the DMFT equations were solved numerically for the paramagnetic phase and a semielliptic DOS [150–152]. In the following discussion the bandwidth is equal to 4, and time is thus measured in units of $4\hbar$/bandwidth, e.g., on the order of femto-seconds for a bandwidth on the order of eV. An interesting question is whether such an isolated system can thermalize due to the many-body interaction alone, i.e., whether at least some of its properties are the same as for an equilibrium system with the same energy [169]. Before the quench the momentum distribution is a step

function, and the Fermi surface discontinuity $\Delta n(t)$ remains nonzero for a finite time after the quench. For quenches to $U \lesssim 2.5$ the momentum discontinuity first reaches a so-called prethermalization plateau for $t \lesssim 5$ due to the vicinity of the integrable point at $U = 0$. This plateau in $\Delta n(t)$ is given to good accuracy by $2Z - 1$ [151], where Z is the Fermi-liquid quasiparticle weight *in equilibrium* at zero temperature and for interaction U. This value and also the transient behavior at short times is precisely predicted by second-order unitary perturbation theory in U [170, 171]. On the other hand, the double occupation essentially relaxes to its thermal value on this timescale, showing that the potential energy (and therefore also the kinetic energy) relax quicker than the occupation of individual states. For large U the behavior is different, showing strong so-called collapse-and-revival oscillations with approximate frequency $2\pi/U$. They stem from the vicinity of the atomic limit (i.e., zero hopping amplitude), for which the propagator $\mathrm{e}^{-\mathrm{i}Ht}$ is exactly periodic with period $2\pi/U$ [172]. For finite hopping (small compared to U) these oscillations are damped and decay on timescales of order $\hbar$/bandwidth. The oscillations of the double occupation are not centered at its thermal value, but rather at a different value that can be derived from strong-coupling perturbation theory [151]. The situation is thus similar to that at small coupling in the sense that the relaxation to the thermal state is delayed because the system is stuck in a metastable state close to an integrable point. However, both the weak-coupling prethermalization plateau in $\Delta n(t)$ as well as the strong-coupling oscillations vanish in a narrow region of interaction parameters U near 3.2 [151]. For quenches of U to approximately this value the system thermalizes rapidly: Both the momentum distribution and thus also the double occupation (due to energy conservation after quench) relax to their thermal values. In fact the retarded nonequilibrium Green function relaxes to the corresponding equilibrium function [150], so that all observables that can be obtained from it tend to the thermal value predicted by equilibrium statistical mechanics. In other words, for quenches to interaction values in the vicinity of $U \approx 3.2$ the isolated system indeed thermalizes rapidly due to the many-body interactions. Many interesting questions remain open in this context, e.g., how thermalization depends on the parameters of the system. We refer to the reviews [143, 144] for further discussion.

7.9 Summary and Outlook

Due to the intensive international research over the last 2 decades the DMFT has quickly developed into a powerful method for the investigation of electronic systems with strong correlations. It provides a comprehensive, non-perturbative and thermodynamically consistent approximation scheme for the investigation of finite-dimensional systems (in particular for dimension $d = 3$), and is particularly useful for the study of problems where perturbative approaches are inapplicable. For this reason the DMFT has now become the standard mean-field theory for fermionic correlation problems, including cold atoms in optical lattices [173–175]. The study

of models in nonequilibrium using an appropriate generalization of DMFT has become yet another fascinating new research area [148, 150–152, 155–168].

Until a few years ago research into correlated electron systems concentrated on homogeneous bulk systems. DMFT investigations of systems with internal or external inhomogeneities such as thin films and multi-layered nanostructures are still very new [174, 176–181]. They are particularly important in view of the novel types of functionalities of such systems, which may have important applications in electronic devices. Here the DMFT and its non-local extensions [100–102] will certainly become very useful.

In particular, the development of the ab initio band-structure calculation technique referred to as LDA + DMFT has proved to be a breakthrough in the investigation of electronically correlated materials. It had already provided important insights into the spectral and magnetic properties of correlated electron materials [49,52,94–97]. Clearly, this approach has a great potential for further developments. Indeed, it is not hard to foresee that the LDA + DMFT framework will eventually develop into a comprehensive ab initio approach which is able to describe, and even predict, the properties of complex correlated materials.

Acknowledgments We thank Jim Allen, Vladimir Anisimov, Nils Blümer, Ralf Bulla, Liviu Chioncel, Theo Costi, Vlad Dobrosavljević, Peter van Dongen, Martin Eckstein, Volker Eyert, Florian Gebhard, Karsten Held, Walter Hofstetter, Vaclav Janiš, Anna Kauch, Stefan Kehrein, Georg Keller, Gabi Kotliar, Jan Kuneš, Ivan Leonov, Walter Metzner, Michael Moeckel, Igor Nekrasov, Thomas Pruschke, Xinguo Ren, Shigemasa Suga, Götz Uhrig, Martin Ulmke, Ruud Vlaming, Philipp Werner, and Unjong Yu for valuable collaborations. Support by the Deutsche Forschungsgemeinschaft through TRR 80 and FOR 1346 is gratefully acknowledged. KB was also supported by the grant N N202 103138 of the Polish Ministry of Science and Education.

References

1. J.H. de Boer, E.J.W. Verwey, Proc. Phys. Soc. **49**, 59 (1937)
2. N.F. Mott, R. Peierls, Proc. Phys. Soc. A **49**, 72 (1937)
3. M. Imada, A. Fujimori, Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
4. M.C. Gutzwiller, Phys. Rev. Lett. **10**, 59 (1963)
5. J. Hubbard, Proc. Roy. Soc. Lond. **A276**, 238 (1963)
6. J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963)
7. E. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
8. D. Vollhardt, in *Lectures on the Physics of Strongly Correlated Systems XIV*, ed. by A. Avella, F. Mancini. AIP Conference Proceedings, vol. 1297 (American Institute of Physics, Melville, 2010), p. 339
9. D. Vollhardt, in *Correlated Electron Systems*, ed. by V.J. Emery (World Scientific, Singapore, 1993), p. 57
10. R.J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic, London, 1982)
11. C. Itzykson, J.M. Drouffe, *Statistical Field Theory* (Cambridge University Press, Cambridge, 1989)
12. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
13. U. Wolff, Nucl. Phys. B **225**, 391 (1983)
14. E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)

15. W. Metzner, Z. Phys. B **77**, 253 (1989)
16. E. Müller-Hartmann, Z. Phys. B **76**, 211 (1989)
17. H. Schweitzer, G. Czycholl, Solid State Comm. **69**, 171 (1989)
18. H. Schweitzer, G. Czycholl, Z. Phys. B **83**, 93 (1991)
19. U. Brandt, C. Mielsch, Z. Phys. B **75**, 365 (1989)
20. P.G.J. van Dongen, F. Gebhard, D. Vollhardt, Z. Phys. **76**, 199 (1989)
21. F. Kajzar, J. Friedel, J. Phys. (Paris) **39**, 397 (1978)
22. G. Treglia, F. Ducastelle, D. Spanjaard, Phys. Rev. B **22**, 6472 (1980)
23. G. Bulk, R.J. Jelitto, Phys. Rev. B **41**, 413 (1990)
24. J.M. Luttinger, J.C. Ward, Phys. Rev. **118**, 1417 (1960)
25. V. Janiš, Z. Phys. B **83**, 227 (1991)
26. V. Janiš, D. Vollhardt, Int. J. Mod. Phys. B **6**, 731 (1992)
27. B. Velický, S. Kirkpatrick, H. Ehrenreich, Phys. Rev. **175**, 745 (1968)
28. F. Yonezawa, K. Morigaki, Suppl. Prog. Theor. Phys. **53**, 1 (1973)
29. R.J. Elliot, J.A. Krumhansl, P.L. Leath, Rev. Mod. Phys. **46**, 465 (1974)
30. R. Vlaming, D. Vollhardt, Phys. Rev. B **45**, 4637 (1992)
31. A. Georges, G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
32. M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
33. A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
34. M. Fabrizio, in *Lectures on the Physics of Highly Correlated Electron Systems XI*, ed. by A. Avella, F. Mancini. AIP Conference Proceedings, vol. 918 (American Institute of Physics, Melville, 2007), p. 3
35. J.E. Hirsch, R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986)
36. J.K. Freericks, V. Zlatić, Rev. Mod. Phys. **75**, 1333 (2003)
37. R. Bulla, Phys. Rev. Lett. **83**, 136 (1999)
38. R. Bulla, T.A. Costi, T. Pruschke, Rev. Mod. Phys. **80**, 395 (2008)
39. M. Karski, C. Raas, G.S. Uhrig, Phys. Rev. B **77**, 075116 (2008)
40. M. Caffarel, W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
41. Q. Si, M.J. Rozenberg, G. Kotliar, A.E. Ruckenstein, Phys. Rev. Lett. **72**, 2761 (1994)
42. M.J. Rozenberg, G. Moeller, G. Kotliar, Mod. Phys. Lett. B **8**, 535 (1994)
43. M.J. Rozenberg, X.Y. Zhang, G. Kotliar, Phys. Rev. Lett. **69**, 1236 (1992)
44. A. Georges, W. Krauth, Phys. Rev. Lett. **69**, 1240 (1992)
45. A.N. Rubtsov, V.V. Savkin, A.I. Lichtenstein, Phys. Rev. B **72**, 035122 (2005)
46. P. Werner, A. Comanac, L. de' Medici, M. Troyer, A.J. Millis, Phys. Rev. Lett. **97**, 076405 (2006)
47. K. Haule, Phys. Rev. B **75**, 155113 (2007)
48. X.Y. Zhang, M.J. Rozenberg, G. Kotliar, Phys. Rev. Lett. **70**, 1666 (1993)
49. K. Held, I.A. Nekrasov, G. Keller, V. Eyert, N. Blümer, A.K. McMahan, R.T. Scalettar, T. Pruschke, V.I. Anisimov, D. Vollhardt, Psi-k Newsletter **56**, 65 (2003); Reprinted in Phys. Status Solidi B **243**, 2599 (2006)
50. N.E. Bickers, D.J. Scalapino, Ann. Phys. (NY) **193**, 206 (1989)
51. A.I. Lichtenstein, M.I. Katsnelson, J. Phys. Condens. Matter **11**, 1037 (1999)
52. G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006)
53. V. Drchal, V. Janiš, J. Kudrnovský, V.S. Oudovenko, X. Dai, K. Haule, G. Kotliar, J. Phys. Condens. Matter **17**, 61 (2005)
54. D.E. Logan, M.T. Glossop, J. Phys. Condens. Matter **12**, 985 (2000)
55. A. Kauch, K. Byczuk, ArXiv:0912.4278 (2009)
56. V. Janiš, P. Augustinský, Phys. Rev. B **75**, 165108 (2007)
57. T. Pruschke, M. Jarrell, J.K. Freericks, Adv. Phys. **44**, 187 (1995)
58. A. Georges, in *Lectures on the Physics of Highly Correlated Electron Systems VIII*, ed. by A. Avella, F. Mancini. AIP Conference Proceedings, vol. 715 (American Institute of Physics, Melville, 2004), p. 3
59. G. Kotliar, D. Vollhardt, Phys. Today **57**(3), 53 (2004)

60. N.F. Mott, Rev. Mod. Phys. **40**, 677 (1968)
61. N.F. Mott, *Metal-Insulator Transitions*, 2nd edn. (Taylor and Francis, London, 1990)
62. F. Gebhard, *The Mott Metal-Insulator Transition* (Springer, Berlin, 1997)
63. D.B. McWhan, J.P. Remeika, Phys. Rev. B **2**, 3734 (1970)
64. D.B. McWhan, A. Menth, J.P. Remeika, W.F. Brinkman, T.M. Rice, Phys. Rev. B **7**, 1920 (1973)
65. T.M. Rice, D.B. McWhan, IBM J. Res. Develop. **14**, 251 (1970)
66. J. Hubbard, Proc. Roy. Soc. Lond. **A281**, 401 (1964)
67. W.F. Brinkman, T.M. Rice, Phys. Rev. B **2**, 4302 (1970)
68. K. Byczuk, M. Kollar, K. Held, Y.F. Yang, I.A. Nekrasov, T. Pruschke, D. Vollhardt, Nat. Phys. **3**, 168 (2007)
69. A. Toschi, M. Capone, C. Castellani, K. Held, Phys. Rev. Lett. **102**, 076402 (2009)
70. C. Raas, P. Grete, G.S. Uhrig, Phys. Rev. Lett. **102**, 076406 (2009)
71. A. Lanzara, P.V. Bogdanov, X.J. Zhou, S.A. Kellar, D.L. Feng, E.D. Lu, T. Yoshida, H. Eisaki, A. Fujimori, K. Kishio, J.I. Shimoyama, T. Noda, S. Uchida, Z. Hussain, Z.X. Shen, Nature **412**, 510 (2001)
72. Z.X. Shen, A. Lanzara, S. Ishihara, N. Nagaosa, Philos. Mag. B **82**, 1349 (2002)
73. H. He, Y. Sidis, P. Bourges, G.D. Gu, A. Ivanov, N. Koshizuka, B. Liang, C.T. Lin, L.P. Regnault, E. Schoenherr, B. Keimer, Phys. Rev. Lett. **86**, 1610 (2001)
74. J. Hwang, T. Timusk, G.D. Gu, Nature **427**, 714 (2004)
75. A. Hofmann, X.Y. Cui, J. Schäfer, S. Meyer, P. Höpfner, C. Blumenstein, M. Paul, L. Patthey, E. Rotenberg, J. Bünemann, F. Gebhard, T. Ohm, W. Weber, R. Claessen, Phys. Rev. Lett. **102**, 187204 (2009)
76. R. Bulla, T.A. Costi, D. Vollhardt, Phys. Rev. B **64**, 045103 (2001)
77. S.K. Mo, H.D. Kim, J.W. Allen, G.H. Gweon, J.D. Denlinger, J.H. Park, A. Sekiyama, A. Yamasaki, S. Suga, P. Metcalf, K. Held, Phys. Rev. Lett. **93**, 076404 (2004)
78. M.J. Rozenberg, R. Chitra, G. Kotliar, Phys. Rev. Lett. **83**, 3498 (1999)
79. J. Joo, V. Oudovenko, Phys. Rev. B **64**, 193102 (2001)
80. N. Blümer, *Metal-Insulator Transition and Optical Conductivity in High Dimensions* (Shaker, Aachen, 2003)
81. T. Pruschke, Prog. Theor. Phys. Suppl. **160**, 274 (2005)
82. D. Vollhardt, P. Wölfle, *The Superfluid Phases of Helium 3* (Taylor and Francis, London, 1990)
83. H. Park, K. Haule, G. Kotliar, Phys. Rev. Lett. **101**, 186403 (2008)
84. P. Hohenberg, W. Kohn, Phys. Rev. **136B**, 864 (1964)
85. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965)
86. R.O. Jones, O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
87. V.I. Anisimov, J. Zaanen, O.K. Andersen, Phys. Rev. B **44**, 943 (1991)
88. V.I. Anisimov, F. Aryasetiawan, A.I. Lichtenstein, J. Phys. Cond. Matter **9**, 767 (1997)
89. V.I. Anisimov, A.I. Poteryaev, M.A. Korotin, A.O. Anokhin, G. Kotliar, J. Phys.: Cond. Matt. **9**, 7359 (1997)
90. A.I. Lichtenstein, M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)
91. I.A. Nekrasov, K. Held, N. Blümer, A.I. Poteryaev, V.I. Anisimov, D. Vollhardt, Eur. Phys. J. B **18**, 55 (2000)
92. K. Held, I.A. Nekrasov, N. Blümer, V.I. Anisimov, D. Vollhardt, Int. J. Mod. Phys. B **15**, 2611 (2001)
93. K. Held, I.A. Nekrasov, G. Keller, V. Eyert, N. Blümer, A.K. McMahan, R. Scalettar, T. Pruschke, V. Anisimov, D. Vollhardt, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, ed. by J. Grotendorst, D. Marks, A. Muramatsu. NIC Series Volume, vol. 10, (NIC Directors, Forschungszentrum Jülich, Berlin, 2002), p. 175
94. A.I. Lichtenstein, M.I. Katsnelson, G. Kotliar, in *Electron Correlations and Materials Properties*, ed. by A. Gonis, N. Kioussis, M. Ciftan (Kluwer, New York, 2002), p. 428
95. K. Held, Adv. Phys. **56**, 829 (2007)

96. M.I. Katsnelson, V.Y. Irkhin, L. Chioncel, A.I. Lichtenstein, R.A. de Groot, Rev. Mod. Phys. **80**, 315 (2008)
97. V. Anisimov, Y. Izyumov, *Electronic Structure of Correlated Materials*, Springer Series in Solid-State Sciences, vol. 163 (Springer, Berlin, 2010)
98. J. Kuneš, I. Leonov, M. Kollar, K. Byczuk, V.I. Anisimov, D. Vollhardt, Eur. Phys. J. Special Topics **180**, 5 (2010)
99. M. Jarrell, J.E. Gubernatis, Phys. Rep. **269**, 133 (1996)
100. M. Potthoff, Adv. Solid State Phys. **45**, 135 (2005)
101. T. Maier, M. Jarrell, T. Pruschke, M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
102. K. Held, A.A. Katanin, A. Toschi, Prog. Theor. Phys. Suppl. **176**, 117 (2008)
103. A. Fujimori, I. Hase, H. Namatame, Y. Fujishima, Y. Tokura, H. Eisaki, S. Uchida, K. Takegahara, F.M.F. de Groot, Phys. Rev. Lett. **69**, 1796 (1992)
104. A. Sekiyama, H. Fujiwara, S. Imada, S. Suga, H. Eisaki, S.I. Uchida, K. Takegahara, H. Harima, Y. Saitoh, I.A. Nekrasov, G. Keller, D.E. Kondakov, A.V. Kozhevnikov, T. Pruschke, K. Held, D. Vollhardt, V.I. Anisimov, Phys. Rev. Lett. **93**, 156402 (2004)
105. I.A. Nekrasov, G. Keller, D.E. Kondakov, A.V. Kozhevnikov, T. Pruschke, K. Held, D. Vollhardt, V.I. Anisimov, Phys. Rev. B **72**, 155106 (2005)
106. E. Pavarini, S. Biermann, A. Poteryaev, A.I. Lichtenstein, A. Georges, O.K. Andersen, Phys. Rev. Lett. **92**, 176403 (2004)
107. I.H. Inoue, I. Hase, Y. Aiura, A. Fujimori, K. Morikawa, T. Mizokawa, Y. Haruyama, T. Maruyama, Y. Nishihara, Physica C **235–240**, 1007 (1994)
108. P.A. Lee, T.V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985)
109. N.F. Mott, Proc. Phys. Soc. Lond., Sect. A **62**, 415 (1949)
110. P.W. Anderson, Phys. Rev. **109**, 1492 (1958)
111. V. Janiš, D. Vollhardt, Phys. Rev. B **46**, 15712 (1992)
112. V. Janiš, M. Ulmke, D. Vollhardt, Europhys. Lett. **24**, 287 (1993)
113. M. Ulmke, V. Janiš, D. Vollhardt, Phys. Rev. B **51**, 10411 (1995)
114. V. Dobrosavljević, G. Kotliar, Phys. Rev. B **50**, 1430 (1994)
115. M.C.O. Aguiar, V. Dobrosavljević, E. Abrahams, G. Kotliar, Phys. Rev. B **73**, 115117 (2006)
116. M. Potthoff, M. Balzer, Phys. Rev. B **75**, 125112 (2007)
117. V. Janiš, D. Vollhardt, Phys. Rev. B **46**, 15712 (1992)
118. V. Dobrosavljević, G. Kotliar, Phys. Rev. Lett. **78**, 3943 (1997)
119. V. Dobrosavljević, A.A. Pastor, B.K. Nikolić, Europhys. Lett. **62**, 76 (2003)
120. F.X. Bronold, A. Alvermann, H. Fehske, Phil. Mag. **84**, 637 (2004)
121. K. Byczuk, W. Hofstetter, D. Vollhardt, Phys. Rev. Lett. **94**, 056404 (2005)
122. K. Byczuk, W. Hofstetter, D. Vollhardt, Phys. Rev. Lett. **102**, 146403 (2009)
123. K. Byczuk, M. Ulmke, D. Vollhardt, Phys. Rev. Lett. **90**, 196403 (2003)
124. K. Byczuk, M. Ulmke, Eur. Phys. J. B **45**, 449 (2005)
125. K. Byczuk, W. Hofstetter, D. Vollhardt, Phys. Rev. B **69**, 045112 (2004)
126. P. Lombardo, R. Hayn, G.I. Japaridze, Phys. Rev. B **74**, 085116 (2006)
127. M. Potthoff, M. Balzer, Phys. Rev. B **75**, 125112 (2007)
128. U. Yu, K. Byczuk, D. Vollhardt, Phys. Rev. Lett. **100**, 246401 (2008)
129. R.J. Anglin, W. Ketterle, Nature **416**, 211 (2002)
130. M. Greiner, O. Mandel, T. Esslinger, T.W. Hänsch, I. Bloch, Nature **415**, 39 (2002)
131. M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sen, U. Sen, Adv. Phys. **56**, 243 (2007)
132. I. Bloch, J. Dalibard, W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008)
133. D. Jaksch, C. Bruder, J.I. Cirac, C.W. Gardiner, P. Zoller, Phys. Rev. Lett. **81**, 3108 (1998)
134. A. Isacsson, M.C. Cha, K. Sengupta, S.M. Girvin, Phys. Rev. B **72**, 184507 (2005)
135. S.D. Huber, E. Altman, H.P. Büchler, G. Blatter, Phys. Rev. B **75**, 085106 (2007)
136. A. Micheli, G.K. Brennen, P. Zoller, Nature **2**, 341 (2006)
137. M.P.A. Fisher, P.B. Weichman, G. Grinstein, D.S. Fisher, Phys. Rev. B **40**, 546 (1989)
138. K. Byczuk, D. Vollhardt, Phys. Rev. B **77**, 235106 (2008)
139. A. Hubener, M. Snoek, W. Hofstetter, Phys. Rev. B **80**, 245109 (2009)

140. W.J. Hu, N.H. Tong, Phys. Rev. B **80**, 245110 (2009)
141. P. Anders, E. Gull, L. Pollet, M. Troyer, P. Werner, Phys. Rev. Lett. **105**, 096402 (2010)
142. K. Byczuk, D. Vollhardt, Ann. Phys. (Berlin) **18**, 622 (2009)
143. J. Dziarmaga, Adv. Phys. **59**, 1063 (2010)
144. A. Polkovnikov, K. Sengupta, A. Silva, M. Vengalattore, Rev. Mod. Phys. **83**, 863 (2011)
145. L. Perfetti, P.A. Loukakos, M. Lisowski, U. Bovensiepen, H. Berger, S. Biermann, P.S. Cornaglia, A. Georges, M. Wolf, Phys. Rev. Lett. **97**, 067402 (2006)
146. S. Wall, D. Brida, S.R. Clark, H.P. Ehrke, D. Jaksch, A. Ardavan, S. Bonora, H. Uemura, Y. Takahashi, T. Hasegawa, H. Okamoto, G. Cerullo, A. Cavalleri, Nat. Phys. **7**, 114 (2011)
147. P. Schmidt, H. Monien, arXiv:cond-mat/0202046
148. V. Turkowski, J.K. Freericks, Phys. Rev. B **71**, 085104 (2005)
149. R. van Leeuwen, N.E. Dahlen, G. Stefanucci, C.O. Almbladh, U. von Barth, in *Time-Dependent Density Functional Theory*, ed. by M.A.L. Marques, C.A. Ullrich, F. Nogueira, A. Rubio, K. Burke, E.K.U. Gross. Lecture Notes in Physics, vol. 706 (Springer, Berlin, 2006)
150. M. Eckstein, M. Kollar, P. Werner, Phys. Rev. B **81**, 115131 (2010)
151. M. Eckstein, M. Kollar, P. Werner, Phys. Rev. Lett. **103**, 056403 (2009)
152. M. Eckstein, P. Werner, Phys. Rev. B **82**, 115115 (2010)
153. M. Eckstein, A. Hackl, S. Kehrein, M. Kollar, M. Moeckel, P. Werner, F.A. Wolf, Eur. Phys. J. Special Topics **180**, 217 (2010)
154. J.K. Freericks, H.R. Krishnamurthy, T. Pruschke, Phys. Rev. Lett. **102**, 136401 (2009)
155. M. Eckstein, M. Kollar, Phys. Rev. B **78**, 245113 (2008)
156. B. Moritz, T.P. Devereaux, J.K. Freericks, Phys. Rev. B **81**, 165112 (2010)
157. M. Eckstein, M. Kollar, Phys. Rev. B **78**, 205119 (2008)
158. J.K. Freericks, V.M. Turkowski, V. Zlatić, Phys. Rev. Lett. **97**, 266408 (2006)
159. M.T. Tran, Phys. Rev. B **78**, 125103 (2008)
160. J.K. Freericks, Phys. Rev. B **77**, 075109 (2008)
161. A.V. Joura, J.K. Freericks, T. Pruschke, Phys. Rev. Lett. **101**, 196401 (2008)
162. N. Tsuji, T. Oka, H. Aoki, Phys. Rev. B **78**, 235124 (2008)
163. N. Tsuji, T. Oka, H. Aoki, Phys. Rev. Lett. **103**, 047403 (2009)
164. M. Eckstein, T. Oka, P. Werner, Phys. Rev. Lett. **105**, 146404 (2010)
165. N. Tsuji, T. Oka, P. Werner, H. Aoki, Phys. Rev. Lett. **106**, 236401 (2011)
166. M. Eckstein, M. Kollar, Phys. Rev. Lett. **100**, 120404 (2008)
167. M. Eckstein, M. Kollar, New J. Phys. **12**, 055012 (2010)
168. N. Eurich, M. Eckstein, P. Werner, Phys. Rev. B **83**, 155122 (2011)
169. M. Rigol, V. Dunjko, M. Olshanii, Nature **452**, 854 (2008)
170. M. Moeckel, S. Kehrein, Phys. Rev. Lett. **100**, 175702 (2008)
171. M. Moeckel, S. Kehrein, Ann. Phys. **324**, 2146 (2009)
172. M. Greiner, O. Mandel, T.W. Hänsch, I. Bloch, Nature **419**, 51 (2002)
173. A. Rapp, G. Zarand, C. Honerkamp, W. Hofstetter, Phys. Rev. Lett. **98**, 160405 (2007)
174. M. Snoek, I. Titvinidze, C. Toke, K. Byczuk, W. Hofstetter, New J. Phys. **10**, 093008 (2008)
175. U. Schneider, L. Hackermüller, S. Will, T. Best, I. Bloch, T.A. Costi, R.W. Helmes, D. Rasch, A. Rosch, Science **322**, 1520 (2008)
176. J.K. Freericks, *Transport in Multilayered Nanostructures – The Dynamical Mean-Field Approach* (Imperial College Press, London, 2006)
177. M. Potthoff, W. Nolting, Phys. Rev. B **59**, 2549 (1999)
178. M. Takizawa, H. Wadati, K. Tanaka, M. Hashimoto, T. Yoshida, A. Fujimori, A. Chikamtsu, H. Kumigashira, M. Oshima, K. Shibuya, T. Mihara, T. Ohnishi, M. Lippmaa, M. Kawasaki, H. Koinuma, S. Okamoto, A.J. Millis, Phys. Rev. Lett. **97**, 057601 (2006)
179. L. Chen, J.K. Freericks, Phys. Rev. B **75**, 1251141 (2007)
180. K. Byczuk, in *Condensed Matter Physics in the Prime of the 21st Century: Phenomena, Materials, Ideas, Methods*, ed. by J. Jedrzejewski (World Scientific, Singapore, 2008), p. 1
181. R.W. Helmes, T.A. Costi, A. Rosch, Phys. Rev. Lett. **100**, 056403 (2008)

Chapter 8
Cluster Perturbation Theory

David Sénéchal

Abstract Cluster perturbation theory (CPT) is a simple approximation scheme that applies to lattice models with local interactions, like the Hubbard model, or models where the local interaction is predominant. It proceeds by tiling the lattice into identical, finite-size clusters, solving these clusters exactly and treating the inter-cluster hopping terms at first order in strong-coupling perturbation theory. This review will focus on the kinematical aspects of CPT, in particular the periodization procedure, and on the practical implementation of CPT using an exact diagonalization solver for the cluster. Applications of CPT will be briefly reviewed.

8.1 Introduction: CPT in a Nutshell

Cluster perturbation theory (CPT) is a simple approximation scheme that applies to lattice models with local interactions, like the Hubbard model, or models where the local interaction is predominant. It can be viewed as a cluster extension of strong-coupling perturbation theory [1], limited to lowest order [2]. Its features are found in more sophisticated approaches that rely on a tiling of the lattice into clusters, like the variational cluster approximation (VCA) and the cellular dynamical mean-field theory (C-DMFT).

CPT was first introduced in [3], and later, independently, in [4], where it was related to the strong-coupling perturbation theory and where a periodization scheme was introduced.

CPT has been used mostly to calculate the spectral function of strongly correlated models, where it has the advantage of allowing for a continuum of wavevectors (useful in particular in momentum-dependent plots) while at the same time

D. Sénéchal (✉)
Département de physique, Université de Sherbrooke, Sherbrooke, QC, Canada J1K 2R1
e-mail: david.senechal@usherbrooke.ca

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_8,

capturing short-range dynamical correlations. Most notable is its application to the understanding of the pseudogap phenomenon in the normal state of high-T_c materials[20], but it has been applied to a variety of situations (see Sect. 8.8).

Let us establish some notation by writing the lattice model as the sum of a one-body part H_0 and an interaction part H_1 ($H = H_0 + H_1$), where

$$H_0 = \sum_{\alpha,\beta} t_{\alpha\beta} c_\alpha^\dagger c_\beta \qquad H_1 = \sum_i H_{1,i}\,. \tag{8.1}$$

The greek indices α, β label distinct one-electron states in a Wannier basis, and correspond to a aggregation of site, spin and possibly band indices: we write $\alpha = (i, \sigma)$, where i labels a site on the Bravais lattice γ and σ represents spin and band indices. N_σ is the number of states associated with each lattice site (for instance, $N_\sigma = 2$ for the one-band Hubbard model). H_1 is assumed to be a sum of local interactions, each confined to a lattice site $\boldsymbol{r}_i \in \gamma$. We will use boldface symbols for matrices acting in the space of one-electron states: $\boldsymbol{t}$ will stand for the matrix with elements $t_{\alpha\beta}$, and likewise for the Green function $\boldsymbol{G}(\omega)$, and so on.

CPT, like other quantum cluster methods, proceeds by dividing the lattice γ into a superlattice Γ of identical clusters of L sites each. The lattice Hamiltonian H is written as $H = H' + H_V$, where H' is the cluster Hamiltonian, obtained by severing the hopping terms between different clusters, which are put into H_V:

$$H' = \sum_{\alpha,\beta} t'_{\alpha\beta} c_\alpha^\dagger c_\beta + H_U \qquad H_V = \sum_{\alpha,\beta} V_{\alpha\beta} c_\alpha^\dagger c_\beta\,, \tag{8.2}$$

where $\boldsymbol{t}'$ contains the one-body terms that are confined to a given cluster, and $\boldsymbol{V}$ the inter-cluster terms.[1] H_V is treated as a perturbation of the cluster part H'. In what follows, primes ($'$) will be used to distinguish quantities associated with the cluster from the corresponding quantities associated with the full lattice.

CPT is an approximation scheme for the one-electron Green function $\boldsymbol{G}(\omega)$. It can be viewed as an application of the strong-coupling perturbation theory developed by Pairault et al. [1, 5], in which the one-body part H_0 was considered a perturbation of the interaction part H_1. But the only true requirement of that approach is that the perturbation be a one-body operator. Thus, strong-coupling perturbation theory can be applied to the case where H_V is a perturbation over H'; this was carried out to lowest order in H_V in [2, 4]. The result for the one-electron Green function $\boldsymbol{G}(\omega)$ in terms of the unperturbed Green function $\boldsymbol{G}'(\omega)$ and the perturbation $\boldsymbol{V}$ is

$$\boldsymbol{G}^{-1}(\omega) = \boldsymbol{G}'^{-1}(\omega) - \boldsymbol{V}\,. \tag{8.3}$$

[1] In the VCA, $\boldsymbol{V}$ is the difference between the lattice Hamiltonian H and the reference Hamiltonian H', and as such may also contain intra-cluster terms.

A site index i can be mapped to a pair (m, a), where m labels clusters within the superlattice and a labels sites within a given cluster. The cluster Green function $\boldsymbol{G}'$ is diagonal in m, whereas $\boldsymbol{V}$ is essentially off-diagonal in m. The CPT approximation explicitly breaks translation invariance on the original lattice γ, down to translation invariance on the superlattice $\Gamma \in \gamma$ only. The residual translation invariance makes it more convenient to use a basis of one-electron states labelled by a wavevector $\tilde{\boldsymbol{k}}$ in the Brillouin zone of the superlattice, instead of the cluster index m (this is explain at length in Sect. 8.2). We will commonly use the basis of one-electron states labelled by $(\tilde{\boldsymbol{k}}, a, \sigma)$. This is a mixed Fourier-position space representation, in which the Fourier transform has been carried out at large distances only. The CPT Green function (8.3) is diagonal in $\tilde{\boldsymbol{k}}$, but not in a (cluster indices):

$$\boldsymbol{G}^{-1}(\tilde{\boldsymbol{k}}, \omega) = \boldsymbol{G}'^{-1}(\omega) - \boldsymbol{V}(\tilde{\boldsymbol{k}}). \tag{8.4}$$

The matrices appearing in the above formula are now of the order LN_σ (the number of one-electron states in the cluster), i.e. they are matrices with indices (a, σ) only. $\boldsymbol{G}'$ is independent of $\tilde{\boldsymbol{k}}$, whereas $\boldsymbol{V}$ is frequency independent.

The basic CPT relation (8.4) may also be expressed in terms of the self-energy $\boldsymbol{\Sigma}$ of the cluster Hamiltonian:

$$\boldsymbol{G}^{-1}(\tilde{\boldsymbol{k}}, \omega) = \boldsymbol{G}_0{}^{-1}(\tilde{\boldsymbol{k}}, \omega) - \boldsymbol{\Sigma}(\omega), \tag{8.5}$$

where $\boldsymbol{G}_0(\tilde{\boldsymbol{k}}, \omega)$ is the Green function derived from the non-interacting part H_0 of the lattice Hamiltonian. This follows simply from the relations:

$$\boldsymbol{G}'^{-1} = \omega - \boldsymbol{t}' - \boldsymbol{\Sigma} \tag{8.6}$$

$$\boldsymbol{G}_0{}^{-1} = \omega - \boldsymbol{t}' - \boldsymbol{V}, \tag{8.7}$$

where $\boldsymbol{t}'$ is the restriction to the cluster of the hopping matrix (chemical potential included). It is in the form (8.5) that CPT was first proposed [3].

Since the form of the inter-hopping term $\boldsymbol{V}(\tilde{\boldsymbol{k}})$ is known analytically, the CPT Green function (8.4) is readily computed once the cluster Green function $\boldsymbol{G}'^{-1}(\omega)$ is known. The latter must be computed numerically, for instance by an exact diagonalization of the cluster Hamiltonian H'. It will be shown in Sect. 8.3 that the CPT Green function has the same analytical properties as the exact Green function: its poles are located on the real axis and its residues obey the basic sum rules. The CPT Green function can be used to compute averages of one-body operators and some thermodynamic quantities. After a periodization procedure (see Sect. 8.5), it provides an approximation to the spectral function, useful in particular when comparing with ARPES experiments. The CPT Green function pervades the VCA (see Chap. 9.5) and the C-DMFT (see Chap. 10.7).

CPT has the following characteristics:

1. Although it may be derived using the strong-coupling perturbation theory, it is exact in the non-interacting limit, since the self-energy disappears in that case.
2. It is also exact in the strong-coupling limit $t_{ij} \to 0$.
3. It provides an approximate lattice Green function for a continuum of wavevectors. Hence its usefulness in comparing with ARPES data. Even though CPT does not have the self-consistency of DMFT approaches, at fixed computing resources it allows for the best momentum resolution (likewise for the VCA).
4. Although formulated as a lowest-order result of the strong-coupling perturbation theory, it is not controlled by including higher-order terms in that perturbation expansion – this would be extremely difficult – but rather by increasing the cluster size.
5. It cannot describe broken-symmetry states, since it is based on the solution of a single cluster that is too small to support spontaneously broken symmetries. Broken-symmetry states can be described by the VCA and the C-DMFT, which can both be viewed as extensions or refinements of the CPT (alternately, one could view the CPT as a special case of VCA without variational parameters).
6. Clusters in CPT have open boundary conditions. When periodic boundary conditions are used by adding the missing hopping terms in H' while subtracting them in H_V, the results obtained are unsatisfactory (see Fig. 5 of [2]). This is also shown by treating the missing hopping terms as variational parameters [6]. However, a periodic CPT (or PCPT) has been developed [7] that connects with the dynamic cluster approximation (DCA), itself based on periodic clusters.

This review of CPT will proceed as follows: Section 8.2 will introduce the notation and the various Fourier transforms used to relate the different bases of one-body states used in CPT. Section 8.3 will describe the Lehmann representation of the cluster Green function $\boldsymbol{G}'$ and of the CPT Green function $\boldsymbol{G}$. Section 8.4 will mostly review the exact diagonalization solver used in most applications of CPT. Section 8.5 will describe various schemes used to recover a wavevector-resolved Green function out of CPT, particularly useful when comparing with ARPES data. Section 8.6 will explain how to calculate various quantities of interest from the CPT Green function. Section 8.7 will show some results obtained on the Hubbard model, and Sect. 8.8 will review applications of CPT to other models.

8.2 Cluster Kinematics

Cluster methods are based on a cluster decomposition of the model, i.e. on a tiling of the original lattice γ with identical clusters of L sites each. Mathematically, this corresponds to introducing a superlattice Γ, whose sites, labelled by latin indices (e.g., $m, n, \dots$), form a subset of the lattice γ. This superlattice is generated by D basis vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_D$ belonging to γ: every site $\boldsymbol{r}_m$ of the superlattice may be expressed as an integer linear combination of these basis vectors. Associated

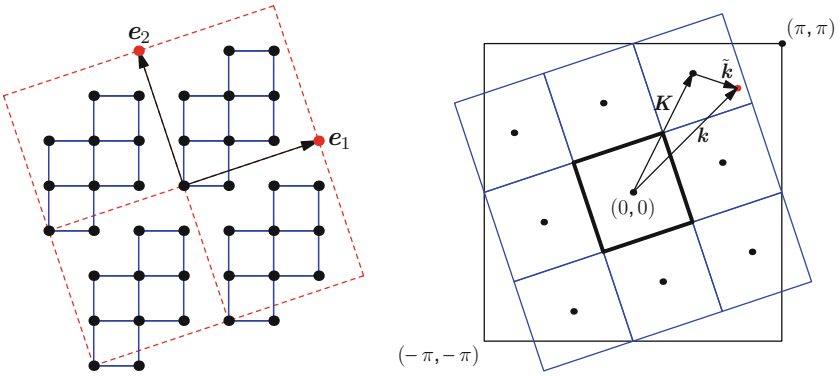

Fig. 8.1 *Left panel*: Tiling of the square lattice with 10-site clusters and superlattice vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. *Right panel*: A wavevector $\boldsymbol{k}$ has a unique decomposition $\boldsymbol{k} = \tilde{\boldsymbol{k}} + \boldsymbol{K}$, where $\boldsymbol{K}$ is one of the L elements of the reciprocal superlattice that belongs to the original Brillouin zone BZ_γ

with each site of Γ is a cluster of L sites, whose shape is not uniquely determined by the superlattice structure (for instance, see Fig. 8.1). The cluster sites will be labelled by latin indices $(a, b, \dots)$. Each site $\boldsymbol{r}_i$ of the original lattice γ can be expressed in a unique way as a combination of a superlattice vector $\boldsymbol{r}_m$ and of a site $\boldsymbol{r}_a$ within the cluster: $\boldsymbol{r}_i = \boldsymbol{r}_m + \boldsymbol{r}_a$.[2] This means that the index i may in fact be replaced by a compound index (m, a), and we have the following equivalence between summations:

$$\sum_i \quad \longleftrightarrow \quad \sum_{m \in \Gamma} \sum_a . \tag{8.8}$$

The number of sites in the cluster is simply the ratio of the unit cell volumes of the two lattices. In $D = 3$, this is

$$L = \frac{V_\Gamma}{V_\gamma} = |(\boldsymbol{e}_1 \wedge \boldsymbol{e}_2) \cdot \boldsymbol{e}_3| \tag{8.9}$$

(the above formulae can be adapted to $D = 2$ by setting $\boldsymbol{e}_3 = (0, 0, 1)$).

The Brillouin zone of the original lattice, denoted BZ_γ, contains L points belonging to the reciprocal superlattice Γ^*. The Brillouin zone of the superlattice, BZ_Γ, also called the *reduced Brillouin zone*, has a volume L times smaller than that of the original Brillouin zone. Any wavevector $\boldsymbol{k}$ of the original Brillouin zone can be uniquely expressed as

$$\boldsymbol{k} = \boldsymbol{K} + \tilde{\boldsymbol{k}}, \tag{8.10}$$

[2]For simplicity, we will suppress the spin and band indices σ in this section, but the whole discussion is trivially generalized to the case where there are many electron states per lattice site.

where $\boldsymbol{K}$ belongs both to the reciprocal superlattice *and* to BZ_γ, and $\tilde{\boldsymbol{k}}$ belongs to BZ_Γ (see Fig. 8.1). Thus, we have the equivalent summations

$$\sum_{\boldsymbol{k}} \quad \longleftrightarrow \quad \sum_{\tilde{\boldsymbol{k}}} \sum_{\boldsymbol{K}} . \tag{8.11}$$

Going between momentum space and real space, by discrete Fourier transforms, can be done either directly ($i \leftrightarrow \boldsymbol{k}$) or independently for cluster and superlattice indices ($m \leftrightarrow \tilde{\boldsymbol{k}}$ and $a \leftrightarrow \boldsymbol{K}$):

$$\begin{aligned}
f_j &= \frac{1}{\sqrt{N}} \sum_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j} f(\boldsymbol{k}) & f(\boldsymbol{k}) &= \frac{1}{\sqrt{N}} \sum_{j} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j} f_j \\
f_m &= \sqrt{\frac{L}{N}} \sum_{\tilde{\boldsymbol{k}}} \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{k}}\cdot\boldsymbol{r}_m} f(\tilde{\boldsymbol{k}}) & f(\tilde{\boldsymbol{k}}) &= \sqrt{\frac{L}{N}} \sum_{m} \mathrm{e}^{-\mathrm{i}\tilde{\boldsymbol{k}}\cdot\boldsymbol{r}_m} f_m, \\
f_a &= \frac{1}{\sqrt{L}} \sum_{\boldsymbol{K}} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_a} f_{\boldsymbol{K}} & f_{\boldsymbol{K}} &= \frac{1}{\sqrt{L}} \sum_{a} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_a} f_a
\end{aligned} \tag{8.12}$$

where f stands for a generic one-index quantity, and N is the total number of sites in the lattice γ, which we suppose to be very large (the lattices γ and Γ being then formally periodic).[3] These discrete Fourier transforms close by virtue of the following identities

$$\frac{1}{N} \sum_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j} = \delta_{\boldsymbol{r}_j} \qquad \frac{1}{N} \sum_{j} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j} = \Delta_\gamma(\boldsymbol{k}) \tag{8.13}$$

$$\frac{L}{N} \sum_{\tilde{\boldsymbol{k}}} \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{k}}\cdot\boldsymbol{r}_m} = \delta_{\boldsymbol{r}_m} \qquad \frac{L}{N} \sum_{m} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_m} = \Delta_\Gamma(\tilde{\boldsymbol{k}}) \tag{8.14}$$

$$\frac{1}{L} \sum_{\boldsymbol{K}} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_a} = \delta_{\boldsymbol{r}_a} \qquad \frac{1}{L} \sum_{a} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_a} = \Delta_\gamma(\boldsymbol{K}), \tag{8.15}$$

where $\delta_{\boldsymbol{r}}$ is a Kronecker delta ($= 1$ if and only if $\boldsymbol{r} = 0$ within a discrete set), and the Δ's are the so-called Laue functions:

$$\Delta_\gamma(\boldsymbol{k}) = \sum_{\boldsymbol{Q}\in\gamma^*} \delta_{k+\boldsymbol{Q}} \qquad \Delta_\Gamma(\tilde{\boldsymbol{k}}) = \sum_{\boldsymbol{P}\in\Gamma^*} \delta_{\tilde{\boldsymbol{k}}+\boldsymbol{P}}. \tag{8.16}$$

[3] Dependence on quasi-continuous indices, like $\boldsymbol{k}$ and $\tilde{\boldsymbol{k}}$, will be indicated by parentheses instead of subscripts. This notation may rightfully be deemed capricious, since the labels i and $\boldsymbol{k}$ take the same number N of values, but we adopt it nonetheless as it helps reminding us that the values of the labels are closely separated.

Laue functions are used instead of Kronecker deltas in momentum space because of the possibility of Umklapp processes. Note especially that even though

$$\delta_{\boldsymbol{k}} = \delta_{\tilde{\boldsymbol{k}}}\delta_{\boldsymbol{K}} \qquad (\boldsymbol{k} = \tilde{\boldsymbol{k}} + \boldsymbol{K}), \tag{8.17}$$

the same does not hold for the Laue functions:

$$\Delta_\gamma(\boldsymbol{k}) \neq \Delta_\Gamma(\tilde{\boldsymbol{k}})\Delta_\gamma(\boldsymbol{K}). \tag{8.18}$$

Instead we have the following relations:

$$\Delta_\Gamma(\tilde{\boldsymbol{k}}) = \sum_{\boldsymbol{K}} \Delta_\gamma(\tilde{\boldsymbol{k}} + \boldsymbol{K}) \tag{8.19}$$

$$\Delta_\gamma(\boldsymbol{k}) = \Delta_\gamma(\tilde{\boldsymbol{k}} + \boldsymbol{K}) = \delta_{\tilde{\boldsymbol{k}}}\Delta_\gamma(\boldsymbol{K}) \tag{8.20}$$

which reflect the arbitrariness in the choice of Brillouin zone of the superlattice.[4]

A one-index quantity like the destruction operator $c_j = c_{m,a}$ can be represented in a variety of ways, through partial Fourier transforms:

$$c_a(\tilde{\boldsymbol{k}}) = \sqrt{\frac{L}{N}} \sum_m \mathrm{e}^{-\mathrm{i}\tilde{\boldsymbol{k}}\cdot\boldsymbol{r}_m} c_{m,a} \tag{8.21}$$

$$c_{m,\boldsymbol{K}} = \frac{1}{\sqrt{L}} \sum_a \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_a} c_{m,a} \tag{8.22}$$

$$c_{\boldsymbol{K}}(\tilde{\boldsymbol{k}}) = \frac{1}{\sqrt{N}} \sum_{m,a} \mathrm{e}^{-\mathrm{i}(\tilde{\boldsymbol{k}}\cdot\boldsymbol{r}_m + \boldsymbol{K}\cdot\boldsymbol{r}_a)} c_{m,a} \tag{8.23}$$

$$c(\boldsymbol{k}) = \frac{1}{\sqrt{N}} \sum_i \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i} c_i \tag{8.24}$$

The last two representations are not identical, since the phases in the two cases differ by $\tilde{\boldsymbol{k}} \cdot \boldsymbol{r}_a$. They are related by a L-dimensional unitary matrix:

$$c(\tilde{\boldsymbol{k}} + \boldsymbol{K}) = \sum_{\boldsymbol{K}'} \Lambda_{\boldsymbol{K}\boldsymbol{K}'}(\tilde{\boldsymbol{k}}) c_{\boldsymbol{K}'}(\tilde{\boldsymbol{k}}), \tag{8.25}$$

where

$$\Lambda_{\boldsymbol{K}\boldsymbol{K}'}(\tilde{\boldsymbol{k}}) = \frac{1}{L} \sum_a \mathrm{e}^{-\mathrm{i}\boldsymbol{r}_a\cdot(\tilde{\boldsymbol{k}} + \boldsymbol{K} - \boldsymbol{K}')}. \tag{8.26}$$

[4] We use the term *Brillouin zone* in a rather liberal manner, as a complete and irreducible set of wavevectors, and not as the Wigner–Seitz cell of the reciprocal lattice.

A two-index quantity like the hopping matrix t_{ij} may thus have a number of different representations. Due to translation invariance on the lattice, this matrix is diagonal when expressed in momentum space: $t(\boldsymbol{k},\boldsymbol{k}') = \varepsilon(\boldsymbol{k})\delta_{\boldsymbol{k},\boldsymbol{k}'}$, $\varepsilon_{\boldsymbol{k}}$ being the dispersion relation:

$$t_{ij} = \frac{1}{N}\sum_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)}\varepsilon_{\boldsymbol{k}}. \tag{8.27}$$

However, we will generally use the mixed representation:

$$t_{ab}(\tilde{\boldsymbol{k}}) = \sum_m \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{k}}\cdot\boldsymbol{r}_m} t_{ij} \qquad \begin{matrix} i = (0,a) \\ j = (m,b) \end{matrix}. \tag{8.28}$$

For instance, if we tile the one-dimensional lattice with clusters of length $L = 2$, the nearest-neighbor hopping matrix, corresponding to the dispersion relation $\varepsilon(k) = -2t\cos(k)$, has the following mixed representation:

$$\boldsymbol{t}(\tilde{\boldsymbol{k}}) = -t\begin{pmatrix} 0 & 1+\mathrm{e}^{-2\mathrm{i}\tilde{k}} \\ 1+\mathrm{e}^{2\mathrm{i}\tilde{k}} & 0 \end{pmatrix} \tag{8.29}$$

8.3 Lehmann Representation of the Green Function

The zero-temperature cluster Green function $G'_{ab}(\omega)$ has the following expression, as a function of the complex-valued frequency ω:

$$G'_{ab}(\omega) = G'_{ab,e}(\omega) + G'_{ab,h}(\omega) \tag{8.30}$$

$$G'_{ab,e}(\omega) = \langle\Omega|c_a \frac{1}{\omega - H + E_0} c_b^\dagger|\Omega\rangle \tag{8.31}$$

$$G'_{ab,h}(\omega) = \langle\Omega|c_b^\dagger \frac{1}{\omega + H - E_0} c_a|\Omega\rangle, \tag{8.32}$$

where $|\Omega\rangle$ is the ground state of H'. The above expression is used as the starting point of the Green function's numerical computation (see Sects. 8.4.3 and 8.4.4).

By inserting completeness relations in (8.30), one finds the Lehmann representation:

$$G'_{ab}(\omega) = \sum_m \frac{\langle\Omega|c_a|m\rangle\langle m|c_b^\dagger|\Omega\rangle}{\omega - E_m + E_0} + \sum_n \frac{\langle\Omega|c_b^\dagger|n\rangle\langle n|c_a|\Omega\rangle}{\omega + E_n - E_0}. \tag{8.33}$$

The two sums are over different sets of eigenstates, in the spaces with one more and one less electron, respectively. Let us introduce the notation

$$Q^{(e)}_{am} = \langle \Omega | c_a | m \rangle \qquad Q^{(h)}_{an} = \langle \Omega | c^\dagger_a | n \rangle \tag{8.34}$$

as well as $\omega^{(e)}_m = E_m - E_0 > 0$ and $\omega^{(h)}_n = -E_n + E_0 < 0$ to write

$$G'_{ab}(\omega) = \sum_m \frac{Q^{(e)}_{am} Q^{(e)*}_{bm}}{\omega - \omega^{(e)}_m} + \sum_n \frac{Q^{(h)}_{an} Q^{(h)*}_{bn}}{\omega - \omega^{(h)}_n}. \tag{8.35}$$

The $Q^{(e)}_{am}$ forms a $L \times N^{(e)}$ matrix, where $N^{(e)}$ is the number of states $|m\rangle$ that give a non-zero contribution to the first sum above. Likewise, the $Q^{(h)}_{am}$ forms a $L \times N^{(h)}$ matrix. Let $M = N^{(e)} + N^{(h)}$ and let us introduce a $L \times M$ matrix $\boldsymbol{Q}$ by joining vertically the matrix $\boldsymbol{Q}^{(h)}$ below the matrix $\boldsymbol{Q}^{(e)}$, and let ω_r denote the elements of the concatenated sets $\{\omega^{(e)}_m\}$ and $\{\omega^{(h)}_n\}$. Then we can write

$$G'_{ab}(\omega) = \sum_r \frac{Q_{ar} Q^*_{br}}{\omega - \omega_r}. \tag{8.36}$$

If we introduce the diagonal matrix $\Lambda_{rs} = \delta_{rs}\omega_r$ and

$$\boldsymbol{g}(\omega) = \frac{1}{\omega - \boldsymbol{\Lambda}}, \tag{8.37}$$

then we have the matrix expression

$$\boldsymbol{G}(\omega) = \boldsymbol{Q}\boldsymbol{g}(\omega)\boldsymbol{Q}^\dagger. \tag{8.38}$$

This is a very general representation of the exact cluster Green function. If we restore spin and band indices, the matrix $\boldsymbol{Q}$ becomes $N_\sigma L \times M$ instead of $L \times M$.

The band Lanczos method (Sect. 8.4.4) will provide a truncated Lehmann representation of the Green function, in which the number M of columns of the matrix $\boldsymbol{Q}$ is small, but with essentially the same properties as the exact matrix.

8.3.1 The Lehmann Representation and the CPT Green Function

The CPT Green function (8.4) also has a Lehmann representation of the form (8.38). Following [8], we write

$$\begin{aligned} \boldsymbol{G}(\tilde{\boldsymbol{k}}, \omega) &= \frac{1}{(\boldsymbol{Q}\boldsymbol{g}(\omega)\boldsymbol{Q}^\dagger)^{-1} - \boldsymbol{V}(\tilde{\boldsymbol{k}})} \\ &= \boldsymbol{Q}\boldsymbol{g}(\omega)\boldsymbol{Q}^\dagger + (\boldsymbol{Q}\boldsymbol{g}(\omega)\boldsymbol{Q}^\dagger)\boldsymbol{V}(\boldsymbol{Q}\boldsymbol{g}(\omega)\boldsymbol{Q}^\dagger) + \cdots \end{aligned}$$

$$= \boldsymbol{Q}\Big(\boldsymbol{g}(\omega) + \boldsymbol{g}(\omega)(\boldsymbol{Q}^\dagger \boldsymbol{V}\, \boldsymbol{Q})\boldsymbol{g}(\omega) + \cdots\Big)\boldsymbol{Q}^\dagger$$

$$= \boldsymbol{Q}\frac{1}{\omega - \boldsymbol{L}(\tilde{\boldsymbol{k}})}\boldsymbol{Q}^\dagger, \tag{8.39}$$

where $\boldsymbol{L}(\tilde{\boldsymbol{k}}) = \boldsymbol{\Lambda} + \boldsymbol{Q}^\dagger \boldsymbol{V}(\tilde{\boldsymbol{k}})\boldsymbol{Q}$. The poles of $\boldsymbol{G}(\tilde{\boldsymbol{k}}, \omega)$ are those of $[\omega - \boldsymbol{L}(\tilde{\boldsymbol{k}})]^{-1}$, which we denote as $\omega_r(\tilde{\boldsymbol{k}})$. They are simply the eigenvalue of the $N \times N$ matrix $\boldsymbol{L}(\tilde{\boldsymbol{k}})$.

Let $\boldsymbol{U}(\tilde{\boldsymbol{k}})$ be the matrix that diagonalizes $\boldsymbol{L}(\tilde{\boldsymbol{k}})$:

$$\boldsymbol{U}(\tilde{\boldsymbol{k}})\boldsymbol{L}(\tilde{\boldsymbol{k}})\boldsymbol{U}^\dagger(\tilde{\boldsymbol{k}}) = \tilde{\boldsymbol{\Lambda}}(\tilde{\boldsymbol{k}}). \tag{8.40}$$

Then

$$\boldsymbol{G}(\tilde{\boldsymbol{k}}, \omega) = \tilde{\boldsymbol{Q}}(\tilde{\boldsymbol{k}})\frac{1}{\omega - \tilde{\boldsymbol{\Lambda}}(\tilde{\boldsymbol{k}})}\tilde{\boldsymbol{Q}}^\dagger(\tilde{\boldsymbol{k}}), \tag{8.41}$$

where $\tilde{\boldsymbol{Q}}(\tilde{\boldsymbol{k}}) = \boldsymbol{Q}\boldsymbol{U}(\tilde{\boldsymbol{k}})$. This has the same form as (8.36) and constitutes a Lehmann representation for the CPT Green function.

The representations (8.36) or (8.41) ensure the positivity of the cluster Green function and the CPT Green function, respectively, i.e. the positive character of the corresponding spectral functions. Indeed, the local (cluster) spectral weight is

$$A_a(\omega) = -2 \lim_{\eta \to 0} \operatorname{Im} G'_{aa}(\omega + i\eta) \tag{8.42}$$

and

$$G'_{aa}(\omega) = \sum_r \frac{|Q_{ar}|^2}{\omega - \omega_r}. \tag{8.43}$$

This expression has poles on the real axis only with positive residues, and this guarantees that the corresponding spectral function $A_a(\omega)$ is positive. Moreover, the properties $\boldsymbol{Q}\boldsymbol{Q}^\dagger = \mathbf{1}$ and $\tilde{\boldsymbol{Q}}(\tilde{\boldsymbol{k}})\tilde{\boldsymbol{Q}}^\dagger(\tilde{\boldsymbol{k}}) = \mathbf{1}$ ensure that the spectral functions obtained from the cluster or the CPT Green functions are normalized.

An essential characteristic of CPT emerges from this Lehmann representation: the poles of the CPT Green function, given by the diagonal matrix $\tilde{\boldsymbol{\Lambda}}(\tilde{\boldsymbol{k}})$, disperse as a function of reduced wavevector.

8.4 The Impurity Solver

Applying the CPT, or any other quantum cluster approach, requires the ability to compute the Green function of a single, finite cluster. The procedure used to do so is commonly called the *impurity solver*, an expression borrowed from DMFT methods, in which the cluster is formally treated like an impurity embedded in an

effective medium. Since CPT is often used to obtain spectral information in terms of real frequencies, the most common impurity solver used with CPT is the exact diagonalization (ED) solver, which has the advantage of providing real-frequency information. Of course, the ED solver has the disadvantage of being limited to small clusters (16 orbitals or less) given currently available computational resources. In this section we will review the exact diagonalization method, based on the Lanczos algorithm.

The first step of the exact diagonalization procedure is the choice of a basis of state vectors, and the construction of the Hamiltonian in this basis – this may or may not involve keeping the Hamiltonian matrix in memory. Then the ground state is found in a quasi-exact way by an iterative method such as the Lanczos algorithm. The Green function is thereafter calculated by similar means to be described below. The main difficulty in executing the method is the large memory required, which grows exponentially with the number of degrees of freedom. The main implementation difficulty is to optimize the method by coding states belonging to a given irreducible representation of the cluster's point group.

To simplify the discussion, we will focus on the one-band Hubbard model.

8.4.1 Coding of the Basis States

The first step in the exact diagonalization procedure is to define a coding scheme for the quantum basis states. Let us first describe how this is done without taking point group symmetries into account. A basis state may be specified by the occupation number n_α ($= 0$ or 1) of electrons in the orbital labelled α and has the following expression in terms of creation operators:

$$(c^\dagger_{1\uparrow})^{n_{1\uparrow}} \cdots (c^\dagger_{L\uparrow})^{n_{L\uparrow}} (c^\dagger_{1\downarrow})^{n_{1\downarrow}} \cdots (c^\dagger_{L\downarrow})^{n_{L\downarrow}} |0\rangle, \tag{8.44}$$

where the order in which the creation operators are applied is important, but a matter of convention. If the number of orbitals is smaller than or equal to 32, the string of n_α's forms the binary representation of a 32-bit unsigned integer b, which can be split into spin up and spin down parts:

$$b = b_\uparrow + 2^L b_\downarrow. \tag{8.45}$$

There are 2^{2L} such states, but not all are relevant, since the Hubbard Hamiltonian is block-diagonal : The number of electrons of a given spin ($N_\uparrow$ and $N_\downarrow$) is conserved and commutes with the Hamiltonian H'. Therefore, the exact diagonalization is to be performed in a sector (i.e., a subspace) of the total Hilbert space with fixed values of $N_\uparrow$ and $N_\downarrow$. This space has the tensor product structure

$$V = V_{N_\uparrow} \otimes V_{N_\downarrow} \tag{8.46}$$

and has a dimension $d = d(N_\uparrow)d(N_\downarrow)$, where

$$d(N_\sigma) = \frac{L!}{N_\sigma!(L-N_\sigma)!} \tag{8.47}$$

is the number of ways to distribute N_σ electrons among L sites.

Note that the ground state $|\Omega\rangle$ of the Hamiltonian generally belongs to the sector $N_\uparrow = N_\downarrow$. For a half-filled, zero-spin system ($N_\uparrow = N_\downarrow = L/2$), this translates into $d = (L!/(L/2)!^2)^2$, which behaves like $4^L/L$ for large L: The size of the eigenproblem grows exponentially with system size. In contrast, the non-interacting problem can be solved only by concentrating on one-electron states. For this reason, the exact diagonalization of the Hubbard Hamiltonian is restricted to systems of the order of 16 sites or less.

In practice, a generic state vector is represented by a d-component array of double precision numbers. In order to apply or construct the Hamiltonian acting on such vectors, we need a way to translate the label of a basis state (an integer i from 0 to $d-1$), into the binary representation (8.44). A generic and safe way to accomplish this is to build a two-way look-up table that tabulates the correspondence between consecutive integer labels and the binary representation of the basis state (for instance, by using the STL `map` container).

The next step is to construct the Hamiltonian matrix. The particular structure of the Hubbard model Hamiltonian brings a considerable simplification in the simple case studied here. Indeed, the Hamiltonian has the form

$$H' = K_\uparrow \otimes 1 + 1 \otimes K_\downarrow + V_{\text{int}}, \tag{8.48}$$

where $K_\uparrow$ only acts on up electrons and $K_\downarrow$ on down electrons, and where the Coulomb repulsion term V_{int} is diagonal in the occupation number basis. Thus, storing the Hamiltonian in memory is not a problem: the diagonal V_{int} is stored (an array of size d), and the kinetic energy K_σ (a matrix having a small fraction of d_σ^2 elements) is stored in sparse form. Constructing this matrix, formally expressed as

$$K = \sum_{a,b} t_{ab} c_a^\dagger c_b, \tag{8.49}$$

needs some care with the signs. Basically, two basis states $|b_\sigma\rangle$ and $|b'_\sigma\rangle$ are connected with this matrix if their binary representations differ at two positions a and b. The matrix element is then $(-1)^{M_{ab}} t_{ab}$, where M_{ab} is the number of occupied sites between a and b, i.e. assuming $a < b$,

$$M_{ab} = \sum_{c=a+1}^{b-1} n_c. \tag{8.50}$$

For instance, the two states (10010110) and (10011100) with $L = 8$ are connected with the matrix element $+t_{46}$, where the sites are numbered from 0 to $L - 1$.

Calculating the Hubbard interaction is straightforward: a bit-wise AND is applied to the up and down parts of a binary state ($b_\uparrow$ & $b_\downarrow$ in C or C++) and the number of set bits of the result is the number of doubly occupied sites in that basis state.

The above procedure is slightly more complicated if the Hamiltonian contains one-body terms that mix up and down electrons, such as spin flip terms or a superconducting pairing source term. Such terms arise in applying the VCA or the C-DMFT with an ED solver.

8.4.2 The Lanczos Algorithm for the Ground State

Once we have a representation of the basis vectors and Hamiltonian H', we can proceed to the ED by finding the ground state of H', using the Lanczos algorithm. Generally, the Lanczos method [9] is used when one needs the extreme eigenvalues of a matrix too large to be fully diagonalized (e.g., with the Householder algorithm). The method is iterative and involves only multiply-add's from the matrix. This means in particular that the matrix does not necessarily have to be constructed explicitly, since only its action on a vector is needed. In some extreme cases where it is practical to do so, the matrix elements can be calculated "on the fly," and this allows us to save the memory associated with storing the matrix itself.

The basic idea behind the Lanczos method is to build a projection $\mathscr{H}$ of the full Hamiltonian matrix H' onto a Krylov subspace. Starting with a (random) state $|\phi_0\rangle$, a Krylov subspace is spanned by the iterated application of H':

$$\mathscr{K} = \text{span}\left\{|\phi_0\rangle, H'|\phi_0\rangle, H'^2|\phi_0\rangle, \cdots, H'^M|\phi_0\rangle\right\} \tag{8.51}$$

the generating vectors above are not mutually orthogonal, but a sequence of mutually orthogonal vectors can be built from the following recursion relation

$$|\phi_{n+1}\rangle = H'|\phi_n\rangle - a_n|\phi_n\rangle - b_n^2|\phi_{n-1}\rangle, \tag{8.52}$$

where

$$a_n = \frac{\langle\phi_n|H|\phi_n\rangle}{\langle\phi_n|\phi_n\rangle} \qquad b_n^2 = \frac{\langle\phi_n|\phi_n\rangle}{\langle\phi_{n-1}|\phi_{n-1}\rangle} \qquad b_0 = 0 \tag{8.53}$$

and we set the initial conditions $b_0 = 0$, $|\phi_{-1}\rangle = 0$. At any given step, only three state vectors are kept in memory (ϕ_{n+1}, ϕ_n and ϕ_{n-1}). On the basis of normalized states $|n\rangle = |\phi_n\rangle/\sqrt{\langle\phi_n|\phi_n\rangle}$, the projected Hamiltonian has the tridiagonal form

$$\begin{pmatrix} a_0 & b_1 & 0 & 0 & \cdots & 0 \\ b_1 & a_1 & b_2 & 0 & \cdots & 0 \\ 0 & b_2 & a_2 & b_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_N \end{pmatrix}. \tag{8.54}$$

Such a matrix is readily diagonalized by fast methods special to tridiagonal matrices. Many in-house implementations of the method stop the iterations when the lowest eigenvalue E_0 changes by less than some tolerance (e.g., 10^{-14}) from one iteration to the next. However, a much safer convergence criterion is to stop when the estimated Ritz residual $|H'|\Omega\rangle - E_0|\Omega\rangle|$ is smaller than a preset tolerance [9].

The ground state energy E_0 and wave-function $|\Omega\rangle$ are very well approximated by the lowest eigenvalue and the corresponding eigenvector of the tridiagonal matrix, which are obtained by standard methods. This provides us with the ground state $|\Omega\rangle$ in the reduced basis $\{|\phi_n\rangle\}$. But we need the ground state in the original basis, and this requires retracing the Lanczos iterations a second time – for the $|\phi_n\rangle$ are not stored in memory – and constructing the ground state progressively at each iteration from the known coefficients $\langle\Omega|\phi_n\rangle$.

The Lanczos procedure is simple and efficient. The convergence is fast if the lowest eigenvalue E_0 is well separated from the next one (E_1). It slows down if $E_1 - E_0$ is small. If the ground state is degenerate ($E_1 = E_0$), the procedure will converge to a vector of the ground state subspace, a different one each time the initial state $|\phi_0\rangle$ is changed. The number M of iterations required varies between a few tens and $\sim$200, depending on the system size and on $E_1 - E_0$.

Note that the sequence of Lanczos vectors $|\phi_n\rangle$ is in principle orthogonal, as this is guaranteed by the three-way recursion relation (8.52). However, numerical error will introduce "orthogonality leaks," and after a few tens of iterations the Lanczos basis will become over-complete in the Krylov subspace. This will translate in multiple copies of the ground state eigenvalue in the tridiagonal matrix (8.54), which should not be taken as a true degeneracy. However, as long as one is only interested in the ground state and not in the multiplicity of the lowest eigenvalues, this is not a problem.

8.4.3 The Lanczos Algorithm for the Green Function

We will first describe a Lanczos algorithm for calculating the Green function (8.30), that provides a continued-fraction representation of its frequency dependence. In the next subsection, we will instead present an alternate method based on the Band Lanczos algorithm, that provides a Lehmann representation of the Green function and that is both faster and more memory intensive.

Consider first the function $G'_{aa,e}(\omega)$. One needs to know the action of $(\omega - H' + E_0)^{-1}$ on the state $|\phi_a\rangle = c_a^\dagger|\Omega\rangle$, and then to calculate

$$G'_{aa,e} = \langle\phi_a|\frac{1}{\omega - H' + E_0}|\phi_a\rangle. \tag{8.55}$$

As with any generic function of H', this one can be expanded in powers of H':

$$\frac{1}{z - H'} = \frac{1}{z} + \frac{1}{z^2}H' + \frac{1}{z^3}H'^2 + \cdots \tag{8.56}$$

and the action of this operator can be evaluated exactly at order H'^M in a Krylov subspace (8.51). Thus we again resort to the Lanczos algorithm: A Lanczos sequence is calculated from the initial, normalized state $|\phi_0\rangle = |\phi_a\rangle/b_0$, where $b_0^2 = \langle\phi_a|\phi_a\rangle$. This sequence generates a tridiagonal representation of H', albeit in a different Hilbert space sector: that with $N_\uparrow + 1$ up-spin electrons and $N_\downarrow$ down-spin electrons. Once the preset maximum number of Lanczos steps, or a near zero value of b_n, has been reached, the tridiagonal representation (8.54) may then be used to calculate (8.55). This is the matrix element $b_0^2[(\omega - H' + E_0)^{-1}]_{00}$ (the first element of the inverse of a tridiagonal matrix), which has a simple continued fraction form [10]:

$$G'_{aa,e}(\omega) = \cfrac{b_0^2}{\omega - a_0 - \cfrac{b_1^2}{\omega - a_1 - \cfrac{b_2^2}{\omega - a_2 - \cdots}}}. \tag{8.57}$$

Once the arrays $\{a_n\}$ and $\{b_n\}$ are known, evaluating the Green function reduces to the calculation of a truncated continued fraction, which can be done recursively in M steps, starting from the bottom floor of the fraction.

Consider next the case $a \neq b$. The continued fraction representation applies only when the same state $|\phi\rangle$ appears on the two sides of (8.55). If $a \neq b$ we may instead use the following trick: we define the combination

$$G'^{+}_{ab,e}(\omega) = \langle\Omega|(c_a + c_b)\frac{1}{\omega - H' + E_0}(c_a + c_b)^\dagger|\Omega\rangle. \tag{8.58}$$

Using the symmetry $G'_{ab,e}(\omega) = G'_{ba,e}(\omega)$, this leads to

$$G'_{ab,e}(\omega) = \frac{1}{2}(G'^{+}_{ab,e}(\omega) - G'_{aa,e}(\omega) - G'_{bb,e}(\omega)), \tag{8.59}$$

where $G'^{+}_{ab,e}$ can be calculated in the same way as $G'_{aa,e}$, i.e. with a simple continued fraction. We proceed likewise for $G'^{+}_{ab,h}(\omega)$. The cluster Green function is thus

encoded in $L(L+1)$ continued fractions, whose coefficients are stored in memory, so that $\boldsymbol{G}'(\omega)$ can be computed on demand for any complex frequency ω.

8.4.4 The Band Lanczos Algorithm for the Green Function

An alternate way of calculating the cluster Green function is to apply the *band Lanczos* procedure [11]. This is a generalization of the Lanczos procedure where the Krylov subspace is spanned not by one, but by many states. Let us assume that up and down spins are decoupled, so that the Green function is $L \times L$ block diagonal. The L states $|\phi_a\rangle = c_a^\dagger |\Omega\rangle$ are first constructed, and then one builds the projection $\mathscr{H}$ of H' on the Krylov subspace spanned by

$$\Big\{ |\phi_1\rangle, \ldots, |\phi_L\rangle, H'|\phi_1\rangle, \ldots, H'|\phi_L\rangle, \ldots, H'^M|\phi_1\rangle, \ldots, H'^M|\phi_L\rangle \Big\}. \tag{8.60}$$

A Lanczos basis $\{|n\rangle\}$ is constructed by the successive application of H' and orthonormalization with respect to the previous $2L$ basis vectors. In principle, each new basis vector $|n\rangle$ is already automatically orthogonal to basis vectors $|1\rangle$ through $|n-2L-1\rangle$, although "orthogonality leaks" arise eventually and may be problematic. A practical rule of thumb to avoid these problems is to control the number M of iterations by the convergence of the lowest eigenvalue of $\mathscr{H}$. Independently of this, one must be careful about potential redundant basis vectors in the Krylov subspace, which must be properly "deflated" [11]. The number of states R in the Krylov subspace at convergence is typically between 100 and 300, depending on the system size. The $R \times R$ matrix $\mathscr{H}$, which has a tridiagonal structure in the ordinary Lanczos method, now has a band structure made of L diagonals around the central diagonal. It is then a simple matter to obtain a Lehmann representation of the Green function in the Krylov subspace (see Sect. 8.3) by calculating the projections Q_{ar} of $|\phi_a\rangle$ on the eigenstates of $\mathscr{H}$ (the inner products of the $|\phi_a\rangle$'s with the Lanczos vectors are calculated as the latter are constructed). The Green function can then be expressed in a Lehmann representation (8.36). The two contributions $G'_{ab,e}$ and $G'_{ab,h}$ to the Green function are computed separately, and the corresponding matrices $\boldsymbol{Q}$ and $\boldsymbol{\Lambda}$ are simply concatenated to form the complete $\boldsymbol{Q}$- and $\boldsymbol{\Lambda}$-matrices, which are then stored and allow again for a quick calculation of the Green function as a function of the complex frequency ω. The $L \times R$ matrix $\boldsymbol{Q}$ has the property

$$\boldsymbol{Q}\boldsymbol{Q}^\dagger = \mathbf{1}_{L\times L}. \tag{8.61}$$

This holds even if the Lehmann representation is obtained from a subspace and not the full space, and is simply a consequence of the anti-commutation relations $\{c_a, c_b^\dagger\} = \delta_{ab}$.

The band Lanczos method requires more memory than the usual Lanczos method, since $2L + 1$ vectors must simultaneously be kept in memory, compared to 3 for the simple Lanczos method. On the other hand, it is faster since all pairs (a, b) are covered in a single procedure, compared to $L(L + 1)/2$. Thus, we gain a factor L^2 in speed at the cost of a factor L in memory. Another advantage is that it provides a Lehmann representation of the Green function.

8.4.5 Cluster Symmetries

The exact diagonalization procedure may be optimized by taking advantage of the symmetries of the cluster Hamiltonian, in particular those coming from cluster geometry. If the Hamiltonian is invariant under a discrete group $\mathfrak{G}$ of symmetry operations and $|\mathfrak{G}|$ denotes the number of such elements (the order of the group), the dimension of the largest Hilbert space needed can be reduced by a factor of almost $|\mathfrak{G}|$, and the number of state vectors needed in the band Lanczos method reduced by the same factor. The corresponding speed and memory gain are appreciable. The price to pay is a higher complexity in coding the basis states. Since calculating matrix elements "on the fly" becomes then more time consuming, it is reasonable to store the Hamiltonian in memory. It is assumed here that open boundary conditions are used, and therefore there is no translation symmetry within the cluster; thus we are concerned with point groups, not space groups.

Let us start with a simple example: a cluster invariant with respect to a single inversion or a single rotation by π. One may think of a one-dimensional cluster, for instance, with a left–right inversion. The corresponding symmetry group is C_2, with two elements: the identity e and the inversion ι. The group C_2 contains two irreducible representations, noted A and B, containing states that are respectively even and odd with respect to ι. Because the Hamiltonian is invariant under inversion: $H' = \iota^{-1} H' \iota$, eigenvectors of H' will be either even or odd, i.e., belong either to the A or to the B representation. Likewise, the Hamiltonian will have no matrix elements between states belonging to different representations. In order to take advantage of this fact, one needs to construct a basis containing only states of a given representation. The occupation number basis states $|b\rangle$ introduced above (or binary states, as we will call them) are no longer adequate. For the simple group C_2, one should rather consider the even and odd combinations $|b\rangle \pm \iota|b\rangle$ (and some of these combinations may vanish). Yet we still need a scheme to label the different basis states and have a quick access to their occupation number representation, which allows us to compute matrix elements. We will now briefly describe how this can be done (a more detailed discussion can be found, e.g., in [12]).

We will restrict the discussion to groups generated by a finite number of Z_2 operations. Such groups are Abelian and their irreducible representations, all strictly one-dimensional, are labelled by the parity $(\pm)$ of each generator. For instance, this includes the group C_{2v}, generated by a reflexions about the x and y axes, with the four irreducible representations:

$$A_1 = (+,+) \qquad A_2 = (-,-) \qquad B_1 = (+,-) \qquad B_2 = (-,+). \tag{8.62}$$

In this class of groups, the number of group elements $|\mathfrak{G}|$ coincides with the number of irreducible representations (there is an isomorphism between the two), and group multiplication coincides with tensor products. Each group element forms its own conjugacy class and the group character $\chi_g^{(\alpha)}$ of element g in the representation α is either 1 or -1. Under the action of the group $\mathfrak{G}$, each binary state generates an "orbit" of binary states, whose length is the order $|\mathfrak{G}|$ of the group or a divisor thereof. To such an orbit corresponds at most one state in the irreducible representation labelled α, given by the corresponding projection operator:

$$|\psi\rangle = \frac{1}{|\mathfrak{G}|} \sum_g \chi_g^{(\alpha)} g|b\rangle \tag{8.63}$$

(this state may vanish). We can then select a representative binary state for each orbit (e.g., the one associated with the smallest binary representation) and use it as a label for the state $|\psi\rangle$. We still need an index function $B(i)$ which provides the representative binary state for each consecutive label i. The reverse correspondence $i = I(b)$ is stored in a tree structure, such as the one provided by the STL `map` container.

Once the basis has been constructed, one needs to construct a matrix representation of the Hamiltonian in that representation. Given two states $|\psi_1\rangle$ and $|\psi_2\rangle$, represented by the binary states $|b_1\rangle$ and $|b_2\rangle$, it is a simple matter to show that the matrix element is

$$\langle\psi_2|H'|\psi_1\rangle = \frac{1}{|\mathfrak{G}|} \sum_g \chi_h^{(\alpha)} \phi_g(b) \langle gb_2|H'|b_1\rangle, \tag{8.64}$$

where the phase $\phi_g(b)$ is defined by the relation

$$g|b\rangle = \phi_g(b)|gb\rangle. \tag{8.65}$$

In the above relation, $|gb\rangle$ is the binary state obtained by applying the symmetry operation g to the occupation numbers forming b, whereas the phase $\phi_g(b)$ is the product of signs collected from all the permutations of creation operators needed to go from b to gb. Equation (8.64) is used as follows to construct the Hamiltonian matrix: First, the Hamiltonian can be written as $H' = \sum_r H'_r$, where H'_r is a hopping term between specific sites, or a diagonal term like the interaction. One then loops over all b_1's. For each b_1, and each term H'_r, one constructs the single binary state $H'_r|b_1\rangle$. One then finds the representative b_2 of that binary state, by applying on it all possible symmetry operations until g is found such that $|gb_2\rangle = H'_r|b_1\rangle$. During this operation, the phase $\phi_g(b)$ must also be collected. Then the matrix element (8.64) is added to the list of stored matrix elements. Since each term H'_r individually is not invariant under the group, there will be more matrix elements

generated than there should be, i.e. there will be cancellations between different matrix elements associated with the same pair (b_1, b_2) and produced by the different H'_r's. For this reason, it is useful to first store all matrix elements associated with a given b_1 in an intermediate location in order for the cancellations to take effect, and then to store the cleaned up "column" labelled by b_1 to its definitive storage location. Needless to say, one should only store the row and column indices of each element of a given value.

8.4.6 Green Functions Using Cluster Symmetries

Most of the time, the ground state lies in the trivial (symmetric) representation. However, taking advantage of symmetries in the calculation of the Green function requires all the irreducible representations to be included in the calculation. The Green function matrix $\boldsymbol{G}'$ will be block-diagonal in a basis of states belonging to representations of $\mathfrak{G}$. To calculate it, one should use symmetry eigen-operators

$$c_\rho^{(\alpha)} = \sum_a M_{\rho a}^{(\alpha)} c_a \tag{8.66}$$

such that $c_\rho^{(\alpha)}$ transforms under representation α, and ρ labels the different possibilities. For instance, for a linear cluster of length 4 and an inversion symmetry that maps the sites (1234) into (4321), these operators are

$$\begin{aligned} c_1^{(A)} &= c_1 + c_4 \quad & c_1^{(B)} &= c_1 - c_4 \\ c_2^{(A)} &= c_2 + c_3 \quad & c_2^{(B)} &= c_2 - c_3 \end{aligned} . \tag{8.67}$$

For each representation, one may use the Band Lanczos procedure and obtain a Lehmann representation $Q_{\rho r}^{(\alpha)}$ for the associated Green function $G_{\rho\sigma}^{(\alpha)}(\omega)$. If the ground state is in representation α and the operators $c_\rho^{(\beta)}$ of representation β are used, the Hilbert space sector to work with will be the tensor product representation $\alpha \otimes \beta$, which poses no problem at all when all irreps are one-dimensional, but would bring additional complexity if the ground states were in a multi-dimensional representation. Finally, one may bring together the different pieces, by building a $L \times L$ matrix $M_{\rho a}$ that is the vertical concatenation of the various rectangular matrices $M_{\rho a}^{(\alpha)}$, and returning to the usual $\boldsymbol{Q}$-matrix representation

$$Q_{ar} = (\boldsymbol{M}^{-1})_{a\rho} Q_{\rho r}. \tag{8.68}$$

Using cluster symmetries for the Green function saves a factor $|\mathfrak{G}|$ in memory because of the reduction of the Hilbert space dimension, and an additional factor of $|\mathfrak{G}|$ since the number of input vectors in the band Lanczos procedure is also divided

by $|\mathfrak{G}|$. Typically then, most of the memory will be used to store the Hamiltonian matrix.

It might appear at this stage that a better way to proceed would be to start the whole diagonalization procedure (including the coding of basis states) with the help of the proper symmetric orbitals labelled by the irreducible representation (ρ, β) instead of site indices a. The problem comes from the interaction term in H_1': it is much simpler when expressed in terms of the c_a's and $c_a^\dagger$'s; using any other basis would lead to a proliferation of matrix elements in such a large number that it would dominate memory and computing time requirements.

8.4.7 Other Solvers

A problem with the Lanczos method is the breakdown of orthogonality between Lanczos vectors due to numerical errors. This is not a problem when computing the ground state, but affects the accuracy of the Green function in both the Lanczos and band Lanczos methods.

A way out of this is the Kernel polynomial method (KPM) [13]. Like the Lanczos and band Lanczos methods, it is an iterative procedure using matrix vector multiplications that can yield a representation of the Green function, this time in terms of Chebyshev polynomials. Its advantage lies in its numerical stability: orthogonality at each iteration is not important, and, therefore, the accuracy of the spectral function thus calculated increases safely with the number of iterations. Its disadvantage is that it provides the Green function at real frequencies only, whereas the Lanczos method provides a representation for the whole complex plane of frequencies. Since integrals of the Green function are routinely performed along the imaginary axis in order to compute physical quantities, the KPM is less useful in this context. However, it is the most accurate method for obtaining the spectral function. See [13] for a detailed description of the method and for an example of its application to the spectral function of the one-dimensional Hubbard model using CPT.

Nothing in CPT is specific to the ED solver. A Quantum Monte Carlo (QMC) solver could be used just as well, with the usual caveats that the cluster Green function would be affected by statistical errors and available as a function of imaginary time only (or a finite number of Matsubara frequencies). Thus, using a QMC solver is not the most natural one to compute the spectral function, lest one is comfortable with using the maximum entropy method to perform an approximate analytic continuation towards real frequencies. This was accomplished in [14, 15] and applied to a monolayer of C_{60} molecules. A first attempt at using the more recent continuous-time quantum Monte Carlo (CTQMC) method in the more general context of VCA was carried out in [16]. We refer the reader to these works for a detailed discussion.

8.5 Periodization

The CPT Green function (8.4) breaks translation invariance on the original lattice γ. This means that it is not diagonal when expressed in terms of full wavevectors: $\boldsymbol{G} \to G(\boldsymbol{k}, \boldsymbol{k}')$. Because of the residual superlattice translation invariance, $\boldsymbol{k}'$ and $\boldsymbol{k}$ must map to the same wavevector of the superlattice Brillouin zone (or reduced Brillouin zone) and differ by an element of the reciprocal superlattice: $\boldsymbol{k}' - \boldsymbol{k} \in \Gamma^*$.

A periodization procedure is an ad hoc prescription to recover a Green function diagonal in $\boldsymbol{k}$. The most direct periodization procedure, the so-called G-scheme, was proposed in [4]. It simply discards the off-diagonal piece of the Green function when expressed in the $\boldsymbol{k} = (\boldsymbol{K}, \tilde{\boldsymbol{k}})$ basis. This is done as follows: In the $(\boldsymbol{K}, \tilde{\boldsymbol{k}})$ basis, the matrix $\boldsymbol{G}$ has the following form (see (8.12)):

$$G_{\boldsymbol{K}\boldsymbol{K}'}(\tilde{\boldsymbol{k}}, \omega) = \frac{1}{L} \sum_{a,b} \mathrm{e}^{-\mathrm{i}(\boldsymbol{K}\cdot\boldsymbol{r}_a - \boldsymbol{K}'\cdot\boldsymbol{r}_b)} G_{ab}(\tilde{\boldsymbol{k}}, \omega). \tag{8.69}$$

This form can be further converted to the full wavevector basis ($\boldsymbol{k} = \boldsymbol{K} + \tilde{\boldsymbol{k}}$) by the use of the unitary matrix $\boldsymbol{\Lambda}$ of (8.25):

$$\begin{aligned} G(\tilde{\boldsymbol{k}} + \boldsymbol{K}, \tilde{\boldsymbol{k}} + \boldsymbol{K}') &= \left(\boldsymbol{\Lambda}(\tilde{\boldsymbol{k}}) \boldsymbol{G} \boldsymbol{\Lambda}^\dagger(\tilde{\boldsymbol{k}})\right)_{\boldsymbol{K}\boldsymbol{K}'} \\ &= \frac{1}{L^2} \sum_{a,b,\boldsymbol{K}_1,\boldsymbol{K}_1'} \mathrm{e}^{-\mathrm{i}(\tilde{\boldsymbol{k}}+\boldsymbol{K}-\boldsymbol{K}_1)\cdot\boldsymbol{r}_a} \mathrm{e}^{\mathrm{i}(\tilde{\boldsymbol{k}}+\boldsymbol{K}'-\boldsymbol{K}_1')\cdot\boldsymbol{r}_b} G_{\boldsymbol{K}_1\boldsymbol{K}_1'} \\ &= \frac{1}{L} \sum_{a,b} \mathrm{e}^{-\mathrm{i}(\tilde{\boldsymbol{k}}+\boldsymbol{K})\cdot\boldsymbol{r}_a} \mathrm{e}^{\mathrm{i}(\tilde{\boldsymbol{k}}+\boldsymbol{K}')\cdot\boldsymbol{r}_b} G_{ab}(\tilde{\boldsymbol{k}}, \omega). \end{aligned} \tag{8.70}$$

Now let us set $\boldsymbol{K} = \boldsymbol{K}'$. We can also freely replace $\tilde{\boldsymbol{k}}$ by $\boldsymbol{k} = \tilde{\boldsymbol{k}} + \boldsymbol{K}$ in $G_{ab}(\tilde{\boldsymbol{k}}, \omega)$, since $\boldsymbol{V}(\tilde{\boldsymbol{k}})$ is unchanged when $\tilde{\boldsymbol{k}}$ is shifted by a reciprocal superlattice vector, because of translation invariance on the superlattice. We end up with the following simple formula:

$$G_{\mathrm{per}}(\boldsymbol{k}, \omega) = \frac{1}{L} \sum_{a,b} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b)} G_{ab}(\boldsymbol{k}, \omega) \qquad \text{(G-scheme)} \tag{8.71}$$

which has the straightforward appearance of a residual Fourier transform.

The G-scheme is natural, as the density of states $N(\omega)$ is the trace of the imaginary part of the Green function:

$$\begin{aligned} N(\omega) &= -\frac{2}{N} \operatorname{Im} \operatorname{tr} \boldsymbol{G}(\omega) = -\frac{2}{N} \operatorname{Im} \sum_i G_{ii}(\omega) \\ &= -\frac{2}{N} \operatorname{Im} \sum_{\boldsymbol{k}} G(\boldsymbol{k}, \omega), \end{aligned} \tag{8.72}$$

and the spectral function $A(\boldsymbol{k},\omega)$, as a partial trace, involves only the diagonal part. Indeed, it is a simple matter to show from the anti-commutation relations that the frequency integral of the Green function is the unit matrix:

$$-2\,\mathrm{Im}\int\frac{\mathrm{d}\omega}{2\pi}\,\boldsymbol{G}(\omega)=\mathbf{1}. \tag{8.73}$$

This being representation independent, it follows that the frequency integral of the imaginary part of the off-diagonal components of the Green function vanishes.

Another possible prescription is the so-called Σ-scheme, which periodizes the self-energy $\boldsymbol{\Sigma}$ instead of the Green function [17, 18]. In other words, one computes the momentum dependent self-energy

$$\Sigma_{\mathrm{per}}(\boldsymbol{k},\omega)=\frac{1}{L}\sum_{a,b}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a-\boldsymbol{r}_b)}\,\Sigma_{ab}(\omega) \tag{8.74}$$

and recovers the Green function (and the spectral function) with Dyson's equation:

$$G_{\mathrm{per}}(\boldsymbol{k},\omega)^{-1}=\omega-t(\boldsymbol{k})-\Sigma_{\mathrm{per}}(\boldsymbol{k},\omega)\qquad(\Sigma\text{-scheme}), \tag{8.75}$$

where $t(\boldsymbol{k})$ is the dispersion relation. This is appealing since $\boldsymbol{\Sigma}$ is an irreducible quantity, as opposed to $\boldsymbol{G}$. However, as shown in Fig. 8.2, the Σ-scheme leads to unphysical results in the spectral function. For instance, it leads to non-zero spectral weight within the Mott gap in the one-dimensional Hubbard model, for all values of U (the Mott transition occurs at $U\to\infty$ instead of $U=0$, the correct result).

Yet another prescription is the M-scheme, in which the first lattice cumulant of the Green function is periodized [19]. In practice, this proceeds as follows: The matrix of one-body terms is split into diagonal and off-diagonal parts:

$$\boldsymbol{t}(\tilde{\boldsymbol{k}})=\boldsymbol{t}^{\mathrm{diag}}(\tilde{\boldsymbol{k}})+\boldsymbol{t}^{\mathrm{off}}(\tilde{\boldsymbol{k}}). \tag{8.76}$$

We then proceed exactly like in the G-scheme, but without the off-diagonal piece of $\boldsymbol{t}$. In other words, we periodize the quantity

$$\boldsymbol{G}^{\mathrm{diag}}(\tilde{\boldsymbol{k}},\omega)=\Big(\omega-\boldsymbol{t}^{\mathrm{diag}}(\tilde{\boldsymbol{k}})-\boldsymbol{\Sigma}(\omega)\Big)^{-1} \tag{8.77}$$

as in (8.71) and (8.74) and obtain $G^{\mathrm{diag}}_{\mathrm{per}}(\boldsymbol{k},\omega)$. We then express $\boldsymbol{t}^{\mathrm{off}}(\tilde{\boldsymbol{k}})$ in the full Fourier representation ($t^{\mathrm{off}}(\boldsymbol{k})$) and finally construct the periodized Green function

$$G_{\mathrm{per}}(\boldsymbol{k},\omega)=\Big[G^{\mathrm{diag}}_{\mathrm{per}}(\boldsymbol{k},\omega)^{-1}-t^{\mathrm{off}}(\boldsymbol{k})\Big]^{-1}\qquad(\text{M-scheme}). \tag{8.78}$$

The three periodization schemes described above are compared in Fig. 8.2, which shows the spectral function of the one-dimensional Hubbard model computed from

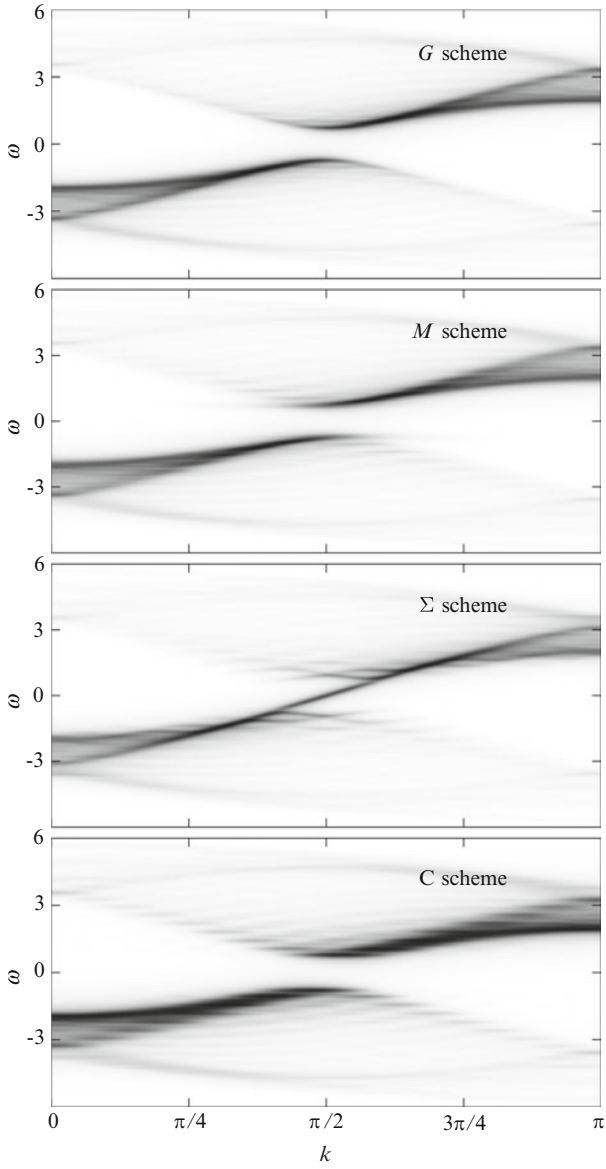

Fig. 8.2 CPT spectral function of the one-dimensional, half-filled Hubbard model with $U = 4$, $t = 1$, in various periodization schemes, from a 16-site cluster. See text for comments

the G, Σ and M-schemes. The bottom panel also shows the effect of periodizing only the cluster Green function $\boldsymbol{G}'(\omega)$, without applying CPT (what we call the C-scheme). As we can see, the G-scheme reproduces the expected feature of the spectral function, with a direct Mott gap at $k = \pi/2$. The Σ-scheme has a

substantial spectral weight dispersing across the Mott gap and is completely wrong in this case. The M-scheme is not so different from the G-scheme, except that the minimum in the dispersion of the main spectral feature is not precise at $k = \pi/2$, as if the Mott gap as indirect. Finally, the C-scheme shows no dispersion of any of the peaks, just a modulation of their weight as a function of wavevector, as expected. We conclude that the G-scheme is preferable, as well as the simplest and most natural periodization scheme.

Note, however, that expectation values of one-body operators and other thermodynamic quantities can (and should) be computed from the CPT Green function (8.4) without periodization.

The periodized Green functions have the correct analytic properties and lead to a positive and normalized spectral weight. Indeed, let us consider any matrix function $G_{ab}(\omega)$ with a Lehmann representation:

$$G_{ab}(\omega) = \sum_r \frac{Q_{ar} Q^*_{br}}{\omega - \omega_r} \tag{8.79}$$

that we periodize in the way of (8.71):

$$\begin{aligned} G(\boldsymbol{k}, \omega) &= \frac{1}{L} \sum_{a,b} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_b)} G_{ab} \\ &= \frac{1}{L} \sum_{\alpha,\beta} \sum_r \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_a} Q_{ar} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_b} Q^*_{br} \frac{1}{\omega - \omega_r} \\ &= \sum_r \frac{|\psi_r(\boldsymbol{k})|^2}{\omega - \omega_r}, \end{aligned} \tag{8.80}$$

where

$$\psi_r(\boldsymbol{k}) = \frac{1}{L} \sum_a \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_a} Q_{ar}. \tag{8.81}$$

The above expression manifestly leads to a positive spectral weight, since all residues are real and all poles on the real axis (periodization only affects the residues, not the location of the poles). Moreover, since the matrix Q is affected by a unitary transformation (the discrete Fourier transform), the trace is not affected:

$$\sum_r |\psi_r(\boldsymbol{k})|^2 = \sum_{a,r} |Q_{ar}|^2 = \boldsymbol{Q}\boldsymbol{Q}^\dagger = 1 \tag{8.82}$$

and the normalization of the spectral function is not affected. If G_{ab} is the cluster Green function, this corresponds to the C-scheme; if it is the CPT Green function (8.4), this corresponds to the G-scheme; if it is the cluster self-energy Σ_{ab}, then

the resulting periodized self-energy also has the correct analytic properties and so does the Σ-scheme Green function. Last, the M-scheme too has the correct analytic properties, since removing and reinserting later the off-diagonal part $\boldsymbol{t}^{\text{off}}(\tilde{\boldsymbol{k}})$ do not affect the analytic properties of the Green function.

8.6 Computing Physical Quantities

In many applications that use the CPT Green function, whether in VCA or C-DMFT, it is necessary to compute various physical quantities, which often can be expressed as the expectation value of a one-body operator of the general form

$$\mathcal{O} = s_{\alpha\beta} c_\alpha^\dagger c_\beta. \tag{8.83}$$

In the simple case of the number of electrons, the matrix $\boldsymbol{s}$ is diagonal:

$$s_{i\sigma,j\sigma'} = \delta_{ij}\delta_{\sigma\sigma'} \tag{8.84}$$

but one could be interested in the expectation value of the kinetic energy, or some order parameter.

We will show in this section how to practically calculate the density $\bar{\mathcal{O}} = s_{\alpha\beta}\langle c_\alpha^\dagger c_\beta\rangle/N$ with the help of a Green function defined along the imaginary axis, as obtained for instance from the ED solver and the CPT (we assume zero temperature).

From the Lehmann representation of the Green function, we see that $\langle c_\alpha^\dagger c_\beta\rangle$ is given by the integral of the Green function along a contour $C_<$ surrounding the negative real frequency axis counterclockwise:

$$\langle c_\beta^\dagger c_\alpha\rangle = \int_{C_<} \frac{\mathrm{d}z}{2\pi i} G_{\alpha\beta}(z). \tag{8.85}$$

Therefore, the expectation value we are looking for is

$$\bar{\mathcal{O}} = \frac{1}{N} s_{\beta\alpha}\langle c_\beta^\dagger c_\alpha\rangle = \frac{1}{N}\int_{C_<} \frac{\mathrm{d}z}{2\pi i}\,\mathrm{tr}\,[\boldsymbol{s}\boldsymbol{G}(z)]\,. \tag{8.86}$$

The trace includes a sum over lattice sites, spin and band. In the $(a, \tilde{\boldsymbol{k}})$ basis, this becomes

$$\bar{\mathcal{O}} = \frac{1}{N}\sum_{\tilde{\boldsymbol{k}}}\int_{C_<} \frac{\mathrm{d}z}{2\pi i}\,\mathrm{tr}\left[\boldsymbol{s}(\tilde{\boldsymbol{k}})\boldsymbol{G}(\tilde{\boldsymbol{k}}, z)\right], \tag{8.87}$$

where we assumed that the matrix $\boldsymbol{s}$ is diagonal in $\tilde{\boldsymbol{k}}$ (translation invariance over the superlattice).

Next, let us consider the asymptotic behavior of the Green function as $z \to \infty$: $\boldsymbol{G}(z) \to \mathbf{1}/z$. This allows us to modify (8.87) as follows:

$$\bar{\mathcal{O}} = \frac{1}{N} \sum_{\tilde{\boldsymbol{k}}} \int_{C_<} \frac{\mathrm{d}z}{2\pi i} \left\{ \operatorname{tr} \left[\boldsymbol{s}(\tilde{\boldsymbol{k}}) \boldsymbol{G}(\tilde{\boldsymbol{k}}, z) \right] - \frac{\operatorname{tr} \boldsymbol{s}(\tilde{\boldsymbol{k}})}{z - p} \right\}, \tag{8.88}$$

where $p > 0$ (in practice, we use $p \sim 1$). The term we added does not contribute, since its unique pole lies outside the contour. However, it modifies the asymptotic behavior of the integrand, which now decays as $1/z^2$. This allows us to replace the contour $C_<$ by an integral along the imaginary axis, plus an infinite semi-circle that does not contribute, since the integrand falls faster than $1/z$.

Consider next the part of the contour $C_<$ that lies above the real axis, and let us follow this contour clockwise and call it C. Let C' be the mirror image of C below the real axis, followed counterclockwise. To each z and $\mathrm{d}z$ of C correspond the mirror images z^* and $\mathrm{d}z^*$ on C', so that

$$I[C'] = \int_{C'} \mathrm{d}z\, f(z) = \int_C \mathrm{d}z^*\, f(z^*). \tag{8.89}$$

If, in addition, the integrand is such that $f(z^*) = f^*(z)$, then

$$I[C'] = \int_C \mathrm{d}z^*\, f^*(z) = \left(\int_C \mathrm{d}z\, f(z) \right)^* = I^*[C]. \tag{8.90}$$

The integral of $f(z)$ along the counterclockwise contour $C_<$ would then be

$$I[C_<] = I[C'] - I[C] = I^*[C] - I[C] = -2i \operatorname{Im} I[C]. \tag{8.91}$$

One of the properties of the Green function is its hermiticity: $G_{\alpha\beta}(z^*) = G^*_{\beta\alpha}(z)$. In the mixed Fourier representation, this is rather expressed as $\boldsymbol{G}(\tilde{\boldsymbol{k}}, z^*) = \boldsymbol{G}^\dagger(-\tilde{\boldsymbol{k}}, z)$. We also assume that $\boldsymbol{s}$ is Hermitian: $\boldsymbol{s}(\tilde{\boldsymbol{k}}) = \boldsymbol{s}^\dagger(-\tilde{\boldsymbol{k}})$ so that the expectation value is real. This means that the integrand of the expectation value respects the condition $f(z^*) = f^*(z)$.

Finally, the expectation value becomes

$$\bar{\mathcal{O}} = \frac{1}{N} \sum_{\tilde{\boldsymbol{k}}} \int_0^\infty \frac{\mathrm{d}\omega}{\pi} \operatorname{Re} \left\{ \operatorname{tr} \left[\boldsymbol{s}(\tilde{\boldsymbol{k}}) \boldsymbol{G}(\tilde{\boldsymbol{k}}, i\omega) \right] - \frac{\operatorname{tr} \boldsymbol{s}(\tilde{\boldsymbol{k}})}{i\omega - p} \right\}. \tag{8.92}$$

This expression can be used in practice (i.e. numerically) to compute the desired quantity.

Let us make an important remark on the computation of expectation values with the periodized Green functions described in the previous section. The result (8.86) is quite general and could formally be expressed as

$$\bar{\mathcal{O}} = \mathrm{Tr}\,(\boldsymbol{s}\boldsymbol{G}), \tag{8.93}$$

where the trace (Tr) stands for a functional trace, i.e. includes an integral over frequency as well as a trace over site and band indices. The above expression is basis-independent; in the full wavevector basis of one-particle states (see, e.g., (8.70)), the frequency summand would take the following form:

$$\frac{1}{N}\sum_{\tilde{\boldsymbol{k}},\tilde{\boldsymbol{k}}',\boldsymbol{K},\boldsymbol{K}'} \mathrm{Re}\left\{s(\tilde{\boldsymbol{k}}+\boldsymbol{K},\tilde{\boldsymbol{k}}'+\boldsymbol{K}')G(\tilde{\boldsymbol{k}}'+\boldsymbol{K}',\tilde{\boldsymbol{k}}+\boldsymbol{K},i\omega)\right\}. \tag{8.94}$$

If the operator $\mathcal{O}$ is translation invariant, as it usually is, then

$$s(\tilde{\boldsymbol{k}}+\boldsymbol{K},\tilde{\boldsymbol{k}}'+\boldsymbol{K}') = \delta_{\boldsymbol{K}\boldsymbol{K}'}\delta_{\tilde{\boldsymbol{k}}\tilde{\boldsymbol{k}}'}s(\boldsymbol{k}) = \delta_{\boldsymbol{k}\boldsymbol{k}'}s(\boldsymbol{k}) \tag{8.95}$$

and the above reduces to

$$\sum_{\boldsymbol{k}} \mathrm{Re}\left\{s(\boldsymbol{k})G_{\mathrm{per}}(\boldsymbol{k},i\omega)\right\}, \tag{8.96}$$

where $G_{\mathrm{per}}(\boldsymbol{k},i\omega)$ is the Green function (8.71), periodized in the G-scheme. This means that expectation values of translation invariant, one-body operators, computed in the G-scheme of periodization, coincide with those computed without periodization, i.e. with (8.92). This does not hold for the other periodization schemes, as it crucially depends on our discarding the off-diagonal elements of $\boldsymbol{G}$ in the full wavevector basis, which is possible because we take the trace of $\boldsymbol{G}$ against a matrix $\boldsymbol{s}$ that is itself diagonal in that basis. The G-scheme is therefore the best periodization scheme, in the sense of expectation values of one-body operators.

8.7 Results on the Hubbard Model

The CPT was devised at first to be applied on the Hubbard model, in particular to calculate an approximate spectral function that could be compared with ARPES data. The main advantage of CPT in this context is that it provides momentum-resolved spectral information. This is particularly useful in investigating the pseudogap observed in high-temperature superconductors in their normal phase, as was done in [20] with a one-band Hubbard model appropriate to cuprates (in particular YBCO). Figure 8.3, borrowed from [20], shows the spectral function

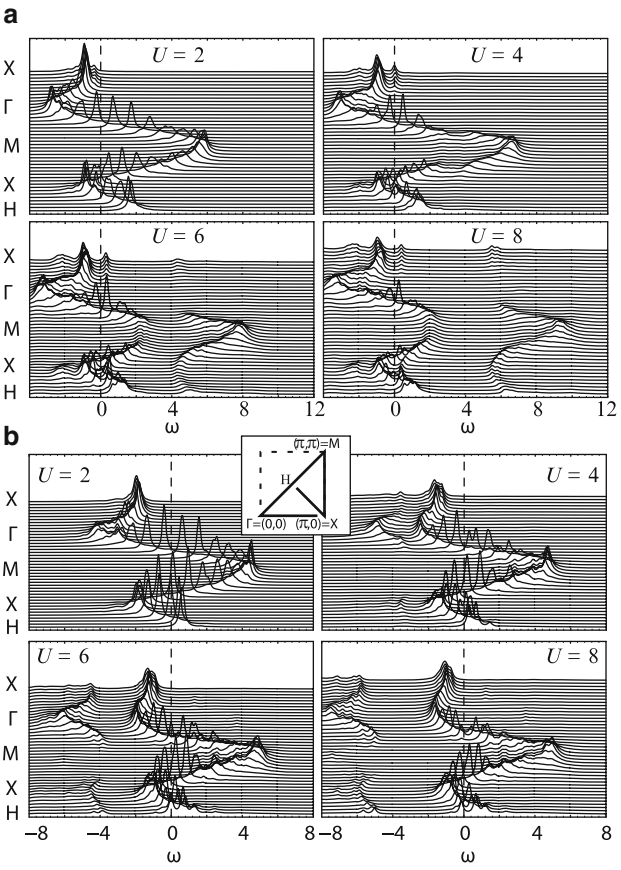

Fig. 8.3 Spectral weight for wavevectors along the high-symmetry directions shown in the inset. The band parameters are $t = 1$, $t' = -0.3$ and $t'' = 0.2$. (**a**): CPT calculations on a 3×4 cluster with 10 electrons on the cluster (17% hole doped). (**b**): the same, with 14 electrons (17% electron doped). From [20]

obtained from CPT for the hole- and electron-doped one-band Hubbard model, at various values of U, on a 3×4 cluster. The separation of the two Hubbard bands as U increases is clearly visible, as well as the suppression of the quasi-particle weight along the X–H direction in the hole-doped case, and along the diagonal in the electron-doped case. This weight suppression constitutes the pseudogap, i.e., a gap that opens only along a certain portion of the Fermi surface. The pseudogap is also clearly seen from plots of the spectral function as a function of wavevector (Fig. 8.4). The intersection of the anti-ferromagnetic zone boundary (white diagonal) with the non-interacting Fermi surface (black curve) defines "hot spots" where the scattering of quasi-particles by short-range anti-ferromagnetic fluctuations depletes the spectral weight. This compares well with ARPES results for electron-doped (Fig. 3 of [21]) and hole-doped (Fig. 8 of [22]) cuprates.

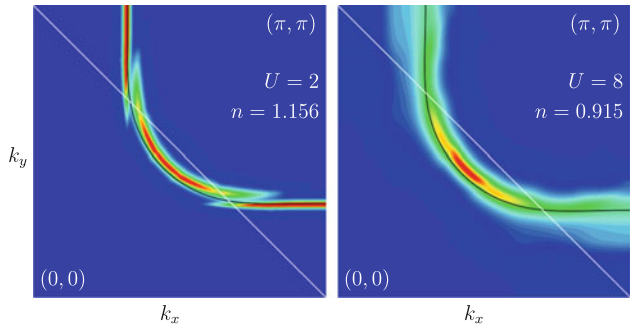

Fig. 8.4 Density plot of the spectral function at the Fermi level, in the first quadrant of the Brillouin zone, for an electron doped system at weak coupling (*left*), and a hole-doped system at stronger coupling (*right*). The non-interacting Fermi surface is shown in *black*, and the antiferromagnetic zone boundary in *white*. A 4×4 cluster was used

Comparisons with the known results from the one-dimensional Hubbard model are particularly instructive. We have shown the spectral function in various periodization schemes for the half-filled, one-dimensional Hubbard model on Fig. 8.2. On that figure (top panel), the spin charge separation is clearly visible, especially far from k_F, where the energy difference between the holon and spinon is largest. Access to a continuum of wavevectors by CPT allows for a natural smoothing of the computed density of states (DoS), compared to the same quantity obtained from the cluster Green function alone. Figure 8.5 shows results for clusters of sizes 4 and 16. The CPT density of states, when extrapolated to zero Lorenzian broadening η, display a set of (sometimes overlapping) bands that correspond to the dispersing poles of the CPT Green function. Overall, the CPT DoS, when calculated with a natural value of η, is less spiky than the corresponding cluster DoS. Non-zero spectral weights at the tails of the distribution and within the Mott gap are due to the finite value of η.

Let us now consider the spectral gap at half-filling. Estimates of this gap from a 12-site cluster, and extrapolations to infinite size are shown in Fig. 8.6 (see caption for details). We see that the CPT provides a better estimate of the gap than the cluster alone, even more so when an infinite-size extrapolation can be computed. The behavior as $U \to 0$ merits attention: cluster data alone, at a fixed cluster size, do not show the vanishing of the gap, but the CPT result does. However, both the raw cluster and CPT gaps tend to 0 as $U \to 0$ when an extrapolation is performed, although the gap does not go to zero nearly as fast as it should, from the exact solution. Long-range fluctuations (much longer than the cluster size) are likely important to set the correct gap value in this regime.

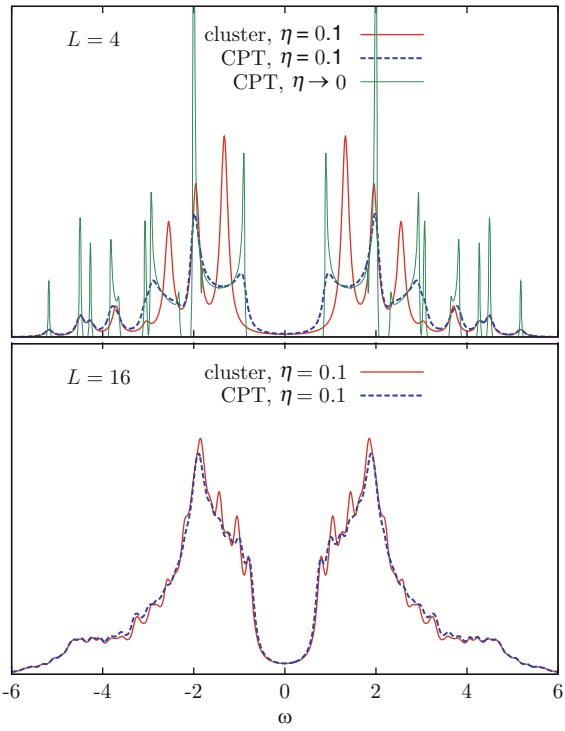

Fig. 8.5 Density of States (DoS) for the half-filled, one-dimensional Hubbard model ($U = 4, t = 1$). *Top panel*: results for the 4-site cluster; the DoS on the cluster is shown with a Lorenzian broadening $\eta = 0.1$, as well as the DoS from the CPT Green function with the same η. In addition, an extrapolation to $\eta \to 0$ is shown. *Bottom panel*: same, for the 16-site cluster (no extrapolation shown)

8.8 Applications to Other Models

8.8.1 Multi-Band Hubbard Models

Hubbard models with more than one band can be treated with CPT just like the one-band model. Of course, the computational burden depends on the total number of orbitals on the cluster, and therefore a three-band model on a 4-site cluster will require the same resources as a one-band model on a 12-site cluster, everything else being equal. The CPT was first applied to the three-band Hubbard model for high-temperature superconductors in [24] and [2]. The three-band Hubbard model was studied more recently using VCA [25]. The periodic Anderson model (or Kondo lattice) studied in [26] with the VCA can be viewed as a two-band Hubbard model in which only one band is correlated. Likewise, the CPT was applied in [27] to a two-band Hubbard model used to model an organic polymer coupled to a quantum wire: one (uncorrelated) band describing the wire and the other (correlated) describing

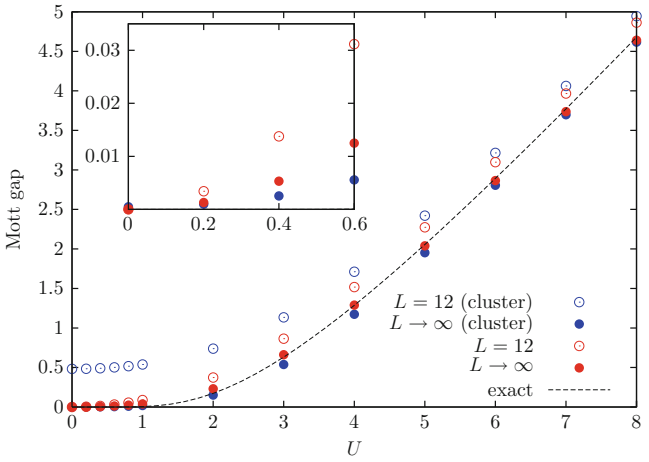

Fig. 8.6 Mott gap of the half-filled, one-dimensional Hubbard model, as a function of U. The *open blue circles* are obtained by looking at the lowest energy level of a 12-site cluster, in the sectors with 12 and 13 electrons respectively. The *filled blue circles* are obtained by extrapolating this estimate to $L \to \infty$ from $L = 4, 6, 8, 10$ and 12. The *open red circles* are obtained from the Lehmann representation (8.41) of the CPT Green function based on a 12-site cluster, at $k = \pi/2$. The *filled red circles* are obtained by extrapolating this estimate to $L \to \infty$ from $L = 4, 6, 8, 10$ and 12. The *dashed line* is the known exact result from the Lieb and Wu solution [23]

the polymer, the goal of this work being to explore the possibility of a spintronics device in which spin channels could be controlled by a gate instead of an external magnetic field.

8.8.2 *t–J or Spin Models*

The basic assumption behind CPT is that the interaction part of the Hamiltonian does not have any inter-cluster terms (8.1). This excludes, among others, the $t - J$ model, which has exchange-like interaction and correlated hopping, and the extended Hubbard model with its nearest-neighbor interactions. Of course this is only a restriction if we insist on the strong-coupling perturbation theory character of the basic formula (8.3). However, that formula, considered as a ad hoc approximation, may provide a good approximation to the lattice Green function even if the Hamiltonian does not have the form (8.1), especially considering that CPT is not controlled by computing higher orders in strong-coupling perturbation theory, but rather by increasing the cluster size.

For instance, the $t-J$ model is a strong-coupling effective model for the Hubbard model, with $J = 4t^2/U$. Therefore the cluster Green function of the $t - J$ model is expected to be close to that of the Hubbard model if $J \ll t$. In [28, 29], the $t - J$ model was used in conjunction with CPT to calculate the spectral function in the presence of stripes, with $J = 0.4t$. The advantage of the $t - J$ model lies in its smaller Hilbert space, following the trace over states with doubly occupied sites.

Purely spin models, like the Heisenberg model, are not covered by the strong perturbation theory of [1]. But CPT has recently been extended to such models [30] by generalizing to CPT techniques used in spin systems.

8.8.3 Extended Hubbard Models

Treating extended interactions in CPT is possible provided an additional approximation is made, in the form of a mean-field treatment of the interactions residing on inter-cluster bonds. Sometimes there are no inter-cluster bonds: this is the case of monolayers of fullerene molecules (C_{12} and C_{20}) studied in [15]. There, the extended interaction is treated exactly within the fullerene molecule which makes up the cluster, whereas the inter-molecular hybridization is treated by CPT. A similar path is followed in [31], where the spectral functions of a $j = \frac{5}{2}$ multiplet of $5f$ atomic states are computed using CPT. The cluster is then made of a single lattice site, with 6 atomic states and the Coulomb repulsion between states attached to different sites is neglected. Likewise, the inter-band Coulomb interaction introduced in the electron–hole Hubbard model [32] fits perfectly the standard CPT framework.

The studies of stripes carried out in [29,33] using CPT did not involve an explicit treatment of the Coulomb interaction, as the stripes were supposed static. On the other hand, [34] did carry out such a mean-field decomposition to study charge ordering in the extended Hubbard model with the help of the VCA. Let us discuss how such a mean-field decomposition can be carried out. The extended interaction for the whole lattice may be written as

$$H_{\text{ext}} = \sum_{(ij)} V_{ij} n_i n_j , \tag{8.97}$$

where the sum is carried of pairs (i, j) of sites and $n_i = n_{i\uparrow} + n_{i\downarrow}$ is the electron density at site i (we suppose a one-band model for simplicity). Let us write the Coulomb matrix V_{ij} as a sum of an intra-cluster part $V_{ij}^{(1)}$ and an inter-cluster part $V_{ij}^{(2)}$. The latter is then treated in the Hartree approximation and factorized:

$$H_{\text{ext}}^{\text{MF}} = \sum_{(ij)} V_{ij}^{(1)} n_i n_j + \sum_{(ij)} V_{ij}^{(2)} \big(n_i \bar{n}_j + n_j \bar{n}_i - \bar{n}_i \bar{n}_j \big), \tag{8.98}$$

where the Weiss fields $\bar{n}_i$ can be set by the self-consistency condition $\bar{n}_i = \langle n_i \rangle$. Since the matrix $\boldsymbol{V}^{(2)}$ is real and symmetric, it can be diagonalized by an orthogonal matrix: $\boldsymbol{V}^{(2)} = \boldsymbol{O}\boldsymbol{\Lambda}\boldsymbol{O}^T$. One then defines eigen-operators $O_a = \sum_j O_{aj} n_j$ and eigen-fields $h_a = \lambda_a \sum_j O_{aj} \bar{n}_j$ such that

$$H_{\text{ext}}^{\text{MF}} = \sum_{(ij)} V_{ij}^{(1)} n_i n_j + \sum_a \left\{ h_a O_a - \frac{h_a^2}{2\lambda_a} \right\} . \tag{8.99}$$

The self-consistency condition is then $h_a = \lambda_a \langle O_a \rangle$.

Some of the operators O_a break spatial symmetries, and their introduction hints at the possibility of treating broken symmetry states within CPT itself. But the VCA or DMFT approaches are much better approaches to broken symmetries and should be pursued instead. In particular, the mean-field decomposition described above applies just as well to VCA and C-DMFT.

8.8.4 Phonons

The CPT can also be applied to phonon-mediated interactions, again provided that the phonon interaction be local, which restricts us to the Holstein or the Holstein–Hubbard model. Zhao et al. [35] applies CPT to the one-dimensional Holstein model, which involves dispersionless phonons only:

$$H = -t \sum_i \left(c_{i\sigma}^\dagger c_{i\sigma} + \text{H.c.} \right) + \omega_0 \sum_i b_i^\dagger b_i - \gamma \sum_i n_i (b_i^\dagger + b_i), \qquad (8.100)$$

where b_i annihilates a phonon at site i, ω_0 is the phonon frequency and γ the electron–phonon coupling. From a computational point of view, the difficulty with phonons is that they require an infinite-dimensional Hilbert space, and thus some truncation is needed. Zhao et al. [35] uses an optimized phonon approach [36–38] for that purpose. At fixed ω_0, this systems undergoes a transition from a normal state to a charge-density state wave (CDW) at a certain critical coupling g_c. The CDW state has a gap clearly visible in the CPT spectrum [35]. This is a case of spontaneous breaking of a discrete symmetry that occurs at the same time in the phonon and electron system.

Adding a Hubbard interaction to (8.100) yields the Holstein–Hubbard model. As shown in [39], the electron–phonon interaction has a sizeable effect on the electron spectral function, eventually suppressing the holon peak visible in the absence of phonons, while keeping the spin peak relatively intact. The authors argue that the ARPES spectra of the quasi-one-dimensional organic conductor TTF-TCNQ reported in [40] cannot be correctly interpreted with the Hubbard model alone, but requires the introduction of phonons.

Acknowledgments The author would like to thank the following people for discussions which, over the years, have strengthened and widened his understanding of quantum cluster methods: M. Civelli, G. Kotliar, B. Kyung, M. Jarrell, Th. Maier, S. Okamoto, D. Plouffe, M. Potthoff, A-M. Tremblay, and C. Weber. Computational resources for this review were provided by RQCHP and Compute Canada.

References

1. S. Pairault, D. Sénéchal, A.M.S. Tremblay, Phys. Rev. Lett. **80**, 5389 (1998)
2. D. Sénéchal, D. Perez, D. Plouffe, Phys. Rev. B **66**, 075129 (2002)
3. C. Gros, R. Valenti, Phys. Rev. B **48**, 418 (1993)

4. D. Sénéchal, D. Perez, M. Pioro-Ladrière, Phys. Rev. Lett. **84**, 522 (2000)
5. S. Pairault, D. Sénéchal, A.M. Tremblay, Eur. Phys. J. B **16**, 85 (2000)
6. M. Potthoff, M. Aichhorn, C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003)
7. M.T. Tran, Phys. Rev. B **74**(15), 155121 (2006)
8. M. Aichhorn, E. Arrigoni, M. Potthoff, W. Hanke, Phys. Rev. B **74**(23), 235117 (2006)
9. A. Ruhe, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (SIAM, PA, 2000), chap. 4.1
10. E. Dagotto, Rev. Mod. Phys. **66**(3), 763 (1994)
11. R. Freund, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (SIAM, PA, 2000), chap. 4.6
12. N. Laflorencie, D. Poilblanc, in *Quantum Magnetism*. Lecture Notes in Physics, vol. 645 (Springer, Berlin, 2004), pp. 227–252
13. A. Weiße, G. Wellein, A. Alvermann, H. Fehske, Rev. Mod. Phys. **78**(1), 275 (2006)
14. F. Lin, E.S. Sørensen, C. Kallin, A.J. Berlinsky, Phys. Rev. B **75**(7), 075112 (2007)
15. F. Lin, E.S. Sørensen, C. Kallin, A.J. Berlinsky, J. Phys. Condens. Matter **19**(45), 456206 (2007)
16. G. Li, W. Hanke, A.N. Rubtsov, S. Bäse, M. Potthoff, Phys. Rev. B **80**(19), 195118 (2009)
17. G. Kotliar, S. Savrasov, G. Pálsson, G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
18. M. Capone, M. Civelli, S.S. Kancharla, C. Castellani, G. Kotliar, Phys. Rev. B **69**(19), 195105 (2004)
19. T.D. Stanescu, G. Kotliar, Phys. Rev. B **74**(12), 125110 (2006)
20. D. Sénéchal, A.M.S. Tremblay, Phys. Rev. Lett. **92**, 126401 (2004)
21. N. Armitage et al., Phys. Rev. Lett. **88**, 257001 (2002)
22. F. Ronning et al., Phys. Rev. B **67**, 165101 (2003)
23. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**(25), 1445 (1968)
24. C. Dahnken, E. Arrigoni, W. Hanke, J. Low Temp. Phys. **126**(3-4), 949 (2002)
25. E. Arrigoni, M. Aichhorn, M. Daghofer, W. Hanke, New J. Phys. **11**(5), 055066 (2009)
26. S.H. ans S Kudo, T. Shirakawa, Y. Ohta, J. Phys. Conf. Ser. **150**, 042060 (2009)
27. W.Z. Wang, Phys. Rev. B **73**(23), 235325 (2006)
28. M. Zacher, R. Eder, E. Arrigoni, W. Hanke, Int. J. Modern Phys. B **14**(29-31), 3783 (2000)
29. M. Zacher, R. Eder, E. Arrigoni, W. Hanke, Phys. Rev. Lett. **85**(12), 2585 (2000)
30. A.S. Ovchinnikov, I.G. Bostrem, V.E. Sinitsyn, Theor. Math. Phys. **162**, 179 (2010)
31. F. Pollmann, G. Zwicknagl, Phys. Rev. B **73**(3), 035121 (2006)
32. K. Asano, T. Nishida, T. Ogawa, Phys. Status Solidi (B) **245**, 2729 (2008)
33. M. Zacher, R. Eder, E. Arrigoni, W. Hanke, Phys. Rev. B **65**(4), 045109 (2002)
34. M. Aichhorn, H.G. Evertz, W. von der Linden, M. Potthoff, Phys. Rev. B **70**(23), 235107 (2004)
35. H. Zhao, C.Q. Wu, H.Q. Lin, Phys. Rev. B **71**(11), 115201 (2005)
36. C. Zhang, E. Jeckelmann, S.R. White, Phys. Rev. Lett. **80**(12), 2661 (1998)
37. C. Zhang, E. Jeckelmann, S.R. White, Phys. Rev. B **60**(20), 14092 (1999)
38. A. Weiße, H. Fehske, G. Wellein, A.R. Bishop, Phys. Rev. B **62**(2), R747 (2000)
39. W.Q. Ning, H. Zhao, C.Q. Wu, H.Q. Lin, Phys. Rev. Lett. **96**(15), 156402 (2006)
40. M. Sing, U. Schwingenschlögl, R. Claessen, P. Blaha, J. Carmelo, L. Martelo, P. Sacramento, M. Dressel, C. Jacobsen, Phys. Rev. B **68**(12), 125111 (2003)

Chapter 9
Dynamical Cluster Approximation

H. Fotso, S. Yang, K. Chen, S. Pathak, J. Moreno, M. Jarrell, K. Mikelsons, E. Khatami, and D. Galanakis

Abstract The dynamical cluster approximation (DCA) is a method which systematically incorporates nonlocal corrections to the dynamical mean-field approximation. Here we present a pedagogical discussion of the DCA by describing it as a Φ-derivable coarse-graining approximation in k-space, which maps an infinite lattice problem onto a periodic finite-sized cluster embedded in a self-consistently determined effective medium. We demonstrate the method by applying it to the two-dimensional Hubbard model. From this application, we show evidences of the presence of a quantum critical point (QCP) at a finite doping underneath the superconducting dome. The QCP is associated with the second-order terminus of a line of first order phase separation transitions. This critical point is driven to zero temperature by varying the band parameters, generating the QCP. The effect of the proximity of the QCP to the superconducting dome is also discussed.

9.1 Introduction

Some of the most exotic properties of materials, including high-temperature superconductivity, colossal magnetoresistance, and heavy Fermion and non-Fermi liquid

H. Fotso (✉) · S. Yang · K. Chen · S. Pathak · J. Moreno · M. Jarrell
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803, USA
e-mail: herbert.fotso@gmail.com; yangphysics@gmail.com; kchen5@lsu.edu; pathaksandeep@gmail.com; moreno@lsu.edu; jarrellphysics@gmail.com

K. Mikelsons · E. Khatami
Department of Physics, Georgetown University, Washington, DC 20057, USA
e-mail: karlis@physics.georgetown.edu; ehsankhatami@gmail.com

D. Galanakis
Division of Physics and Applied Physics, Nanyang Technological University of Singapore, 21 Nanyang Link, Singapore 637371
e-mail: dimitris.galanakis@gmail.com

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_9,

behaviors, are due to strong electronic correlations. The materials which display these properties are characterized by either narrow electronic bands or compact orbitals with large angular momentum in the valence shell. In either case, the potential energy associated with some of these electronic degrees of freedom is comparable to or larger than their electronic kinetic energy (bandwidth), which invalidates conventional perturbative approaches. Thus, we resort to mean-field approximations, which we solve numerically so that diagrams to all orders are included, and we use them to study simplified models representing these systems.

For example, the Hubbard model [1–3] is the simplest model of a correlated electronic lattice system. Together with the $t - J$ model, they are thought to at least qualitatively describe some of the properties of transition metal oxides, and high-temperature superconductors [4]. The periodic Anderson model along with various Kondo lattice models has been proposed to describe both the actinide and lanthanide heavy fermion systems and the Kondo insulators. The Holstein model incorporates the essential physics of strongly interacting electrons and phonons. All of these model Hamiltonians contain at least two major ingredients: a local interaction term and a nonlocal hopping term. For example, the Hubbard model Hamiltonian (see Fig. 9.1) is

$$H = -t \sum_{\langle j,k\rangle\sigma} (c^{\dagger}_{j\sigma} c_{k\sigma} + c^{\dagger}_{k\sigma} c_{j\sigma}) + \epsilon \sum_{j} (n_{j\uparrow} + n_{j\downarrow}) + U \sum_{j} (n_{j\uparrow} - 1/2)(n_{j\downarrow} - 1/2)\,, \tag{9.1}$$

where $c^{\dagger}_{j\sigma}$ ($c_{j\sigma}$) creates (destroys) an electron at a site j with a spin σ, $n_{i\sigma} = c^{\dagger}_{i\sigma} c_{i\sigma}$, t is the nearest neighbor hopping which sets the unit of energy and U is the on-site Coulomb repulsion between the electrons.

However, except for special limits, even such a simplified model like (9.1) cannot be solved exactly. For example, for the Hubbard model, no exact solutions exist except in one dimension, where the knowledge is in fact rather complete [5–7]. The periodic Anderson model is only solvable in the limit where the orbital degeneracy diverges [8], and the Holstein model is only solvable in the Eliashberg–Migdal limit where vertex corrections may be neglected. Clearly a new approach to these models is needed if nontrivial exact solutions are desired.

Metzner and Vollhardt [9, 10], Kuramoto [11, 12], and Müeller-Hartmann [13] suggested such a new approach based on a mean-field theory which becomes exact when the dimensionality $d = \infty$. The resulting formalism neglects dynamical

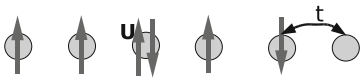

Fig. 9.1 Cartoon of the Hubbard model, characterized by a single band with near-neighbor hopping t, and local repulsion U

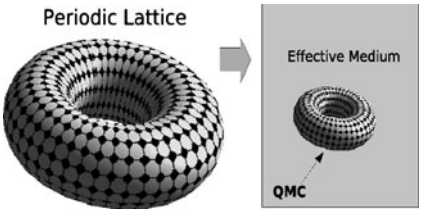

Fig. 9.2 Quantum cluster approaches, like the DMFA and DCA, map the infinite lattice problem onto a self-consistently embedded cluster problem

intersite correlations while retaining the important local dynamical correlations. The resulting formalism is called the dynamical mean field approximation (DMFA) since it may be employed in any dimension, but is only exact on infinite-dimensional lattices. In finite dimensions, the dynamical cluster approximation (DCA) is used to systematically study nonlocal corrections to the DMFA [14, 15]. Quantum cluster approaches, such as the DMFA and DCA, work by mapping an infinite periodic lattice onto a self-consistently embedded cluster problem, as illustrated in Fig. 9.2. Correlations up to the cluster size are treated explicitly, while those at longer length scales are treated in a mean-field level. The DMFA/DCA cluster problem may be solved by a variety of methods; however, quantum Monte Carlo (QMC) has been the first numerically exact solver employed [16] and remains the most powerful and adaptable method.

In this article, we will present a pedagogical discussion of the DCA and its relationship to the DMFA. In Sect. 9.2 we will first derive the DMFA as a coarse-graining approximation, extend this logic to derive the DCA and also provide a derivation from the Φ functional. In Sect. 9.3, we describe how physical quantities are calculated in this formalism. In Sect. 9.4 we will discuss applications of the DCA showing how it is used to find evidence for a quantum critical point (QCP) underneath the superconducting dome and to investigate the nature of this QCP and finally to study the relationship between superconductivity and the QCP.

9.2 The Dynamical Mean Field and Cluster Approximations

9.2.1 The Dynamical Mean-Field Approximation

The DCA algorithm can be derived in analogy with the DMFA. The DMFA is a local approximation which was used by Kuramoto in perturbative calculations as a simplification of the k-summations which render the problem intractable [11, 12]. But it was after the work of Metzner and Vollhardt [9] and Müller-Hartmann [13], who showed that this approximation becomes exact in the limit of infinite dimensions, that it received extensive attention. In this approximation, one neglects the spatial dependence of the self-energy, retaining only its variation with time. See the reviews by Pruschke et al. [17] and Georges et al. [18] for a more extensive treatment.

In this section, we will show that it is possible to re-interpret the DMFA as a coarse-graining approximation. The DMFA consists of mapping the original lattice problem to a self-consistent impurity problem. This is equivalent to averaging the Green functions used to calculate the irreducible diagrammatic insertions over the Brillouin zone. An important consequence of this averaging is that the self-energy and the irreducible vertices of the lattice are independent of the momentum. Hence, they are those of the impurity.

Müller-Hartmann [13] showed that this coarse-graining becomes exact in the limit of infinite dimensions. For Hubbard-like models, the properties of the bare vertex are completely characterized by the Laue function Δ which expresses the momentum conservation at each vertex. In a conventional diagrammatic approach

$$\begin{aligned}\Delta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) &= \sum_{\mathbf{r}} \exp\left[\mathrm{i}\mathbf{r} \cdot (\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4)\right] \\ &= N\delta_{\mathbf{k}_1+\mathbf{k}_2,\mathbf{k}_3+\mathbf{k}_4}, \end{aligned} \tag{9.2}$$

where $\mathbf{k}_1$ and $\mathbf{k}_2$ ($\mathbf{k}_3$ and $\mathbf{k}_4$) are the momenta entering (leaving) each vertex through its Green function legs. However, as the dimensionality $D \to \infty$ Müller-Hartmann showed that the Laue function reduces to [13]

$$\Delta_{D\to\infty}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = 1 + \mathcal{O}(1/D). \tag{9.3}$$

The DMFA assumes the same Laue function, $\Delta_{\mathrm{DMFA}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = 1$, even in the context of finite dimensions. Thus, the conservation of momentum at internal vertices is neglected and we may freely sum over the internal momentum labels of each Green function leg. This leads to a collapse of the momentum dependent contributions and only local terms remain.

This argument may then be applied to the generating functional Φ, which is the sum over all closed connected compact graphs constructed from the dressed Green's function G and the bare interaction.The second-order contribution to Φ for a Hubbard-like model is illustrated in Fig. 9.3. The self-energy Σ may be obtained from a functional derivative of Φ with respect to the Green's function G, which effectively removes one of the Green's function lines (Fig. 9.4).

The perturbative series for Φ, Σ and the irreducible vertices Γ in the DMFA are identical to those of the corresponding impurity model, so that conventional

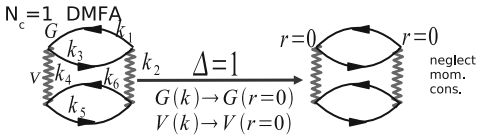

Fig. 9.3 The second-order contribution to the generating functional Φ. As we apply the DMFA coarse-graining approximation, (9.3), Φ becomes a functional of the local Green's function and interaction

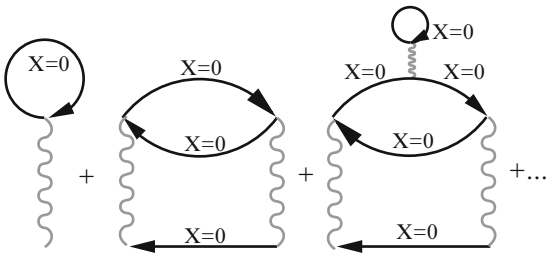

Fig. 9.4 The DMFA self-energy contains only local corrections. See, e.g., the third graph. To prevent overcounting these contributions, the local self-energy must be excluded, cf. (9.4), from the Green's function line used in most cluster solvers

impurity solvers may be used. However, since most impurity solvers can be viewed as methods that sum all the graphs, not just the skeletal ones, it is necessary to exclude Σ from the local propagator input, $\mathcal{G}$, to the impurity solver, in order to avoid overcounting local self-energy contributions. Therefore, in Matsubara frequencies

$$\mathcal{G}(i\omega_n)^{-1} = G(i\omega_n)^{-1} + \Sigma(i\omega_n), \tag{9.4}$$

where $i\omega_n = (2n+1)\pi T$, $\Sigma(i\omega_n)$ is the self-energy and $G(i\omega_n)$ the full local Green's function. Hence, in the local approximation, the Hubbard model has the same diagrammatic expansion as an Anderson impurity with a bare local propagator $\mathcal{G}(i\omega_n; \Sigma)$, which is determined self-consistently.

An algorithm constructed from this approximation is the following: (1) An initial guess for $\Sigma(i\omega_n)$ is chosen (usually from the perturbation theory). (2) $\Sigma(i\omega_n)$ is used to calculate the corresponding local Green's function

$$G(i\omega_n) = \int \mathrm{d}\eta \frac{\rho^0(\eta)}{i\omega_n - (\eta - \mu) - \Sigma(i\omega_n)}, \tag{9.5}$$

where ρ^0 is the noninteracting density of states, and μ is the chemical potential. (3) Starting from $G(i\omega_n)$ and $\Sigma(i\omega_n)$ used in the second step, the host Green's function $\mathcal{G}(i\omega_n)^{-1} = G(i\omega_n)^{-1} + \Sigma(i\omega_n)$ is calculated, which serves as the bare Green's function of the impurity model. (4) Starting with $\mathcal{G}(i\omega_n)$, the local Green's function $G(i\omega_n)$ is obtained using the QMC method (or another technique). (5) Using the QMC output for the cluster Green's function $G(i\omega_n)$ and the host Green's function $\mathcal{G}(i\omega_n)$ from the third step, a new $\Sigma(i\omega_n) = \mathcal{G}(i\omega_n)^{-1} - G(i\omega_n)^{-1}$ is calculated, which is then used in step (2) to reinitialize the process. Steps (2)–(5) are repeated until convergence is reached. If in step (4) the QMC algorithm of Hirsch and Fye [19, 20] is used to compute the local Green's function $G(\tau)$ or other physical quantities in imaginary time, local dynamical quantities are then calculated by analytically continuing the corresponding imaginary-time quantities using the maximum-entropy method (MEM) [21].

9.2.2 The Dynamical Cluster Approximation

Like the DMFA, the DCA may be intuitively motivated with a coarse-graining transformation. In the DMFA, the propagators used to calculate Φ and its derivatives were coarse-grained over the entire Brillouin zone, leading to local (momentum independent) irreducible quantities. In the DCA, we wish to relax this condition, and systematically restore momentum conservation and nonlocal corrections. Thus, in the DCA, the reciprocal space of the lattice which contains N points is divided into N_c cells of identical linear size Δk. The coarse-graining transformation is set by averaging the Green function within each cell. If $N_c = 1$ the original lattice problem is mapped to an impurity problem (DMFA). If N_c is larger than one, then nonlocal corrections of length $\approx \pi/\Delta k$ to the DMFA are introduced. Provided that the propagators are sufficiently weakly momentum dependent, this is a good approximation. If N_c is chosen to be small, the cluster problem can be solved using conventional techniques such as the QMC, the non-crossing approximation (NCA) or the fluctuation exchange approximation (FLEX). This averaging process also establishes a relationship between the systems of size N and N_c. A simple choice, which will be discussed in Sect. 9.2.3, is to equate the irreducible quantities (self-energy, irreducible vertices) of the cluster to those in the lattice.

This coarse-graining procedure and the relationship of the DCA to the DMFA is illustrated by a microscopic diagrammatic derivation of the DCA. The DCA systematically restores the momentum conservation at internal vertices relinquished by the DMFA. The Brillouin zone is divided into $N_c = L^D$ cells of linear size $\Delta k = 2\pi/L$ (cf. Fig. 9.5 for $N_c = 8$). Each cell is represented by a cluster momentum $\mathbf{K}$ in the center of the cell. We require that momentum conservation be (partially) observed for momentum transfers between cells, i.e., for momentum transfers larger than Δk, but neglected for momentum transfers within a cell, i.e., less than Δk. This requirement can be established by using the Laue function [15]

$$\Delta_{\mathrm{DCA}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = N_c \delta_{\mathbf{M}(\mathbf{k}_1)+\mathbf{M}(\mathbf{k}_2),\mathbf{M}(\mathbf{k}_3)+\mathbf{M}(\mathbf{k}_4)}, \tag{9.6}$$

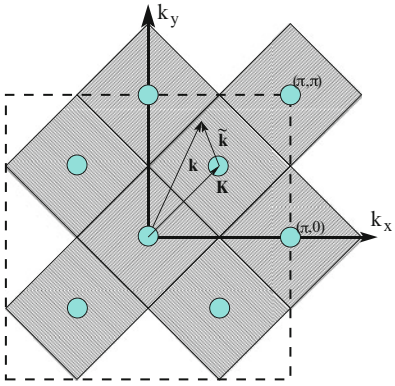

Fig. 9.5 Coarse-graining cells for $N_c = 8$ (differentiated by alternating fill patterns) that partition the first Brillouin zone (*dashed line*). Each cell is centered on a cluster momentum $\mathbf{K}$ (*filled circles*). To construct the DCA cluster, we map a generic momentum in the zone such as $\mathbf{k}$ to the nearest cluster point $\mathbf{K} = \mathbf{M}(\mathbf{k})$ so that $\tilde{\mathbf{k}} = \mathbf{k} - \mathbf{K}$ remains in the cell around $\mathbf{K}$

where $\mathbf{M}(\mathbf{k})$ is a function which maps $\mathbf{k}$ onto the momentum label $\mathbf{K}$ of the cell containing $\mathbf{k}$ (see Fig. 9.5). This choice for the Laue function systematically interpolates between the exact result, (9.2), which it recovers when $N_c \to N$ and the DMFA result, (9.3), which it recovers when $N_c = 1$. With this choice of the Laue function the momenta of each internal leg may be freely summed over the cell.

This is illustrated for the second-order term in the generating functional in Fig. 9.6. Each internal leg $G(\mathbf{k})$ in a diagram is replaced by the coarse-grained Green function $\bar{G}(\mathbf{M}(\mathbf{k}))$, defined by

$$\bar{G}(\mathbf{K}) \equiv \frac{N_c}{N} \sum_{\tilde{\mathbf{k}}} G(\mathbf{K} + \tilde{\mathbf{k}}), \tag{9.7}$$

where N is the number of points of the lattice, N_c is the number of cluster $\mathbf{K}$ points, and the $\tilde{\mathbf{k}}$ summation runs over the momenta of the cell about the cluster momentum $\mathbf{K}$ (see Fig. 9.5). The diagrammatic consequences for the generating functional and its derivatives are unchanged; however, the complexity of the problem is greatly reduced since $N_c \ll N$.

9.2.3 Φ Derivability

The coarse-graining approximation can be applied to the generating functional Φ. The generating functional is the sum over all of the closed connected compact diagrams, such as the one shown in Fig. 9.6. It is defined as

$$\Phi(G) = \sum_{l,\sigma} p_l tr\left[\Sigma_\sigma^l G_\sigma\right]. \tag{9.8}$$

The trace indicates summation over frequency, momentum and spin. Here, Σ_σ^l is the set of irreducible self-energy diagrams of lth order in the interaction, G_σ is the dressed Green function related to Σ_σ by the Dyson equation $G_\sigma^{-1} = G_\sigma^{0-1} - \Sigma_\sigma$,

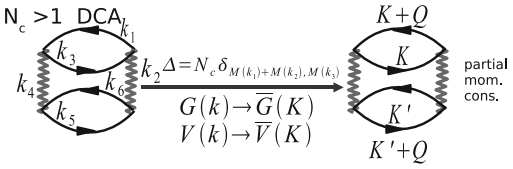

Fig. 9.6 A second-order term in the generating functional of the Hubbard model. Here the *undulating line* represents the interaction U, and on the LHS (RHS) the *solid line* the lattice (coarse-grained) single-particle Green functions. When the DCA Laue function is used to describe momentum conservation at the internal vertices, the momenta collapse onto the cluster momenta and each lattice Green function is replaced by the coarse-grained result

where G^0 is the noninteracting Green function, p_l is a counting factor equal to the number of occurrences of G_σ in each term (for Hubbard-like models, $p_l = 1/l$). The free energy F can be expressed as

$$F = -k_B T \left(\Phi\left(G\right) - \mathrm{tr}\left[\Sigma_\sigma G_\sigma\right] - \mathrm{tr}\ln\left[-G_\sigma\right]\right). \tag{9.9}$$

With the above definition, it holds that $\Sigma_\sigma = \delta\Phi/\delta G_\sigma$, as required for a "$\Phi$-derivable" theory, and the free energy is stationary under variations of G. In addition, the irreducible vertex function is obtained by a second variation of Φ, $\Gamma_{\sigma\sigma'} = \delta^2\Phi/\delta G_\sigma \delta G_{\sigma'} = \delta\Sigma_\sigma/\delta G_{\sigma'}$.

The DCA can be microscopically motivated by the choice of the Laue function Δ_{DCA} (1.6). Within this formalism, the effect of the chosen Laue function is the replacement of the self-energy Σ_σ and the irreducible vertex $\Gamma_{\sigma,\sigma'}$ by the corresponding coarse-grained quantities (indicated here by the bars). Consider for example the Schwinger–Dyson equation relating the self-energy to the two-particle reducible vertex $T^{(2)}$, $\Sigma = GGGT^{(2)}$. The vertices connecting the Green function to $T^{(2)}$ do not preserve momentum conservation within the cells about the cluster momentum due to the DCA Laue function. Consequently, the lattice Green function G_σ is replaced by the coarse-grained Green function $\bar{G}_\sigma$. The external momentum label $(\mathbf{k})$ of the self-energy is in principle still a lattice momentum; however, the self-energy will only depend on $\mathbf{k}$ through the function $\mathbf{M}(\mathbf{k})$. If we use this self-energy in the calculation of its contribution to the Φ functional, the Laue function on the vertices will "reduce" both the self-energy as well as the closing Green function to their corresponding coarse-grained expressions. Consequently, the DCA Φ functional reads

$$\Phi_{\mathrm{DCA}}\left(G\right) = \sum_l p_l tr\left[\bar{\Sigma}_\sigma^l \bar{G}_\sigma\right]. \tag{9.10}$$

In correspondence to the lattice system,

$$\frac{\delta\Phi_{\mathrm{DCA}}}{\delta\bar{G}_\sigma} = \bar{\Sigma}_\sigma = \frac{\delta\Phi_{\mathrm{DCA}}}{\delta G_\sigma}, \tag{9.11}$$

where the second equality follows since the variation $\delta/\delta G_\sigma$ corresponds to cutting a Green function line , so that $\delta\bar{G}_{\sigma\mathbf{K}}/\delta G'_\sigma\mathbf{k}' = \delta_{\mathbf{K},\mathbf{M}(\mathbf{k}')}\delta_{\sigma,\sigma'}$. It follows that the DCA estimate of the free energy is

$$F_{\mathrm{DCA}} = -k_B T \left(\Phi_{\mathrm{DCA}} - \mathrm{tr}\left[\Sigma_\sigma G_\sigma\right] - \mathrm{tr}\ln\left[-G_\sigma\right]\right), \tag{9.12}$$

F_{DCA} is stationary with respect to $\mathbf{G}_\sigma$ when

$$\frac{-1}{k_B T}\frac{\delta F_{\mathrm{DCA}}}{\delta G_\sigma(\mathbf{k})} = \bar{\Sigma}_\sigma(\mathbf{M}(\mathbf{k})) - \Sigma_\sigma(\mathbf{k}) = 0, \tag{9.13}$$

which means that $\Sigma(\mathbf{k}) = \bar{\Sigma}_\sigma(\mathbf{M}(\mathbf{k}))$ is the proper approximation for the lattice self-energy corresponding to $\Phi_{\rm DCA}$. The corresponding lattice single-particle propagator is then given by

$$G(\mathbf{k}, z) = \frac{1}{z - (\epsilon_{\mathbf{k}} - \mu) - \bar{\Sigma}(\mathbf{M}(\mathbf{k}), z)}, \tag{9.14}$$

where $\epsilon_{\mathbf{k}}$ is the quasiparticle energy, and μ the chemical potential. A similar procedure is used to construct the two-particle quantities needed to determine the phase diagram or the nature of the dominant fluctuations that can eventually destroy the high-temperature ground state. This procedure is a generalization of the method of calculating response functions in the DMFA [16, 22].

The introduction of the momentum dependence in the DCA self-energy allows one to detect some precursors to transitions which are absent in the DMFA; but for the actual determination of the nature of the instability, one needs to compute the response functions. These susceptibilities are thermodynamically defined as second derivatives of the free energy with respect to external fields. $\Phi_{\rm DCA}(G)$ and $\bar{\Sigma}_\sigma$, and hence $F_{\rm DCA}$ depend on these fields only through G_σ and the bare G^0_σ. Following Baym and Kadanoff [23], it is easy to verify that, the approximation

$$\Gamma_{\sigma,\sigma'} \approx \bar{\Gamma}_{\sigma,\sigma'} \equiv \delta\bar{\Sigma}_\sigma / \delta G_{\sigma'} \tag{9.15}$$

yields the same estimate that would be obtained from the second derivative of $F_{\rm DCA}$ with respect to the applied field. For example, the first derivative of the free energy with respect to a spatially homogeneous external magnetic field h is the magnetization,

$$m = \mathrm{Tr}\left[\sigma G_\sigma\right]. \tag{9.16}$$

The susceptibility is given by the second derivative,

$$\chi = \frac{\partial m}{\partial h} = \mathrm{Tr}\left[\sigma \frac{\partial G_\sigma}{\partial h}\right]. \tag{9.17}$$

We substitute $G_\sigma = \left(G^{0-1}_\sigma - \bar{\Sigma}_\sigma\right)^{-1}$, and evaluate the derivative,

$$\chi = \mathrm{Tr}\left[\sigma \frac{\partial G_\sigma}{\partial h}\right] = \mathrm{Tr}\left[G^2_\sigma \left(1 + \sigma \frac{\partial \bar{\Sigma}_\sigma}{\partial G_{\sigma'}} \frac{\partial G_{\sigma'}}{\partial h}\right)\right]. \tag{9.18}$$

We can generalize this argument to include the staggered susceptibility by identifying $\chi_{\sigma,\sigma'} = \sigma \frac{\partial G_{\sigma'}}{\partial h}$, and $\chi_{\rm st} = Tr[\chi_{\sigma,-\sigma}]$ and $\chi^0_\sigma = G^2_\sigma$. By collecting all the terms within both traces, and sum over the cell momenta $\tilde{\mathbf{k}}$, we obtain the two-particle Dyson's equation

$$\begin{aligned} &2\left(\bar{\chi}_{\sigma,\sigma} - \bar{\chi}_{\sigma,-\sigma}\right) \\ &\quad = 2\bar{\chi}^0_\sigma + 2\bar{\chi}^0_\sigma \left(\bar{\Gamma}_{\sigma,\sigma} - \bar{\Gamma}_{\sigma,-\sigma}\right)\left(\bar{\chi}_{\sigma,\sigma} - \bar{\chi}_{\sigma,-\sigma}\right). \end{aligned} \tag{9.19}$$

We see that again it is the irreducible quantity, i.e., the vertex function, for which the cluster and lattice quantities are equal.

9.2.4 Algorithm

A variety of techniques may be used to sum the cluster diagrams in order to calculate the cluster self-energy, Σ_c, and the cluster vertex function, Γ_c. In the past, we have used the QMC [24], the non-crossing approximation [25] or the Fluctuation-Exchange approximation [26]. Here, we will mainly use QMC techniques. Since QMC is systematically exact; i.e., it effectively sums all diagrams to all orders, care must be taken when defining the initial Green function (the solid lines in Fig. 9.6) to avoid overcounting diagrams on the cluster. For example, to fourth order and higher in perturbation theory for the self-energy, nontrivial self-energy corrections enter in the diagrammatic expansion for the cluster self energy of the Hubbard model. To avoid overcounting these contributions, we must first subtract off the self-energy corrections on the cluster from the Green function line used to calculate Σ_c and its functional derivatives. This cluster-excluded Green function is given by

$$\frac{1}{\mathcal{G}(\mathbf{K}, z)} = \frac{1}{\bar{G}(\mathbf{K}, z)} + \Sigma_c(\mathbf{K}, z) \tag{9.20}$$

which is the coarse-grained Green function with correlations on the cluster excluded. Since $\Sigma_c(\mathbf{K}, z)$ is not known a priori, it must be determined self-consistently, starting from an initial guess, usually from the perturbation theory. This guess is used to calculate $\bar{G}$ from (9.7). $\mathcal{G}(\mathbf{K}, z)$ is then calculated with (9.20), and it is used to initialize the QMC calculation. The QMC estimate for the cluster self energy is then used to calculate a new estimate for $\bar{G}(\mathbf{K})$ using (9.7). The corresponding $\mathcal{G}(\mathbf{K})$ is used to reinitialize the procedure which continues until $G_c = \bar{G}$ and the self-energy converges to the desired accuracy.

One of the difficulties encountered in earlier attempts to include nonlocal corrections to the DMFA was that these methods were not causal [27, 28]. The spectral weight was not conserved and the imaginary parts of the one-particle retarded Green functions and self-energies were not negative definite as required by causality. The DCA algorithm presented in this subsection does not present these problems. This algorithm is fully causal as shown by Hettler et al. [15]. They analyze the different steps of the self-consistent loop and found that none of them breaks the causality of the Green functions. Starting from the QMC block, one can see that if the input $\mathcal{G}$ is causal, since the QMC algorithm is essentially exact, the output G_c will also be causal. Then the corresponding $\Sigma_c(\mathbf{K}, i\omega_n)$ is causal. This in turn ensures that the coarse-grained Green function $\bar{G}(\mathbf{K}, i\omega_n)$ also fulfills causality. The only nontrivial operation which may break causality is the calculation of $\mathcal{G}(\mathbf{K}, i\omega_n)$. Hettler et al. [15] used a geometric proof to show that even this part of the loop respects causality.

9.3 Physical Quantities

Most experiments measure quantities which can be expressed as reducible one or two-particle Green's functions. As discussed above, the appropriate way to calculate these quantities is to first extract the corresponding irreducible quantity from the cluster calculation, and then use it to calculate the reducible quantity. For example, to calculate the single-particle Green's function (relevant for angle-resolved photoemission spectroscopy) we first extract the cluster self-energy and use the Dyson equation to construct the lattice Green's function. To calculate the phase diagram, we calculate the irreducible vertices in the different scattering channels Γ, and insert them into the Bethe–Salpeter equations for the lattice. In this subsection we will provide more details about the relationship between the lattice and cluster two-particle Green's functions and describe how a lattice susceptibility may be calculated efficiently.

9.3.1 Particle–Hole Channel

As a specific example, we will describe the calculation of the two-particle particle–hole Green's function

$$\begin{aligned}\chi_{\sigma,\sigma'}(q,k,k') = \int_0^\beta \int_0^\beta \int_0^\beta \int_0^\beta \mathrm{d}\tau_1 \mathrm{d}\tau_2 \mathrm{d}\tau_3 \mathrm{d}\tau_4 \\ \times \mathrm{e}^{\mathrm{i}\left((\omega_n+\nu_n)\tau_1-\omega_n\tau_2+\omega_{n'}\tau_3-(\omega_{n'}+\nu_n)\tau_4\right)} \\ \times \langle T_\tau c^\dagger_{\mathbf{k}+\mathbf{q}\sigma}(\tau_1) c_{\mathbf{k}\sigma}(\tau_2) c^\dagger_{\mathbf{k}'\sigma'}(\tau_3) c_{\mathbf{k}'+\mathbf{q}\sigma'}(\tau_4)\rangle,\end{aligned}$$

where we adopt the conventional notation [29] $k = (\mathbf{k}, i\omega_n)$, $k' = (\mathbf{k}, \omega'_n)$, $q = (\mathbf{q}, \nu_n)$ and T_τ is the time ordering operator.

$\chi_{\sigma,\sigma'}(q,k,k')$ and $\Gamma_{\sigma,\sigma'}(q,k,k')$ are related to each other through the Bethe–Salpeter equation (Fig. 9.7):

$$\begin{aligned}\chi_{\sigma,\sigma'}(q,k,k') = \chi^0_{\sigma,\sigma'}(q,k,k') + \chi^0_{\sigma,\sigma''}(q,k,k'') \\ \times \Gamma_{\sigma'',\sigma'''}(q,k'',k''')\chi_{\sigma''',\sigma'}(q,k''',k'),\end{aligned} \tag{9.21}$$

where frequency labels have been suppressed, and $\Gamma_{\sigma,\sigma'}(q,k,k')$ is the two-particle irreducible vertex which is the analogue of the self-energy, $\chi^0_{\sigma,\sigma'}(q,k,k'')$ is the noninteracting susceptibility constructed from a pair of fully dressed single-particle Green's functions. As usual, a summation is to be made for repeated indices.

We now make the DCA substitution $\Gamma_{\sigma,\sigma'}(q,k,k') \to \Gamma_{\mathrm{c}\sigma,\sigma'}(q,\mathbf{M}(k),\mathbf{M}(k'))$ in (9.21). We ultimately want to sum over all k and k' to calculate the susceptibility at q. Note that after the DCA substitution only the bare and dressed two-particle

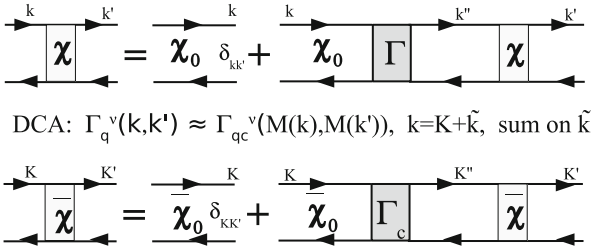

Fig. 9.7 The Bethe–Salpeter equation in the DCA. We approximate the lattice irreducible vertex Γ^{ν} by the $\Gamma_{\mathrm{c}}^{\nu}$ from the DCA cluster and coarse-grain over the momentum $\tilde{\mathbf{k}}$. The remaining equation is a function only of the cluster momenta $\mathbf{K}$ and may be solved by inversion

Green's functions χ depend upon the momenta $\tilde{\mathbf{k}}$ within a cell. Since χ and χ^0 in the product on the RHS of (9.21) share no common momentum labels, we may now freely sum over the momenta $\tilde{\mathbf{k}}$ within a cell, yielding

$$\bar{\chi}_{\sigma,\sigma'}(q,K,K') = \bar{\chi}^0_{\sigma,\sigma'}(q,K,K') + \bar{\chi}^0_{\sigma,\sigma''}(q,K,K'') \times \Gamma_{\mathrm{c}\sigma'',\sigma'''}(q,K'',K''')\bar{\chi}_{\sigma''',\sigma'}(q,K''',K'). \tag{9.22}$$

By coarse-graining the Bethe–Salpeter equation, we have greatly reduced its complexity; each of the matrices above is sufficiently small that they may be easily manipulated using standard techniques.

In contrast with the single-particle case where the coarse-grained quantities are identical to those of the cluster, $\chi_{\mathrm{c}\sigma,\sigma'}(q,K,K')$ is not equal to $\bar{\chi}_{\sigma,\sigma'}(q,K,K')$. This is because the self-consistency is made only at the single-particle level. Unlike the single-particle case where both $\Sigma(K)$ and $\bar{G}(K)$ are directly calculated, only the cluster susceptibility is calculated by the cluster solver, neither $\Gamma_{\sigma,\sigma'}(q,K,K')$ nor the coarse-grained susceptibility $\bar{\chi}_{\sigma,\sigma'}(q,K,K')$ is calculated during the self-consistency. Instead, the coarse-grained noninteracting susceptibility $\bar{\chi}^0_{\sigma,\sigma'}(q,K,K')$ is calculated in a separate program after the DCA converges using the following relation

$$\bar{\chi}^0_{\sigma,\sigma'}[(\mathbf{q},i\nu_n);(\mathbf{K},i\omega_n);(\mathbf{K}',i\omega_n')] = \delta_{\sigma,\sigma'}\delta_{\mathbf{K},\mathbf{K}'}\delta_{\omega_n,\omega_n'}\frac{N_{\mathrm{c}}}{N}\sum_{\tilde{\mathbf{k}}} G_{\sigma}(\mathbf{K}+\tilde{\mathbf{k}},i\omega_n) \times G_{\sigma}(\mathbf{K}+\tilde{\mathbf{k}}+\mathbf{q},i\omega_n+\nu_n). \tag{9.23}$$

The vertex function is extracted by inverting the cluster two-particle Bethe–Salpeter equation

$$\chi_{\mathrm{c}\sigma,\sigma'}(q,K,K') = \chi^0_{\mathrm{c}\sigma,\sigma'}(q,K,K') + \chi^0_{\mathrm{c}\sigma,\sigma''}(q,K,K'') \times \Gamma_{\mathrm{c}\sigma'',\sigma'''}(q,K'',K''')\chi_{\mathrm{c}\sigma''',\sigma'}(q,K''',K'). \tag{9.24}$$

If we combine (9.24) and (9.22), then the coarse-grained susceptibility may be obtained after the elimination of $\Gamma(\mathbf{q}, \mathbf{K}, \mathbf{K}')$ between the two equations. It reads

$$\bar{\chi}^{-1} = \chi_{\mathrm{c}}^{-1} - \chi_{\mathrm{c}}^{0^{-1}} + \bar{\chi}^{0^{-1}},$$

where, for example, $\bar{\chi}$ is the matrix formed from $\bar{\chi}_{\sigma,\sigma'}(\mathbf{q}, \mathbf{K}, \mathbf{K}')$ for fixed $\mathbf{q}$. The charge (ch) and spin (sp) susceptibilities $\chi_{\mathrm{ch,sp}}(\mathbf{q}, T)$ are deduced from $\bar{\chi}$

$$\chi_{\mathrm{ch,sp}}(q, T) = \frac{(k_{\mathrm{B}}T)^2}{N_{\mathrm{c}}^2} \sum_{KK'\sigma\sigma'} \lambda_{\sigma\sigma'} \bar{\chi}_{\sigma,\sigma'}(q, K, K'), \tag{9.25}$$

where $\lambda_{\sigma\sigma'} = 1$ for the charge channel and $\lambda_{\sigma\sigma'} = \sigma\sigma'$ for the spin channel.

9.3.2 Particle–Particle Channel

The calculation of susceptibilities in the particle–particle channel is essentially identical to the above equation. The exception to this rule occurs when we calculate susceptibilities for transitions to states of lower symmetry than the lattice symmetry. For example, in order to obtain the pairing susceptibility of the desired symmetry (s, p, d), the two-particle Green's function must be multiplied by the corresponding form factors $g(\mathbf{k})$ and $g(\mathbf{k}')$. In the study of the Hubbard model below, we will be particularly interested in $g(\mathbf{k}) = 1$ (s wave), $g(\mathbf{k}) = \cos(k_x) + \cos(k_y)$ (extended s wave) and $g(\mathbf{k}) = \cos(k_x) - \cos(k_y)$ ($d_{x^2-y^2}$ wave). These symmetries have been evoked as possible candidates for the superconducting ground state of cuprate superconductors (Fig. 9.8).

These factors modify the Bethe–Salpeter equations

$$\begin{aligned} g(\mathbf{k})\chi(q,k,k')g(\mathbf{k}') = {} & g(\mathbf{k})\chi^0(q,k,k')g(\mathbf{k}') + g(\mathbf{k})\chi^0(q,k,k'') \\ & \times \Gamma(q,k'',k''') \times \chi(q,k''',k')g(\mathbf{k}'), \end{aligned} \tag{9.26}$$

where

$$\begin{aligned} \chi(q,k,k') = {} & \int_0^\beta \int_0^\beta \int_0^\beta \int_0^\beta \mathrm{d}\tau_1 \mathrm{d}\tau_2 \mathrm{d}\tau_3 \mathrm{d}\tau_4 \\ & \times \mathrm{e}^{\mathrm{i}\left((\omega_n+\nu_n)\tau_1 - \omega_n\tau_2 + \omega_{n'}\tau_3 - (\omega_{n'}+\nu_n)\tau_4\right)} \\ & \times \langle T_\tau c^\dagger_{\mathbf{k}+\mathbf{q}\sigma}(\tau_1) c^\dagger_{-\mathbf{k}-\sigma}(\tau_2) c_{-\mathbf{k}'-\sigma}(\tau_3) c_{\mathbf{k}'+\mathbf{q}\sigma}(\tau_4) \rangle. \end{aligned} \tag{9.27}$$

On the LHS, we have dropped the spin indices since we will consider only opposite-spin pairing. Equation (9.26) cannot be easily solved if it is coarse-grained, since this will partially convolve $\chi(q, k, k')$ with *two* factors of g on the LHS and

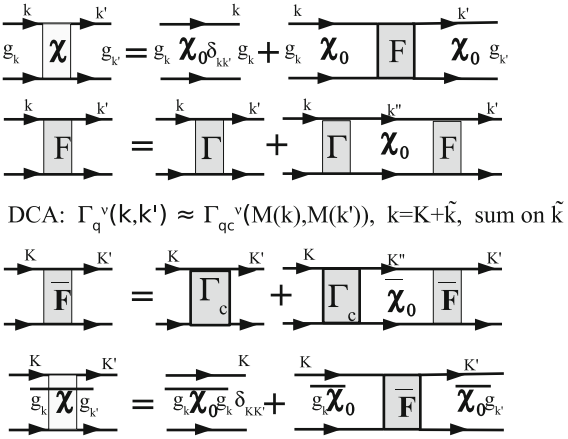

Fig. 9.8 Calculation of particle–particle projected susceptibilities. Often we want to calculate a projected particle–particle susceptibility (e.g., d-wave, with $g_{\mathbf{k}} = \cos(k_x) - \cos(k_y)$). Here the Bethe–Salpeter equation is rewritten in terms of the reducible vertex F. We approximate the lattice irreducible vertex Γ^ν by the Γ_c^ν from the DCA cluster and coarse-grain over the $\tilde{\mathbf{k}}$. Then the projected bare bubbles are calculated, and the remaining equation is a function of the cluster $\mathbf{K}$ only and may be solved by inversion

one factor on the RHS. Hence for the pairing susceptibilities, or for any situation where nontrivial form factors must be used, we use the equivalent equation involving the reducible vertex F (instead of the irreducible vertex Γ)

$$\begin{aligned} g(\mathbf{k})\chi(q,k,k')g(\mathbf{k}') = {} & g(\mathbf{k})\chi^0(q,k,k')g(\mathbf{k}') \\ & + g(\mathbf{k})\chi^0(q,k,k'') \\ & \times F(q,k'',k''')\chi^0(q,k''',k')g(\mathbf{k}'), \end{aligned} \tag{9.28}$$

where

$$\begin{aligned} F(q,k,k') = {} & \Gamma(q,k,k') \\ & + \Gamma(q,k,k'')\chi^0(q,k'',k''')\Gamma(q,k''',k') + \cdots \end{aligned} \tag{9.29}$$

We define

$$\Pi_{g,g}(q,k,k') = g(\mathbf{k})\chi(q,k,k')g(\mathbf{k}') \tag{9.30}$$

$$\Pi^0_{g,g}(q,k,k') = g(\mathbf{k})\chi^0(q,k,k')g(\mathbf{k}') \tag{9.31}$$

$$\Pi^0_{g}(q,k,k') = g(\mathbf{k})\chi^0(q,k,k')\,. \tag{9.32}$$

The remaining steps of the calculation are similar to the particle–hole case. We invert the cluster particle–particle Bethe–Salpeter equation with $g = 1$ for the

cluster, in order to extract Γ_c. We then coarse-grain (9.29) and use Γ_c to calculate the coarse-grained $\bar{F} = \Gamma_c \left(1 - \bar{\chi}^0 \Gamma_c\right)^{-1}$. We then coarse-grain (9.28), and use the coarse-grained $\bar{F}$ to calculate the coarse-grained $\bar{\Pi}_{g,g}$

$$\bar{\Pi}_{g,g}(q, K, K') = \bar{\Pi}^0_{g,g}(q, K, K') + \bar{\Pi}^0_g(q, K, K'')\bar{F}(q, K'', K''')\bar{\Pi}^0_g(q, K''', K'). \quad (9.33)$$

The pairing susceptibility of a desired symmetry is given by

$$P_g(q, T) = \frac{(k_B T)^2}{N_c^2} \sum_{K,K'} \bar{\Pi}_{gg}(q, K, K'). \quad (9.34)$$

9.4 DCA and Quantum Criticality in the Hubbard Model

9.4.1 Evidence of the Quantum Critical Point at Optimal Doping

The phase diagram of the hole-doped cuprates exhibits some unusual properties including a pseudogap (PG) at low doping and unusual metallic behavior at higher doping. This has lead researchers to postulate the existence of a QCP at optimal doping in the cuprates phase diagram. Some investigators have also argued that the PG is related with the establishment of order [30–35], and the optimal doping is in the proximity of the QCP associated to this order. Others have argued that the QCP is located at the transition from the non-Fermi liquid (NFL) to the Fermi liquid (FL) ground state with no order established in the PG region [36]. We use the DCA to explore the presence of this QCP in the two-dimensional Hubbard model [37]. Investigating the single-particle properties, we find further evidence for the QCP and determine that it is the terminus of a V-shaped marginal Fermi liquid (MFL) region separating the NFL PG region from the FL region at high doping.

In this section we analyze several physical quantities using the known forms of the self-energy in the MFL and the FL regions, as well as an ansatz in the region beyond but near the QCP, when the system crosses over from MFL to FL. Within the DCA we can evaluate $Z_0(\mathbf{k}) = (1 - \mathrm{Im}\Sigma(\mathbf{k}, i\omega_0)/\omega_0)^{-1}$, where $\omega_0 = \pi T$ is the lowest Fermion Matsubara frequency. For a well behaved self-energy, $\lim_{T\to 0} Z_0(\mathbf{k}) = Z(\mathbf{k})$ is the quasiparticle renormalization factor. In this problem, the relevant low energy scales are the antiferromagnetic exchange energy J near half-filling, the PG temperature T^* in the PG region, and the effective Fermi energy T_X at higher doping. From the previously described analysis [38], we find that $J_{\mathrm{eff}} \approx 0.44t$ for $n = 0.95$ and $n = 1$. We extract T^* and T_X from fits to the data[37], as presented in the figures below, where data for $T \ll J_{\mathrm{eff}}$ are included in each case. T^* can be also determined from the peak in the susceptibility (see Fig. 9.9).

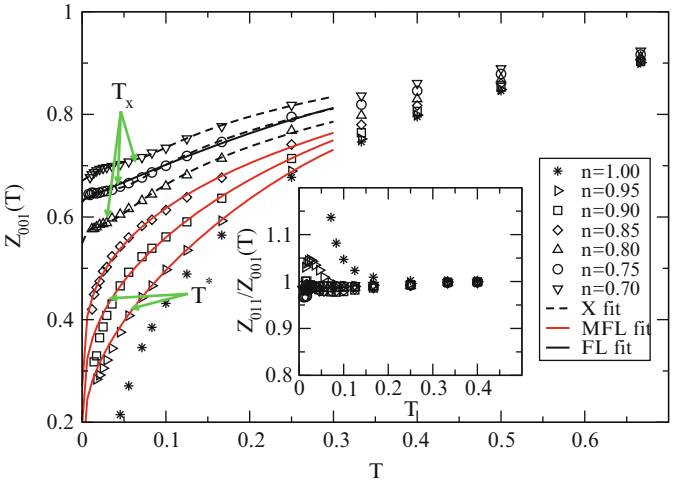

Fig. 9.9 (Taken from [37]) The single-particle density of states in the pseudogap region for various temperatures with $n = 0.95$, $U = 6t = 1.5$, $W = 8t = 2$. The unit of energy is set to $4t = 1$. Inset: The bulk, $\mathbf{Q} = 0$, cluster susceptibility for the same parameters. The PG in the DOS begins to develop at roughly the same temperature T^* which identifies the peak susceptibility

The quasiparticle fraction is calculated with $\mathbf{k}$ on the Fermi surface (FS) as defined by the maximum along the $(1, 1)$ and $(0, 1)$ directions of $|\nabla n(\mathbf{k})|$. This FS is slightly different from the one identified using the spectral function $A(\mathbf{k}, \omega = 0)$ [39] when $n > 0.85$. However, the quasiparticle weights $Z \to 0$ everywhere on both Fermi surfaces (and shows a similar anisotropy on both). So, our conclusions do not depend on whether $|\nabla n(\mathbf{k})|$ or $A(\mathbf{k}, \omega = 0)$ is used to identify the FS. Since we are interested in the crossover from PG to FL behavior, and the PG is stronger along the $(0, 1)$ direction, we present detailed results and analysis for the $(0, 1)$ direction only. The quasiparticle fraction along the $(0, 1)$ direction, Z_{001}, is shown in the main panel in Fig. 9.10 for different fillings.

As the filling n increases through $n = 0.85$, the low-temperature Matsubara quasiparticle data changes its behavior. The data have a negative curvature at all T for $n > 0.85$; while for $n < 0.85$, the data have a negative curvature at high T and develops a weak positive curvature at lower T. The change in curvature of the low temperature data for $n < 0.85$ is easily understood as a crossover to an FL region. On the other hand, the MFL always has a negative curvature. So at the transition between FL and MFL, a region of positive curvature is found at $T \approx T_X$. The ratio of the quasiparticle fraction at the FS along the (01) and the (11) direction, Z_{011}/Z_{001}, plotted in the inset of Fig. 9.10 as a function of temperature for different fillings shows that the conclusions from the above analysis are not specific to the direction $(0, 1)$. The ratio is seen to be essentially the same for all fillings at the QCP, indicating that Z is essentially isotropic at the QCP, and becomes progressively more anisotropic as we dope into the PG region. Furthermore, Z calculated at $\mathbf{k} = (0, \pi)$

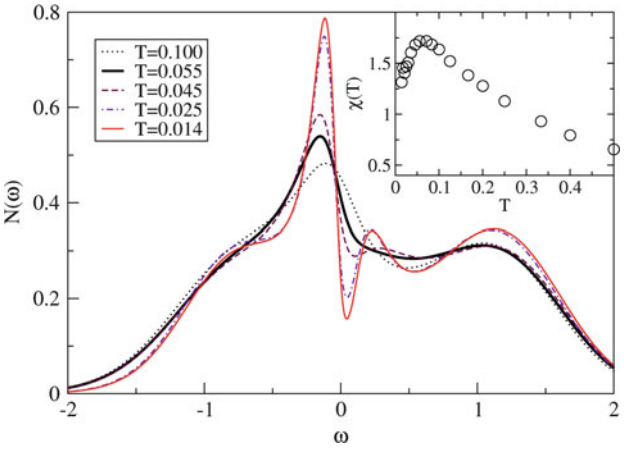

Fig. 9.10 (Taken from [37]) Matsubara quasiparticle fraction $Z_0(\mathbf{k})$ versus temperature T evaluated with $\mathbf{k}$ on the Fermi surface along the $(0, 1)$ direction for different fillings n when $U = 6t$ and the bandwidth $W = 8t$. The unit of energy is such that $4t = 1$. The *lines* represent fits in the region $T < 0.3$ to either the MFL form, for $n \geq 0.85$, or the crossover form (X), for $n < 0.85$. The *arrows* indicate the values T_X extracted from the crossover fit or T^* (cf. Figs. 9.9 and 9.11). Note that the data for $n = 0.85$ fit the MFL nearly perfectly, while the data for $n > 0.85$ are poorly fit by the MFL for $T < T^*$ because, due to the formation of the pseudogap, the MFL temperature dependence is too slow to provide a good fit. The data for $n = 0.75$ were also fit by the FL form; however, the fit is clearly worse than that obtained by the crossover form. *Inset*: The ratio, Z_{011}/Z_{001}, is plotted as a function of temperature for different fillings. The ratio is essentially the same for all fillings at the QCP, indicating that Z is essentially isotropic, and becomes progressively more anisotropic as we dope into the PG region

(not shown) is qualitatively the same as that calculated along the 01 direction on the FS. Therefore, the QCP, which separates the low-temperature FL phase from the PS region, cannot be an artifact of the interpolation nor due to the change of the Fermi wavevector with filling. Rather, it is due to a dramatic change in the nature of the self-energy for momenta near the FS.

The PG region, $n > 0.85$, is further characterized by exploring the temperature dependence of the density of states (DOS) and the bulk, $\mathbf{Q} = 0$, spin susceptibility of the cluster, as shown in Fig. 9.9 and its inset, respectively. A concomitant depression appears in the low energy DOS at temperatures below the energy, T^*, of the peak in the susceptibility. The corresponding $Z_{001}(T)$ is well fit with the MFL form for $T > T^*$, while it fits poorly for $T < T^*$ (Fig. 9.10) due to the formation of the PG.

In Fig. 9.11, we show the relevant temperatures near the QCP, T_X and T^*. T_X is determined from the fits while T^* is determined from the peak in the susceptibility and the initial appearance of the PG in the DOS as shown in Fig. 9.9. Here, T_c is the superconducting critical temperature determined in [40] from the divergence of the pairing susceptibilities as discussed in Sect. 9.3.2.

Further evidence of the presence of the QCP separating the FL region from the NFL PG region can be obtained by studying the thermodynamics of the system [41].

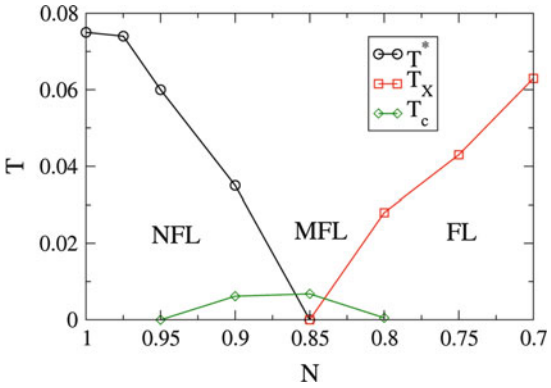

Fig. 9.11 (Taken from [37]) The pseudogap temperature T^*, identified from the peak in the susceptibility and the emergence of the PG in the DOS shown in Fig. 9.9. The FL to MFL crossover temperature identified by fits to the Matsubara quasiparticle data shown in Fig. 9.10. T_c is the superconducting critical temperature determined in [40] from the divergence of the pairing susceptibilities as discussed in Sect. 9.3.2. The unit of energy is set to $4t = 1$

The Hubbard model can be rewritten as:

$$H = \sum_{\mathbf{k}\sigma} \epsilon^0_{\mathbf{k}} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}, \tag{9.35}$$

where $\epsilon^0_{\mathbf{k}} = -2t\left(\cos k_x + \cos k_y\right)$ is the tight binding dispersion. The quadratic part of the Hamiltonian, referred to as the kinetic energy, and the potential energy may be calculated as [42]

$$E_{\mathrm{k}} = \frac{T}{N} \sum_{\omega_n, \mathbf{k}, \sigma} \epsilon^0_{\mathbf{k}} G_\sigma(\mathbf{k}, i\omega_n) \tag{9.36}$$

$$E_{\mathrm{p}} = \frac{T}{2N} \sum_{\omega_n, \mathbf{k}, \sigma} \Sigma_\sigma(\mathbf{k}, i\omega_n) G_\sigma(\mathbf{k}, i\omega_n)\,, \tag{9.37}$$

Both E_{k} and E_{p} are expected to exhibit a leading T^2 low temperature behavior in the FL region and $T^2 \ln T$ behavior [43] in the MFL region. continuous time quantum Monte Carlo (CTQMC) is used to solve the cluster problem and the energies are calculated using (9.36) and (9.37). CTQMC avoids systematic errors and prevents them from accumulating between different temperatures in the calculation of the entropy given by partial integration:

$$S(\beta, n) = S(0, n) + \beta E(\beta, n) - \int_0^\beta E(\beta', n) \mathrm{d}\beta', \tag{9.38}$$

where $S(0,n) = -n \ln \frac{n}{2} - (2-n) \ln \left(1 - \frac{n}{2}\right)$, n is the filling, $\beta = 1/T$ and E is the total energy obtained by summing up E_{p} and E_{k} [44]. Since DCA preserves thermodynamical consistency [45], our entropy results also satisfy the Maxwell relation

$$\left(\frac{\partial S}{\partial n}\right)_{T,U} = -\left(\frac{\partial \mu}{\partial T}\right)_{U,n}, \tag{9.39}$$

where μ is the chemical potential.

The behavior of the numerically calculated potential energy (E_{p}) and kinetic energy (E_{k}) is consistent with the analytical expressions in the FL and MFL regions. However, we find that the characteristic energy scales of the FL and PG vanish at the QC doping where the MFL behavior persists to the lowest accessible temperature. This is consistent with the existence of a QCP at zero temperature between the FL and PG regions. To illustrate this we fit the total energy away from half-filling to the form:

$$E(T) = E(0) + Af(T)T^2 + B\,(1 - f(T))\,T^2 \ln \frac{T}{\Omega}, \tag{9.40}$$

where $f(T) = 1/\left(\exp\left((T - T_X)/\theta\right) + 1\right)$ describes the crossover from the MFL to the quadratic behavior, characteristic of an FL or presumably a PG region. A, B, θ, T_X and Ω are the fitting parameters of the QMC energy data, as shown in Fig. 9.12. The fit is indistinguishable from the data for all fillings at low T. In agreement with the previous estimates, these fits indicate that $T_{\mathrm{FL}} \approx 0.15t$ for $n = 0.70$ and $T^* \approx 0.24t$ for $n = 0.95$.

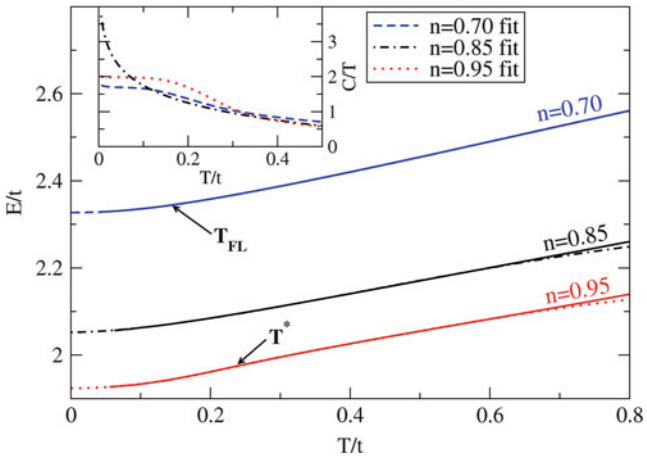

Fig. 9.12 (Taken from [41]) Total energy per site, E, versus temperature for different fillings. The data are fit to a crossover form of the energy, (9.40) (*dashed lines*). The values of T_X determined from the fit are indicated as T_{FL} for $n = 0.70$ and T^* for $n = 0.95$. In the inset, the specific heat calculated from the fit is plotted versus temperature

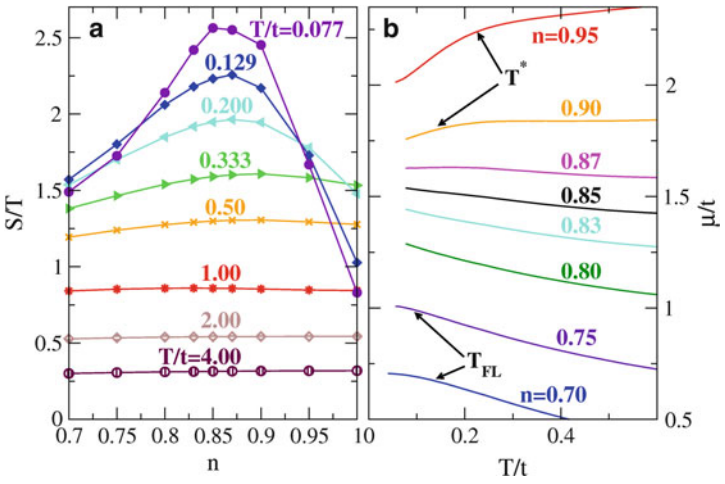

Fig. 9.13 (Taken from [41]) *Left panel*: filling dependence of S/T showing emergence of a peak at $n = 0.85$ at low temperatures. *Right panel*: Chemical potential versus temperature for a range of fillings with PG and FL energy scales shown as T^* and $T_{\rm FL}$ for $n = 0.95$ and $n = 0.70$, respectively. Note that the position of the maximum of entropy in the *left panel* corresponds to $\partial\mu/\partial T = 0$ in the *right panel*. As the temperature is lowered, the maximum of entropy shifts towards lower n, causing a local particle–hole symmetry for $n = 0.85$ at low T (see text)

The calculation of the specific heat is known to be a very difficult problem. It usually involves a fit of $E(T)$ to a regularized (smooth) functional form [46, 47]. In the present case, we already have an excellent fit, so C/T is simply obtained from a derivative of the fit divided by temperature. For $n = 0.70$, at low temperatures, C/T is flat in T, as one expects for an FL. The data in the PG region, $n = 0.95$, also show this behavior, but, at the critical filling, $n = 0.85$, the data show a weak divergence at low T consistent with quantum critical behavior [41].

The behavior of the entropy per site near the critical filling as the system is cooled confirms the physics seen in C/T with no need for a fit or a numerical derivative. With decreasing T, the entropy is more strongly quenched in the FL and PG regions than in the MFL region, creating a maximum in S/T at $n = 0.85$ and low temperature (see Fig. 9.13a). The persistent rise of S/T at critical doping as $T \to 0$ is consistent with the increase to C/T. The near overlap for $n < 0.85$ of the low temperature S/T at different temperatures also agrees with the constant C/T indicative of a FL.

Equation (9.39) indicates that a local maximum in S/T versus n corresponds to a flat chemical potential as a function of temperature. For this reason, the critical filling at low T can be identified from the temperature dependence of μ for different fillings. This is shown in Fig. 9.13b, where one can see that the near temperature independence of μ at $n = 0.90$ for $0.25t < T < 0.50t$ evolves into a broad maximum centered around $T = 0.15t$ for $n = 0.87$ which presumably moves to $n = 0.85$ at low enough temperatures. These observations are consistent with the

evolution of the maximum in S/T versus n as the temperature is lowered from $0.50t$ to $0.08t$ (see Fig. 9.13a). A stationary chemical potential can be the signature of a local particle–hole symmetry, in analogy with the half-filled case. This is consistent with the observation of near particle–hole symmetry in the cuprates in the proximity of optimal doping [36].

9.4.2 *Nature of the Quantum Critical Point in the Hubbard Model*

A systematic study of the phase diagram of the Hubbard model as a function of additional control parameters allows us to identify the nature of the QCP in the cuprates. We use an extended Hubbard model where the tight binding dispersion is modified to include t', the hopping between next-nearest neighbors. The dispersion is then $\epsilon^0_{\mathbf{k}} = -2t\left(\cos k_x + \cos k_y\right) - 4t'\left(\cos k_x \cos k_y - 1\right)$. Our results suggest that the QCP is the zero-temperature limit of a line of second-order phase separation transition as shown schematically in Fig. 9.14 [48]. The control parameter for this transition is t'.

To illustrate this, we calculate the filling, n, versus μ and the compressibility (or bulk charge susceptibility), $\mathrm{d}n/\mathrm{d}\mu$, by taking its numerical derivative. To connect with previous results, simulations were performed with $U = 6t$ (Fig. 9.15a), but, as discussed previously [48], the region of divergent charge fluctuations is larger and more accessible for $U = 8$ and cluster size $N_c = 8$. For this reason, we also present results for these parameters where additional studies have been conducted

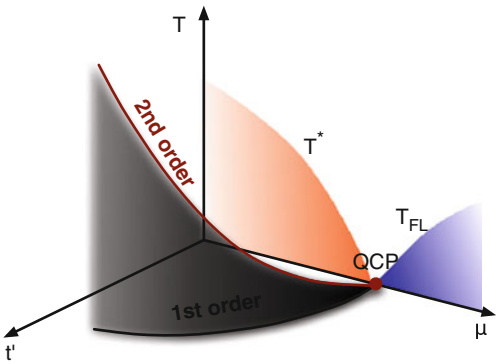

Fig. 9.14 (Taken from [48]) Schematic phase diagram of the 2D Hubbard model in the temperature (T), chemical potential (μ) and next-near-neighbor hopping (t') space. For $t' > 0$ the first-order phase separation transition terminates at a second-order critical point at doping n_c and temperature T_{ps}. The line of second-order critical points (T_{ps}, n_c) approaches the QCP on the $t' = 0$ plane. This is the critical point separating the pseudogap (PG) from the Fermi liquid (FL) region

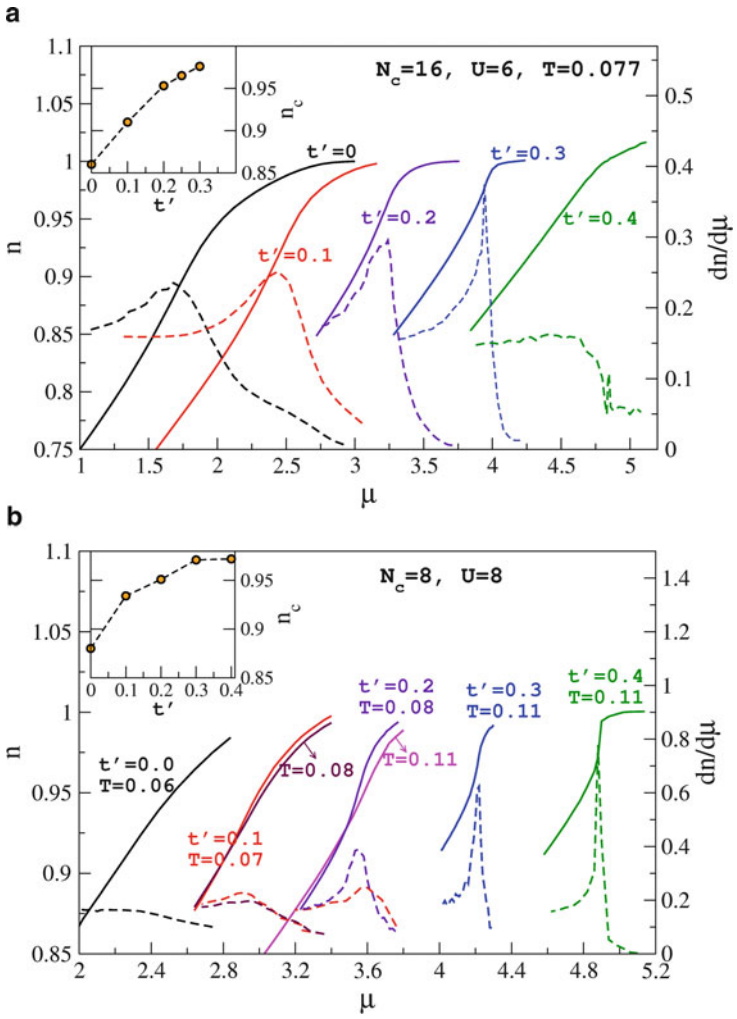

Fig. 9.15 (Taken from [48]) Filling, n (*solid lines*), and compressibility, $dn/d\mu$ (*dashed lines*), plotted versus chemical potential, μ, for various values of t' for (**a**) $U = 6$, $N_c = 16$ and $T = 0.077$ and (**b**) $U = 8$, $N_c = 8$ at different temperatures. The unit of energy is t. The critical filling, where the compressibility peaks, is plotted in the corresponding inset. In (**a**), when $t' \to 0$, the peak in the charge susceptibility is located at the QCP identified previously [37]

(cf. Fig. 9.15b). Figure 9.15a shows n versus μ for $U = 6t$, $T = 0.077t$ and t' ranging from 0.0 to 0.4. The filling n increases monotonically with μ and shows a pronounced flat region associated with the Mott gap, especially for $t' < 0.4$. An inflection appears in $n(\mu)$ at finite doping and becomes more pronounced as t' increases. It translates into a peak in the susceptibility that becomes sharper and moves closer to half-filling as t' is increased. The peak in the susceptibility and the

plateau in $n(\mu)$ near half-filling disappear for $t' > 0.3$. In the inset, we plot n_c, the value of the critical filling at the peak as a function of t'. For $t' = 0$, $n_c = 0.86$ is in agreement with the filling of the QCP ($n_c = 0.85$) found previously for these parameters [37,41]. These results suggest an association between the QCP and charge fluctuations.

For temperatures below a critical temperature T_c, the filling is observed to develop a hysteresis as a function of μ. As mentioned before, the DCA equations are solved self-consistently starting with an initial guess for the self-energy, usually zero, the result from a higher temperature or that of perturbation theory. The solution is generally unique and independent from the initial guess for doping away from a critical doping δ_c, such as 0 or 10% doping. However, we find that for a critical chemical potential μ_c, if the initial self-energy is that corresponding to the undoped solution ($n = 1$), then n versus μ will look as the upper curve (squares) in Fig. 9.16, whereas if it is that of a large doping solution ($n < 1$), n versus μ will be described by the lower curve (circles) in Fig. 9.16. The fully converged self-energy from a previous point is used to initialize the calculation in both cases.

To further investigate the association between the QCP and charge fluctuations, we study the behavior of the bulk charge susceptibility, $\chi_c(\mathbf{Q} = 0, T)$, and its divergence as $t' \to 0$. We follow the line of second order critical points of these first order transitions as t' changes using χ_c as shown in Fig. 9.17. We plot the inverse

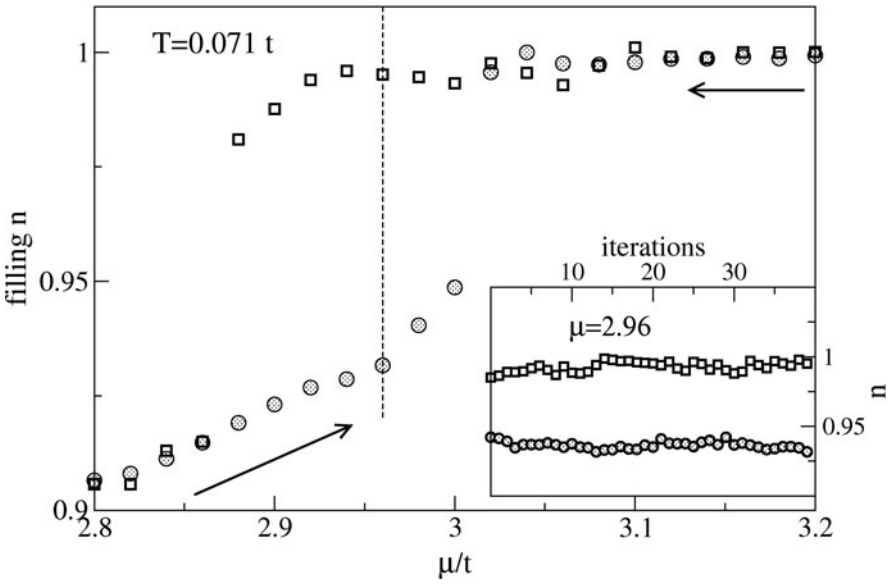

Fig. 9.16 (Taken from [49]) Filling n versus chemical potential for $t' = 0.3t$ at $T = 0.071\ t$ and $N_c = 8$. Two solutions describing a hysteresis are found, one incompressible with $n \approx 1$ (*squares*) and other a doped one (*circles*). *Inset*: stability of the two solutions versus DCA iterations when $\mu = 2.96t$ (middle of the hysteresis, corresponding to the *dotted line* in the main figure)

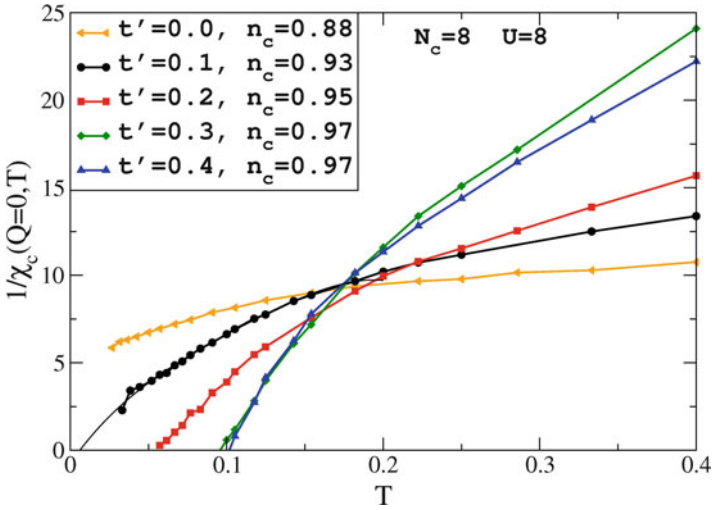

Fig. 9.17 (Taken from [48]) Inverse bulk charge susceptibility versus temperature when $U = 8t$, $N_c = 8$ for several values of t'. The unit of energy is t. The values of the critical filling n_c shown in the legend correspond to the maximum of the low temperature compressibility, or the filling where it first diverges

charge susceptibility at n_c as a function of temperature for different values of t' and $U = 8t$, $N_c = 8$. The critical filling n_c shown in the legend is the filling where the compressibility either diverges or is peaked at the lowest accessible temperature. The temperature of the second-order critical point is found to increase and move towards half-filling as t' is increased. However, in these results, unlike those of $U = 6t$ (Fig. 9.15a), the critical point seems to avoid half-filling even for $t' = 0.4t$. As can be seen in the persistence of the flat region in $n(\mu)$ near $n = 1$ for $t' = 0.4t$, the stronger Coulomb interaction $U = 8t$ also appears to strengthen the Mott gap for this value of t' (Fig. 9.15b).

The evidence discussed here strongly favors an interpretation involving a QCP as opposed to a simple crossover from the FL to an NFL as the filling increases towards one. The fits to the quasiparticle fraction $Z_0(\mathbf{k})$, the behavior of the DOS, the $T^2 lnT$ behavior of the kinetic and potential energies, the peak in S/T which sharpens as T falls, and the logarithmic behavior of the specific heat are clear signatures of quantum criticality. The results also show that the QCP is the terminal point of a line of second-order critical points associated with first-order phase separation transitions. The critical temperature is driven to zero as $t' \to 0$.

9.4.3 *Relationship Between Superconductivity and the Quantum Critical Point*

The critical doping at which the QCP is identified appears to be in close proximity to the optimal superconducting doping, as found in the Hubbard model [48] or in

the $t-J$ model [50]. Although this proximity might indicate that the QCP enhances pairing, the mechanism of such an enhancement remains unclear. Using DCA, we attempt to separate two different scenarios about the role of the QCP in the superconducting mechanism [40]. The *first*, the quantum critical BCS (QCBCS) proposed by She and Zaanen [51], argues that the presence of the QCP leads to a replacement of the BCS logarithmic divergence of the pairing bubble by an algebraic divergence, resulting in a stronger pairing instability and a higher critical temperature compared to the BCS result for the same pairing interaction. The *second* scenario suggests that the pairing interaction is mediated by remnant fluctuations [52, 53]. This would be translated into a strongly enhanced pairing interaction in the vicinity of the QCP. We find that near the QCP, the pairing interaction depends monotonically on the doping and shows no special feature, whereas the pairing susceptibility acquires an algebraic dependence on the temperature. These findings are consistent with the first scenario.

The superconducting transition temperature T_c for a conventional BCS superconductor is determined by the condition $V\chi_0'(\omega=0)=1$, where χ_0' is the real part of the $q=0$ bare pairing susceptibility and V is the strength of the pairing interaction. The transition is driven by the divergence of $\chi_0'(\omega=0)$ which for a FL is given by $\chi_0'(T)\propto N(0)\ln(\omega_D/T)$, where $N(0)$ is the single-particle DOS at the Fermi surface and ω_D is the phonon Debye cutoff frequency. This leads to the well-known BCS superconducting transition formula, $T_c=\omega_D\exp[-1/(N(0)V)]$. We will use the same T_c equation $V\chi_0'(\omega=0)=1$ to analyze our results for the Hubbard model and look for the possibility that $\chi_0'(\omega=0)\sim 1/T^{\alpha}$.

This analysis starts with the Bethe–Salpeter equation for the pairing channel:

$$\chi(Q)_{P,P'}=\chi_0(Q)_P\delta_{P,P'}+\sum_{P''}\chi(Q)_{P,P''}\Gamma(Q)_{P'',P'}\chi_0(Q)_{P'}, \tag{9.41}$$

where χ is the dynamical susceptibility, $\chi_0(Q)_P$ $[=-G(P+Q)G(-P)]$ is the bare susceptibility, which is constructed from G, the dressed one-particle Green's function, Γ is the vertex function, and indices $P^{[\ldots]}$ and external index Q denote both momentum and frequency as discussed in Sect. 9.3.1. The divergence of the susceptibility is detected by solving the eigenvalue equation $\Gamma\chi_0\phi=\lambda\phi$ [54] for fixed Q. By decreasing the temperature the leading eigenvalue λ increases reaching one at a temperature T_c, where the system undergoes a phase transition.

In order to be able to identify whether χ_0 or Γ dominates at the phase transition, we will make a BCS approximation and project them onto the d-wave pairing channel, which was found to be dominant [55, 56]. For χ_0, the d-wave projection is given by

$$\chi_{0d}(\omega)=\frac{\sum_k\chi_0(\omega,q=0)_k g_d(k)^2}{\sum_k g_d(k)^2}, \tag{9.42}$$

where $g_d(k)=(\cos(k_x)-\cos(k_y))$ is the d-wave form factor. For the pairing strength we use the d-wave projection:

$$V_d = \frac{\sum_{k,k'} g_d(k)\Gamma_{k,k'} g_d(k')}{\sum_k g_d(k)^2} \tag{9.43}$$

using Γ at the lowest Matsubara frequency [57].

We further explore the different contributions to the strength of the pairing vertex V_d by performing an exact decomposition of Γ into its different cross-channels and projecting out the d-wave contribution of each contribution [57, 58]. The cross-channels include the fully irreducible vertex Λ, the charge channel ($S = 0$) contribution Φ_c and the spin channel ($S = 1$) contribution Φ_s. This follows from the fact that the vertex can be written as

$$\Gamma = \Lambda + \Phi_c + \Phi_s. \tag{9.44}$$

The d-wave projection as described above then gives

$$V_d = V_d^{\Lambda} + V_d^{c} + V_d^{m}. \tag{9.45}$$

We use $U = 6t$ ($4t = 1$) for both the $N_c = 12$ and $N_c = 16$ clusters. For these clusters, as found in the previous section, evidences of the QCP are observed around a doping of $\delta \approx 0.15$. In Fig. 9.18, we show the eigenvalues λ for different channels (magnetic, charge, and pairing) at the critical doping as a function of temperature. The results, for both cluster sizes, indicate a superconducting transition around $T_c = 0.007$ and an enhancement of the charge susceptibility with decreasing temperatures as can be expected from the QCP that arises as a terminus of a line of second-order phase separation transitions.

In Fig. 9.19, we present the strength of V_d as a function of doping for a range of temperatures. Here it is found that V_d decreases monotonically with increasing

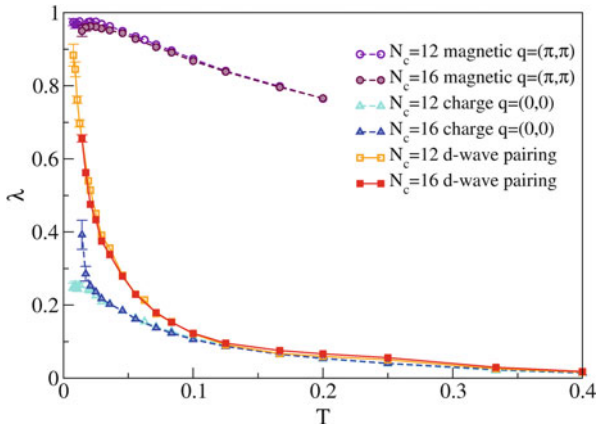

Fig. 9.18 (Taken from [40]) Plots of leading eigenvalues for different channels at the critical doping for $N_c = 12$ and $N_c = 16$ site clusters. The energy is set to $4t = 1$

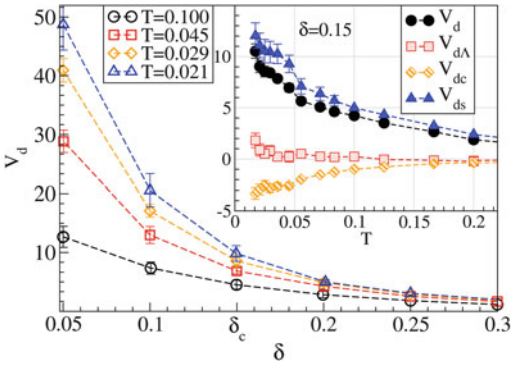

Fig. 9.19 (Taken from [40]) Plots of V_d, the strength of the d-wave pairing interaction, for various temperatures with $U = 6t$ and $N_c = 16$, where $4t = 1$. V_d decreases monotonically with doping, and shows no feature at the critical doping. In the inset are plots of the contributions to V_d from the charge V_d^c and spin V_d^s cross-channels and from the fully irreducible vertex V_d^Λ versus T at the critical doping. As the temperature is lowered, $T \ll J \approx 0.11 = 0.44t$, the contribution to the pairing interaction from the spin channel is clearly dominant

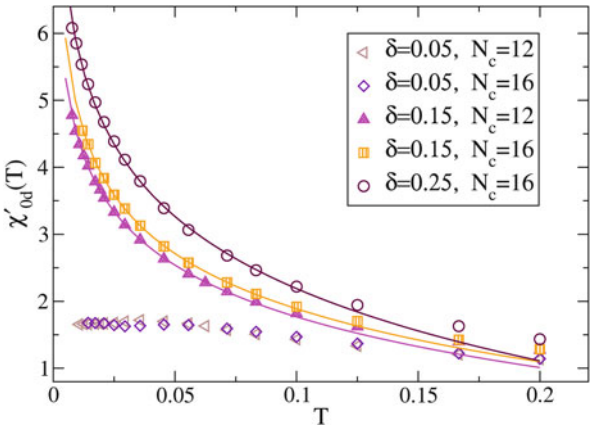

Fig. 9.20 (Taken from [40]) Plots of $\chi'_{0d}(\omega = 0)$, the real part of the bare d-wave pairing susceptibility at zero frequency, versus temperature at three characteristic values of the doping. The *solid lines* are fits to $\chi'_{0d}(\omega = 0) = B/\sqrt{T} + A\ln(\omega_c/T)$ for $T < J$. In the underdoped case ($\delta = 0.05$), $\chi'_{0d}(\omega = 0)$ does not grow with decreasing temperature. At the critical doping ($\delta = \delta_c = 0.15$), $\chi'_{0d}(\omega = 0)$ shows power-law behavior with $B = 0.04$ for the 12 site and $B = 0.09$ for the 16-site clusters (in both $A = 1.04$ and $\omega_c = 0.5$). In the overdoped region ($\delta = 0.25$), a log divergence is found, with $B = 0$ obtained from the fit

doping as seen in a previous study [59]. V_d does not show any feature at the critical doping $\delta_c = 0.15$. This effectively rules out the second scenario mentioned above. The different components in the inset suggest that V_d at the QCP originates predominantly from the spin channel. This behavior is similar to what was found previously while studying the pairing interaction [59].

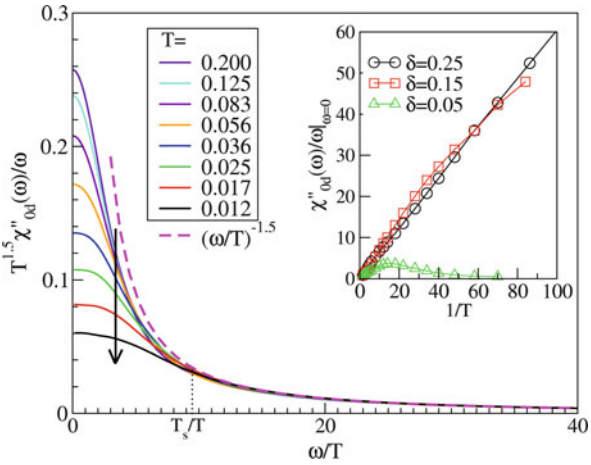

Fig. 9.21 (Taken from [40]) Plots of $T^{1.5}\chi''_{0d}(\omega)/\omega$ versus ω/T at the QC doping ($\delta = 0.15$) for $N_c = 16$. The *arrow* denotes the direction of decreasing temperature. The *curves* coincide for $\omega/T > 9 \approx (4t/J)$ defining a scaling function $H(\omega/T)$, corresponding to a contribution to $\chi'_{0d}(T) = \frac{1}{\pi}\int d\omega \chi''_{0d}(w)/w \propto 1/\sqrt{T}$ as found in Fig. 9.20. For $\omega/T > 9 \approx (4t/J)$, $H(\omega/T) \approx (\omega/T)^{-1.5}$ (*dashed line*). On the x-axis, we add the label $T_s/T \approx (4t/J)$, where T_s represents the energy scale where curves start deviating from H. The inset shows the unscaled zero-frequency result $\chi''_{0d}(\omega)/\omega\big|_{\omega=0}$ plotted versus inverse temperature

Unlike the pairing strength V_d, the d-wave pairing susceptibility χ_{0d} exhibits significantly different behaviors in the different doping regions around the QCP (Fig. 9.20). In the underdoped region ($\delta = 0.05$), also known as the pseudogap or NFL region, $\chi'_{0d}(\omega = 0)$ saturates to a finite value at low temperatures. At the critical doping, however, it diverges quickly when the temperature decreases, roughly following a $1/\sqrt{T}$ power-law behavior. The overdoped or FL region shows the expected log divergence.

To further understand the temperature-dependence of the d-wave pairing susceptibility at the quantum critical filling, we investigate $T^{1.5}\chi''_{0d}(\omega)/\omega$ and plot it in Fig. 9.21 as a function of ω/T. Scaled to this form, the curves from different temperatures collapse on each other so that

$$\frac{T^{1.5}\chi''_{0d}(\omega)}{\omega} = H\left(\frac{\omega}{T}\right) \approx \left(\frac{\omega}{T}\right)^{-1.5} \tag{9.46}$$

for $\omega/T \gtrsim 9 \approx 4t/J$, $J \approx 0.44t$. For $0 < \omega/T < 4t/J$, the BCS behavior appears and the curves deviate from $H(x)$ with $\chi''_{0d}(\omega)/\omega\big|_{\omega=0}$ weakly sublinear in $1/T$ as shown in the inset. Away from critical doping, the curves do not show such a collapse (not shown). $\chi''_{0d}(\omega)/\omega$ goes to zero with decreasing temperature (inset) in the underdoped region ($\delta = 0.05$) while it develops a narrow peak at low ω of width $\omega \approx T_X$ and height $\propto 1/T$ in the FL region (inset).

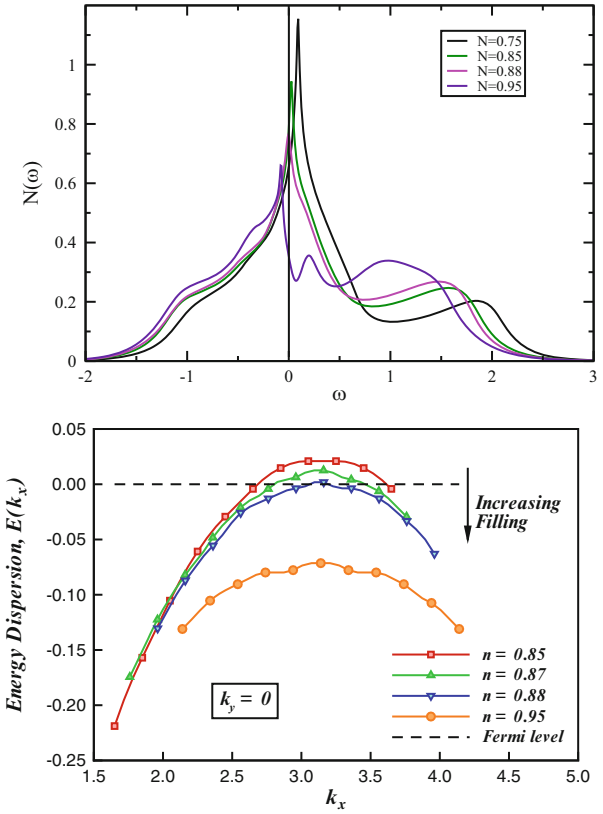

Fig. 9.22 (*Top*) Dependence of the density of states $N(\omega)$ with frequency ω for various fillings, where $U = 6t$, $4t = 1$, $N_c = 16$, and $T = 1/58$. The peak in $N(\omega)$ moves through $\omega = 0$ at the quantum critical filling and exhibits particle hole symmetry for small energies. (*Bottom*) Single-particle dispersion around the Fermi vector $\mathbf{k}_F$ along the antinodal direction $(\pi, 0)$. As the filling is increased, the saddle point in the dispersion moves through the Fermi level at the quantum critical filling, consistent with the density of states picture

Our results for χ_{0d} and V_d provide some understanding of the previously found [48] superconducting dome. As the doping increases, the pairing vertex V_d falls monotonically whereas $\chi'_{0d}(T)$ is strongly suppressed in the low doping or pseudogap region and enhanced at the critical and higher doping. These facts alone could lead to a dome-shaped region of superconductivity. Additionally, the algebraic divergence of the pairing susceptibility $\chi'_{0d}(T)$, as seen in Fig. 9.20, causes superconductivity to be strongly enhanced near the QCP so that one might expect $T_c \propto (V_d B)^2$, with $B = \frac{1}{\pi}\int dx H(x)$, to replace the conventional BCS form in the FL region. This evidence supports the QCBCS scenario proposed by She and Zaanen [51].

Another interesting feature of the QCP is the proximity to the Fermi level of a van Hove singularity (vHS) in the single-particle dispersion. Many authors have

suggested that a vHS in the single-particle DOS at the Fermi level will enhance the superconducting transition temperature [60–63]. In an early DMFA study of the Hubbard model by Majumdar and Krishnamurthy [64], it was found that if the bare dispersion is modified such that the vHS is at the Fermi level for the noninteracting case, it remains pinned to the Fermi level for the interacting case. The noninteracting model studied here with $t' = 0$ has a vHS at the Fermi energy at half filling. The *new finding*, using the DCA method, is that the interactions produce a self-energy with sufficient momentum dependence so that the renormalized dispersion yields a vHS at the Fermi level at a finite filling. The interacting model has a gap at half-filling and as the filling is increased through the QCP, a vHS crosses the Fermi level near the QCP filling. This can be seen both in the DOS (top) and flat region in the energy dispersion (bottom) of Fig. 9.22. Unfortunately, the momentum dependence of the self-energy alone is not sufficient to reproduce the scaling of $\chi''_{0d}(\omega)$ found in Fig. 9.21. This scaling points towards a deeper origin of the enhanced divergence of the bare pairing polarization, such as that suggested by She and Zaanen in their quantum critical BCS scenario[51].

9.5 Conclusion

We have presented a pedagogical introduction to the DCA. We have described how coarse-graining methods can be used to derive both the DMFA and the DCA, which map the lattice to a self-consistently embedded cluster problem. We also showed how DMFA and DCA can be derived from a Φ functional. The DMFA is a local approximation while the DCA incorporates systematic nonlocal corrections. We have showed how the DCA is used to study the Hubbard model and the evidence it provides for the presence of a QCP underneath the superconducting dome. This QCP is the terminus of a line of second order phase separation transitions. Finally we have seen that the QCP may explain the relatively high superconducting critical temperature through an enhancement of the pairing susceptibility at the critical doping. In brief, the DCA provides an efficient tool to study correlated electron systems.

Acknowledgments We would like to thank S. Kivelson, Michael Ma, A. Macridin, P. Phillips, E. W. Plummer, D. J. Scalapino, C. Sen, J.-H. She, A. M. Tremblay, C. Varma, I. Vekhter, N. S. Vidhyadhiraja, M. Vojta, and J. Zaanen for useful conversations. This research was supported by NSF DMR-0312680, DMR-0706379, NSF OISE-0952300, DOE CMSN DE-FG02-04ER46129, and by the DOE SciDAC grant DE-FC02-06ER25792. This research used resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

1. J. Hubbard, Proc. R. Soc. A**2**76, 238(1963)
2. M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159(1963)

3. J. Kanamori, Prog. Theor. Phys. **30**, 257(1963)
4. Proceedings of the International Conference on Itinerant-Electron Magnetism. Physica B+C **9**1 (1977); for a review about the theory of the Hubbard model see also D. Vollhardt, Rev. Mod. Phys. **56**, 99(1984)
5. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
6. H. Frahm, V.E. Korepin, Phys. Rev. B **42**, 10533 (1990)
7. N. Kawakami, S.-K. Yang, Phys. Rev. Lett. **65**, 2309 (1990)
8. G. Czycholl, J. Mag. Mag. Mat. **47,48**, 350 (1985)
9. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
10. P.G.J. van Dongen, D. Vollhardt, Phys. Rev. Lett. **65**, 1663 (1990)
11. Y. Kuramoto, Theory of Heavy Fermions and Valence Fluctuations: Proceedings of the Eighth Taniguchi Symposium, Shima Kanko, Japan, April 10–13, 1985 Eds T. Kasuya, T. Saso (Springer Series in Solid-State Sciences Vol 62) 152–162
12. C. Kim, Y. Kuramoto and T. Kasuya, J. Phys. Soc. Japan **5**9, 2414 (1990)
13. E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)
14. M.H. Hettler, A.N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, H.R. Krishnamurthy, Phys. Rev. B **58**, 7475 (1998)
15. M.H. Hettler, M. Mukherjee, M. Jarrell, H.R. Krishnamurthy, Phys. Rev. B **61**, 12739 (2000)
16. M. Jarrell Phys. Rev. Lett. **69**, 168 (1992)
17. T. Pruschke, M. Jarrell, J.K. Freericks, Adv. in Phys. **42**, 187 (1995)
18. A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
19. R.M. Fye, J.E. Hirsch, Phys. Rev. B **38**, 433 (1988)
20. M. Jarrell, Th. Maier, C. Huscroft, S. Moukouri, Phys. Rev. B **64**, 195130/1–23 (2001)
21. M. Jarrell, J.E. Gubernatis, Phys. Rep. **269**, 3 (1996)
22. V. Zlatić, B. Horvatić, Sol. St. Comm. **75**, 263 (1990)
23. G. Baym, L.P. Kadanoff, Phys. Rev. **124**, 287, (1961)
24. C. Huscroft, M. Jarrell, Th. Maier, S. Moukouri, A.N. Tahvildarzadeh, Phys. Rev. Lett., **86**, 3691 (2001)
25. Th. Maier, M. Jarrell, Th. Pruschke, J. Keller, Eur. Phys. J. B **13**, 613 (2000)
26. N.E. Bickers, D.J. Scalapino, S.R. White, Phys. Rev. Lett. **62**, 961 (1989)
27. P.G.J. van Dongen, Phys. Rev. B **50**, 14016 (1994)
28. A. Schiller, K. Ingersent, Phys. Rev. Lett. **75**, 113 (1995)
29. A.A. Abrikosov, L.P. Gorkov, I.E. Dzyalishinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York, 1975)
30. S. Chakravarty, R.B. Laughlin, D.K. Morr, C. Nayak, Phys. Rev. B **63**, 094503 (2001)
31. S. Kivelson, E. Fradkin, V. Emery, Nature (London) **393**, 550 (1998)
32. M. Vojta, Y. Zhang, S. Sachdev, Phys. Rev. B **62**, 6721 (2000)
33. S. Sachdev, Rev. Mod. Phys. **75**, 913 (2003)
34. S. Sachdev, Nat. Phys. **4**, 173 (2008)
35. C.M. Varma, Phys. Rev. Lett. **83**, 3538 (1999)
36. S. Chakraborty, D. Galanakis, P. Phillips, Phys. Rev. B **82**, 214503 (2010)
37. N.S. Vidhyadhiraja, A. Macridin, C. Sen, M. Jarrell, M. Ma, Phys. Rev. Lett. **1**02, 206407 (2009)
38. A. Macridin, M. Jarrell, T. Maier, D.J. Scalapino, Phys. Rev. Lett., **99**, 237001 (2007)
39. Th. A. Maier, Th. Pruschke, M. Jarrell, Phys. Rev. B **66**, 075102 (2002)
40. S.-X. Yang, H. Fotso, S.-Q. Su, D. Galanakis, E. Khatami, J.-H. She, J. Moreno, J. Zaanen, M. Jarrell, Phys. Rev. Lett˙**1**06, 047004 (2011)
41. K. Mikelsons, E. Khatami, D. Galanakis, A. Macridin, J. Moreno, M. Jarrell, Phys. Rev. B **80**, 140505(R) (2009)
42. A.L. Fetter, J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971)
43. M. Crisan, C.P. Moca, J. Supercond. **9**, 49 (1996)
44. F. Werner, O. Parcollet, A. Georges, S.R. Hassan, Phys. Rev. Lett. **9**5, 056401 (2005)
45. Th. Maier, M. Jarrell, Th. Pruschke, M.H. Hettler, Rev. Mod. Phys. **7**7, 1027 (2005)

46. C. Huscroft, R. Gass, M. Jarrell, Phys. Rev. B **6**1, 9300 (2000)
47. A.K. McMahan, C. Huscroft, R.T. Scalettar, E.L. Pollock, J. Comp. Aided Mater. Des. **5**, 131 (1998)
48. E. Khatami, K. Mikelsons, D. Galanakis, A. Macridin, J. Moreno, R.T. Scalettar, M. Jarrell, Phys. Rev. B **81**, 201101(R) (2010)
49. A. Macridin, M. Jarrell, Th. Maier, Phys. Rev. B **7**4, 085104 (2006)
50. K. Haule, G. Kotliar, Phys. Rev. B **7**6, 092503 (2007)
51. J.-H. She, J. Zaanen, Phys. Rev. B **8**0, 184518 (2009)
52. C. Castellani et al., Z. Phy. B **103**, 137–144 (1997)
53. E.-G. Moon, A. Chubukov, J. Low Temp. Phys. **1**61, 263 (2010)
54. N. Bulut et al., Phys. Rev. B **4**7, 6157(R) (1993)
55. M. Jarrell et al., EuroPhys. Lett. **56**, 563 (2001)
56. T.A. Maier et al., Phys. Rev. Lett. **95**, 237001 (2005)
57. T.A. Maier et al., Phys. Rev. Lett. **96**, 047005 (2006)
58. S.-X. Yang, H. Fotso, J. Liu, T.A. Maier, K. Tomko, E.F. DAzevedo, R.T. Scalettar, T. Pruschke, M. Jarrell, Phys. Rev. E **80**, 046706 (2009)
59. T.A. Maier et al., Phys. Rev. B **76**, 144516 (2007)
60. J. Labbè, J. Bok, Europhys. Lett. **3**, 22 (1987)
61. J. Friedel, J. Phys. (Paris) **48**, 10 (1987)
62. D.M. Newns, H.R. Krishnamurthy, P.C. Pattnaik, C.C. Tsuei, C.L. Kane, Phys. Rev. Lett. **69**, 8 (1992)
63. R.S. Markiewicz, Phys. Chem. Solids **58**, 8 1179–1310 (1997)
64. P. Majumdar, H.R. Krishnamurthy, cond-mat/9604057v1 (1996)

Chapter 10
Self-Energy-Functional Theory

Michael Potthoff

Abstract Self-energy-functional theory is a formal framework which allows us to derive non-perturbative and thermodynamically consistent approximations for lattice models of strongly correlated electrons from a general dynamical variational principle. The construction of the self-energy functional and the corresponding variational principle is developed within the path-integral formalism. Different cluster mean-field approximations, such as the variational cluster approximation and cluster extensions of dynamical mean-field theory, are derived in this context and their mutual relationship and internal consistency are discussed.

10.1 Motivation

The method of Green's functions and diagrammatic perturbation theory [1] represents a powerful approach to study systems of interacting electrons in the thermodynamical limit. Several interesting phenomena, such as spontaneous magnetic order, correlation-driven metal–insulator transitions or high-temperature superconductivity, however, emerge in systems where electron correlations are strong. Rather than starting from the non-interacting Fermi gas as the reference point around which the perturbative expansion is developed, a local perspective appears to be more attractive for strongly correlated electron systems, in particular for prototypical lattice models with local interaction, such as the famous Hubbard model [2–4]:

$$H = \sum_{ij\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + \frac{U}{2} \sum_{i\sigma} n_{i\sigma} n_{i-\sigma}. \tag{10.1}$$

M. Potthoff (✉)
I. Institut für Theoretische Physik, Universität Hamburg, Jungiusstr. 9, 20355 Hamburg, Germany
e-mail: michael.potthoff@physik.uni-hamburg.de

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_10,

The local part of the problem, i.e., the Hubbard atom, can be solved easily since its Hilbert space is small. It is, therefore, tempting to start from the atomic limit and to treat the rest of the problem, the "embedding" of the atom into the lattice, in some approximate way. The main idea of the so-called Hubbard-I approximation [2] is to calculate the one-electron Green's function from the Dyson equation where the self-energy is approximated by the self-energy of the atomic system. This is one of the most simple embedding procedures. It already shows that the language of diagrammatic perturbation theory, Green's functions and diagrammatic objects, such as the self-energy, can be very helpful to construct an embedding scheme.

The Hubbard-I approach turns out to be a too crude approximation to describe the above-mentioned collective phenomena. One of its advantages, however, is that it offers a perspective for systematic improvement: Nothing prevents us to start with a more complicated "atom" and employ the same trick: We consider a partition of the underlying lattice with L sites (where $L \to \infty$) into L/L_c disconnected clusters consisting of L_c sites each. If L_c is not too large, the self-energy of a single Hubbard cluster is accessible by standard numerical means [5] and can be used as an approximation in the Dyson equation to get the Green's function of the full model. This leads to the cluster perturbation theory (CPT) [6, 7].

The CPT can also be motivated by treating the Hubbard interaction U and the inter-cluster hopping V as a perturbation of the system of disconnected clusters with intra-cluster hopping t'. The CPT Green's function is then obtained by summing the diagrams in perturbation theory to all orders in U and V but neglecting vertex corrections which intermix U and V interactions.

While these two ways of deriving CPT are equivalent, one aspect of the former is interesting: Taking the self-energy from some reference model (the cluster) is reminiscent of dynamical mean-field theory (DMFT) [8–10], where the self-energy of an impurity model approximates the self-energy of the lattice model. This provokes the question whether both, the CPT and the DMFT, can be understood in single unifying theoretical framework.

This question is one motivation for the topic of this chapter on self-energy-functional theory (SFT) [11–14]. Another one is that there are certain deficiencies of the CPT. While CPT can be seen as a cluster mean-field approach since correlations beyond the cluster extensions are neglected, it not self-consistent, i.e., there is no feedback of the resulting Green's function on the cluster to be embedded (some ad hoc element of self-consistency is included in the original Hubbard-I approximation). In particular, there is no concept of a Weiss mean field and, therefore, the CPT cannot describe different phases of a thermodynamical system nor phase transitions. Another related point is that the CPT provides the Green's function only but no thermodynamical potential. Different ways to derive e.g., the free energy from the Green's function [1, 15, 16] give inconsistent results.

To overcome these deficiencies, a self-consistent cluster-embedding scheme has to be set up. Ideally, this results from a variational principle for a general thermodynamical potential which is formulated in terms of dynamical quantities as e.g., the self-energy or the Green's function. The variational formulation should

ensure the internal consistency of corresponding approximations and should make contact with the DMFT. This sets the goals of SFT and also the plan of this chapter.

10.2 Self-Energy Functional

10.2.1 Hamiltonian, Grand Potential and Self-Energy

We consider a system of electrons in thermodynamical equilibrium at a temperature T and chemical potential μ. The Hamiltonian of the system $H = H(\boldsymbol{t}, \boldsymbol{U}) = H_0(\boldsymbol{t}) + H_1(\boldsymbol{U})$ consists of a non-interacting part specified by one-particle parameters $\boldsymbol{t}$ and an interaction part with interaction parameters $\boldsymbol{U}$:

$$H_0(\boldsymbol{t}) = \sum_{\alpha\beta} t_{\alpha\beta}\, c_\alpha^\dagger c_\beta,$$

$$H_1(\boldsymbol{U}) = \frac{1}{2} \sum_{\alpha\beta\gamma\delta} U_{\alpha\beta\delta\gamma}\, c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta. \tag{10.2}$$

The index α refers to an arbitrary set of quantum numbers labelling an orthonormal basis of one-particle states $|\alpha\rangle$. As is apparent from the form of H, the total particle number N is conserved.

The grand potential of the system is given by $\Omega_{\boldsymbol{t},\boldsymbol{U}} = -T \ln Z_{\boldsymbol{t},\boldsymbol{U}}$, where

$$Z_{\boldsymbol{t},\boldsymbol{U}} = \mathrm{tr}\rho_{\boldsymbol{t},\boldsymbol{U}} \quad \text{with} \quad \rho_{\boldsymbol{t},\boldsymbol{U}} = \exp\left(-\frac{(H(\boldsymbol{t},\boldsymbol{U}) - \mu N)}{T}\right) \tag{10.3}$$

is the partition function. The dependence of the partition function (and of other quantities discussed below) on the parameters $\boldsymbol{t}$ and $\boldsymbol{U}$ is made explicit through the subscripts.

The one-particle Green's function $G_{\alpha\beta}(\omega) \equiv \langle\langle c_\alpha; c_\beta^\dagger \rangle\rangle$ of the system is the main object of interest. It will provide the static expectation value of the one-particle density matrix $c_\alpha^\dagger c_\beta$ and the spectrum of one-particle excitations related to a photoemission experiment [17]. The Green's function can be defined for complex ω via its spectral representation:

$$G_{\alpha\beta}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}z\, \frac{A_{\alpha\beta}(z)}{\omega - z}, \tag{10.4}$$

where the spectral density $A_{\alpha\beta}(z) = \int_{-\infty}^{\infty} \mathrm{d}t\, \exp(\mathrm{i}zt) A_{\alpha\beta}(t)$ is the Fourier transform of

$$A_{\alpha\beta}(t) = \frac{1}{2\pi} \langle [c_\alpha(t), c_\beta^\dagger(0)]_+ \rangle, \tag{10.5}$$

which involves the anticommutator of an annihilator and a creator with a Heisenberg time dependence $O(t) = \exp(\mathrm{i}(H - \mu N)t)\, O \exp(-\mathrm{i}(H - \mu N)t)$.

Due to the thermal average, $\langle \mathcal{O} \rangle = \mathrm{Tr}(\rho_{\boldsymbol{t},\boldsymbol{U}} \mathcal{O})/Z_{\boldsymbol{t},\boldsymbol{U}}$, the Green's function depends on $\boldsymbol{t}$ and $\boldsymbol{U}$ and is denoted by $G_{\boldsymbol{t},\boldsymbol{U},\alpha\beta}(\omega)$. For the diagram technique employed below, we need the Green's function on the imaginary Matsubara frequencies $i\omega_n \equiv i(2n+1)\pi T$ with integer n [1]. In the following the elements $G_{\boldsymbol{t},\boldsymbol{U},\alpha\beta}(i\omega_n)$ are considered to form a matrix $\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}}$, which is diagonal with respect to n.

The "free" Green's function $\boldsymbol{G}_{\boldsymbol{t},0}$ is obtained for $\boldsymbol{U} = 0$, and its elements are given by:

$$G_{\boldsymbol{t},0,\alpha\beta}(i\omega_n) = \left(\frac{1}{i\omega_n + \mu - \boldsymbol{t}} \right)_{\alpha\beta} . \tag{10.6}$$

Therewith, we can define the self-energy via Dyson's equation

$$\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}} = \frac{\mathbf{1}}{\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}}}, \tag{10.7}$$

i.e., $\boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}} = \boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{G}^{-1}_{\boldsymbol{t},\boldsymbol{U}}$. The full meaning of this definition becomes clear within the context of diagrammatic perturbation theory [1].

Here, we like to list some important properties of the self-energy only: (1) via Dyson's equation, it determines the Green's function. (2) The self-energy has a spectral representation similar to (10.4). (3) In particular, the corresponding spectral function (matrix) is positive definite, and the poles of $\boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}}$ are on the real axis [18]. (4) $\Sigma_{\alpha\beta}(\omega) = 0$ if α or β refers to one-particle orbitals that are non-interacting, i.e., if α or β does not occur as an entry of the matrix of interaction parameters $\boldsymbol{U}$. Those orbitals or sites are called non-interacting. This property of the self-energy is clear from its diagrammatic representation. (5) If α refers to the sites of a Hubbard-type model with local interaction, the self-energy can generally be assumed to be more local than the Green's function. This is corroborated e.g., by explicit calculations using the weak-coupling perturbation theory [19–21] and by the fact that the self-energy is purely local on infinite-dimensional lattices [8, 22].

10.2.2 Luttinger–Ward Functional

We would like to distinguish between dynamic quantities, like the self-energy, which is frequency-dependent and related to the (one-particle) excitation spectrum, on the one hand, and static quantities, like the grand potential and its derivatives with respect to μ, T, etc. which are related to the thermodynamics, on the other. A link between static and dynamic quantities is needed to set up a variational principle which gives the (dynamic) self-energy by requiring a (static) thermodynamical potential to be stationary. There are several such relations [1, 15, 16]. The

Luttinger–Ward (LW) functional $\widehat{\Phi}_{\boldsymbol{U}}[\boldsymbol{G}]$ provides a special relation with several advantageous properties [23]:

1. $\widehat{\Phi}_{\boldsymbol{U}}[\boldsymbol{G}]$ is a functional. Functionals $\widehat{A} = \widehat{A}[\cdots]$ are indicated by a hat and should be distinguished clearly from physical quantities A.

2. The domain of the LW functional is given by "the space of Green's functions." This has to be made more precise later.

3. If evaluated at the exact (physical) Green's function, $\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}}$, of the system with Hamiltonian $H = H(\boldsymbol{t}, \boldsymbol{U})$, the LW functional gives a quantity

$$\widehat{\Phi}_{\boldsymbol{U}}[\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}}] = \Phi_{\boldsymbol{t},\boldsymbol{U}}, \tag{10.8}$$

   which is related to the grand potential of the system via:

$$\Omega_{\boldsymbol{t},\boldsymbol{U}} = \Phi_{\boldsymbol{t},\boldsymbol{U}} + \mathrm{Tr}\ln \boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}} - \mathrm{Tr}(\boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}}\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}}). \tag{10.9}$$

   Here the notation $\mathrm{Tr}\boldsymbol{A} \equiv T\sum_n \sum_\alpha \mathrm{e}^{\mathrm{i}\omega_n 0^+} A_{\alpha\alpha}(i\omega_n)$ is used. 0^+ is a positive infinitesimal.

4. The functional derivative of the LW functional with respect to its argument is:

$$\frac{1}{T}\frac{\delta\widehat{\Phi}_{\boldsymbol{U}}[\boldsymbol{G}]}{\delta \boldsymbol{G}} = \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{U}}[\boldsymbol{G}]. \tag{10.10}$$

   Clearly, the result of this operation is a functional of the Green's function again. This functional is denoted by $\widehat{\boldsymbol{\Sigma}}$ since its evaluation at the physical (exact) Green's function $\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}}$ yields the physical self-energy:

$$\widehat{\boldsymbol{\Sigma}}[\boldsymbol{G}_{\boldsymbol{t},\boldsymbol{U}}] = \boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}}. \tag{10.11}$$

5. The LW functional is "universal": The functional relation $\widehat{\Phi}_{\boldsymbol{U}}[\cdots]$ is completely determined by the interaction parameters $\boldsymbol{U}$ (and does not depend on $\boldsymbol{t}$). This is made explicit by the subscript. Two systems (at the same chemical potential μ and temperature T) with the same interaction $\boldsymbol{U}$ but different one-particle parameters $\boldsymbol{t}$ (on-site energies and hopping integrals) are described by the same Luttinger–Ward functional. Using (10.10), this implies that the functional $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{U}}[\boldsymbol{G}]$ is universal, too.

6. Finally, the LW functional vanishes in the non-interacting limit:

$$\widehat{\Phi}_{\boldsymbol{U}}[\boldsymbol{G}] \equiv 0 \quad \text{for} \quad \boldsymbol{U} = 0. \tag{10.12}$$

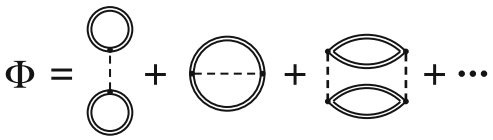

Fig. 10.1 Original definition of the Luttinger–Ward functional $\widehat{\Phi}_U[\boldsymbol{G}]$, see [23]. *Double lines*: fully interacting propagator $\boldsymbol{G}$. *Dashed lines*: interaction $\boldsymbol{U}$

10.2.3 Diagrammatic Derivation

In the original paper by Luttinger and Ward [23] it is shown that $\widehat{\Phi}_U[\boldsymbol{G}]$ can be constructed order by order in diagrammatic perturbation theory. The functional is obtained as the limit of the infinite series of closed renormalized skeleton diagrams, i.e., closed diagrams without self-energy insertions and with all free propagators replaced by fully interacting ones (see Fig. 10.1). There is no known case where this skeleton-diagram expansion could be summed up to get a closed form for $\widehat{\Phi}_U[\boldsymbol{G}]$. Therefore, the explicit functional dependence is unknown even for the most simple types of interactions like the Hubbard interaction.

Using the classical diagrammatic definition of the LW functional, the properties (1)–(6) listed in the previous section are easily verified: By construction, $\widehat{\Phi}_U[\boldsymbol{G}]$ is a functional of $\boldsymbol{G}$ which is universal (properties (1), (2) and (5)). Any diagram in the series depends on $\boldsymbol{U}$ and on $\boldsymbol{G}$ *only*. Particularly, it is independent of $\boldsymbol{t}$. Since there is no zeroth-order diagram, $\widehat{\Phi}_U[\boldsymbol{G}]$ trivially vanishes for $\boldsymbol{U} = 0$, this proves (6).

Diagrammatically, the functional derivative of $\widehat{\Phi}_U[\boldsymbol{G}]$ with respect to $\boldsymbol{G}$ corresponds to the removal of a propagator from each of the Φ diagrams. Taking care of topological factors [1, 23], one ends up with the skeleton-diagram expansion of the self-energy (4). Therefore, (10.11) is obtained in the limit of this expansion.

Equation (10.9) can be derived by a coupling-constant integration [23]. Alternatively, it can be verified by integrating over μ: We note that the μ derivatives of the l.h.s and of the r.h.s of (10.9) are equal for any fixed interaction strength $\boldsymbol{t}$, $\boldsymbol{U}$ and T, namely, $(\partial/\partial\mu)(\Phi_{\boldsymbol{t},U} + \mathrm{Tr}\ln \boldsymbol{G}_{\boldsymbol{t},U} - \mathrm{Tr}\,\boldsymbol{\Sigma}_{\boldsymbol{t},U}\boldsymbol{G}_{\boldsymbol{t},U}) = \mathrm{Tr}\,\boldsymbol{G}_{\boldsymbol{t},U}^{-1}(\partial \boldsymbol{G}_{\boldsymbol{t},U}/\partial\mu) - \mathrm{Tr}\,\boldsymbol{G}_{\boldsymbol{t},U}(\partial\boldsymbol{\Sigma}_{\boldsymbol{t},U}/\partial\mu) = -\mathrm{Tr}\ \boldsymbol{G}_{\boldsymbol{t},U} = -\langle N\rangle_{\boldsymbol{t},U} = \partial\Omega_{\boldsymbol{t},U}/\partial\mu$. Here, we have used (10.3) in the last step and (10.7), (10.8) and (10.10) before. $\langle N\rangle_{\boldsymbol{t},U}$ is the grand-canonical average of the total particle-number operator. Integration over μ then yields (10.9). Note that the equation holds trivially for $\mu \to -\infty$, i.e., for $\langle N\rangle_{\boldsymbol{t},U} \to 0$ since $\boldsymbol{\Sigma}_{\boldsymbol{t},U} = 0$ and $\Phi_{\boldsymbol{t},U} = 0$ in this limit.

10.2.4 Derivation Using the Path Integral

For the diagrammatic derivation it has to be assumed that the skeleton-diagram series is convergent. It is, therefore, interesting to see how the LW functional can

be defined and how its properties can be verified within a path-integral formulation. This is non-perturbative. The path-integral construction of the LW functional was first given in [24].

Using Grassmann variables [16] $\xi_\alpha(i\omega_n) = T^{1/2}\int_0^{1/T} \mathrm{d}\tau\ \mathrm{e}^{\mathrm{i}\omega_n\tau}\xi_\alpha(\tau)$ and $\xi^*_\alpha(i\omega_n) = T^{1/2}\int_0^{1/T} \mathrm{d}\tau\ \mathrm{e}^{-\mathrm{i}\omega_n\tau}\xi^*_\alpha(\tau)$, the elements of the Green's function are given by $G_{\boldsymbol{t},U,\alpha\beta}(i\omega_n) = -\langle\xi_\alpha(i\omega_n)\xi^*_\beta(i\omega_n)\rangle_{\boldsymbol{t},U}$. The average

$$G_{\boldsymbol{t},U,\alpha\beta}(i\omega_n) = \frac{-1}{Z_{\boldsymbol{t},U}}\int d\xi d\xi^* \xi_\alpha(i\omega_n)\xi^*_\beta(i\omega_n)\exp\left(A_{\boldsymbol{t},U,\xi\xi^*}\right) \tag{10.13}$$

is defined with the help of the action $A_{\boldsymbol{t},U,\xi\xi^*} = A_{\boldsymbol{t},\xi\xi^*} + A_{\boldsymbol{U},\xi\xi^*}$, where

$$A_{\boldsymbol{t},U,\xi\xi^*} = \sum_{n,\alpha\beta}\xi^*_\alpha(i\omega_n)((i\omega_n+\mu)\delta_{\alpha\beta} - t_{\alpha\beta})\xi_\beta(i\omega_n) + A_{\boldsymbol{U},\xi\xi^*} \tag{10.14}$$

and

$$A_{\boldsymbol{U},\xi\xi^*} = -\frac{1}{2}\sum_{\alpha\beta\gamma\delta}U_{\alpha\beta\delta\gamma}\int_0^{1/T}\mathrm{d}\tau\ \xi^*_\alpha(\tau)\xi^*_\beta(\tau)\xi_\gamma(\tau)\xi_\delta(\tau). \tag{10.15}$$

This is the standard path-integral representation of the Green's function [16].

The action can be considered as the physical action that is obtained when evaluating the *functional*

$$\widehat{A}_{\boldsymbol{U},\xi\xi^*}[\boldsymbol{G}_0^{-1}] = \sum_{n,\alpha\beta}\xi^*_\alpha(i\omega_n)G^{-1}_{0,\alpha\beta}(i\omega_n)\xi_\beta(i\omega_n) + A_{\boldsymbol{U},\xi\xi^*} \tag{10.16}$$

at the (matrix inverse of the) physical free Green's function, i.e.,

$$A_{\boldsymbol{t},\boldsymbol{U},\xi\xi^*} = \widehat{A}_{\boldsymbol{U},\xi\xi^*}[\boldsymbol{G}^{-1}_{\boldsymbol{t},0}]\ . \tag{10.17}$$

Using this, we define the functional

$$\widehat{\Omega}_{\boldsymbol{U}}[\boldsymbol{G}_0^{-1}] = -T\ln\widehat{Z}_{\boldsymbol{U}}[\boldsymbol{G}_0^{-1}] \tag{10.18}$$

with

$$\widehat{Z}_{\boldsymbol{U}}[\boldsymbol{G}_0^{-1}] = \int d\xi d\xi^*\exp\left(\widehat{A}_{\boldsymbol{U},\xi\xi^*}[\boldsymbol{G}_0^{-1}]\right). \tag{10.19}$$

The functional dependence of $\widehat{\Omega}_{\boldsymbol{U}}[\boldsymbol{G}_0^{-1}]$ is determined by $\boldsymbol{U}$ only, i.e., the functional is universal. Obviously, the physical grand potential is obtained when inserting the physical inverse free Green's function $\boldsymbol{G}^{-1}_{\boldsymbol{t},0}$:

$$\widehat{\Omega}_U[\boldsymbol{G}_{t,0}^{-1}] = \Omega_{t,U} \,. \tag{10.20}$$

The functional derivative of (10.18) leads to another universal functional:

$$\frac{1}{T}\frac{\delta\widehat{\Omega}_U[\boldsymbol{G}_0^{-1}]}{\delta\boldsymbol{G}_0^{-1}} = -\frac{1}{\widehat{Z}_U[\boldsymbol{G}_0^{-1}]}\frac{\delta\widehat{Z}_U[\boldsymbol{G}_0^{-1}]}{\delta\boldsymbol{G}_0^{-1}} \equiv -\widehat{\mathcal{G}}_U[\boldsymbol{G}_0^{-1}] \,, \tag{10.21}$$

with the property

$$\widehat{\mathcal{G}}_U[\boldsymbol{G}_{t,0}^{-1}] = \boldsymbol{G}_{t,U} \,. \tag{10.22}$$

This is easily seen from (10.13).

The strategy to be pursued is the following: $\widehat{\mathcal{G}}_U[\boldsymbol{G}_0^{-1}]$ is a universal ($\boldsymbol{t}$ independent) functional and can be used to construct a universal relation $\boldsymbol{\Sigma} = \widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]$ between the self-energy and the one-particle Green's function – independent from the Dyson equation (10.7). Using this and the universal functional $\widehat{\Omega}_U[\boldsymbol{G}_0^{-1}]$, a universal functional $\widehat{\Phi}_U[\boldsymbol{G}]$ can be constructed, the derivative of which essentially yields $\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]$ and that also obeys all other properties of the diagrammatically constructed LW functional.

To start with, consider the equation

$$\widehat{\mathcal{G}}_U[\boldsymbol{G}^{-1} + \boldsymbol{\Sigma}] = \boldsymbol{G} \,. \tag{10.23}$$

This is a relation between the variables $\boldsymbol{\Sigma}$ and $\boldsymbol{G}$, which for a given $\boldsymbol{G}$ may be solved for $\boldsymbol{\Sigma}$. This defines a functional $\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]$, i.e.

$$\widehat{\mathcal{G}}_U[\boldsymbol{G}^{-1} - \widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]] = \boldsymbol{G} \,. \tag{10.24}$$

For a given Green's function $\boldsymbol{G}$, the self-energy $\boldsymbol{\Sigma} = \widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]$ is defined to be the solution of (10.23). From the Dyson equation (10.7) and (10.22) it is obvious that the relation (10.23) is satisfied for the physical $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{t,U}$ and the physical $\boldsymbol{G} = \boldsymbol{G}_{t,U}$ of the system with Hamiltonian $H_{t,U}$:

$$\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}_{t,U}] = \boldsymbol{\Sigma}_{t,U} \,. \tag{10.25}$$

This construction simplifies the original presentation in [24]. The discussion on the existence and the uniqueness of possible solutions of the relation (10.23) given there applies accordingly to the present case.

With the help of the functionals $\widehat{\Omega}_U[\boldsymbol{G}_0^{-1}]$ and $\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]$, the Luttinger–Ward functional is obtained as:

$$\widehat{\Phi}_U[\boldsymbol{G}] = \widehat{\Omega}_U[\boldsymbol{G}^{-1} + \widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]] - \mathrm{Tr}\ln\boldsymbol{G} + \mathrm{Tr}(\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]\boldsymbol{G}) \,. \tag{10.26}$$

Let us check property (4). Using (10.21) one finds for the derivative of the first term:

$$\frac{1}{T}\frac{\delta\widehat{\Omega}_U[\boldsymbol{G}^{-1}+\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]]}{\delta G_{\alpha\beta}} = -\sum_{\gamma\delta}\widehat{\mathcal{G}}_{U,\gamma\delta}[\boldsymbol{G}^{-1}+\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]]\left(\frac{\delta G^{-1}_{\gamma\delta}}{\delta G_{\alpha\beta}}+\frac{\delta\widehat{\Sigma}_{U,\gamma\delta}[\boldsymbol{G}]}{\delta G_{\alpha\beta}}\right) \tag{10.27}$$

and, using (10.24),

$$\frac{1}{T}\frac{\delta\widehat{\Omega}_U[\boldsymbol{G}^{-1}+\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]]}{\delta G_{\alpha\beta}} = -\sum_{\gamma\delta}G_{\gamma\delta}\left(\frac{\delta G^{-1}_{\gamma\delta}}{\delta G_{\alpha\beta}}+\frac{\delta\widehat{\Sigma}_{U,\gamma\delta}[\boldsymbol{G}]}{\delta G_{\alpha\beta}}\right). \tag{10.28}$$

Therewith,

$$\frac{1}{T}\frac{\delta\widehat{\Phi}_U[\boldsymbol{G}]}{\delta G_{\alpha\beta}} = G^{-1}_{\beta\alpha}-\sum_{\gamma\delta}G_{\gamma\delta}\frac{\delta\widehat{\Sigma}_{U,\gamma\delta}[\boldsymbol{G}]}{\delta G_{\alpha\beta}}+\frac{1}{T}\frac{\delta}{\delta G_{\alpha\beta}}\left(-\mathrm{Tr}\ln\boldsymbol{G}+\mathrm{Tr}(\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]\boldsymbol{G})\right) \tag{10.29}$$

and, finally,

$$\frac{1}{T}\frac{\delta\widehat{\Phi}_U[\boldsymbol{G}]}{\delta G_{\alpha\beta}(i\omega_n)} = \widehat{\Sigma}_{U,\beta\alpha}(i\omega_n)[\boldsymbol{G}]\,, \tag{10.30}$$

where, as a reminder, the frequency dependence has been reintroduced.

The other properties are also verified easily. (1) and (2) are obvious. (3) follows from (10.20), (10.22) and (10.25) and the Dyson equation (10.7). The universality of the LW functional (5) is ensured by the presented construction. It involves universal functionals only. Finally, (6) follows from $\widehat{\mathcal{G}}_{U=0}[\boldsymbol{G}^{-1}] = \boldsymbol{G}$ which implies (via (10.24)) $\widehat{\boldsymbol{\Sigma}}_{U=0}[\boldsymbol{G}] = 0$, and with $\widehat{\Omega}_{U=0}[\boldsymbol{G}^{-1}] = \mathrm{Tr}\ln\boldsymbol{G}$ we get $\widehat{\Phi}_{U=0}[\boldsymbol{G}] = 0$.

10.2.5 Variational Principle

The functional $\boldsymbol{\Sigma}_U[\boldsymbol{G}]$ can be assumed to be invertible *locally* provided that the system is not at a critical point for a phase transition (see also [11]). This allows us to construct the Legendre transform of the LW functional:

$$\widehat{F}_U[\boldsymbol{\Sigma}] = \widehat{\Phi}_U[\widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}]]-\mathrm{Tr}(\boldsymbol{\Sigma}\widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}])\,. \tag{10.31}$$

Here, $\widehat{\boldsymbol{G}}_U[\widehat{\boldsymbol{\Sigma}}_U[\boldsymbol{G}]] = \boldsymbol{G}$. With (10.30) we immediately find

$$\frac{1}{T}\frac{\delta\widehat{F}_U[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}} = -\widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}]\,. \tag{10.32}$$

We now define the self-energy-functional:

$$\widehat{\Omega}_{t,U}[\boldsymbol{\Sigma}] = \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}_{t,0}^{-1}-\boldsymbol{\Sigma}} + \widehat{F}_U[\boldsymbol{\Sigma}]\,. \tag{10.33}$$

Its functional derivative is easily calculated:

$$\frac{1}{T}\frac{\delta\widehat{\Omega}_{t,U}[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}} = \frac{1}{\boldsymbol{G}_{t,0}^{-1}-\boldsymbol{\Sigma}} - \widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}]\,. \tag{10.34}$$

The equation

$$\widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}] = \frac{1}{\boldsymbol{G}_{t,0}^{-1}-\boldsymbol{\Sigma}} \tag{10.35}$$

is a (highly non-linear) conditional equation for the self-energy of the system $H = H_0(\boldsymbol{t}) + H_1(\boldsymbol{U})$. Equations (10.7) and (10.25) show that it is satisfied by the physical self-energy $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{t,U}$. Note that the l.h.s of (10.35) is independent of $\boldsymbol{t}$ but depends on $\boldsymbol{U}$ (universality of $\widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}]$), while the r.h.s is independent of $\boldsymbol{U}$ but depends on $\boldsymbol{t}$ via $\boldsymbol{G}_{t,0}^{-1}$. The obvious problem of finding a solution of (10.35) is that there is no closed form for the functional $\widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}]$. Solving (10.35) is equivalent to a search for the stationary point of the grand potential as a functional of the self-energy:

$$\frac{\delta\widehat{\Omega}_{t,U}[\boldsymbol{\Sigma}]}{\delta\boldsymbol{\Sigma}} = 0\,. \tag{10.36}$$

This is the starting point for the SFT.

10.2.6 Approximation Schemes

Up to this point we have discussed exact relations only. It is clear, however, that it is generally impossible to evaluate the self-energy-functional (10.33) for a given $\boldsymbol{\Sigma}$ and that one has to resort to approximations. Three different types of approximation strategies may be distinguished:

In a *type-I approximation* one derives the Euler equation $\delta\widehat{\Omega}_{t,U}[\boldsymbol{\Sigma}]/\delta\boldsymbol{\Sigma} = 0$ first and then chooses (a physically motivated) simplification of the equation afterwards to render the determination of $\boldsymbol{\Sigma}_{t,U}$ possible. This is most general but also questionable a priori, as normally the approximated Euler equation no longer derives from some approximate functional. This may result in thermodynamical inconsistencies.

A *type-II approximation* modifies the form of the functional dependence, $\widehat{\Omega}_{t,U}[\cdots] \to \widehat{\Omega}^{(1)}_{t,U}[\cdots]$, to get a simpler one that allows for a solution of the

resulting Euler equation $\delta\widehat{\Omega}^{(1)}_{\boldsymbol{t},\boldsymbol{U}}[\boldsymbol{\Sigma}]/\delta\boldsymbol{\Sigma} = 0$. This type is more particular and yields a thermodynamical potential consistent with $\boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}}$. Generally, however, it is not easy to find a sensible approximation of a functional form.

Finally, in a *type-III approximation* one restricts the domain of the functional which must then be defined precisely. This type is most specific and, from a conceptual point of view, should be preferred as compared to type-I or type-II approximations as the exact functional form is retained. In addition to conceptual clarity and thermodynamical consistency, type-III approximations are truly systematic since improvements can be obtained by an according extension of the domain.

Examples for the different cases can be found e.g., in [14]. The classification of approximation schemes is hierarchical: Any type-III approximation can always be understood as a type-II one, and any type-II approximations as type-I, but not vice versa. This does not mean, however, that type-III approximations are superior as compared to type-II and type-I ones. They are conceptually more appealing but do not necessarily provide "better" results. One reason to consider self-energy functionals instead of functionals of the Green's function (see [25,26], for example) is to derive the DMFT as a type-III approximation.

10.3 Variational Cluster Approach

The central idea of SFT is to make use of the universality of (the Legendre transform of) the Luttinger–Ward functional to construct type-III approximations. Consider the self-energy-functional (10.33). Its first part consists of a simple explicit functional of $\boldsymbol{\Sigma}$ while its second part, i.e., $\widehat{F}_{\boldsymbol{U}}[\boldsymbol{\Sigma}]$, is unknown but depends on $\boldsymbol{U}$ only.

10.3.1 Reference System

Due to this universality of $\widehat{F}_{\boldsymbol{U}}[\boldsymbol{\Sigma}]$, one has

$$\widehat{\Omega}_{\boldsymbol{t}',\boldsymbol{U}}[\boldsymbol{\Sigma}] = \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t}',0} - \boldsymbol{\Sigma}} + \widehat{F}_{\boldsymbol{U}}[\boldsymbol{\Sigma}] \tag{10.37}$$

for the self-energy-functional of a so-called "reference system." As compared to the original system of interest, the reference system is given by a Hamiltonian $H' = H_0(\boldsymbol{t}') + H_1(\boldsymbol{U})$ with the same interaction part $\boldsymbol{U}$ but modified one-particle parameters $\boldsymbol{t}'$. The reference system has different microscopic parameters but is assumed to be in the same macroscopic state, i.e., at the same temperature T and the same chemical potential μ. By a proper choice of its one-particle part, the problem

posed by the reference system H' can be much simpler than the original problem posed by H. We assume that the self-energy of the reference system $\boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}$ can be computed exactly, e.g., by some numerical technique.

Combining (10.33) and (10.37), one can eliminate the unknown functional $\widehat{F}_{\boldsymbol{U}}[\boldsymbol{\Sigma}]$:

$$\widehat{\Omega}_{\boldsymbol{t},\boldsymbol{U}}[\boldsymbol{\Sigma}] = \widehat{\Omega}_{\boldsymbol{t}',\boldsymbol{U}}[\boldsymbol{\Sigma}] + \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}} - \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t}',0} - \boldsymbol{\Sigma}}\,. \tag{10.38}$$

It appears that this amounts to a shift of the problem only as the self-energy-*functional* of the reference system again contains the full complexity of the problem. In fact, except for the trivial case $\boldsymbol{U} = 0$, the functional dependence of $\widehat{\Omega}_{\boldsymbol{t}',\boldsymbol{U}}[\boldsymbol{\Sigma}]$ is unknown – even if the reference system is assumed to be solvable, i.e., if the self-energy $\boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}$, the Green's function $\boldsymbol{G}_{\boldsymbol{t}',\boldsymbol{U}}$ and the grand potential $\Omega_{\boldsymbol{t}',\boldsymbol{U}}$ of the reference system are available.

However, inserting the self-energy of the reference system $\boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}$ into the self-energy-functional of the original one, and using $\widehat{\Omega}_{\boldsymbol{t}',\boldsymbol{U}}[\boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}] = \Omega_{\boldsymbol{t}',\boldsymbol{U}}$ and the Dyson equation of the reference system, we find:

$$\widehat{\Omega}_{\boldsymbol{t},\boldsymbol{U}}[\boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}] = \Omega_{\boldsymbol{t}',\boldsymbol{U}} + \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}} - \mathrm{Tr}\ln \boldsymbol{G}_{\boldsymbol{t}',\boldsymbol{U}}\,. \tag{10.39}$$

This is a remarkable result. It shows that an *exact* evaluation of the self-energy-functional of a non-trivial original system is possible, at least for certain self-energies. This requires us to solve a reference system with the same interaction part.

10.3.2 Domain of the Self-Energy-Functional

Equation (10.39) provides an explicit expression of the self-energy-functional $\widehat{\Omega}_{\boldsymbol{t},\boldsymbol{U}}[\boldsymbol{\Sigma}]$. This is suitable to discuss the domain of the functional precisely. Take $\boldsymbol{U}$ to be fixed. We define the space of $\boldsymbol{t}$-representable self-energies as

$$\mathcal{S}_{\boldsymbol{U}} = \{\boldsymbol{\Sigma} \mid \exists \boldsymbol{t} : \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\boldsymbol{t},\boldsymbol{U}}\}. \tag{10.40}$$

This definition of the domain is very convenient since it ensures the correct analytical and causal properties of the variable $\boldsymbol{\Sigma}$.

We can now formulate the result of the preceeding section in the following way. Consider a set of reference systems with $\boldsymbol{U}$ fixed but different one-particle parameters $\boldsymbol{t}'$, i.e., a space of one-particle parameters $\mathcal{T}'$. Assume that the reference

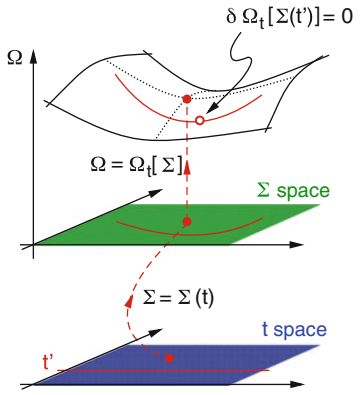

Fig. 10.2 Schematic illustration for the construction of consistent approximations within the SFT. The grand potential is considered as a functional of the self-energy which is parametrized by the one-particle parameters $\boldsymbol{t}$ of the Hamiltonian ($\boldsymbol{U}$ is fixed). At the physical self-energy, Ω is stationary (*filled red circles*). The functional dependence of Ω on $\boldsymbol{\Sigma}$ is not accessible in the entire space of self-energies (Σ space) but on a restricted subspace of "trial" self-energies parametrized by a subset of one-particle parameters $\boldsymbol{t}'$ (*solid red lines*) corresponding to a "reference system," i.e., a manifold of systems with the same interaction part but a one-particle part given by $\boldsymbol{t}'$ which renders the solution possible. The grand potential can be evaluated exactly on the submanifold of reference systems. A stationary point on this submanifold represents the approximate self-energy and the corresponding approximate grand potential (*open circle*)

system with $H' = H_{\boldsymbol{t}',U}$ can be solved exactly for any $\boldsymbol{t}' \in \mathcal{T}'$. Then, the self-energy-functional $\widehat{\Omega}_{\boldsymbol{t},U}[\boldsymbol{\Sigma}]$ can be evaluated exactly on the subspace

$$\mathcal{S}'_U = \{\boldsymbol{\Sigma} \mid \exists \boldsymbol{t}' \in \mathcal{T}' : \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\boldsymbol{t}',U}\} \subset \mathcal{S}_U . \tag{10.41}$$

This fact can be used to construct type-III approximations, see Fig. 10.2.

10.3.3 Construction of Cluster Approximations

A certain approximation is defined by a choice of the reference system or actually by a manifold of reference systems specified by a manifold of one-particle parameters $\mathcal{T}'$. As an example consider Fig. 10.3. The original system is given by the one-dimensional Hubbard model with nearest-neighbor hopping t and Hubbard interaction U. A possible reference system is given by switching off the hopping between clusters consisting of $L_c = 2$ sites each. The hopping within the cluster t' is arbitrary, this defines the space $\mathcal{T}'$. The self-energies in $\mathcal{S}'$, the corresponding Green's functions and grand potentials of the reference system can obviously be calculated easily since the degrees of freedom are decoupled spatially. Inserting

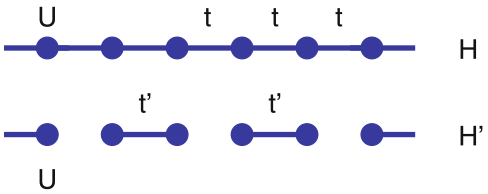

Fig. 10.3 Variational cluster approximation (VCA) for the Hubbard model. *Top*: representation of the original one-dimensional Hubbard model H with nearest-neighbor hopping t and Hubbard interaction U. *Bottom*: reference system H' consisting of decoupled clusters of $L_{\rm c} = 2$ sites each with intra-cluster hopping t' as a variational parameter

these quantities in (10.39) yields the SFT grand potential as a function of $\boldsymbol{t}'$:

$$\Omega(\boldsymbol{t}') \equiv \widehat{\Omega}_{\boldsymbol{t},U}[\boldsymbol{\Sigma}_{\boldsymbol{t}',U}]\,. \tag{10.42}$$

This is no longer a functional but an ordinary function of the variational parameters $\boldsymbol{t}' \in \mathcal{T}'$. The final task then consists in finding a stationary point $\boldsymbol{t}'_{\rm opt}$ of this function:

$$\frac{\partial\Omega(\boldsymbol{t}')}{\partial\boldsymbol{t}'} = 0 \qquad \text{for } \boldsymbol{t}' = \boldsymbol{t}'_{\rm opt}\,. \tag{10.43}$$

In the example considered this is a function of a single variable t' (we assume t' to be the same for all clusters). Note that not only the reference system (in the example the isolated cluster) defines the final result but also the lattice structure and the one-particle parameters of the original system. These enter $\Omega(\boldsymbol{t}')$ via the free Green's function $G_{\boldsymbol{t},0}$ of the original system. In the first term on the r.h.s of (10.39) we just recognize the CPT Green's function $1/(\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}_{\boldsymbol{t}',U})$. The approximation generated by a reference system of disconnected clusters is called variational cluster approximation (VCA).

An example for the results of a numerical calculation is given in Fig. 10.4, see also [27]. The calculation has been performed for the one-dimensional particle–hole symmetric Hubbard model at half-filling and zero temperature. The figure shows the numerical results for the optimal nearest-neighbor intra-cluster hopping t' as obtained from the VCA for a reference system with disconnected clusters consisting of $L_{\rm c} = 10$ sites each. The hopping t' is assumed to be the same for all pairs of nearest neighbors. In principle, one could vary all one-particle parameters that do not lead to a coupling of the clusters to get the optimal result. In most cases, however, it is necessary to restrict oneself to a small number of physically motivated variational parameters to avoid complications arising from a search for a stationary point in a high-dimensional parameter space. For the example discussed here, the parameter space $\mathcal{T}'$ is one-dimensional only. This is the most simple choice but more elaborate approximations can be generated easily. The flexibility to construct approximations of different quality and complexity must be seen as one of the advantages of the VCA and of the SFT in general.

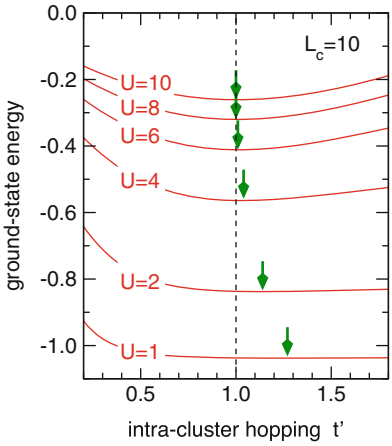

Fig. 10.4 (Taken from [27].) SFT ground-state energy per site, i.e., $(\Omega(t') + \mu\langle N\rangle)/L$, as a function of the intra-cluster nearest-neighbour hopping t' for $L_c = 10$ and different U ($\mu = U/2$) at zero temperature. *Arrows* indicate the respective optimal t'. The energy scale is fixed by $t = 1$

As can be seen from the figure, a non-trivial result, namely $t'_{\rm opt} \neq t = 1$, is found for the optimal value of $t'_{\rm opt}$. We also notice that $t_{\rm opt} > t$. The physical interpretation is that switching off the inter-cluster hopping, which generates the approximate self-energy, can partially be compensated for by enhancing the intra-cluster hopping. The effect is the more pronounced the smaller is the cluster size L_c. Furthermore, it is reasonable that in the case of a stronger interaction and thus more localized electrons, switching off the inter-cluster hopping is less significant. This can be seen in Fig. 10.4: The largest optimal hopping $t'_{\rm opt}$ is obtained for the smallest U.

On the other hand, even a "strong" approximation for the self-energy (measured as a strong deviation of $t'_{\rm opt}$ from t) becomes irrelevant in the weak-coupling limit because the self-energy must vanish for $U = 0$. Generally, we note that the VCA becomes exact in the limit $\boldsymbol{U} = 0$: In (10.38) the first and the third terms on the r.h.s cancel each other and we are left with

$$\widehat{\Omega}_{\boldsymbol{t},U=0}[\boldsymbol{\Sigma}] = \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}}\,. \tag{10.44}$$

Since the trial self-energy has to be taken from a reference system with the same interaction part, i.e., $\boldsymbol{U} = 0$ and thus $\boldsymbol{\Sigma} = 0$, the limit becomes trivial. For weak but finite $\boldsymbol{U}$, the SFT grand potential becomes flatter and flatter until for $\boldsymbol{U} = 0$ the $\boldsymbol{t}'$ dependence is completely irrelevant.

The VCA is also exact in the atomic limit or, more general and again trivial, in the case that there is no restriction on the trial self-energies: $\mathcal{S}' = \mathcal{S}$. In this case, $\boldsymbol{t}'_{\rm opt} = \boldsymbol{t}$ solves the problem, i.e., the second and the third term on the r.h.s of (10.39) cancel each other and $\widehat{\Omega}_{\boldsymbol{t},U}[\boldsymbol{\Sigma}_{\boldsymbol{t}',U}] = \Omega_{\boldsymbol{t}',U}$ for $\boldsymbol{t}' = \boldsymbol{t}$.

Cluster-perturbation theory (CPT) can be understood as being identical with the VCA provided that the SFT expression for the grand potential is used and that no parameter optimization at all is performed. As can be seen from Fig. 10.4, there is

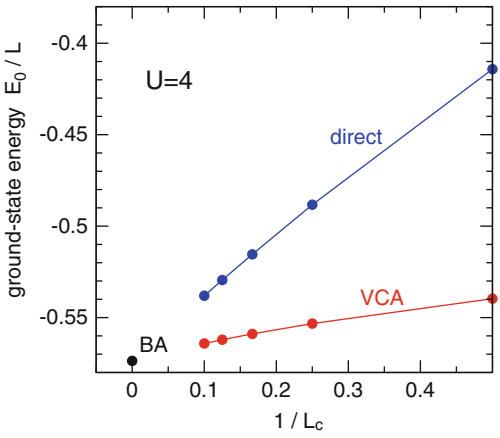

Fig. 10.5 (Taken from [27]). Optimal VCA ground-state energy per site for $U = 4$ for different cluster sizes L_c as a function of $1/L_c$ compared to the exact (BA) result and the direct cluster approach

a gain in binding energy due to the optimization of t', i.e., $\Omega(t'_{\rm opt}) < \Omega(t)$. This means that the VCA improves on the CPT result.

Figure 10.5 shows the ground-state energy (per site), i.e., the SFT grand potential at zero temperature constantly shifted by μN at the stationary point, as a function of the inverse cluster size $1/L_c$. The dependence turns out to be quite regular and allows us to recover the exact Bethe-Ansatz result (BA) [28] by extrapolation to $1/L_c = 0$. The VCA result represents a considerable improvement as compared to the "direct" cluster approach where E_0 is simply approximated by the ground-state energy of an isolated Hubbard chain (with open boundary conditions). Convergence to the exact result is clearly faster within the VCA. Note that the direct cluster approach, opposed to the VCA, is not exact for $U = 0$.

In the example discussed so far a single variational parameter was taken into account only. More parameters can be useful for different reasons. For example, the optimal self-energy provided by the VCA as a real-space cluster technique artificially breaks the translational symmetry of the original lattice problem. Finite-size effects are expected to be the most pronounced at the cluster boundary. This suggests us to consider all intra-cluster hopping parameters as independent variational parameters or at least the hopping at the edges of the chain.

The result is shown in Fig. 10.6. We find that the optimal hopping varies between different nearest neighbours within a range of less than 10%. At the chain edges the optimal hopping is enhanced to compensate the loss of itinerancy due to the switched-off inter-cluster hopping within the VCA. With increasing distance to the edges, the hopping quickly decreases. Quite generally, the third hopping parameter is already close to the physical hopping t. Looking at the $L_c = 10$ results where all (five) different hopping parameters have been varied independently (orange circles), one can see the hopping to slightly oscillate around the bulk value reminiscent of surface Friedel oscillations.

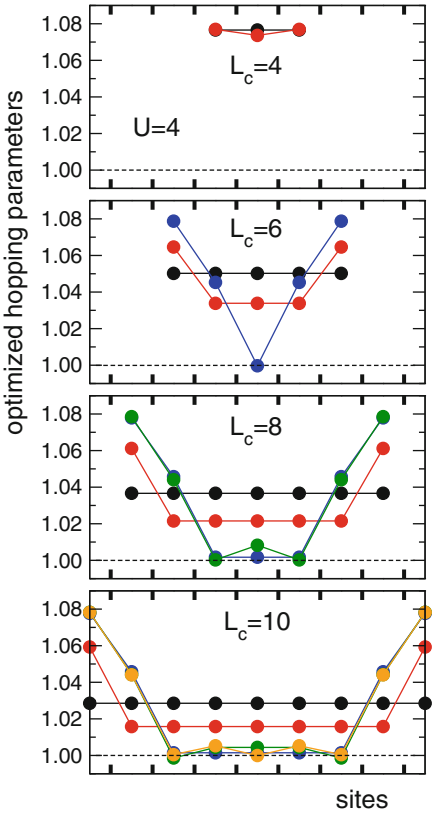

Fig. 10.6 (Taken from [27].) Optimized hopping parameters for the reference systems shown in Fig. 10.3 but for larger clusters with L_c sites each as indicated. VCA results for $U = 4$, $t = 1$ at half-filling and temperature $T = 0$. *Black*: hopping assumed to be uniform. *Red*: two hopping parameters varied independently, the hopping at the two cluster edges and the "bulk" hopping. *Blue*: hopping at the edges, next to the edges and bulk hopping varied. *Green*: four hopping parameters varied. *Orange*: five hopping parameters varied

The optimal SFT grand potential is found to be lower for the inhomogeneous cases as compared to the homogeneous (black) one. Generally, the more variational parameters are taken into account the higher is the decrease of the SFT grand potential at optimal parameters. However, the binding-energy gain due to inhomogeneous hopping parameters is much smaller compared to the gain obtained with a larger cluster.

Considering an additional hopping parameter t_{pbc} linking the two chain edges as a variational parameter, i.e., clusters with periodic boundary conditions always give a minimal SFT grand potential at $t_{\mathrm{pbc}} = 0$ (instead of a stationary point at $t_{\mathrm{pbc}} = 1$). This implies that open boundary conditions are preferred (see also [13]).

10.4 Consistency of Approximations

10.4.1 Analytical Structure of the Green's Function

Constructing approximations within the framework of a dynamical variational principle means that, besides an approximate thermodynamical potential, approximate expressions for the self-energy and the one-particle Greens function are obtained. This raises the question whether their correct analytical structure is respected in an approximation. For approximations obtained from SFT this is easily shown to be the case in fact.

The physical self-energy $\Sigma_{\alpha\beta}(\omega)$ and the physical Green's function $G_{\alpha\beta}(\omega)$ are analytical functions in the entire complex ω plane except for the real axis and have a spectral representation (see (10.4)) with non-negative diagonal elements of the spectral function.

This trivially holds for the SFT self-energy $\Sigma_{\boldsymbol{t}',\boldsymbol{U},\alpha\beta}(\omega)$ since by construction $\Sigma_{\boldsymbol{t}',\boldsymbol{U},\alpha\beta}(\omega)$ is the *exact* self-energy of a reference system. The SFT Green's function is obtained from the SFT self-energy and the free Green's function of the original model via Dyson's equation:

$$\boldsymbol{G} = \frac{1}{\boldsymbol{G}_{\boldsymbol{t},0}^{-1} - \boldsymbol{\Sigma}_{\boldsymbol{t}',\boldsymbol{U}}} \,. \tag{10.45}$$

It is easy to see that it is analytical in the complex plane except for the real axis. To verify that it has a spectral representation with non-negative spectral function, we can equivalently consider the corresponding retarded quantity $\boldsymbol{G}_{\mathrm{ret}}(\omega) = \boldsymbol{G}(\omega + i0^+)$ for real frequencies ω and verify that $\boldsymbol{G}_{\mathrm{ret}} = \boldsymbol{G}_{\mathrm{R}} - i\,\boldsymbol{G}_{\mathrm{I}}$ with $\boldsymbol{G}_{\mathrm{R}}$, $\boldsymbol{G}_{\mathrm{I}}$ Hermitian and $\boldsymbol{G}_{\mathrm{I}}$ non-negative:

We can assume that $\boldsymbol{G}_{0,\mathrm{ret}} = \boldsymbol{G}_{0,\mathrm{R}} - i\,\boldsymbol{G}_{0,\mathrm{I}}$ with $\boldsymbol{G}_{0,\mathrm{R}}$, $\boldsymbol{G}_{0,\mathrm{I}}$ Hermitian and $\boldsymbol{G}_{0,\mathrm{I}}$ non-negative. Since for Hermitian matrices $\boldsymbol{A}$, $\boldsymbol{B}$ with $\boldsymbol{B}$ non-negative, one has $1/(\boldsymbol{A} \pm i\,\boldsymbol{B}) = \boldsymbol{X} \mp i\,\boldsymbol{Y}$ with $\boldsymbol{X}$, $\boldsymbol{Y}$ Hermitian and $\boldsymbol{Y}$ non-negative (see [12]), we find $\boldsymbol{G}_{0,\mathrm{ret}}^{-1} = \boldsymbol{P}_{\mathrm{R}} + i\,\boldsymbol{P}_{\mathrm{I}}$ with $\boldsymbol{P}_{\mathrm{R}}$, $\boldsymbol{P}_{\mathrm{I}}$ Hermitian and $\boldsymbol{P}_{\mathrm{I}}$ non-negative. Furthermore, we have $\boldsymbol{\Sigma}_{\mathrm{ret}} = \boldsymbol{\Sigma}_{\mathrm{R}} - i\,\boldsymbol{\Sigma}_{\mathrm{I}}$ with $\boldsymbol{\Sigma}_{\mathrm{R}}$, $\boldsymbol{\Sigma}_{\mathrm{I}}$ Hermitian and $\boldsymbol{\Sigma}_{\mathrm{I}}$ non-negative. Therefore,

$$\boldsymbol{G}_{\mathrm{ret}} = \frac{1}{\boldsymbol{P}_{\mathrm{R}} + i\,\boldsymbol{P}_{\mathrm{I}} - \boldsymbol{\Sigma}_{\mathrm{R}} + i\,\boldsymbol{\Sigma}_{\mathrm{I}}} = \frac{1}{\boldsymbol{Q}_{\mathrm{R}} + i\,\boldsymbol{Q}_{\mathrm{I}}} \tag{10.46}$$

with $\boldsymbol{Q}_{\mathrm{R}}$ Hermitian and $\boldsymbol{Q}_{\mathrm{I}}$ Hermitian and non-negative.

Note that the proof given here is simpler than corresponding proofs for cluster extensions of the DMFT [29, 30] because the SFT does not involve a "self-consistency condition" which is the main object of concern for possible causality violations.

10.4.2 Thermodynamical Consistency

An advantageous feature of the VCA and of other approximations within the SFT framework is their internal thermodynamical consistency. This is due to the fact that all quantities of interest are derived from an approximate but explicit expression for a thermodynamical potential. In principle the expectation value of any observable should be calculated by

$$\langle A \rangle = \frac{\partial \Omega}{\partial \lambda_A} , \tag{10.47}$$

where $\Omega \equiv \Omega(\boldsymbol{t}')$ is the SFT grand potential (see (10.42)) at $\boldsymbol{t}' = \boldsymbol{t}'_{\text{opt}}$ and λ_A is a parameter in the Hamiltonian of the original system which couples linearly to A, i.e., $H_{\boldsymbol{t},U} = H_0 + \lambda_A A$. This ensures, for example, that the Maxwell relations

$$\frac{\partial \langle A \rangle}{\partial \lambda_B} = \frac{\partial \langle B \rangle}{\partial \lambda_A} \tag{10.48}$$

are respected.

Furthermore, thermodynamical consistency means that expectation values of arbitrary one-particle operators $A = \sum_{\alpha\beta} A_{\alpha\beta} c^\dagger_\alpha c_\beta$ can consistently either be calculated by a corresponding partial derivative of the grand potential on the one hand, or by integration of the one-particle spectral function on the other. As an example we consider the total particle number $N = \sum_\alpha c^\dagger_\alpha c_\alpha$. A priori it is not guaranteed that in an approximate theory the expressions

$$\langle N \rangle = -\frac{\partial \Omega}{\partial \mu} , \tag{10.49}$$

and

$$\langle N \rangle = \sum_\alpha \int_{-\infty}^{\infty} f(z) A_{\alpha\alpha}(z) \mathrm{d}z \tag{10.50}$$

with $f(z) = 1/(\exp(z/T) + 1)$ and the spectral function $A_{\alpha\beta}(z)$ will give the same result.

To prove thermodynamic consistency, we start from (10.49). According to (10.39), there is a two-fold μ dependence of $\Omega \equiv \Omega_{\boldsymbol{t},U}[\boldsymbol{\Sigma}_{\boldsymbol{t}'_{\text{opt}},U}]$: (1) the *explicit* μ dependence due to the chemical potential in the free Green's function of the original model, $\boldsymbol{G}^{-1}_{\boldsymbol{t},0} = \omega + \mu - \boldsymbol{t}$, and (2) an *implicit* μ dependence due to the μ dependence of the self-energy $\boldsymbol{\Sigma}_{\boldsymbol{t}'_{\text{opt}},U}$, the Green's function $\boldsymbol{G}_{\boldsymbol{t}'_{\text{opt}},U}$ and the grand potential $\Omega_{\boldsymbol{t}'_{\text{opt}},U}$ of the reference system:

$$\langle N \rangle = -\frac{\partial \Omega}{\partial \mu_{\text{ex.}}} - \frac{\partial \Omega}{\partial \mu_{\text{im.}}} . \tag{10.51}$$

Note that the implicit μ dependence is due to the chemical potential of the reference system which, by construction, is in the same macroscopic state as the original system *as well as* due to the μ dependence of the stationary point $\boldsymbol{t}'_{\text{opt}}$ itself. The latter can be ignored since

$$\frac{\partial \Omega}{\partial \boldsymbol{t}'} \cdot \frac{\partial \boldsymbol{t}'}{\partial \mu} = 0 \tag{10.52}$$

for $\boldsymbol{t}' = \boldsymbol{t}'_{\text{opt}}$ because of stationarity condition (10.43).

We assume that an overall shift of the one-particle energies $\varepsilon' \equiv t'_{\alpha\alpha}$ is included in the set $\mathcal{T}'$ of variational parameters. Apart from the sign this is essentially the "chemical potential" in the reference system but should be formally distinguished from μ since the latter has a macroscopic thermodynamical meaning and is the same as the chemical potential of the original system which should not be seen as a variational parameter.

The self-energy, the Green's function and the grand potential of the reference system are defined as grand-canonical averages. Hence, their μ dependence due to the grand-canonical Hamiltonian $\mathcal{H}' = H' - \mu N$ is (apart from the sign) the same as their dependence on ε': Consequently, we have:

$$\frac{\partial \Omega}{\partial \mu_{\text{im.}}} = -\frac{\partial \Omega}{\partial \varepsilon'} = 0 \tag{10.53}$$

due to the stationarity condition again.

We are then left with the explicit μ dependence only:

$$\frac{\partial \Omega}{\partial \mu_{\text{ex.}}} = \frac{\partial}{\partial \mu_{\text{ex.}}} \text{Tr} \ln \frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}_{\boldsymbol{t}'_{\text{opt}},U}} = -\text{Tr} \frac{1}{\boldsymbol{G}^{-1}_{\boldsymbol{t},0} - \boldsymbol{\Sigma}_{\boldsymbol{t}'_{\text{opt}},U}} \,. \tag{10.54}$$

Converting the sum over the Matsubara frequencies implicit in the trace Tr into a contour integral in the complex ω plane and using Cauchy's theorem, we can proceed to an integration over real frequencies. Inserting into (10.51), this yields:

$$\langle N \rangle = -\frac{1}{\pi} \text{Im} \int_{-\infty}^{\infty} f(\omega) \, \text{Tr} \frac{1}{\boldsymbol{G}^{-1}_{0,\boldsymbol{t}} - \boldsymbol{\Sigma}_{\boldsymbol{t}',U}} \bigg|_{\omega + i0+} \text{d}\omega \tag{10.55}$$

for $\boldsymbol{t}' = \boldsymbol{t}'_{\text{opt}}$ which is just the average particle number given by (10.50). This completes the proof.

10.4.3 Symmetry Breaking

The above discussion has shown that besides intra-cluster hopping parameters it can also be important to treat one-particle energies in the reference cluster as variational

parameters. In particular, one may consider variational parameters which lead to a lower symmetry of the Hamiltonian.

As an example consider the Hubbard model on a bipartite lattice as the system of interest and disconnected clusters of size L_c as a reference system. The reference-system Hamiltonian shall include, e.g., an additional a staggered magnetic-field term:

$$H'_{\text{fict}} = B' \sum_{i\sigma} z_i (n_{i\uparrow} - n_{i\downarrow}) \,, \tag{10.56}$$

where $z_i = +1$ for sites on sublattice 1, and $z_i = -1$ for sublattice 2. The additional term H'_{fict} leads to a valid reference system as there is no change of the interaction part. We include the field strength B' in the set of variational parameters, $B' \in \mathcal{T}'$.

B' is the strength of a ficticious field or, in the language of mean-field theory, the strength of the internal magnetic field or the Weiss field. This has to be distinguished clearly from an external *physical* field applied to the system with field strength B:

$$H_{\text{phys}} = B \sum_{i\sigma} z_i (n_{i\uparrow} - n_{i\downarrow}) \tag{10.57}$$

This term adds to the Hamiltonian of the original system.

We expect $B'_{\text{opt}} = 0$ in the case of the paramagnetic state and $B = 0$ (and this is easily verified numerically). Consider the B' and B dependence of the SFT grand potential $\Omega(B', B) = \Omega_B[\boldsymbol{\Sigma}_{B'}]$. Here we have suppressed the dependencies on other variational parameters $\boldsymbol{t}'$ and on $\boldsymbol{t}, \boldsymbol{U}$. Due to the stationarity condition, $\partial\Omega(B', B)/\partial B' = 0$, the optimal Weiss field B' can be considered as a function of B, i.e., $B'_{\text{opt}} = B'(B)$. Therefore, we also have:

$$\frac{\mathrm{d}}{\mathrm{d}B} \frac{\partial \Omega(B'(B), B)}{\partial B'} = 0 \,. \tag{10.58}$$

This yields:

$$\frac{\partial^2 \Omega(B'(B), B)}{\partial B'^2} \frac{\mathrm{d}B'(B)}{\mathrm{d}B} + \frac{\partial^2 \Omega(B'(B), B)}{\partial B \partial B'} = 0 \,. \tag{10.59}$$

Solving for $\mathrm{d}B'/\mathrm{d}B$ we find:

$$\frac{\mathrm{d}B'}{\mathrm{d}B} = -\left[\frac{\partial^2 \Omega}{\partial B'^2}\right]^{-1} \frac{\partial^2 \Omega}{\partial B \partial B'} \,. \tag{10.60}$$

This clearly shows that $B' = B'_{\text{opt}}$ has to be interpreted carefully. B' can be much stronger than B if the curvature $\partial^2\Omega/\partial B'^2$ of the SFT functional at the stationary point is small, i.e., if the functional is rather flat as it is the case in the limit $U \to 0$, for example.

From the SFT approximation for the staggered magnetization,

$$m = \sum_{i\sigma} z_i \langle (n_{i\uparrow} - n_{i\downarrow}) \rangle \approx \frac{\mathrm{d}}{\mathrm{d}B} \Omega(B'(B), B) = \frac{\partial \Omega(B'(B), B)}{\partial B} , \qquad (10.61)$$

where the stationarity condition has been used once more, we can calculate the susceptibility,

$$\chi = \frac{\mathrm{d}m}{\mathrm{d}B} = \frac{\partial^2 \Omega(B'(B), B)}{\partial B' \partial B} \frac{\mathrm{d}B'(B)}{\mathrm{d}B} + \frac{\partial^2 \Omega(B'(B), B)}{\partial B^2} . \qquad (10.62)$$

Using (10.60),

$$\chi = \frac{\partial^2 \Omega}{\partial B^2} - \left(\frac{\partial^2 \Omega}{\partial B'^2} \right)^{-1} \left(\frac{\partial^2 \Omega}{\partial B' \partial B} \right)^2 . \qquad (10.63)$$

We see that there are two contributions. The first term is due to the explicit B dependence in the SFT grand potential while the second is due to the implicit B dependence via the B dependence of the stationary point. Equation (10.63) also demonstrates (see [31]) that for the calculation of the paramagnetic susceptibility χ one may first consider spin-independent variational parameters only to find a stationary point. This strongly reduces the computational effort [32]. Once a stationary point is found, partial derivatives according to (10.63) have to be calculated with spin-dependent parameters $\boldsymbol{\lambda}$ in a single final step.

Spontaneous symmetry breaking is obtained at $B = 0$ if there is a stationary point with $B'_{\mathrm{opt}} \neq 0$. Figure 10.7 gives an example for the particle–hole symmetric Hubbard model on the square lattice at half-filling and zero temperature. As a reference system a cluster with $L_{\mathrm{c}} = 10$ sites is considered, and the ficticious staggered

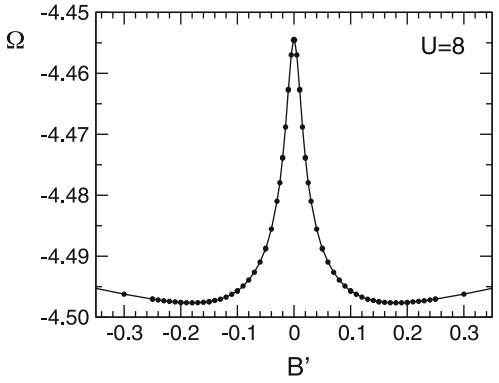

Fig. 10.7 (Taken from [33].) SFT grand potential as a function of the strength of a ficticious staggered magnetic field B'. VCA calculation using disconnected clusters consisting of $L_{\mathrm{c}} = 10$ sites each for the two-dimensional Hubbard model on the square lattice at half-filling, zero temperature, $U = 8$ and nearest-neighbour hopping $t = 1$

magnetic field is taken as the only variational parameter. There is a stationary point at $B' = 0$, which corresponds to the paramagnetic phase. At $B' = 0$ the usual CPT is recovered. The two equivalent stationary points at finite B' correspond to a phase with spontaneous antiferromagnetic order – as expected for the Hubbard model in this parameter regime. The antiferromagnetic ground state is stable as compared to the paramagnetic phase. Its order parameter m is the conjugate variable to the ficticious field. Since the latter is a variational parameter, m can either be calculated by integration of the spin-dependent spectral density or as the derivative of the SFT grand potential with respect to the physical field strength B with the same result. More details are given in [33].

The possibility to study spontaneous symmetry breaking using the VCA with suitably chosen Weiss fields as variational parameters has been exploited frequently in the past. Besides antiferromagnetism [33–36], spiral phases [37], ferromagnetism [32], d-wave superconductivity [38–44], charge order [45,46] and orbital order [47] have been investigated. The fact that an explicit expression for a thermodynamical potential is available allows us to study discontinuous transitions and phase separation as well.

10.4.4 Non-perturbative Conserving Approximations

Continuous symmetries of a Hamiltonian imply the existence of conserved quantities: The conservation of total energy, momentum, angular momentum, spin and particle number is enforced by a not explicitly time-dependent Hamiltonian which is spatially homogeneous and isotropic and invariant under global SU(2) and U(1) gauge transformations. Approximations may artificially break symmetries and thus lead to unphysical violations of conservation laws. Baym and Kadanoff [48, 49] have analyzed under which circumstances an approximation respects the mentioned macroscopic conservation laws. Within diagrammatic perturbation theory it could be shown that approximations that derive from an explicit but approximate expression for the LW functional Φ (Φ-derivable approximations) are "conserving." Examples for conserving approximations are the Hartree–Fock or the fluctuation-exchange approximation [48,50].

The SFT provides a framework to construct Φ-derivable approximations for correlated lattice models which are *non-perturbative*, i.e., do not employ truncations of the skeleton-diagram expansion. Like in weak-coupling conserving approximations, approximations within the SFT are derived from the LW functional, or its Legendre transform $F_U[\boldsymbol{\Sigma}]$. These are Φ-derivable since any type-III approximation can also be seen as a type-II one, see Sect. 10.2.6.

For fermionic lattice models, conservation of energy, particle number and spin have to be considered. Besides the static thermodynamics, the SFT concentrates on the *one-particle* excitations. For the approximate one-particle Green's function, however, it is actually simple to prove directly that the above conservation laws are respected. A short discussion is given in [51].

At zero temperature $T = 0$ there is another non-trivial theorem which is satisfied by any Φ-derivable approximation, namely Luttinger's sum rule [23,52]. This states that at zero temperature the volume in reciprocal space that is enclosed by the Fermi surface is equal to the average particle number. The original proof of the sum rule by Luttinger and Ward [23] is based on the skeleton-diagram expansion of Φ in the exact theory and is straightforwardly transferred to the case of a Φ-derivable approximation. This also implies that other Fermi-liquid properties, such as the linear trend of the specific heat at low T and Fermi-liquid expressions for the $T = 0$ charge and the spin susceptibility are respected by a Φ-derivable approximation.

For approximations constructed within the SFT, a different proof has to be found. One can start with (10.39) and perform the zero-temperature limit for an original system (and thus for a reference system) of a finite size L. The different terms in the SFT grand potential then consist of finite sums. The calculation proceeds by taking the μ-derivative, for $T = 0$, on both sides of (10.39). This yields the following result (see [51] for details):

$$\langle N\rangle = \langle N\rangle' + 2\sum_{\boldsymbol{k}} \Theta(G_{\boldsymbol{k}}(0)) - 2\sum_{\boldsymbol{k}} \Theta(G'_{\boldsymbol{k}}(0))\,. \tag{10.64}$$

Here, $\langle N\rangle$ ($\langle N\rangle'$) is the ground-state expectation value of the total particle number N in the original (reference) system, and $G_{\boldsymbol{k}}(0)$ ($G'_{\boldsymbol{k}}(0)$) is the diagonal elements of the one-electron Green's function $\boldsymbol{G}$ at $\omega = 0$. As Luttinger's sum rule reads

$$\langle N\rangle = 2\sum_{\boldsymbol{k}} \Theta(G'_{\boldsymbol{k}}(0))\,, \tag{10.65}$$

this implies that, within an approximation constructed within the SFT, the sum rule is satisfied if and only if it is satisfied for the reference system, i.e., if $\langle N\rangle' = 2\sum_{\boldsymbol{k}} \Theta(G'_{\boldsymbol{k}}(0))$. This demonstrates that the theorem is "propagated" to the original system irrespective of the approximation that is constructed within the SFT. This propagation also works in the opposite direction. Namely, a possible violation of the exact sum rule for the reference system would imply a violation of the sum rule, expressed in terms of approximate quantities, for the original system.

There are no problems to take the thermodynamic limit $L \to \infty$ (if desired) on both sides of (10.64). The $\boldsymbol{k}$ sums turn into integrals over the unit cell of the reciprocal lattice. For a D-dimensional lattice the $D-1$-dimensional manifold of $\boldsymbol{k}$ points with $G_{\boldsymbol{k}}(0) = \infty$ or $G_{\boldsymbol{k}}(0) = 0$ form Fermi or Luttinger surfaces, respectively. Translational symmetry of the original as well as the reference system may be assumed but is not necessary. In the absence of translational symmetry, however, one has to re-interprete the wave vector $\boldsymbol{k}$ as an index which refers to the elements of the diagonalized Green's function matrix $\boldsymbol{G}$. The exact sum rule generalizes accordingly but can no longer be expressed in terms of a Fermi surface since there is no reciprocal space. It is also valid for the case of a translationally symmetric original Hamiltonian where, due to the choice of a reference system with

reduced translational symmetries, such as employed in the VCA, the symmetries of the (approximate) Green's function of the original system are (artificially) reduced. Since with (10.64) the proof of the sum rule is actually shifted to the proof of the sum rule for the reference system only, we are faced with the interesting problem of the validity of the sum rule for a finite cluster. For small Hubbard clusters with non-degenerate ground state this has been checked numerically with the surprising result that violations of the sum rule appear in certain parameter regimes close to half-filling, see [51]. This leaves us with the question where the proof of the sum rule fails if applied to a system of finite size. This is an open problem that has been stated and discussed in [51, 53] and that is probably related to the breakdown of the sum rule for Mott insulators [54].

10.5 Bath Degrees of Freedom

10.5.1 Motivation and Dynamical Impurity Approximation

Within SFT, an approximation is specified by the choice of the reference system. The reference system must share the same interaction part with the original model and should be amenable to a (numerically) exact solution. These requirements restrict the number of conceivable approximations. So far we have considered a decoupling of degrees of freedom by partitioning a Hubbard-type lattice model into finite Hubbard clusters of L_c sites each, which results in the VCA.

Another element in constructing approximation is to *add* degrees of freedom. Since the interaction part has to be kept unchanged, the only possibility to do that consists in adding new uncorrelated sites (or "orbitals"), i.e., sites where the Hubbard U vanishes. These are called "bath sites." The coupling of bath sites to the correlated sites with finite U in the reference system via a one-particle term in the Hamiltonian is called "hybridization."

Figure 10.8 shows different possibilities. Reference system A yields a trial self-energy which is local $\Sigma_{ij\sigma}(\omega) = \delta_{ij}\,\Sigma(\omega)$ and has the same pole structure as the self-energy of the atomic limit of the Hubbard model. This results in a variant of the Hubbard-I approximation [2]. Reference systems B and C generate VCA. In reference system D an additional bath site is added to the finite cluster. Reference system E generates a local self-energy again but, as compared to A, allows to treat more variational parameters, namely the on-site energies of the correlated and of the bath site and the hybridization between them. We call the resulting approximation a "dynamical impurity approximation" (DIA) with $L_\mathrm{b} = 1$.

The DIA is a mean-field approximation since the self-energy is local which indicates that non-local two-particle correlations, e.g., spin–spin correlations, do not feed back to the one-particle Green function. It is, however, quite different from static mean-field (Hartree–Fock) theory since even on the $L_\mathrm{b} = 1$ level it includes retardation effects that result from processes $\propto V^2$ where the electron

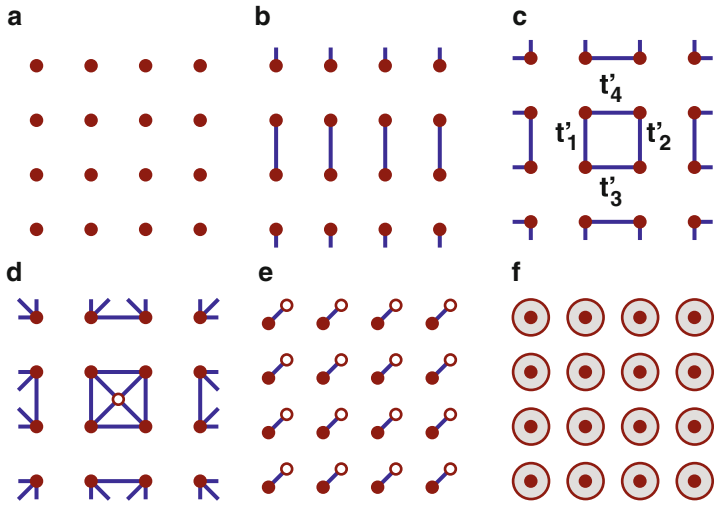

Fig. 10.8 Different possible reference systems with the same interaction as the single-band Hubbard model on a square lattice. *Filled circles*: correlated sites with U as in the Hubbard model. *Open circles*: uncorrelated "bath" sites with $U = 0$. *Lines*: nearest-neighbour hopping. *Big circles*: continuous bath consisting of $L_\mathrm{b} = \infty$ bath sites. Reference systems H'_A, H'_B and H'_C generate variational cluster approximations (VCA), H'_E yields dynamical impurity approximation (DIA), H'_F the DMFT, and H'_D an intermediate approximation (VCA with one additional bath site per cluster)

hops from the correlated to the bath site and back. This improves, as compared to the Hubbard-I approximation, the frequency dependence of the self-energy, i.e., the description of the temporal quantum fluctuations. The two-site ($L_\mathrm{b} = 1$) DIA is the most simple approximation which is non-perturbative, conserving, thermodynamically consistent and which respects the Luttinger sum rule (see [51]). Besides the "atomic" physics that leads to the formation of the Hubbard bands, it also includes in the most simple form the possibility to form a local singlet, i.e., to screen the local magnetic moment on the correlated site by coupling to the local moment at the bath site. The correct Kondo scale is missed, of course. Since the two-site DIA is computationally extremely cheap, it has been employed frequently in the past, in particular to study the physics of the Mott metal–insulator transition [11, 12, 51, 55–59].

10.5.2 Relation to Dynamical Mean-Field Theory

Starting from E and adding more and more bath sites to improve the description of temporal fluctuations, one ends up with reference system F where a continuum of

bath sites ($L_\mathrm{b} = \infty$) is attached to each of the disconnected correlated sites. This generates the "optimal" DIA.

To characterize this approximation, we consider the SFT grand potential given by (10.39) and analyze the stationarity condition (10.43): Calculating the derivative with respect to $\boldsymbol{t}'$, we get the general SFT Euler equation:

$$T \sum_n \sum_{\alpha\beta} \left(\frac{1}{\mathbf{G}^{-1}_{\boldsymbol{t},0}(i\omega_n) - \boldsymbol{\Sigma}_{\boldsymbol{t}',U}(i\omega_n)} - \mathbf{G}_{\boldsymbol{t}',U}(i\omega_n) \right)_{\beta\alpha} \frac{\partial \Sigma_{\boldsymbol{t}',U,\alpha\beta}(i\omega_n)}{\partial \mathbf{t}'} = 0 \,. \tag{10.66}$$

For the physical self-energy $\boldsymbol{\Sigma}_{\boldsymbol{t},U}$ of the original system $H_{\boldsymbol{t},U}$, the equation was fulfilled since the bracket would be zero. Vice versa, since $\boldsymbol{G}_{\boldsymbol{t}',U} = \widehat{\boldsymbol{G}}_U[\boldsymbol{\Sigma}_{\boldsymbol{t}',U}]$, the physical self-energy of $H_{\boldsymbol{t},U}$ is determined by the condition that the bracket be zero. Hence, one can consider the SFT Euler equation to be obtained from the *exact* conditional equation for the "vector" $\boldsymbol{\Sigma}$ in the self-energy space $\mathcal{S}_U$ through *projection* onto the hypersurface of $\mathbf{t}'$ representable trial self-energies $\mathcal{S}'_U$ by taking the scalar product with vectors $\partial \Sigma_{\boldsymbol{t}',U,\alpha\beta}(i\omega_n)/\partial \mathbf{t}'$ tangential to the hypersurface.

Consider now the Hubbard model in particular and the trial self-energies generated by reference system F. Actually, F is a set of disconnected single-impurity Anderson models (SIAMs). Assuming translational symmetry, these impurity models are identical replicas. The self-energy of the SIAM is non-zero on the correlated ("impurity") site only. Hence, the trial self-energies are local and site-independent, i.e., $\Sigma_{ij\sigma}(i\omega_n) = \delta_{ij}\,\Sigma(i\omega_n)$, and thus (10.66) reads:

$$T \sum_n \sum_{i\sigma} \left(\frac{1}{\mathbf{G}_0^{-1}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n)} - \mathbf{G}'(i\omega_n) \right)_{ii\sigma} \frac{\partial \Sigma_{ii\sigma}(i\omega_n)}{\partial \mathbf{t}'} = 0 \,, \tag{10.67}$$

where the notation has been somewhat simplified. The Euler equation would be solved if one-particle parameters of the SIAM and therewith an impurity self-energy can be found that, when inserted into the Dyson equation of the Hubbard model, yields a Green's function, the *local* element of which is equal to the impurity Green's function $\boldsymbol{G}'$. Namely, the bracket in (10.67), i.e., the local (diagonal) elements of the bracket in (10.66), vanishes. Since the "projector" $\partial \Sigma_{ii\sigma}(i\omega_n)/\partial \mathbf{t}'$ is local, this is sufficient.

This way of solving the Euler equation, however, is just the prescription to obtain the self-energy within DMFT [9], and setting the local elements of the bracket to zero is just the self-consistency equation of DMFT. We, therefore, see that reference system F generates the DMFT. It is remarkable that with the VCA and the DMFT quite different approximations can be obtained in one and the same theoretical framework.

Note that for any finite L_b, as for example, with reference system E, it is impossible to satisfy the DMFT self-consistency equation exactly since the impurity Green's function $\boldsymbol{G}'$ has a finite number of poles while the lattice Green's function

$(1/\mathbf{G}_0^{-1}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n))_{ii\sigma}$ has branch cuts. Nevertheless, with the "help" of the projector, it is easily possible to find a stationary point $\boldsymbol{t}_{\rm opt}$ of the self-energy-functional and to satisfy (10.67).

Conceptually, this is rather different from the DMFT exact-diagonalization (DMFT-ED) approach [60] which also solves a SIAM with a finite $L_{\rm b}$ but which approximates the DMFT self-consistency condition. This means that within DMFT-ED an additional ad hoc prescription is necessary which, opposed to the DIA, will violate thermodynamical consistency. However, an algorithmic implementation via a self-consistency cycle to solve the Euler equation is simpler within the DMFT-ED as compared to the DIA [56, 59]. It has, therefore, been suggested to guide improved approximations to the self-consistency condition within DMFT-ED by the DIA [61]. The convergence of results obtained by the DIA to those of full DMFT with increasing number of bath sites $L_{\rm b}$ is usually fast. With respect to the Mott transition in the single-band Hubbard model, $L_{\rm b} = 3$ is sufficient to get almost quantitative agreement with DMFT–QMC results, for example [56].

Real-space DMFT for inhomogeneous systems [62–64] with a local but site-dependent self-energy $\Sigma_{ij\sigma}(i\omega_n) = \delta_{ij}\,\Sigma_i(i\omega_n)$ is obtained from reference system F if the one-particle parameters of the SIAMs at different sites are allowed to be different. This enlarges the space of trial self-energies. In this case we get one local self-consistency equation for each site. The different impurity models can be solved independently from each other in each step of the self-consistency cycle while the coupling between the different sites is provided by the Dyson equation of the Hubbard model.

10.5.3 Cluster Mean-Field Approximations

As a mean-field approach, the DIA does not include the feedback of non-local two-particle correlations on the one-particle spectrum and on the thermodynamics. The DIA self-energy is local and takes into account temporal correlations only. A straightforward idea to include short-range spatial correlations in addition is to proceed to a reference system with $L_{\rm c} > 1$, i.e., system of disconnected finite clusters with $L_{\rm c}$ sites each. The resulting approximation can be termed a cluster mean-field theory since despite the inclusion of short-range correlations, the approximation is still mean-field-like on length scales exceeding the cluster size.

For an infinite number of bath degrees of freedom $L_{\rm b} \to \infty$ attached to each of the $L_{\rm c} > 1$ correlated sites the cellular DMFT [30, 65] is recovered, see Fig. 10.9. Considering a single-band Hubbard model again, this can be seen from the corresponding Euler equation:

$$T \sum_n {\sum_{ij\sigma}}' \left(\frac{1}{\mathbf{G}_0^{-1}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n)} - \mathbf{G}'(i\omega_n) \right)_{ij\sigma} \frac{\partial \Sigma_{ji\sigma}(i\omega_n)}{\partial \mathbf{t}'} = 0\,, \qquad (10.68)$$

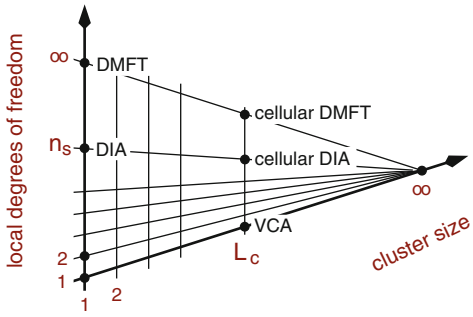

Fig. 10.9 Dynamical impurity and cluster approximations generated by different reference systems within SFT. L_c is the number of correlated sites in the reference cluster. $n_s = 1 + L_b$ is the number of local degrees of freedom where L_b denotes the number of additional bath sites attached to each of the L_c correlated sites. The variational cluster approximation (VCA) is obtained for finite clusters with $L_c > 1$ but without bath sites $n_s = 1$. This generalizes the Hubbard-I-type approximation obtained for $L_c = 1$. The dynamical impurity approximation (DIA) is obtained for $n_s > 1$ but for a single correlated site $L_c = 1$. A continuum of bath sites, $n_s = \infty$ generates DMFT ($L_c = 1$) and cellular DMFT ($L_c > 1$). $L_c = \infty$, irrespective of the number of local degrees of freedom, corresponds to the exact solution

where the prime at the sum over the sites indicates that i and j must belong to the same cluster of the reference system. Namely, $\Sigma_{ij\sigma}(i\omega_n) = 0$ and also the "projector" $\partial\Sigma_{ji\sigma}(i\omega_n)/\partial\mathbf{t}' = 0$ if i and j belong to different clusters. This stationarity condition can be fulfilled if

$$\left(\frac{1}{\mathbf{G}_0^{-1}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n)} - \mathbf{G}'(i\omega_n)\right)_{ij\sigma} = 0\,. \tag{10.69}$$

Note that $\mathbf{G}'(i\omega_n)$ is a matrix which is labeled as $G'_{ij,kl}(i\omega_n)$, where $i, j = 1, \ldots, L_c$ refer to the correlated sites in the cluster while $k, l = 1, \ldots, L_b L_c$ to the bath sites. $G'_{ij}(i\omega_n)$ are the elements of the cluster Green's function on the correlated sites. The condition (10.69) is just the self-consistency condition of the C-DMFT.

As is illustrated in Fig. 10.9, the exact solution can be obtained with increasing cluster size $L_c \to \infty$ either from a sequence of reference systems with a continuous bath $n_s = 1 + L_b = \infty$, corresponding to C-DMFT, or from a sequence with $n_s = 1$, corresponding to VCA, or with a finite, small number of bath sites ("cellular DIA"). Systematic studies of the one-dimensional Hubbard model [13, 27] have shown that the energy gain which is obtained by attaching a bath site is lower than the gain obtained by increasing the cluster. This suggests that the convergence to the exact solution could be faster on the "VCA axis" in Fig. 10.9. For a definite answer, however, more systematic studies, also in higher dimensions, are needed.

In any case bath sites help to get a smooth dependence of physical quantities when varying the electron density or the (physical) chemical potential. The reason is that bath sites also serve as "charge reservoirs," i.e., during a μ scan the ground

state of the reference cluster may stay in one and the same sector characterized by the conserved total particle number in the cluster while the particle number on the correlated sites and the approximate particle number in the original lattice model evolve continuously [27, 32]. This is achieved by a μ-dependent charge transfer between correlated and bath sites. In addition, (at least) a single bath site per correlated site in a finite reference cluster is also advantageous to include the interplay between local (Kondo type) and non-local (antiferromagnetic) singlet formation. This has been recognized to be important in studies of the Mott transition [66] in the two-dimensional and of ferromagnetic order in one-dimensional systems [32]. For studies of spontaneous $U(1)$ symmetry breaking, e.g., d-wave superconductivity in the two-dimensional Hubbard model [38–41], doping dependencies can be investigated without bath sites due to mixing of cluster states with different particle numbers.

10.5.4 Translation Symmetry

For any cluster approximation formulated in real space there is an apparent problem: Due to the construction of the reference system as a set of decoupled clusters, the trial self-energies do not preserve the translational symmetries of the original lattice. Trivially, this also holds if periodic boundary conditions are imposed for the individual cluster. Transformation of the original problem to reciprocal space does not solve the problem either since this also means to transform a local Hubbard-type interaction into a non-local interaction part which basically couples all $\boldsymbol{k}$ points.

There are different ideas to overcome this problem. We introduce a "periodizing" universal functional

$$\widehat{T}[\boldsymbol{\Sigma}]_{ij} = \frac{1}{L}\sum_{i'j'} \delta_{i-j,i'-j'}\,\Sigma_{i'j'}, \tag{10.70}$$

which maps any trial self-energy onto translationally invariant one. In reciprocal space this corresponds to the substitution $\Sigma_{\boldsymbol{k},\boldsymbol{k}'} \to \widehat{T}[\boldsymbol{\Sigma}]_{\boldsymbol{k},\boldsymbol{k}'} = \delta_{\boldsymbol{k},\boldsymbol{k}'}\Sigma_{\boldsymbol{k}}$. Using this, we replace the self-energy-functional of (10.37) by

$$\widehat{\Omega}^{(1)}_{t',U}[\boldsymbol{\Sigma}] = \mathrm{Tr}\ln\frac{1}{\boldsymbol{G}^{-1}_{t',\boldsymbol{0}} - \widehat{\boldsymbol{T}}[\boldsymbol{\Sigma}]} + \widehat{F}_U[\boldsymbol{\Sigma}]\,, \tag{10.71}$$

as suggested in [67]. This new functional is different from the original one. However, as the physical self-energy is supposed to be a translational invariant, it is a stationary point of both, the original and the modified functional. This means that the modified functional can likewise be used as a starting point to construct approximations. It turns out (see [68] for an analogous discussion in the case of disorder) that for a reference system with $L_{\mathrm{c}} > 1$ and $n_{\mathrm{s}} = \infty$, the corresponding

Euler equation reduces to the self-consistency equation of the so-called periodized cellular DMFT (PC-DMFT) [69]. The same modified functional can also be used to construct a periodized VCA, for example.

While the main idea to recover the PC-DMFT is to modify the form of the self-energy-functional, the dynamical cluster approximation (DCA) [29, 70, 71] is obtained with the original functional but a modified hopping term in the Hamiltonian of the original system. We replace $\boldsymbol{t} \to \widetilde{\boldsymbol{t}}$ and consider the functional $\Omega_{\widetilde{\boldsymbol{t}},U}[\boldsymbol{\Sigma}]$. To ensure that the resulting approximations systematically approach the exact solution for cluster size $L_\mathrm{c} \to \infty$, the replacement $\boldsymbol{t} \to \widetilde{\boldsymbol{t}}$ must be controlled by L_c, i.e., it must be exact up to irrelevant boundary terms in the infinite-cluster limit. This is the case for

$$\widetilde{\boldsymbol{t}} = (\boldsymbol{V}\boldsymbol{W})\boldsymbol{U}^\dagger \boldsymbol{t}\boldsymbol{U}(\boldsymbol{V}\boldsymbol{W})^\dagger \,, \tag{10.72}$$

where $\boldsymbol{U}$, $\boldsymbol{V}$, and $\boldsymbol{W}$ are unitary transformations of the one-particle basis. $\boldsymbol{U}$ is the Fourier transformation with respect to the original lattice consisting of L sites ($L \times L$ unitary matrix). $\boldsymbol{W}$ is the Fourier transformation on the cluster ($L_\mathrm{c} \times L_\mathrm{c}$), and $\boldsymbol{V}$ the Fourier transformation with respect to the superlattice consisting of L/L_c supersites given by the clusters ($L/L_\mathrm{c} \times L/L_\mathrm{c}$). The important point is that for any finite L_c the combined transformation is $\boldsymbol{V}\boldsymbol{W} = \boldsymbol{W}\boldsymbol{V} \neq \boldsymbol{U}$, while this becomes irrelevant in the limit $L_\mathrm{c} \to \infty$. The detailed calculation (see [68] for the analogous disorder case) shows that the DCA is recovered for a reference system with $L_\mathrm{c} > 1$ and $n_\mathrm{s} = \infty$, if periodic boundary conditions are imposed for the cluster. The same modified construction can also be used to a get simplified DCA-type approximation without bath sites, for example. This "simplified DCA" is related to the periodized VCA in the same way as the DCA is related to the PC-DMFT. The simplified DCA would represent a variational generalization of a non-self-consistent approximation ("periodic CPT") introduced recently [72].

10.6 Systematics of Approximations

Since the SFT unifies different dynamical approximations within a single formal framework, the question arises how to judge on the relative quality of two different approximations resulting from two different reference systems. This, however, is not straightforward for several reasons. First, it is important to note that a stationary point of the self-energy-functional is not necessarily a minimum but rather a saddle point in general (see [11] for an example). The self-energy-functional is not convex. Actually, despite several recent efforts [26,73,74], there is no functional relationship between a thermodynamical potential and time-dependent correlation functions, Green's functions, self-energies, etc. which is known to be convex.

Furthermore, there is no a priori reason why, for a given reference system, the SFT grand potential at a stationary point should be lower than the SFT grand potential at another one that results from a simpler reference system, e.g., a smaller

cluster. This implies that the SFT does not provide upper bounds to the physical grand potential. There is e.g., no proof (but also no counterexample) that the DMFT ground-state energy at zero temperature must be higher than the exact one. On the other hand, in practical calculations the upper-bound property is usually found to be respected, as can be seen for the VCA in Fig. 10.5, for example. Nevertheless, the non-convexity must be seen as a disadvantage as compared to methods based on wave functions which via the Ritz variational principle are able to provide strict upper bounds.

To discuss how to compare two approximations within SFT, we first have to distinguish between "trivial" and "non-trivial" stationary points for a given reference system. A stationary point is referred to as "trivial" if the one-particle parameters are such that the reference system decouples into smaller subsystems. If, at a stationary point, all degrees of freedom (sites) are still coupled to each other, the stationary point is called "non-trivial." It is possible to prove the following theorem [75]: Consider a reference system with a set of variational parameters $\boldsymbol{t}' = \boldsymbol{t}'' + \boldsymbol{V}$, where $\boldsymbol{V}$ couples two separate subsystems. For example, $\boldsymbol{V}$ could be the inter-cluster hopping between two subclusters which completely decouples the degrees of freedom for $\boldsymbol{V} = 0$ and all $\boldsymbol{t}''$. Then

$$\Omega_{\boldsymbol{t},U}[\boldsymbol{\Sigma}_{\boldsymbol{t}''+\boldsymbol{V}}] = \Omega_{\boldsymbol{t},U}[\boldsymbol{\Sigma}_{\boldsymbol{t}''+0}] + \mathcal{O}(\boldsymbol{V}^2)\ , \tag{10.73}$$

provided that the functional is stationary at $\boldsymbol{\Sigma}_{\boldsymbol{t}'',U}$ *when varying* $\boldsymbol{t}''$ *only* (this restriction makes the theorem non-trivial). This means that going from a more simple reference system to a more complicated one with more degrees of freedom coupled should generate a new non-trivial stationary point with $\boldsymbol{V} \neq 0$ while the "old" stationary point with $\boldsymbol{V} = 0$ being still a stationary point with respect to the "new" reference system. Coupling more and more degrees of freedom introduces more and more stationary points, and none of the "old" ones is "lost."

Consider a given reference system with a non-trivial stationary point and a number of trivial stationary points. An intuitive strategy to decide between two stationary points would be to always take the one with the lower grand potential $\Omega_{\boldsymbol{t},U}[\boldsymbol{\Sigma}_{\boldsymbol{t}',U}]$. A sequence of reference systems (e.g. H'_{A}, H'_{B}, H'_{C}, ...) in which more and more degrees of freedom are coupled and which eventually recovers the original system H itself shall be called a "systematic" sequence of reference systems. For such a systematic sequence, the suggested strategy trivially produces a series of stationary points with monotonously decreasing grand potential. Unfortunately, however, the strategy is useless because it cannot ensure that a systematic sequence of reference systems generates a systematic sequence of approximations as well, i.e., one cannot ensure that the respective *lowest* grand potential in a systematic sequence of reference systems converges to the exact grand potential. Namely, the stationary point with the lowest SFT grand potential could be a trivial stationary point (like the one associated with a very simple reference system only as H'_{A} or H'_{B} in Fig. 10.8, for example). Such an approximation must be considered as poor since the exact conditional equation for the self-energy is projected onto a very low-dimensional space only.

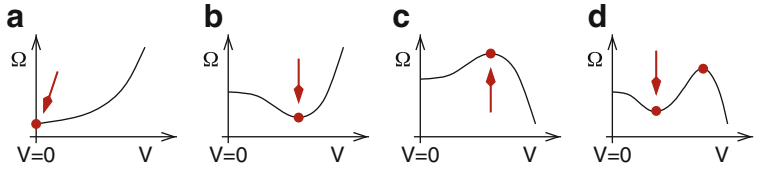

Fig. 10.10 Possible trends of the SFT grand potential Ω as a function of a variational parameter V coupling two subsystems of a reference system. $V = 0$ corresponds to the decoupled case and must always represent a "trivial" stationary point. *Circles* show the stationary points to be considered. The point $V = 0$ has to be disregarded in all cases except for (**a**). The *arrow* marks the respective optimum stationary point according to the rules discussed in the text

Therefore, one has to construct a different strategy which necessarily approaches the exact solution when following up a systematic sequence of reference systems. Clearly, this can only be achieved if the following rule is obeyed: *A non-trivial stationary point is always preferred as compared to a trivial one (R0).* A non-trivial stationary point at a certain level of approximation, i.e., for a given reference system, becomes a trivial stationary point on the next level, i.e., in the context of a "new" reference system that couples at least two different units of the "old" reference system. Hence, by construction, the rule R0 implies that the exact result is approached for a systematic series of reference systems.

Following the rule (R0), however, may lead to inconsistent thermodynamic interpretations in case of a trivial stationary point with a lower grand potential as a non-trivial one. To avoid this, R0 has to be replaced by: *Trivial stationary points must be disregarded completely unless there is no non-trivial one (R1).* This automatically ensures that there is at least one stationary point for any reference system, i.e., at any approximation level.

To maintain a thermodynamically consistent picture in case that there are more than a single *non-trivial* stationary points, we finally postulate: *Among two non-trivial stationary points for the same reference system, the one with lower grand potential has to be preferred (R2).*

The rules are illustrated by Fig. 10.10. Note that the grand potential away from a stationary point does not have a direct physical interpretation. Hence, there is no reason to interpret the solution corresponding to the maximum in Fig. 10.10c) as "locally unstable." The results of [12, see Figs. 2 and 4 therein] nicely demonstrate that with the suggested strategy (R1, R2) one can consistently describe continuous as well as discontinuous phase transitions.

The rules R1 and R2 are unambiguously prescribed by the general demands for systematic improvement and for thermodynamic consistency. There is no acceptable alternative to this strategy. The strategy reduces to the standard strategy (always taking the solution with lowest grand potential) in the case of the Ritz variational principle because here a non-trivial stationary point does always have a lower grand potential as compared to a trivial one.

There are also some consequences of the strategy which might be considered as disadvantageous but must be tolerated: (1) For a sequence of stationary points that are determined by R1 and R2 from a systematic sequence of reference systems, the convergence to the corresponding SFT grand potentials is not necessarily monotonous. (2) Given two different approximations specified by two different reference systems, it is not possible to decide which one should be regarded as superior unless both reference systems belong to the same systematic sequence of reference systems. In Fig. 10.8, one has $H'_{\mathrm{A}} < H'_{\mathrm{B}} < H'_{\mathrm{C}} < H'_{\mathrm{D}}$, where "$<$" stands for "is inferior compared to." Furthermore, $H'_{\mathrm{E}} < H'_{\mathrm{F}}$ and $H'_{\mathrm{A}} < H'_{\mathrm{E}}$ but there is no relation between H'_{B} and H'_{E}, for example.

10.7 Summary

The above discussion has shown that SFT provides a general framework which allows us to construct different dynamical approximations for lattice models of strongly correlated electrons. These approximations derive from a fundamental variational principle, formulated in terms of the grand potential expressed as a functional of the self-energy, by restricting the domain of the functional. This leads to non-perturbative and thermodynamically consistent approximations. The SFT unifies different known approximations in a single theoretical frame and provides new dynamical impurity (DIA) and variational cluster approximations (VCA).

The essential step in the numerical evaluation consists in the calculation of the Green's function or the self-energy of a reference system with the same interaction part as the original model but with spatially decoupled degrees of freedom. Details of the numerical procedure can be found in [11, 12, 27, 32, 76], for example. Typically, exact diagonalization or the (band) Lanczos approach [77, 78] but also quantum Monte–Carlo techniques may be used [79] as a reference-system solver. Since bath sites can be integrated out within Green-function based QMC schemes, QMC as an impurity/cluster solver is the method of choice for finite-temperature DMFT or cluster DMFT approaches, i.e., for reference systems with a continuum of bath sites. At zero temperature, and using reference systems without bath sites or a few bath degrees of freedom only, the SFT provides computationally fast techniques which complement the (cluster) DMFT methods.

Besides applications to Hubbard-type model systems, the VCA has recently been employed to study the correlated electronic structure of real materials, such as NiO [80], CoO, MnO [81], $LaCoO_3$ [31], TiOCl [82], CrO_2 [83], TiN [84] and NiMnSb [85]. Furthermore, the theory has been extended to study Bose systems [67, 86] and the Jaynes–Cummings lattice [87, 88], electron-phonon systems [55], systems with non-local interactions [89], systems with quenched disorder [68] and more.

Acknowledgments The author would like to thank M. Aichhorn, F.F. Assaad, E. Arrigoni, M. Balzer, R. Bulla, C. Dahnken, R. Eder, W. Hanke, A. Hewson, M. Jarrell, M. Kollar, G. Kotliar,

A.I. Lichtenstein, A.J. Millis, W. Nolting, D. Senechal, A.-M.S. Tremblay and D. Vollhardt for cooperations and many helpful discussions. Support by the Deutsche Forschungsgemeinschaft within the SFB 668 (project A14) and FOR 1346 (project P1) is gratefully acknowledged.

References

1. A.A. Abrikosow, L.P. Gorkov, I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New Jersey, 1964)
2. J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963)
3. M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
4. J. Kanamori, Prog. Theor. Phys. (Kyoto) **30**, 275 (1963)
5. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)
6. C. Gros, R. Valenti, Phys. Rev. B **48**, 418 (1993)
7. D. Sénéchal, D. Pérez, M. Pioro-Ladrière, Phys. Rev. Lett. **84**, 522 (2000)
8. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
9. A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
10. G. Kotliar, D. Vollhardt, Phys. Today **57**, 53 (2004)
11. M. Potthoff, Euro. Phys. J. B **32**, 429 (2003)
12. M. Potthoff, Euro. Phys. J. B **36**, 335 (2003)
13. M. Potthoff, M. Aichhorn, C. Dahnken, Phys. Rev. Lett. **91**, 206402 (2003)
14. M. Potthoff, Adv. Solid State Phys. **45**, 135 (2005)
15. A.L. Fetter, J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971)
16. J.W. Negele, H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, Redwood City, 1988)
17. M. Potthoff, in *Theory of electron spectroscopies*, ed. by K. Baberschke, M. Donath, W. Nolting. Band-Ferromagnetism (Springer, Berlin, 2001)
18. J.M. Luttinger, Phys. Rev. **121**, 942 (1961)
19. H. Schweitzer, G. Czycholl, Solid State Commun. **74**, 735 (1990)
20. H. Schweitzer, G. Czycholl, Z. Phys. B **83**, 93 (1991)
21. M. Potthoff, W. Nolting, Z. Phys. B **104**, 265 (1997)
22. E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)
23. J.M. Luttinger, J.C. Ward, Phys. Rev. **118**, 1417 (1960)
24. M. Potthoff, Condens. Mat. Phys. **9**, 557 (2006)
25. R. Chitra, G. Kotliar, Phys. Rev. B **62**, 12715 (2000)
26. R. Chitra, G. Kotliar, Phys. Rev. B **63**, 115110 (2001)
27. M. Balzer, W. Hanke, M. Potthoff, Phys. Rev. B **77**, 045133 (2008)
28. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968)
29. M.H. Hettler, M. Mukherjee, M. Jarrell, H.R. Krishnamurthy, Phys. Rev. B **61**, 12739 (2000)
30. G. Kotliar, S.Y. Savrasov, G. Pálsson, G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
31. R. Eder, Phys. Rev. B **81**(3), 035101 (2010). DOI 10.1103/PhysRevB.81.035101
32. M. Balzer, M. Potthoff, Phys. Rev. B **82**, 174441 (2010)
33. C. Dahnken, M. Aichhorn, W. Hanke, E. Arrigoni, M. Potthoff, Phys. Rev. B **70**, 245110 (2004)
34. A.H. Nevidomskyy, C. Scheiber, D. Senechal, A.M.S. Tremblay, Phys. Rev. B **77**, 064427 (2008)
35. S. Horiuchi, S. Kudo, T. Shirakawa, , Y. Ohta, Phys. Rev. B **78**, 155128 (2008)
36. T. Yoshikawa, M. Ogata, Phys. Rev. B **79**, 144429 (2009)
37. P. Sahebsara, D. Senechal, Phys. Rev. Lett. **100**, 136402 (2008)
38. D. Sénéchal, P.L. Lavertu, M.A. Marois, A.M.S. Tremblay, Phys. Rev. Lett. **94**, 156404 (2005)
39. M. Aichhorn, E. Arrigoni, Europhys. Lett. **72**, 117 (2005)
40. M. Aichhorn, E. Arrigoni, M. Potthoff, W. Hanke, Phys. Rev. B **74**, 024508 (2006)

41. M. Aichhorn, E. Arrigoni, M. Potthoff, W. Hanke, Phys. Rev. B **74**, 235117 (2006)
42. P. Sahebsara, D. Senechal, Phys. Rev. Lett. **97**, 257004 (2006)
43. E. Arrigoni, M. Aichhorn, M. Daghofer, W. Hanke, New J. Phys. **11**, 055066 (2009)
44. P. Sahebsara, D. Senechal, arXiv: 0908.0474 (2009)
45. M. Aichhorn, H.G. Evertz, W. von der Linden, M. Potthoff, Phys. Rev. B **70**, 235107 (2004)
46. M. Aichhorn, E.Y. Sherman, H.G. Evertz, Phys. Rev. B **72**, 155110 (2005)
47. X. Lu, E. Arrigoni, Phys. Rev. B **79**, 245109 (2009)
48. G. Baym, L.P. Kadanoff, Phys. Rev. **124**, 287 (1961)
49. G. Baym, Phys. Rev. **127**, 1391 (1962)
50. N.E. Bickers, D.J. Scalapino, S.R. White, Phys. Rev. Lett. **62**, 961 (1989)
51. J. Ortloff, M. Balzer, M. Potthoff, Euro. Phys. J. B **58**, 37 (2007)
52. J.M. Luttinger, Phys. Rev. **119**, 1153 (1960)
53. J. Kokalj, P. Prelovšek, Phys. Rev. B **75**, 045111 (2007)
54. A. Rosch, Eur. Phys. J. B **59**, 495 (2007)
55. W. Koller, D. Meyer, Y. Ono, A.C. Hewson, Europhys. Lett. **66**, 559 (2004)
56. K. Pozgajcic, preprint cond-mat **0407172** (2004)
57. K. Inaba, A. Koga, S.I. Suga, N. Kawakami, Phys. Rev. B **72**, 085112 (2005)
58. K. Inaba, A. Koga, S.I. Suga, N. Kawakami, J. Phys. Soc. Jpn. **74**, 2393 (2005)
59. M. Eckstein, M. Kollar, M. Potthoff, D. Vollhardt, Phys. Rev. B **75**, 125103 (2007)
60. M. Caffarel, W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
61. D. Sénéchal, Phys. Rev. B **81**, 235125 (2010)
62. M. Potthoff, W. Nolting, Phys. Rev. B **52**, 15341 (1995)
63. R.W. Helmes, T.A. Costi, A. Rosch, Phys. Rev. Lett. **100**, 056403 (2008)
64. M. Snoek, I. Titvinidze, C. Töke, K. Byczuk, W. Hofstetter, New J. Phys. **10**, 093008 (2008)
65. A.I. Lichtenstein, M.I. Katsnelson, Phys. Rev. B **62**, R9283 (2000)
66. M. Balzer, B. Kyung, D. Sénéchal, A.M.S. Tremblay, M. Potthoff, Europhys. Lett. **85**, 17002 (2009)
67. W. Koller, N. Dupuis, J. Phys. Condens. Matter **18**, 9525 (2005)
68. M. Potthoff, M. Balzer, Phys. Rev. B **75**, 125112 (2007)
69. G. Biroli, O. Parcollet, G. Kotliar, Phys. Rev. B **69**, 205108 (2004)
70. M.H. Hettler, A.N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, H.R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998)
71. T. Maier, M. Jarrell, T. Pruschke, J. Keller, Euro. Phys. J. B **13**, 613 (2000)
72. T. Minh-Tien, Phys. Rev. B **74**, 155121 (2006)
73. G. Kotliar, Euro. Phys. J. B **11**, 27 (1999)
74. A.H. Nevidomskyy, A.M.T. D. Sénéchal, Phys. Rev. B **77**, 075105 (2008)
75. M. Potthoff, in *Effective Models for Low-Dimensional Strongly Correlated Systems*. ed. by G. Batrouni, D. Poilblanc (AIP proceedings, Melville, 2006)
76. D. Sénéchal, preprint cond-mat/0806.2690, lecture notes CIFAR – PITP International Summer School on Numerical Methods for Correlated Systems in Condensed Matter, Sherbrooke
77. H.Q. Lin, J.E. Gubernatis, Comput. Phys. **7**, 400 (1993)
78. R. Freund, in *Band Lanczos Method*, ed. by Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst. Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide (SIAM, Philadelphia, 2000)
79. G. Li, W. Hanke, A.N. Rubtsov, S. Bäse, M. Potthoff, Phys. Rev. B **80**, 195118 (2009)
80. R. Eder, Phys. Rev. B **76**, 241103(R) (2007)
81. R. Eder, Phys. Rev. B **78**, 115111 (2008)
82. M. Aichhorn, T. Saha-Dasgupta, R. Valenti, S. Glawion, M. Sing, R. Claessen, Phys. Rev. B **80**, 115129 (2009)
83. L. Chioncel, H. Allmaier, E. Arrigoni, A. Yamasaki, M. Daghofer, M.I. Katsnelson, A.I. Lichtenstein, Phys. Rev. B **75**, 140406 (2007)
84. H. Allmaier, L. Chioncel, E. Arrigoni, Phys. Rev. B **79**, 235126 (2009)
85. H. Allmaier, L. Chioncel, E. Arrigoni, M.I. Katsnelson, A.I. Lichtenstein, Phys. Rev. B **81**, 054422 (2010)

86. M. Knap, E. Arrigoni, W. von der Linden, Phys. Rev. B **81**, 235122 (2010)
87. M. Aichhorn, M. Hohenadler, C. Tahan, P.B. Littlewood, Phys. Rev. Lett. **100**, 216401 (2008)
88. M. Knap, E. Arrigoni, W. von der Linden, Phys. Rev. B **81**, 104303 (2010)
89. N.H. Tong, Phys. Rev. B **72**, 115104 (2005)

Chapter 11
Cluster Dynamical Mean Field Theory

David Sénéchal

Abstract Cluster dynamical mean-field theory is an extension of dynamical mean-field theory (DMFT) where the single-site impurity is replaced with a cluster of sites with open boundary conditions. Compared with single-site DMFT, this takes into account short-range correlations exactly and can probe the presence of broken-symmetry phases such as d-wave superconductivity and antiferromagnetism. This chapter reviews the basic CDMFT procedure, as well as issues related to the use of an exact diagonalization solver for the impurity problem. The QMC solvers are also briefly reviewed, as well as results on the Mott transition and on models for the cuprates.

Cluster dynamical mean-field theory [1, 2] (also called *cellular* dynamical mean-field theory and abbreviated CDMFT) is an extension of dynamical mean field theory (DMFT) where the single "impurity" site is replaced by a finite cluster of lattice sites. The extension of single-site DMFT to clusters is motivated by the importance of short-range antiferromagnetic fluctuations in Hubbard models and the possible existence of d-wave pairing. It allows DMFT to take into account short-range spatial correlations into account; this has an effect even on properties that are not directly related to broken symmetry phases, such as the Mott transition.

In this chapter, we first review the general principles behind DMFT, in a language adapted to the use of clusters. CDMFT, like single-site DMFT, can be understood in the framework of the self-energy functional approach (SFA) (see Chap. 9.5). This will be particularly useful in Sect. 11.2, where issues related to the use of an exact-diagonalization solver for the impurity problem are discussed. In Sect. 11.3 we briefly describe the use of quantum Monte Carlo (QMC) solvers (the Hirsch-Fye and continuous-time methods) for the impurity problem. We then review results

D. Sénéchal (✉)
Département de Physique, Université de Sherbrooke, Sherbrooke, QC, Canada J1K 2R1
e-mail: david.senechal@usherbrooke.ca

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_11,

on the Mott transition and on the physics of high-temperature superconductors. We conclude with some remarks on the differences between CDMFT and the dynamical cluster approximation (DCA), reviewed in Chap. 8.8.

11.1 The CDMFT Procedure

We shall not provide a full introduction to DMFT, as this is very well reviewed in Chap. 7 of this volume. Let us simply summarize the main points while introducing a notation adapted to the use of clusters. We will follow the notation used in Chap. 7.9 as far as kinematics is concerned.

11.1.1 The Effective Hamiltonian

Let us consider a general interacting electron Hamiltonian

$$H = \sum_{\alpha,\beta} t_{\alpha\beta} c_\alpha^\dagger c_\beta + \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} U_{\alpha\beta\delta\gamma} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta, \tag{11.1}$$

where c_α annihilates an electron in the one-particle state α. The composite index α is an aggregation of site, spin and band indices. By convention, the chemical potential will be included in the matrix $\boldsymbol{t}$ of one-body terms. In a functional integral language, the corresponding partition function is

$$Z = \int \prod_\alpha [\mathrm{d}c_\alpha \mathrm{d}c_\alpha^*] \exp -S[c, c^*], \tag{11.2}$$

where the imaginary-time action S is

$$S[c, c^*] = \int_0^\beta \mathrm{d}\tau \left\{ \sum_{\alpha,\beta} c_\alpha^* (\delta_{\alpha\beta} \partial_\tau + t_{\alpha\beta}) c_\beta + H_1(c, c^*) \right\} \tag{11.3}$$

and H_1 is the interaction Hamiltonian.

The starting point of CDMFT is a cluster decomposition of the lattice, as described in Chap. 7.9. The lattice action then takes the form

$$S = \sum_{m \in \Gamma} S^{(m)} + \sum_{m,n \in \Gamma} S^{(m,n)}, \tag{11.4}$$

where $S^{(m)}$ is the restriction of S to the cluster labelled m, and $S^{(m,n)}$ contains terms that involve sites belonging to clusters m and n, typically inter-cluster hopping terms. CDMFT assumes that the effect of the environment of each cluster can be well approximated by an effective action

$$\sum_{n\in\Gamma} S^{(m,n)} \to S^{(m)}_{\text{env.}} \tag{11.5}$$

This effectively decouples the clusters. This contribution from the environment is assumed to be uncorrelated, i.e., to be quadratic in c. Thus, the total effective action for a given cluster takes the general form

$$S_{\text{eff}}[c,c^*] = \int_0^\beta \mathrm{d}\tau\mathrm{d}\tau' \sum_{\alpha,\beta} c^*_\alpha(\tau)\mathscr{G}_0^{-1}(\tau-\tau')_{\alpha\beta}c_\beta(\tau') + \int_0^\beta \mathrm{d}\tau\, H_1(c,c^*), \tag{11.6}$$

where $\mathscr{G}_0$ is the *dynamical mean field*. The indices α,β are now restricted to one cluster, and likewise for the interaction Hamiltonian H_1.

In the frequency domain, the dynamical mean field can be written as

$$\mathscr{G}_0^{-1}(i\omega_n) = i\omega_n - \boldsymbol{t}' - \boldsymbol{\Gamma}(i\omega_n), \tag{11.7}$$

where $\boldsymbol{t}'$ is the restriction of the hopping matrix to the cluster and $\boldsymbol{\Gamma}(i\omega_n)$, the *hybridization function*, represents the dynamical hybridization of the cluster orbitals with their effective environment. This is better expressed in terms of an Anderson impurity model [3]; let us explain.

In order for this effective action to make sense, the dynamical mean field $\mathscr{G}_0$ must be causal. This implies that it must have the analytic properties of a Green function: its poles and zeros must lie on the real axis, its eigenvalues must have positive residues, and in addition $\mathscr{G}_0(i\omega_n)$ must behave like $1/i\omega_n$ at large frequencies. This entails a general spectral representation of $\boldsymbol{\Gamma}$ as

$$\Gamma_{\alpha\beta}(i\omega_n) = \int \mathrm{d}\omega' \,\frac{\theta_\alpha(\omega')\theta^*_\beta(\omega')}{i\omega_n-\omega'} \tag{11.8}$$

or, in a discrete version,

$$\Gamma_{\alpha\beta}(i\omega_n) = \sum_{\mu}^{N_b} \frac{\theta_{\alpha\mu}\theta^*_{\beta\mu}}{i\omega_n-\varepsilon_\mu}, \tag{11.9}$$

where ε_μ and $\theta_{\alpha\mu}$ form a collection of parameters that can be adjusted to fit any causal hybridization function as closely as needed. N_b is the number of poles deemed necessary to adequately represent the hybridization function.

Such a discrete hybridization function is naturally obtained from the following Anderson impurity model:

$$H_{\mathrm{AIM}} = \sum_{\alpha,\beta} t'_{\alpha\beta} c^\dagger_\alpha c_\beta + \sum_{\alpha,\mu} \left(\theta_{\alpha\mu} c^\dagger_\alpha a_\mu + \mathrm{H.c.}\right) + \sum_{\mu}^{N_b} \varepsilon_\mu a^\dagger_\mu a_\mu \tag{11.10}$$

Electrons can hop between the cluster sites labelled α, β and a set of effective orbitals (the *bath*) labelled by μ, with annihilation operators a_μ. Let us demonstrate this simple equivalence. The Green function associated with the noninteracting Anderson model (11.10) is simply

$$\boldsymbol{G}_{\mathrm{full}}(i\omega_n) = \frac{1}{i\omega_n - \boldsymbol{T}}, \tag{11.11}$$

where the full hopping matrix $\boldsymbol{T}$ for the combined cluster and bath system is

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{t}' & \boldsymbol{\theta} \\ \boldsymbol{\theta}^\dagger & \boldsymbol{\varepsilon} \end{pmatrix} \tag{11.12}$$

$\boldsymbol{t}$ is the $L \times L$ hopping matrix within cluster degrees of freedom only, $\boldsymbol{\theta}$ is the $L \times N_b$ hopping matrix between bath and cluster orbitals, and $\boldsymbol{\varepsilon}$ the diagonal $N_b \times N_b$ matrix of bath energies ε_μ. The Green function obtained by tracing out the bath degrees of freedom is simply the restriction of $\boldsymbol{G}_{\mathrm{full}}$ to the cluster degrees of freedom only. The mathematical problem at hand is simply to invert a 2×2 block matrix

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}^{-1}, \tag{11.13}$$

where $A_{11} = \omega - \boldsymbol{t}$, $A_{12} = A^\dagger_{21} = \boldsymbol{\theta}$, $A_{22} = \omega - \boldsymbol{\varepsilon}$ and B_{11} is the Green function we are looking for. By working out the inverse matrix condition, we find in particular that

$$A_{11}B_{11} + A_{12}B_{21} = \mathbf{1} \tag{11.14}$$

$$B_{21} = -A_{22}^{-1} A_{21} B_{11} \tag{11.15}$$

and therefore

$$\left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right) B_{11} = \mathbf{1}\,. \tag{11.16}$$

The Green function $\boldsymbol{G}'$ of the cluster is thus given by

$$\boldsymbol{G}'^{-1} = \omega - \boldsymbol{t}' - \boldsymbol{\theta} \frac{1}{\omega - \boldsymbol{\varepsilon}} \boldsymbol{\theta}^\dagger, \tag{11.17}$$

where we recognize the hybridization function (11.9) in the last term. This completes the demonstration.

The full effective Hamiltonian of each cluster is therefore

$$H' = \sum_{\alpha,\beta} t_{\alpha\beta} c_\alpha^\dagger c_\beta + \sum_{\alpha,\mu} \left(\theta_{\alpha\mu} c_\alpha^\dagger a_\mu + \text{H.c.}\right) + \sum_{\mu}^{N_b} \varepsilon_\mu a_\mu^\dagger a_\mu + H_1 \tag{11.18}$$

Adding the interaction term H_1 contributes a self-energy $\boldsymbol{\Sigma}$ to the inverse Green function, but, for a given set of bath parameters, H_1 does not affect the hybridisation function, since the tracing out of bath degrees of freedom does not involve H_1 in any way. Thus, the interacting Green function of the cluster takes the following form:

$$\begin{aligned} \boldsymbol{G}'^{-1}(i\omega_n) &= i\omega_n - \boldsymbol{t}' - \boldsymbol{\Gamma}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n) \\ &= \mathscr{G}_0^{-1}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n) \end{aligned} \tag{11.19}$$

The self-energy $\boldsymbol{\Sigma}$ will of course depend on the value of the bath parameters, i.e., on the environment. This is an improvement over cluster perturbation theory, in which $\boldsymbol{\Sigma}$ is only affected by the cluster itself.

11.1.2 The Self-Consistency Condition

The approximate Green function of the lattice model that follows from the effective Hamiltonian (11.18) is obtained by borrowing the self-energy from the cluster Green function $\boldsymbol{G}'$ and applying it to the lattice, like in CPT:

$$\boldsymbol{G}^{-1}(\tilde{\boldsymbol{k}}, i\omega_n) = i\omega_n - \boldsymbol{t}(\tilde{\boldsymbol{k}}) - \boldsymbol{\Sigma}(i\omega_n). \tag{11.20}$$

Here $\boldsymbol{t}(\tilde{\boldsymbol{k}})$ is the exact dispersion of the lattice model, but expressed as a partial Fourier transform, i.e., as matrix in cluster indices with a dependence on the reduced wave-vector $\tilde{\boldsymbol{k}}$ defined in the Brillouin zone of the superlattice.

Let us apply a partial Fourier transform to $\boldsymbol{G}(\tilde{\boldsymbol{k}}, i\omega_n)$ back to real-space, to project it onto the cluster:

$$\bar{\boldsymbol{G}}(i\omega_n) = \frac{L}{N} \sum_{\tilde{\boldsymbol{k}}} \left[i\omega_n - \boldsymbol{t}(\tilde{\boldsymbol{k}}) - \boldsymbol{\Sigma}(i\omega_n) \right]^{-1} \tag{11.21}$$

L being the number of sites in the cluster and N the total number of lattice sites (N/L is the number of distinct values of $\tilde{\boldsymbol{k}}$). Ideally, this projected Green function $\bar{\boldsymbol{G}}$ should coincide with the cluster Green function $\boldsymbol{G}'$ calculated from the dynamical mean field $\mathscr{G}$. The condition $\bar{\boldsymbol{G}} = \boldsymbol{G}'$ closes a self-consistency loop, illustrated in

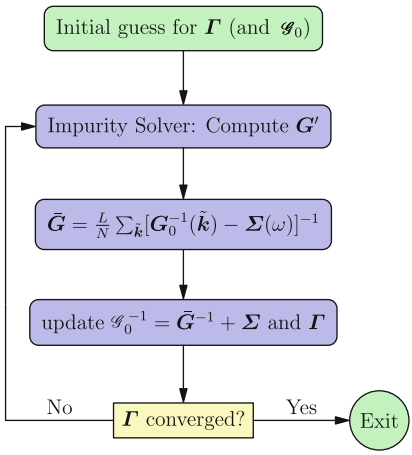

Fig. 11.1 The CDMFT self-consistency loop

Fig. 11.1, that may converge towards an optimal dynamical mean field $\mathscr{G}_0$. Let us summarize the elements of this loop:

1. An initial trial value of the dynamical mean field $\mathscr{G}_0$, or equivalently of the hybridization function $\boldsymbol{\Gamma}$, is selected.
2. The impurity solver is applied and the cluster Green function $\boldsymbol{G}'$ is computed, as well as the associated self-energy $\boldsymbol{\Sigma}$, from (11.19).
3. The CPT Green function (11.20) is computed, as well as its projection $\bar{\boldsymbol{G}}$ onto the cluster, (11.21).
4. The dynamical mean field is updated by substituting $\boldsymbol{G}' \to \bar{\boldsymbol{G}}$ into (11.19):

$$\mathscr{G}_0^{-1}(i\omega_n) \to \bar{\boldsymbol{G}}^{-1}(i\omega_n) + \boldsymbol{\Sigma}(i\omega_n) \tag{11.22}$$

5. One goes back to step 2, until $\mathscr{G}_0$ (or $\boldsymbol{\Gamma}$) converges.

The details of this self-consistency loop depend whether one uses a QMC solver or an exact diagonalization (ED) solver. In particular, since a QMC solver works directly with a discrete representation of the dynamical mean field $\mathscr{G}_0$, step 4 above is rather direct. On the other hand, an ED solver proceeds through the Hamiltonian formulation, and this requires the AIM representation (11.10). Step 4 therefore implies a procedure to infer an updated set of bath parameters $\{\theta_{\alpha\mu}, \varepsilon_\mu\}$ from the updated mean field $\mathscr{G}_0$, or equivalently from the updated hybridization function $\boldsymbol{\Gamma}$. This involves inverting relation (11.9), which can only be done approximately with a finite bath (see Sect. 11.2 below).

11.1.3 The SFA Point of View

Section 9.5 of this volume explains how the SFA may be used to study systems with bath degrees of freedom, such as the AIM representation of the DMFT or CDMFT

systems. Recall that the SFA defines a functional $\Omega[\boldsymbol{\Sigma}]$ of the self-energy, which can be evaluated exactly on a restricted set of self-energies that are the physical self-energies of a family of Hamiltonians H'. The Hamiltonians H' differ from the original Hamiltonian (11.1) in their one-body terms only: they share the same interaction. In the context of CDMFT, the Hamiltonians H' are precisely the AIM representation (11.10), parametrized by the bath parameters collectively written as $(\boldsymbol{\theta}, \boldsymbol{\varepsilon})$. The functional $\Omega[\boldsymbol{\Sigma}]$ reduces, in this restricted space, to a function $\Omega(\boldsymbol{\theta}, \boldsymbol{\varepsilon})$, given by

$$\Omega(\boldsymbol{\theta}, \boldsymbol{\varepsilon}) = \Omega' - \frac{L}{N} T \sum_{\omega_n} \sum_{\tilde{\boldsymbol{k}}} \ln \det \left\{ 1 - \left[\boldsymbol{t}(\tilde{\boldsymbol{k}}) - \boldsymbol{t}' - \boldsymbol{\Gamma}(i\omega_n) \right] \boldsymbol{G}'(i\omega_n) \right\}, \quad (11.23)$$

where Ω' is the grand potential of the AIM (11.10). The optimal bath parameters are found in principle by solving the Euler equations

$$\frac{\partial \Omega}{\partial \boldsymbol{\theta}} = 0 \qquad \frac{\partial \Omega}{\partial \boldsymbol{\varepsilon}} = 0 \quad (11.24)$$

From the above expression for Ω, the Euler equations are

$$\sum_{\omega_n} \mathrm{Tr} \left\{ \left[\boldsymbol{G}'^{-1}(i\omega_n) - \bar{\boldsymbol{G}}^{-1}(i\omega_n) \right] \cdot \frac{\partial \boldsymbol{\Sigma}'(i\omega_n)}{\partial \boldsymbol{\theta}} \right\} = 0. \quad (11.25)$$

and likewise for derivatives with respect to the bath energies $\boldsymbol{\varepsilon}$.

The SFA thus provides an unambiguous and well-justified prescription for setting the parameters $(\boldsymbol{\theta}, \boldsymbol{\varepsilon})$ of the effective medium in DMFT or CDMFT. The condition (11.25) is automatically fulfilled if the CDMFT self-consistency condition $\bar{\boldsymbol{G}} = \boldsymbol{G}'$ is satisfied, thus providing a justification for the self-consistency procedure that is rooted in a dynamic variational principle.

11.2 The Exact Diagonalization Implementation

11.2.1 Working with a Small Bath System: The Distance Function

The ED solver is based on the AIM representation (11.10) of the effective medium. We shall not review here the ED procedure itself: the interested reader is referred to a section of Chap. 7.9 of this volume. The computational size of the ED problem is set by the total number of orbitals in the cluster and bath. In most applications, that number does not exceed 12. Therefore the number of bath parameters in the set $\{\boldsymbol{\theta}, \boldsymbol{\varepsilon}\}$ is rather limited and the self-consistency condition $\bar{\boldsymbol{G}} = \boldsymbol{G}'$ cannot be satisfied exactly. It must instead be approximately satisfied, in some optimal way,

and this is usually done as follows. A merit function is defined:

$$d(\boldsymbol{\theta}, \boldsymbol{\varepsilon}) = \sum_{\omega_n} W(i\omega_n) \operatorname{Tr} \left| \boldsymbol{G}'^{-1}(i\omega_n) - \bar{\boldsymbol{G}}^{-1}(i\omega_n) \right|^2, \tag{11.26}$$

where the dependence on the bath parameters appears in $\boldsymbol{G}'$, through (11.19) and $\mathscr{G}_0$ therein. The sum is taken in principle over all Matsubara frequencies, but in practice a cutoff $i\omega_c$ is imposed. The weights $W(i\omega_n)$ are used to emphasize some frequencies more than others, and their definition is guided more by heuristics than by any rigorous process. Ideally, this function d should vanish, but with a finite bath system it can only be minimized.

Concretely, step 4 of the CDMFT procedure – (11.22), must be implemented as follows: for a given value of $\boldsymbol{\Sigma}$, resulting from a time-consuming ED procedure, the distance function (11.26) must be minimized, in practice using a standard procedure, such as the conjugate gradient method. This minimization does require the frequent evaluation of $\mathscr{G}_0$ through the representation (11.9), but, from a numerical point of view, this is a much lighter task than the ED procedure itself. The minimum of the distance function then defines a new set of bath parameters, from which a new dynamical mean field $\mathscr{G}$ is constructed and the main CDMFT loop summarized in page 346 can proceed.

The ED solver is generally implemented at zero temperature. In that case, the sum over Matsubara frequencies appearing in (11.26) should in principle be replaced by an integral along the imaginary axis. In practice, one uses a small but nonzero "fictitious temperature" β^{-1} in order to evaluate the distance function. The practical implementation of ED-CDMFT thus involves a certain number of "floating" parameters that will quantitatively affect the solutions:

1. The parametrization of the finite bath system itself, i.e., how many parameters $\{\boldsymbol{\theta}, \boldsymbol{\varepsilon}\}$ are actually used and their precise definition.
2. The values of the fictitious temperature β^{-1} and of the cutoff frequency ω_c used in computing the distance function (11.26).
3. The weight $W(i\omega_n)$ appearing in (11.26).

We will discuss the first of these factors in the next section, and the remainder in the subsequent one.

11.2.2 Bath Parametrization

In the most general case, for a cluster containing $2L$ orbitals and $2N_b$ bath orbitals (spin degeneracy included), there would be $2N_b$ bath energies ε_μ and $4LN_b$ hybridization parameters $\theta_{\alpha\mu}$. This means a total of 144 parameters for a typical $(L, N_b) = (4, 8)$ system. Symmetries will greatly reduce that number. For instance, if spin independence is assumed, this is reduced to $N_b(L+1)$. Point group symmetries will reduce that number even further, as we shall now see.

Let us assume that the cluster is invariant under a point group $\mathfrak{G}$, whose elements are permutations of the cluster sites. For instance, $\mathfrak{G} = C_2$ is the group of two elements generated by rotations of π about an axis, or reflexions about a plane; $\mathfrak{G} = C_{2v}$ is the group generated by reflexions about two different planes (e.g., horizontal and vertical). We will assume for simplicity that the group $\mathfrak{G}$ is Abelian, so that all of its irreducible representations are one-dimensional. The square plaquette cluster (2×2) with nearest-neighbor and isotropic next-nearest-neighbor hopping is invariant under the larger group C_{4v}, which is non-Abelian, but we will only use its invariance under the C_{2v} Abelian subgroup. Under these assumptions, we can define linear combinations of cluster orbitals that fall into one-dimensional, irreducible representations of the point group, and the corresponding annihilation operators, that we note $c_\rho^{(r)}$. The index r labels the irreducible representation, and the index ρ labels possible degeneracies within a given representation. For instance, C_{2v} has four irreducible representations that we can label as $(+,+)$, $(+,-)$, $(-,+)$ and $(-,-)$, the signs giving the even or odd character of the orbital under reflexions along the horizontal and vertical axes, respectively. C_2 has two irreducible representations, respectively, even $(+)$ and odd $(-)$ under the group's generator.

In order for the Hamiltonian H' to be invariant under the point group, the bath orbitals also must transform under irreducible representations of the point group. We thus introduce the notation $a_\lambda^{(r)}$ for the bath operators, where λ labels the different realizations of the irreducible representation r within the bath. The cluster Hamiltonian could be rewritten in terms of these eigen-orbitals as

$$H' = \sum_{r,\rho} t'^{(r)}_\rho c^{(r)\dagger}_\rho c^{(r)}_\rho + \sum_{r,\rho,\lambda} \left(\theta^{(r)}_{\rho\lambda} c^{(r)\dagger}_\rho a^{(r)}_\lambda + \text{H.c.} \right) + \sum_{r,\lambda} \varepsilon^{(r)}_\lambda a^{(r)\dagger}_\lambda a^{(r)}_\lambda + H_1 \tag{11.27}$$

(spin indices suppressed). These eigen-orbitals can be chosen so as to make the cluster hopping term diagonal. More important, different irreducible representations cannot be mixed in the hybridization term. This reduces the number of bath parameters from the most general case. For instance, for a 2×2 cluster with spin independence, $N_b = 8$ and C_{2v} symmetry, we can choose to assign two bath sites to each of the four irreducible representations, and this translates into a total of 16 bath parameters. This most general parametric representation of the bath was put forward in [4, 5].

However, many implementations of ED-CDMFT use a somewhat more restrictive, yet more intuitive parametrization of the bath. Consider Fig. 11.2, which illustrates two parametrizations of a four-site bath coupled to a two-site cluster for a one-dimensional problem. On the left panel (a), a representation based on two bath sites per irreducible representation of C_2 is illustrated (the signs appearing over the dashed links apply to the corresponding hybridization amplitudes). On the right (b), each edge of the cluster is associated with its own bath, with equal parameters because of the C_2 symmetry of the system. That representation is clearly a special case of the more general representation (a) when $\varepsilon_1^+ = \varepsilon_2^+ = \varepsilon_1$ and $\varepsilon_1^- = \varepsilon_2^- = \varepsilon_2$.

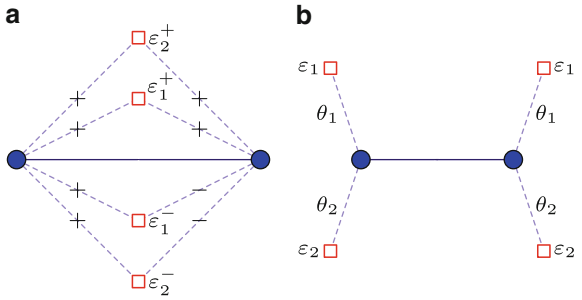

Fig. 11.2 Possible bath configurations for a two-site cluster in one-dimension. *Left* (**a**): parametrization in terms of symmetry eigen-operators. *Right* (**b**): simpler parametrization based on edge sites only. The *blue circles* are real-space cluster sites (not eigen-operators). The *squares* represent bath orbitals

In principle, one should use the most general bath parametrization possible, since this leaves more space to fulfill the CDMFT self-consistency condition. In practice however, the usefulness of the most general parametrization is not obvious. Consider for instance Fig. 11.3, where the most general parametrization was applied to the one-dimensional system illustrated in Fig. 11.2a. The bath parameters ε_1^+ and ε_2^+ are different in the dynamical impurity approximation (i.e., the SFA solution of the problem), but are not differentiated by the distance function used in the CDMFT solution; thus $\varepsilon_1^+ = \varepsilon_2^+$ in the CDMFT solution, thereby providing a solution equivalent to that of the simpler system illustrated in Fig. 11.2b.

Figure 11.4 illustrates a bath configuration associated with a four-site cluster for the two-dimensional Hubbard model. The bath is configured as two "ghost clusters," with sites numbered 1–4 and 5–8 respectively. In studying the normal state, this configuration would involve four independent parameters:

$$\begin{aligned} &\varepsilon_\mu = \varepsilon \quad (\mu = 1,2,3,4) && \varepsilon_\mu = \varepsilon' \quad (\mu = 5,6,7,8) \\ &\theta_{11} = \theta_{22} = \theta_{33} = \theta_{44} = \theta && \theta_{15} = \theta_{26} = \theta_{37} = \theta_{48} = \theta' \end{aligned} \tag{11.28}$$

When probing the antiferromagnetic state, one would in addition introduce a modulation between odd and even sites, increasing the number of independent bath parameters to 8:

$$\begin{aligned} &\varepsilon_1 = \varepsilon_3 = \varepsilon + \delta\varepsilon && \varepsilon_5 = \varepsilon_7 = \varepsilon' + \delta\varepsilon' \\ &\varepsilon_2 = \varepsilon_4 = \varepsilon - \delta\varepsilon && \varepsilon_6 = \varepsilon_8 = \varepsilon' - \delta\varepsilon' \\ &\theta_{11} = \theta_{33} = \theta + \delta\theta && \theta_{15} = \theta_{37} = \theta' + \delta\theta' \\ &\theta_{22} = \theta_{44} = \theta - \delta\theta && \theta_{26} = \theta_{48} = \theta' - \delta\theta' \end{aligned} \tag{11.29}$$

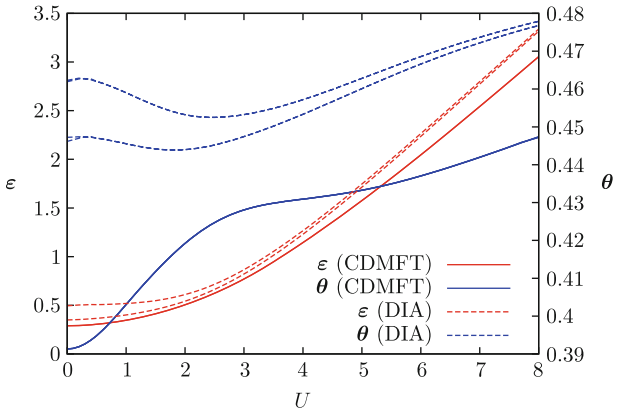

Fig. 11.3 Bath energies (ε_1^+ and ε_2^+, left axis) and hybridizations (θ_1^+ and θ_2^+, right axis) for the system illustrated in Fig. 11.2a, at half-filling, as a function of one-site interaction U. The *full line* is the result of the CDMFT procedure, whereas the *dashed lines* are the DIA results (i.e., using Potthoff's SFA approach). Because of particle hole symmetry, $\varepsilon_1^- = -\varepsilon_2^+$ and $\varepsilon_2^- = -\varepsilon_1^+$ (not shown). The values of ε_1^+ and ε_2^+ are different in the DIA, but they coincide in CDMFT (same for θ_1^+ and θ_2^+)

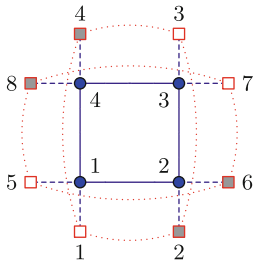

Fig. 11.4 Possible bath configuration for a four-site cluster in two dimensions. The *blue circles* are real-space cluster sites and the *squares* represent bath sites. *Darkened squares* have distinct bath parameters when the antiferromagnetic state is probed. The *dashed curves* represent in-bath pairing terms that are introduced to probe d-wave superconductivity

A possible way to probe d-wave superconductivity is to add singlet pairing terms between bath sites, much as would be done between cluster sites in the variational cluster approach. Each one of these pairing terms has the form

$$\hat{\Delta}_{ij} = a^\dagger_{1\uparrow} a^\dagger_{2\downarrow} + a^\dagger_{2\uparrow} a^\dagger_{1\downarrow} + \text{H.c.} \tag{11.30}$$

and the sum of these terms would be

$$H'_{\text{sc}} = \hat{\Delta}_{12} + \hat{\Delta}_{34} - \hat{\Delta}_{14} - \hat{\Delta}_{23} + \hat{\Delta}_{56} + \hat{\Delta}_{78} - \hat{\Delta}_{58} - \hat{\Delta}_{67} \tag{11.31}$$

The pairing terms just introduced hybridize the bath orbitals between themselves and therefore do not fall into the general form (11.18). In other words, the matrix $\boldsymbol{\varepsilon}$ of (11.12) is no longer diagonal and has an anomalous part. Before applying relation (11.9), one simply needs to diagonalize it and modify the matrix $\boldsymbol{\theta}$ accordingly. The latter would then acquire an anomalous part.

Instead If we were to use a basis of symmetry eigen-operators to parametrize the same system, the pairing terms would show up directly as anomalous hybridizations between cluster and bath. The symmetry group would be C_{2v}, except if superconductivity and antiferromagnetism are probed simultaneously, in which case the cluster-bath symmetry would be reduced to C_2.

11.2.3 The Distance Function

The distance function (11.26) is to some degree arbitrary: there is freedom in the choice of the fictitious temperature β^{-1}, the cutoff ω_c and the weights $W(i\omega_n)$. Some authors also omit the square in (11.26). This arbitrariness reflects the finite size of the bath. At the same time, it completely disappears, along with the need of a distance function, if the bath parameters are set according to the SFA and (11.24) (this is called the dynamical impurity approximation, or DIA, in Chap. 9.5 of this volume).

We will illustrate the effect of this arbitrariness with the following weight functions $W(\omega)$:

$$W(i\omega_n) = 1 \text{ within } \omega_n \in [0, \omega_c] \text{ (sharp cutoff)} \tag{11.32a}$$

$$W(i\omega_n) = 1/\omega_n \text{ (extra weight at low } \omega_n) \tag{11.32b}$$

$$W(i\omega_n) = \operatorname{Tr} |\boldsymbol{\Sigma}^2(i\omega_n)| \text{ within } \omega_n \in [0, \omega_c] \tag{11.32c}$$

All of these functions have a finite support between $\omega = 0$ and some cutoff frequency $i\omega_c$, and are evaluated on a grid of Matsubara frequencies defined by a fictitious temperature $1/\beta$. They ignore high frequencies: although a lot of information is buried in the high-frequency behavior of the Green function (the moments), that information is not weighted uniformly across all Matsubara frequencies. If we use the SFA Euler equation (11.25) as a loose guide, the weight function should be highest where the variation of the self-energy is highest. Under the reasonable assumption that this condition will be generally met when the self-energy itself is largest, the third of the above weight functions is proportional to the square of the self-energy (the trace is there to make it basis independent, like (11.26)). The second of these functions is designed to emphasize very low-frequencies and, under the same assumptions, would best work for a Mott insulator, whose self-energy is singular at low frequencies.

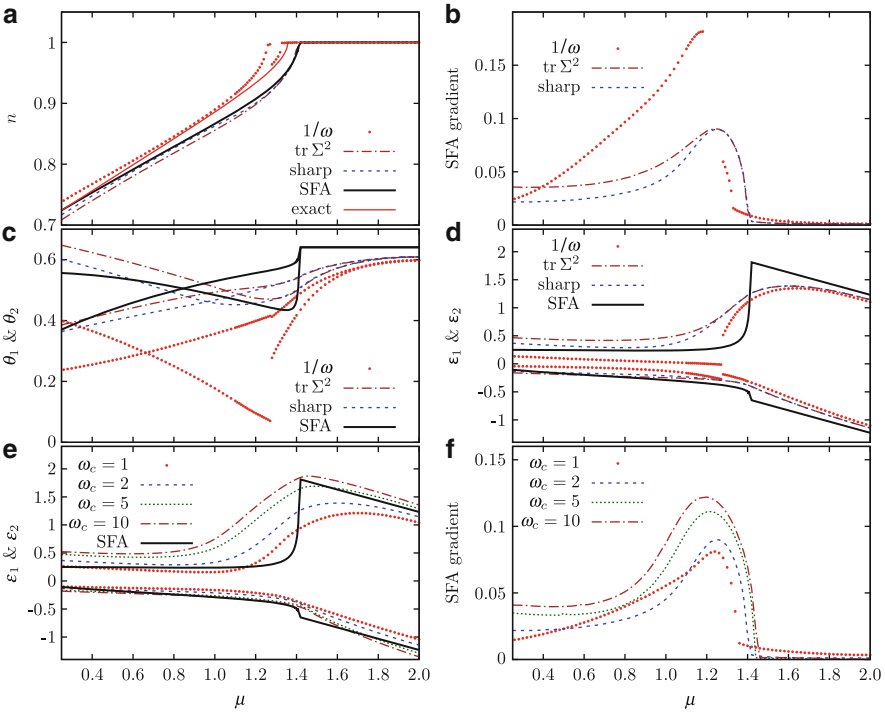

Fig. 11.5 Plots illustrating the CDMFT and DIA solutions from the cluster of Fig. 11.2b, for the one-dimensional Hubbard model, all expressed as a function of the chemical potential μ. Panel (**a**): Electron density n (the exact Lieb and Wu result is also shown). Panel (**b**): SFA gradient associated with the CDMFT solutions of Panel (**a**). Panel (**c**) and (**d**): Corresponding hybridization parameters $\theta_{1,2}$ and bath energies $\varepsilon_{1,2}$ for the weight functions (11.32) and the SFA solution. Panel (**e**) and (**f**): Bath energies $\varepsilon_{1,2}$ for a sharp cutoff (11.32a) with various values of ω_c, and the corresponding values of the SFA gradient. Unless indicated otherwise, the cutoff was set at $\omega_c = 2$ and the fictitious inverse temperature at $\beta = 100$. Taken from [6]

The CDMFT solutions obtained from these different distance functions, as well as those obtained in the DIA, are presented in Fig. 11.5, which we will now explain. All graphs in this figure pertain to the one-dimensional Hubbard model at zero temperature and $U = 4$, $t = 1$, as approximated by the cluster-bath system of Fig. 11.2b. They show various quantities as a function of the chemical potential μ between $\mu = 0$ and the particle–hole symmetric point $\mu = U/2$. Panels (c) and (d) show the bath parameters $\boldsymbol{\theta}$ and $\boldsymbol{\varepsilon}$, for the various weight functions defined in (11.32), as well as those obtained from the DIA (black lines). The striking feature of these two graphs is that the DIA solutions are constant within a range of μ in which the system is exactly half-filled (the bath energies in fact are linear functions of μ, simply because the chemical potential was defined so as to affect the bath electrons as well; but the combinations $\varepsilon_i + \mu$ are indeed constant in this range). This is expected and natural since the physical properties of the system are

μ-independent in the range of μ associated with half-filling (i.e., within the Mott gap). On the other hand, the various CDMFT solutions show a variation of the bath parameters in that range. This cannot be attributed to the finite size of the bath, since the DIA results pertain to the same system. It should instead be attributed to the self-consistency condition $\bar{\boldsymbol{G}} = \boldsymbol{G}'$, which is not exactly equivalent to the more fundamental variational condition (11.24) for a finite bath size.

Why then not use the DIA systematically? Because it is computationally much more complex that CDMFT. Even though the same impurity solver is used in both approaches, the number of times it must be called is much larger in the DIA, and the $\boldsymbol{\varepsilon}$- and $\boldsymbol{\theta}$-dependence of the functional (11.23) apparently makes for a complex landscape on which saddle points are difficult to converge to, except for a small number of bath parameters. For this reason, CDMFT remains a very practical approach. What then should be our guide in choosing the parameters of the distance function (11.26)? One possibility is to try a few, and compare, for each solution, the values of the gradient of the functional (11.23) at each of the CDMFT solutions found:

$$|\nabla \Omega| = \sqrt{\left(\frac{\partial \Omega}{\partial \boldsymbol{\theta}}\right)^2 + \left(\frac{\partial \Omega}{\partial \boldsymbol{\varepsilon}}\right)^2}. \tag{11.33}$$

This is done in Panel (b) of Fig. 11.5. In that example, the sharp cutoff at $\omega_c = 2$ is the one that overall yields the lowest SFA gradient. One could then assert that this solution is closest to the ideal DIA solution. This is more or less confirmed by comparing the bath parameters on panels (c) and (d). Panels (e) and (f) show the same comparisons, this time for different values of the cutoff ω_c, with a sharp cutoff. Again, picking the cutoff that yields the lowest SFA gradient (11.33) also brings us closest to the DIA solution.

Finally, panel (a) of Fig. 11.5 compares the electron density n obtained from the various solutions with the exact result of Lieb and Wu [7]. It is tempting from this graph to assert that the weight function (11.32b) is the best choice, since it is closest to the exact solution, whereas the other weight functions are closer to the DIA solution, which differs from the exact result. However, we only expect convergence towards the exact solution when both the number of sites L and the number of bath sites N_b reach the thermodynamic limit. The closeness of the exact solution to one of the CDMFT solutions is accidental: changing the fictitious temperature β^{-1} would slightly shift that solution on one side or the other.

It is worth noting that all CDMFT solutions, as will as the DIA solution, show an infinite compressibility point ($\partial n/\partial \mu \to \infty$) at some critical value μ_{gap} of the chemical potential, and this is an important property of the exact solution. That property is not recovered when single-site DMFT or P-CDMFT is used [8]. By contrast, it is clearly seen in the QMC-CDMFT solution [9], where the value of μ_{gap} is essentially the same as the one found in the DIA ($1.41t$ for $U = 4$).

11.3 Quantum Monte Carlo Solvers

QMC methods exist in many flavors. The Hirsch-Fye method [10] (HF-QMC) has been used since the beginning of CDMFT [1]. More recently it tends to be replaced by the Continuous-Time QMC (CT-QMC) [11–13]. This is not the place to provide a detailed description of either approach. We will simply provide a cursory introduction to each. A comparison of the performances of the various approaches can be found in [14].

ED and QMC solvers complement each other on many aspects:

1. The QMC solver has access to a practically infinite bath, compared with a maximum size of about $N_b = 8$ in the ED solver. Indeed, the QMC solver does not require in practice the introduction of the Hamiltonian parametrization (11.18) for the bath, since it works directly in the functional representation (11.6) in which the dynamical mean field $\mathscr{G}_0$ may be represented by a frequency-dependent matrix $\mathscr{G}_0(i\omega_n)$ in the space of cluster indices.
2. The QMC solver works at finite temperature, and extrapolations are necessary to recover zero-temperature information. By contrast, the ED solver works primarily at zero temperature; even though a finite-temperature implementation is possible [5], its accuracy is not well controlled.
3. The QMC solver has access to dynamical information in imaginary time only, in the interval $\tau \in [0, \beta]$. Equivalently, this information is confined to the set of Matsubara frequencies $i\omega_n$. Access to approximate real-frequency properties is possible through "analytic" continuation, which is achieved through the uncontrolled maximum entropy method (MEM). On the other hand, the ED solver provides the Green function for any complex or real frequency.
4. The QMC solver is subject to the well-known fermion sign problem. In the two-dimensional Hubbard model, this problem is basically present everywhere except at the particle-hole symmetric point ($\mu = U/2, t' = 0$) and increases in intensity the further one strays from this point, either by increasing doping or the diagonal hopping t'.
5. QMC solvers have access to larger cluster sizes, although this possibility has been exploited more in the DCA than in CDMFT, and with the Hirsch-Fye method. However, the CT-QMC method with hybridization expansion (see below) is limited to small cluster sizes (4–6).
6. On the numerical side, the ED solver requires an amount of memory that grows exponentially with the number of cluster and bath sites, and the Lanczos method scales poorly when parallelized. The QMC solver is easily parallelized and does not require large memory or communication (it scales linearly). But being stochastic, its accuracy grows with the square-root of simulation time. Overall, the QMC solver requires more resources in a typical use, but is more adapted to large supercomputers.

11.3.1 The Hirsch-Fye Method

An introduction to the Hirsch-Fye QMC and its use in CDMFT can be found in [9]. Will we summarize the main points here. We will assume for simplicity that we are dealing with the one-band Hubbard model, with interaction

$$H_1 = U \sum_a n_{a\uparrow} n_{a\downarrow} \tag{11.34}$$

so that the greek index α is now a composite (a, σ) of the site and spin indices. We will in addition assume that the Hamiltonian does not mix or hybridize up and down spins.

The method deals directly with the effective action (11.6), except that the imaginary time segment $[0, \beta]$ must be discretized into N_β time slices. The dynamical mean field $\mathscr{G}_0$ becomes a matrix with cluster and time slice indices, and the effective action (11.6) becomes

$$S_{\text{eff}}[c, c^*] = \Delta\tau^2 \sum_{a,b,\sigma,\tau,\tau'} c^*_{a\sigma\tau} (\mathscr{G}_0^{-1})_{a\tau,b\tau'} c_{b\sigma\tau'} + U\Delta\tau \sum_{a,\tau} n_{a\uparrow\tau} n_{a\downarrow\tau} \tag{11.35}$$

In the above expression the indices τ and τ' are discrete and label the different time slices; the integrals over imaginary time have been replaced by sums, with the slice width $\Delta\tau = \beta / N_\tau$ appearing as a prefactor.

In HF-QMC, the following Hubbard–Stratonovich factorization is applied to the interaction part:

$$\mathrm{e}^{-\Delta\tau U n_{a\uparrow\tau} n_{a\downarrow\tau}} = \mathrm{e}^{-\frac{1}{2}\Delta\tau U [n_{a\uparrow\tau} + n_{a\downarrow\tau}]} \sum_{s_{a\tau}=\pm 1} \mathrm{e}^{\lambda s_{a\tau} [n_{a\uparrow\tau} - n_{a\downarrow\tau}]}, \tag{11.36}$$

where the constant λ is defined by the relation $\cosh\lambda = \mathrm{e}^{\Delta\tau U/2}$. This factorization introduces a set of LN_τ Ising spins $s_{a\tau}$ as auxiliary fields.

Using this decomposition, the action $S_{\text{eff}}[c, c^*]$ becomes quadratic in c, c^* and the partition function can be computed in terms of a propagator $G_\sigma(\{s\})$ which depends on the particular Ising configuration $s_{\alpha\tau}$:

$$Z = \sum_{s_{a\tau}=\pm 1} \int [Dc\, Dc^*] \exp - \left(\sum_{a,b,\sigma,\tau,\tau'} c^*_{a\sigma\tau} [G_\sigma^{-1}(\{s\})]_{a\tau,b\tau'} c_{\beta\sigma\tau'} \right), \tag{11.37}$$

where the inverse propagator G_σ^{-1} reads, for a given Ising configuration $\{s\}$,

$$[G_\sigma^{-1}(\{s\})]_{a\tau,b\tau'} = (\tilde{\mathscr{G}}_0^{-1})_{a\tau,b\tau'} - \sigma\lambda s_{a\tau} \delta_{ab} \Delta_{\tau,\tau'+1}, \tag{11.38}$$

where the "antiperiodic" function $\Delta_{\tau,\tau'+1}$ is defined as 1 if $\tau = \tau' + 1$ and -1 if $\tau = 1$ and $\tau' = N_\tau$ [15]. The dynamical mean field $\tilde{\mathscr{G}}_0$ differs from $\mathscr{G}_0$ by a shift in the chemical potential induced by the Hubbard–Stratonovich factorization. The functional integral over the fermion variables $c_{a\tau}$ results in a simple determinant of the inverse propagator:

$$Z = \sum_{s_{a\tau}=\pm 1} \prod_{\sigma=\pm 1} \det G_\sigma^{-1}(\{s\}). \tag{11.39}$$

The task is reduced to a stochastic evaluation of the sum over Ising configurations, with a determinant to be recalculated for each update of the Ising configuration.

The sum over Ising configurations is carried out by setting up a random walk through configuration space (the Metropolis algorithm). The probability $p(s \to s')$ of transition from a configuration $\{s\}$ to a configuration $\{s'\}$ must obey the principle of detailed balance:

$$\frac{p(s \to s')}{p(s' \to s)} = \frac{\prod_\sigma \det G_\sigma^{-1}(\{s'\})}{\prod_\sigma \det G_\sigma^{-1}(\{s\})}. \tag{11.40}$$

A quick update of the transition probabilities is possible if the Ising configuration $\{s\}$ is updated locally (by flipping one Ising spin at a time) [9, 10].

Once the random walk has reached equilibrium, measurements of physically relevant quantities, such as the cluster Green function $\boldsymbol{G}'$, are performed and statistical errors are estimated using binning methods. The cluster Green function is obtained as a function of time indices τ and τ', as a matter of fact as a function of their difference $\tau - \tau'$. This information must be transformed into a frequency-dependent $\boldsymbol{G}'(i\omega_n)$ to fit into the CDMFT procedure outlined in Fig. 11.1. This is accomplished by (1) constructing a cubic spline interpolation $\boldsymbol{G}'_{\text{inter}}(\tau)$ as a function of continuous imaginary time and (2) computing the Fourier transform

$$\boldsymbol{G}'(i\omega_n) = \int_0^\beta \mathrm{d}\tau\, \boldsymbol{G}'_{\text{inter}}(\tau) \mathrm{e}^{i\omega_n \tau}. \tag{11.41}$$

However, doing this literally is likely to yield the incorrect asymptotic behavior as $i\omega_n \to \infty$, because of the finite value of $N\tau$ and statistical errors. A practical solution is to use a known Green function $\boldsymbol{G}_{\text{ref}}$ (for instance, for the noninteracting problem) and to compute instead

$$\boldsymbol{G}'(i\omega_n) = \boldsymbol{G}_{\text{ref}}(i\omega_n) + \int_0^\beta \mathrm{d}\tau\, \left(\boldsymbol{G}'_{\text{inter}}(\tau) - \boldsymbol{G}_{\text{ref}}(\tau)\right) \mathrm{e}^{i\omega_n \tau}. \tag{11.42}$$

In addition to statistical errors, which are controlled by the binning methods, this approach makes a systematic error related to the finite number N_τ of time slices. In principle, one must repeat the process for a few values of $\Delta\tau$ and stop when averages no longer depend on $\Delta\tau$ within the statistical errors. One can also extrapolate observables to $\Delta\tau \to 0$, knowing that convergence occurs as $(\Delta\tau)^2$ [16].

11.3.2 The Continuous-Time Method

Continuous-time QMC refers to a variety of new-generation QMC methods that do not rely on a fixed discretization of the imaginary-time interval $[0, \beta]$, but rather on the stochastic evaluation of a perturbative expansion. These methods have had a tremendous success in essentially solving most models of bosons [17, 18]. They are fraught with the same sign problem as other QMC approaches when applied to fermions, but this problem is somewhat less severe in impurity-like approaches like CDMFT. In recent years, three flavors of CT-QMC have been developed for impurity problems. Depending on which part of the action is expanded, one speaks of *weak-coupling* [19], *hybridization expansion* [11] or *auxiliary-field* formulations [13].

Generally, CT-QMC assumes a decomposition of the Hamiltonian $H = H_A + H_B$, and evaluates the partition function as a series expansion in powers of H_B. Adopting an interaction representation in which H_A is time-independent, the partition function takes the form

$$\begin{aligned} Z &= \operatorname{Tr} \exp - \left(\beta H_A + \int_0^\beta \mathrm{d}\tau \; H_B(\tau) \right) \\ &= \sum_{k=0}^{\infty} (-1)^k \int_0^\beta \mathrm{d}\tau_1 \cdots \int_{\tau_{k-1}}^\beta \mathrm{d}\tau_k \; \operatorname{Tr} \left(\mathrm{e}^{-\beta H_A} T_\tau H_B(\tau_k) \cdots H_B(\tau_1) \right), \end{aligned} \tag{11.43}$$

where T_τ is the time-ordering operator. We will focus here on the hybridization expansion formulation [11]. In that formulation, H_A contains the cluster and bath parts of Hamiltonian (11.18), whereas H_B contains the hybridization between bath and cluster sites:

$$H_A = \sum_{\alpha,\beta} t_{\alpha\beta} c_\alpha^\dagger c_\beta + H_1 + \sum_{\mu}^{N_b} \varepsilon_\mu a_\mu^\dagger a_\mu = H_{\text{clus}} + H_{\text{bath}} \tag{11.44}$$

$$H_B = H_{\text{hyb}} + H_{\text{hyb}}^\dagger, \qquad H_{\text{hyb}} = \sum_{\alpha,\mu} \theta_{\alpha\mu} c_\alpha^\dagger a_\mu. \tag{11.45}$$

The partition function then takes the form

$$
\begin{aligned}
Z = \sum_{k=0}^{\infty}(-1)^k \int_0^\beta \mathrm{d}\tau_1 \cdots \int_{\tau_{k-1}}^\beta \mathrm{d}\tau_k \int_0^\beta \mathrm{d}\tau'_1 \cdots \int_{\tau'_{k-1}}^\beta \mathrm{d}\tau'_k \\
\sum_{\substack{\alpha_1 \cdots \alpha_k \\ \alpha'_1 \cdots \alpha'_k}} \sum_{\substack{\mu_1 \cdots \mu_k \\ \mu'_1 \cdots \mu'_k}} \theta_{\alpha_1 \mu_1} \cdots \theta_{\alpha_k \mu_k}\, \theta^*_{\alpha'_1 \mu'_1} \cdots \theta^*_{\alpha'_k \mu'_k} \\
\times \operatorname{Tr}\left(\mathrm{e}^{-\beta H_{\mathrm{clus}}} T_\tau\, c_{\alpha_k}(\tau_k) c^\dagger_{\alpha'_k}(\tau'_k) \cdots c_{\alpha_1}(\tau_1) c^\dagger_{\alpha'_1}(\tau'_1) \right) \\
\times \operatorname{Tr}\left(\mathrm{e}^{-\beta H_{\mathrm{bath}}} T_\tau\, a^\dagger_{\mu_k}(\tau_k) a_{\mu'_k}(\tau'_k) \cdots a^\dagger_{\mu_1}(\tau_1) a_{\mu'_1}(\tau'_1) \right),
\end{aligned}
\tag{11.46}
$$

where the first trace is evaluated in the cluster system only, and the second one is evaluated in the bath system only. The latter can be evaluated exactly, since the bath is non-interacting. The result can be expressed in terms of the hybridization function

$$
\Gamma_{\alpha\gamma}(\tau) = \sum_\mu \frac{\theta_{\alpha\mu}\theta^*_{\gamma\mu}}{1+\mathrm{e}^{\beta\varepsilon_\mu}} \times \begin{cases} -\mathrm{e}^{(\tau-\beta)\varepsilon_\mu} \\ \mathrm{e}^{\tau\varepsilon_\mu} \end{cases}, \tag{11.47}
$$

as follows:

$$
\begin{aligned}
&\sum_{\substack{\mu_1 \cdots \mu_k \\ \mu'_1 \cdots \mu'_k}} \theta_{\alpha_1 \mu_1} \cdots \theta_{\alpha_k \mu_k}\, \theta^*_{\alpha'_1 \mu'_1} \cdots \theta^*_{\alpha'_k \mu'_k} \\
&\quad \times \operatorname{Tr}\left(\mathrm{e}^{-\beta H_{\mathrm{bath}}} T_\tau\, a^\dagger_{\mu_k}(\tau_k) a_{\mu'_k}(\tau'_k) \cdots a^\dagger_{\mu_1}(\tau_1) a_{\mu'_1}(\tau'_1) \right) = Z_{\mathrm{bath}} \det \Delta,
\end{aligned}
\tag{11.48}
$$

where Z_{bath} is the bath partition function

$$
Z_{\mathrm{bath}} = \prod_\mu \left(1+\mathrm{e}^{\beta\varepsilon_\mu}\right) \tag{11.49}
$$

and Δ is the $k \times k$ matrix

$$
\Delta_{i,j} = \Gamma_{\alpha_i \gamma_j}(\tau_j - \tau_i). \tag{11.50}
$$

The global partition function then reduces to

$$
\begin{aligned}
Z = Z_{\mathrm{bath}} \sum_{k=0}^{\infty}(-1)^k \int_0^\beta \mathrm{d}\tau_1 \cdots \int_{\tau_{k-1}}^\beta \mathrm{d}\tau_k \int_0^\beta \mathrm{d}\tau'_1 \cdots \int_{\tau'_{k-1}}^\beta \mathrm{d}\tau'_k \sum_{\substack{\alpha_1 \cdots \alpha_k \\ \alpha'_1 \cdots \alpha'_k}} \\
\times \operatorname{Tr}\left(\mathrm{e}^{-\beta H_{\mathrm{clus}}} T_\tau\, c_{\alpha_k}(\tau_k) c^\dagger_{\alpha'_k}(\tau'_k) \cdots c_{\alpha_1}(\tau_1) c^\dagger_{\alpha'_1}(\tau'_1) \right) \det \Delta.
\end{aligned}
\tag{11.51}
$$

Calculating the trace in (11.51) necessitates a knowledge of the exact eigenstates of the cluster system. A complete diagonalization of the cluster Hamiltonian is therefore necessary, and this limits the application of the hybridization expansion method to small clusters (e.g., four sites), much like in the ED solver, except that in this case the bath is infinite. Notice that despite the formal introduction of bath sites in the derivation, in the end the determinant $\det \Delta$ does not rely on a particular discretization of the bath, but only on the hybridization function $\boldsymbol{\Gamma}(i\omega_n)$, or equivalently on the related dynamical mean field $\mathscr{G}(i\omega_n)$.

The calculation of the trace in (11.51) for a given set of times $\{\tau_i, \tau_i'\}$ is usually the computational bottleneck of the method. One way of doing it is to adopt a basis of eigenstates $|m\rangle$ and eigenvalues E_m of H_{clus} and to construct a representation of the various creation and annihilation operators in that basis. The trace in (11.51)is then

$$\text{Tr}\left(\mathrm{e}^{H_{\text{clus}}(\tau_k-\beta)}c_{\alpha_k}\mathrm{e}^{H_{\text{clus}}(-\tau_k+\tau_k')}c^\dagger_{\alpha_k'}\cdots\mathrm{e}^{H_{\text{clus}}(-\tau_1+\tau_1')}c^\dagger_{\alpha_1'}\right), \tag{11.52}$$

where the evolution operators are diagonal.

The integral and sums in (11.51) are carried stochastically. A configuration is defined by a set of times $\{\tau_i, \tau_i'\}$, as set of indices $\{\alpha_i, \alpha_i'\}$ and an expansion order k. Configurations are updated by inserting or deleting times in those lists, thus increasing or decreasing the expansion order k by 1. The weight of each contribution can be positive or negative, depending on the order of operators (this is the source of the sign problem in CT-QMC). Readers are referred to the literature [20, 21] for details on the Monte Carlo evaluation itself (updates, detailed balance and so on).

11.4 The Mott Transition

A key success of dynamical mean-field theory is the picture it provides of the Mott metal–insulator transition (see Chap. 6.5 of this volume). Consider Fig. 11.6, which shows the qualitative phase diagram, on the $U-T$ plane, of the half-filled, particle-hole symmetric Hubbard model. The left panel shows the prediction of single-site DMFT [22]. The Mott transition on the $U-T$ plane is of first order, indicated as a red line in the figure. This first-order line ends at a finite-temperature critical point and at $U_{c1(0)}$, with a region (colored area) where the metallic and insulating phases may coexist. This single-site DMFT picture of the Mott transition has been criticized, mainly on the ground that the absence of feedback of magnetic correlations on single particle excitations yields a nonzero ground-state entropy ($S = N\ln 2$, N being the number of sites) in the Mott phase, and that this exaggerates the stability of the insulating phase at nonzero temperature. Cluster dynamical mean field theory modifies this picture by the addition of short-range correlation effects or, said otherwise, by adding a momentum-dependence to the self-energy. This provides a feedback of short-range antiferromagnetic fluctuations into single-particle properties. However, the main features of the DMFT picture are

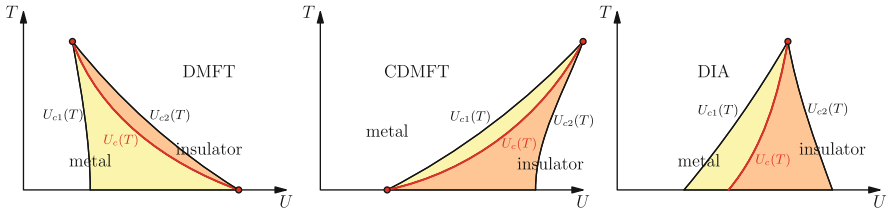

Fig. 11.6 Schematic phase diagram of the half-filled, particle-hole symmetric Hubbard model using single-site DMFT (*left*), CDMFT with a CT-QMC solver on a 2×2 plaquette (*center*) and the Dynamical Impurity Approximation (*right*). At the *red dots*, where the first-order line touches the boundary of the coexistence region, continuous transitions occur

not affected by these refinements. The middle diagram of Fig. 11.6 emerges from a CDMFT study using the continuous-time QMC solver [23]. The essential difference with the single-site result is that the first-order line ends at $U_{c2}(0)$ instead of $U_{c1}(0)$ at zero temperature. The zero-temperature points are the result of an extrapolation, since QMC solvers were used in both cases. The right panel of Fig. 11.6 is a modified scenario inspired from the dynamical impurity approximation (DIA) [24], in which only the zero-temperature axis was actually calculated. It was shown in [24] that the transition is of first order even at zero temperature; two solutions of the Euler equations (11.24) exist, and the transition (red line) occurs when their ground state energies cross. The self-consistency condition $\bar{\mathcal{G}} = \boldsymbol{G}'$, in the bath parametrization used in [24], does not show a first-order transition; the latter can only be seen by solving the more exact conditions (11.24).

Regarding Fig. 11.6, note that the slope of the first-order line is negative in the single-site DMFT solution, but positive in the CDMFT solution. Through the Clausius–Clapeyron equation, this entails that the entropy is larger in the Mott phase than in the metallic phase according to the single-site solution, whereas the opposite is true according to the CDMFT solution. This is another sign that the degeneracy of a single site exaggerates the entropy of the insulating state in DMFT, which leads to an inaccurate description of the Mott transition in single-site DMFT.

Another difference between single-site DMFT and CDMFT lies in the evolution of the density of states across the Mott transition. This is illustrated in Fig. 11.7. In single-site DMFT, the quasiparticle weight Z remains constant while the cental peak in the DOS becomes narrower and narrower until it disappears completely at the transition, leaving a gap already wide open. This can be seen from the left panel of Fig. 11.7[25], or for instance from Fig. 30 of [22]. By contrast, the CDFMT density of states calculated on a four-site plaquette [25, 26] (right panel of Fig. 11.7) shows a drop in Z, until a gap opens up at the transition and increases gradually thereafter. This is the effect of a feedback of short-range correlations on quasiparticle dynamics. Related is the appearance of side bands on either side of the gap that originate from short-range spin dynamics [26].

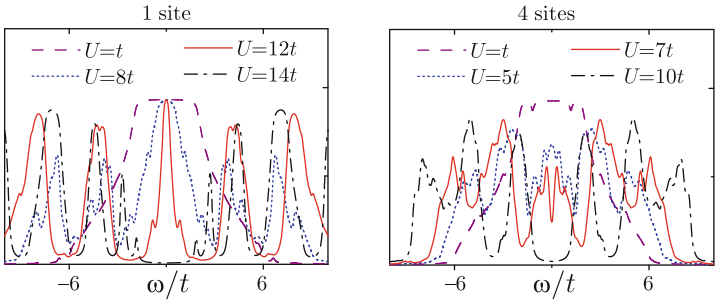

Fig. 11.7 Density of states of the 3D Hubbard model across the Mott transition in single-site DMFT (*left*) and in CDMFT on a 2 × 2 plaquette (data from [25])

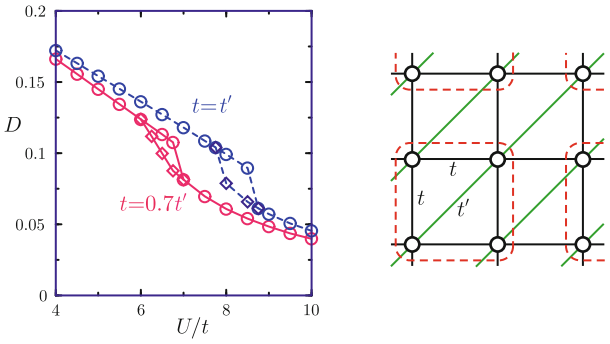

Fig. 11.8 *Left*: Hysteresis curves obtained in [27] for the double occupancy as a function of U for the model illustrated in the *right panel*

However, the square-lattice Hubbard model at half-filling describes an antiferromagnetic insulator. The Mott transition described above is an underlying phenomenon that is revealed only by suppressing antiferromagnetic order in the solutions. A true Mott transition can be seen, however, in systems with magnetic frustration, such as the Hubbard model on a triangular lattice (isotropic or not) at half-filling. The system depicted in the right panel of Fig. 11.8, with a hopping t' along one of the diagonals, provides a minimal modelling of quasi-two-dimensional organic conductors of the BEDT family. If t' is sufficiently close to one, frustration inhibits antiferromagnetism and a Mott transition is observed as a function of U [27–29]. The left panel of Fig. 11.8 shows data taken from [27] for the double occupancy $D = \langle n_\uparrow n_\downarrow \rangle$, obtained by scanning U upwards and downwards. Hysteresis is observed, indicating a first-order transition between a metallic (left) and insulating (right) solution. The Mott insulator in that case is likely a *spin liquid*, i.e., does not support low-energy excitations resulting from a long-range magnetic order. Note that not all types of order have been probed in this system, either with CDMFT or other methods. In particular, above some critical value of U and for t' close to t, the spin

liquid state gives way to a non-collinear magnetic order [30]. Another way of preventing long-range order to reveal the Mott transition is to define the Hubbard model on the Kagomé lattice; CDMFT was applied to this problem in [29, 31].

The Mott transition is not only driven by the interaction strength, but also by doping. The doping-induced Mott transition has been studied in Hirsch-Fye [32] using CT-QMC. A transition between two metallic phases with different electron densities is seen as a function of chemical potential (see Fig. 11.9). The high-density solution describes a correlated metal with smaller double-occupation than the lower-density solution. The transition line coincides with a maximum in the scattering rate, resulting from a strong competition between spin and charge fluctuations. The doping-induced first-order transition observed in [32] ends with a critical line $T(n)$. The critical temperature decreases extremely rapidly with U. Thus, the effects of the Mott transition extend far beyond half-filling.

Related to the Mott transition itself is the pseudogap phenomenon. This is best studied with an ED solver, which provides real-frequency information [5, 26, 28]. In this respect, CDMFT adds nothing spectacular to CPT; indeed, the smaller cluster size in CDMFT, compared to CPT, is not compensated by the benefits of self-consistency (see Chap. 7.9 of this volume).

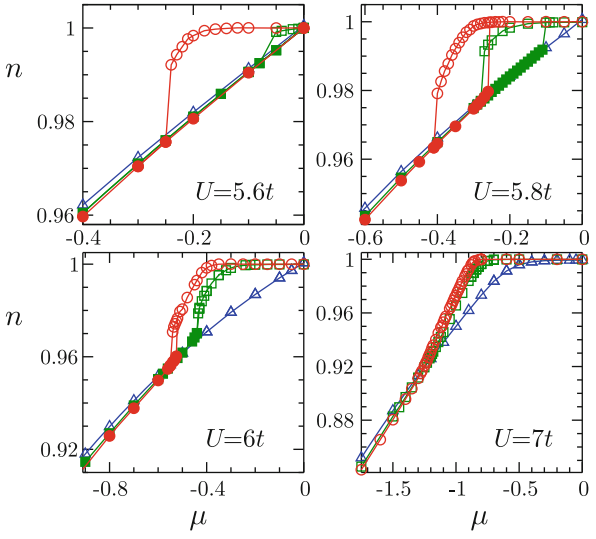

Fig. 11.9 Occupation n as a function of chemical potential μ for the square-lattice Hubbard model, obtained via CDMFT with a CT-QMC solver, from [32]. Several values of U close to the Mott transition are shown, at different temperatures $T = t/10$, $t/25$ and $t/50$ (*triangles*, *squares*, and *circles*, respectively). Two metallic solutions sometimes coexist, and are obtained by sweeping μ in one direction or the other

11.5 Application to the Cuprates

Many of the theoretical approaches described in this volume were motivated by the ambition to explain the origin of high-temperature superconductivity; more precisely, to answer the following question: does the Hubbard model contain the key elements to explain superconductivity in the cuprates? Single-site DMFT alone cannot answer this question, as it lacks the short-range correlation effects needed even to describe d-wave superconductivity. But the question has been addressed by quantum cluster methods: DCA (see Chap. 8.8), the variational cluster approximation (see Chap. 9.5) and CDMFT, in addition to many other methods described in this volume and its sequel.

When probing d-wave superconductivity in CDMFT, the hybridization function (11.9) must contain anomalous terms of the proper symmetry. Within the ED solver, this was at first accomplished as illustrated in Fig. 11.4, i.e., by adding to the bath Hamiltonian d-wave pairing terms between bath orbitals themselves (the red dotted lines of Fig. 11.4), as described in (11.30) and (11.31). In this parametrization, the bath is seemingly made of two "ghost clusters" whose pairing terms mimick the broken symmetry state that could take place on the cluster itself. Note that it is customary in CDMFT not to tamper with the cluster Hamiltonian when probing broken symmetries, contrary to what is done in the variational cluster approximation; the agents of symmetry breaking are concentrated in the bath. The bath Hamiltonian being no longer diagonal in bath orbitals with the addition (11.31), a numerical diagonalization is performed to work with new, diagonal orbitals which restore the usual form (11.9) of the hybridization function. The latter will then have anomalous terms of the correct pairing symmetry and spin.

Any study of d-wave superconductivity within the Hubbard model must also taken into account the possibility of antiferromagnetic order in competition, or in coexistence with, superconductivity. This requires a more general bath parametrization, such as the one given in (11.29), also illustrated in Fig. 11.4. This was done in [33] in the special case of the square-lattice Hubbard model (i.e., with NN hopping only). As seen from Fig. 11.10, not only were solutions with d-wave superconductivity found, but they are rather robust. When antiferromagnetism is not allowed, superconductivity exists all the way to half-filling when U is small enough (below the Mott transition), but is suppressed even at $U = 4t$ when in competition with antiferromagnetism (not shown here).

Kancharla et al. [34] carried a similar analysis on the more realistic one-band model for the cuprates that includes second-neighbour and third-neighbour hopping terms (t' and t''). Figure 11.11 shows that the superconducting order parameter scales like $J = 4t^2/U$, at least in the under-doped region. The right panel of the same figure shows how t' breaks the electron–hole symmetry of the phase diagram and how superconductivity and antiferromagnetism can coexist in a homogeneous solution. Such a mixed state is also found in the variational cluster approximation [35, 36].

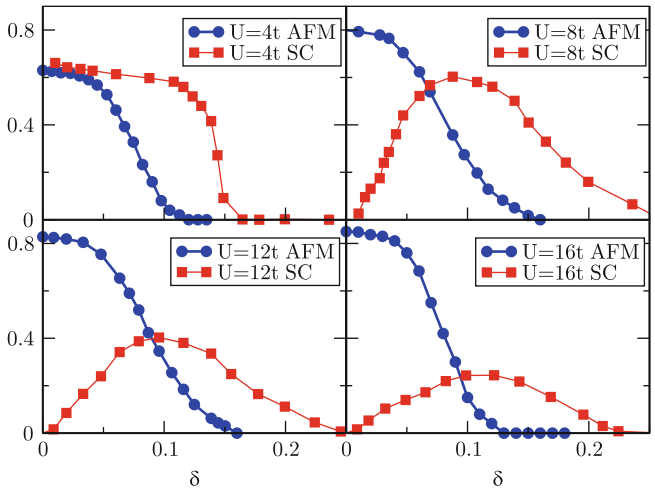

Fig. 11.10 Antiferromagnetic (*circles*) and d-wave (*squares*) order parameters as a function of doping in the square-lattice Hubbard model, from [33]. The two orders were not allowed to compete. The d-wave data is scaled by 10 for clarity

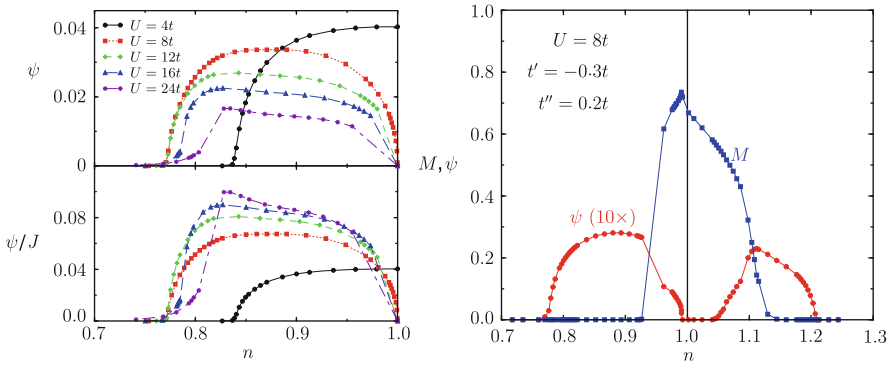

Fig. 11.11 *Left panel, top*: d-wave order parameter ψ as a function of electron density n, for various values of U and $t' = -0.3t$, $t'' = 0.2t$. *Bottom*: the same, scaled by $J = 4t^2/U$. *Right panel*: d-wave (ψ) and antiferromagnetic (M) order parameters vs n, from a common solution where they are allowed to compete. Taken from [34]

The CT-QMC solver was applied to the superconductivity problem in [12], again in the special case $t' = 0$. The d-wave order parameter as a function of doping is illustrated in Fig. 11.12. The CT-QMC allows for an estimate of T_c, but such a computation is very resource-intensive because of critical slowing down. Of course, this T_c has a mean-field in character: The Mermin-Wagner theorem forbids the spontaneous breakdown of continuous symmetries in a purely two-dimensional system at nonzero temperature. In a finite cluster, the long wavelength

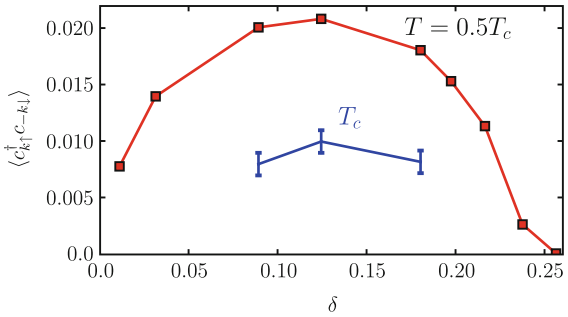

Fig. 11.12 d-wave order parameter vs doping δ from a CT-QMC [12] in the NN Hubbard model. The critical temperature T_c is shown at three values of δ, and the order parameter is shown at a temperature T equal to half the maximum T_c

pair fluctuations that would destabilize a superconducting phase at finite temperature are not at work and cannot be accounted for by a fermionic bath.

A fair criticism of the above results on broken symmetries in the Hubbard model is the lack of finite-size analysis. The existence of broken symmetry phases such as antiferromagnetism and d-wave superconductivity can only be established firmly in the limit of infinite cluster size; an infinite bath-size is not sufficient.[1] In principle infinite-size extrapolations should be performed in order to assess the robustness of CDMFT predictions in the thermodynamic limit. In practice, this requires vast amounts of computing resources, and a solver that can accommodate larger clusters, such as the Hirsch-Fye QMC or the auxiliary-field CT-QMC [13]. This was accomplished within the DCA in [37], with cluster sizes ranging up to $L = 32$ and special attention paid to the cluster shape in relation with the periodic boundary conditions used in DCA. The existence of a Kosterlitz–Thouless transition was confirmed in the square-lattice Hubbard model at $T_c \approx 0.023t$ for $U = 4t$ and $\delta \approx 10\%$. The cost of these computations precluded a wider exploration of parameter space. The assumption underlying current work on superconductivity using cluster approaches is that the thermodynamic limit will bring important renormalizations, but will not qualitatively affect the dependence of the superconducting order upon band parameters, interaction strength or doping. Thus, cluster approaches are important tools in exploring the space of models that can potentially lead to superconductivity or other broken symmetry phases.

To conclude this section, let us briefly mention a few applications of CDMFT to other problems. An interesting issue in systems with attractive interactions is the crossover from a BCS state (weak pairing) to Bose–Einstein Condensation (strong pairing). This was studied in [38] in the context of the attractive Hubbard model. It was observed that for $|U| \lesssim 8t$ the pairing is driven by the potential energy, whereas

[1] We have seen, for instance, that the exact value of the critical chemical potential μ_{gap} in the one-dimensional Hubbard model cannot be recovered with an infinite bath [9].

it is driven by the kinetic energy at strong coupling. By this we mean that the paired state exists because of a gain in potential energy in the first case, and a gain in kinetic energy in the second case. This is relevant to the physics of the cuprates, even though the attractive Hubbard model is used here as an effective, rather than microscopic, model.

Heavy-fermion materials form another large and fascinating class of correlated materials. The hybridization of localized f electrons with conduction bands leads to the formation of quasiparticles that are the source of the Kondo effect. It is expected that there is a critical value of this hybridization below which the f electrons order magnetically, thus suppressing the Kondo effect; this bears the name of *Kondo breakdown*, and was studied with CDMFT in [39] using the periodic Anderson model. Again, the CDMFT results differ from previous single-site DMFT results by a less radical shift of spectral weight across the transition that can be attributed to short-range correlations.

11.6 CDMFT and DCA

CDMFT and the DCA are two cluster extensions of single-site DMFT. Both approaches have revealed antiferromagnetic order and d-wave pairing in the Hubbard model [1, 33, 34, 37, 40–42], and DCA computations have been performed on large enough clusters to show that d-wave superconductivity persists in the thermodynamic limit [37]. We will comment on the differences between the two approaches here, without providing a detailed introduction to DCA, which can be found in Chap. 9 of this volume.

DCA is formulated in reciprocal space, using clusters with periodic boundary conditions, whose self-energies $\Sigma(\boldsymbol{K}, \omega)$ are diagonal in cluster momentum $\boldsymbol{K}$. The impurity solver respects translation invariance, and therefore the hybridization function $\Gamma_{\boldsymbol{K}}$ is also diagonal in cluster momentum. The DCA self-consistency condition reads

$$\frac{1}{i\omega_n - \bar{t}_{\boldsymbol{K}} - \Gamma_{\boldsymbol{K}}(\omega) - \Sigma_{\boldsymbol{K}}(\omega)} = \frac{L}{N} \sum_{\tilde{\boldsymbol{k}}} \frac{1}{i\omega_n - \varepsilon(\tilde{\boldsymbol{k}} + \boldsymbol{K}) - \Sigma_{\boldsymbol{K}}(\omega)}, \tag{11.53}$$

where $\varepsilon(\boldsymbol{k})$ is the lattice dispersion relation and $\bar{t}_{\boldsymbol{K}}$ is a coarse-grained dispersion defined as

$$\bar{t}_{\boldsymbol{K}} = \frac{L}{N} \sum_{\tilde{\boldsymbol{k}}} \varepsilon(\tilde{\boldsymbol{k}} + \boldsymbol{K}) \tag{11.54}$$

that is, the average of the dispersion relation in a reduced Brillouin zone centered around each cluster momentum $\boldsymbol{K}$. Once a solution is found, the lattice Green function is naturally

$$G(\boldsymbol{k},\omega) = \frac{1}{i\omega_n - \varepsilon(\boldsymbol{k}) - \Sigma(\boldsymbol{K}, i\omega_n)}, \tag{11.55}$$

where $\boldsymbol{k} = \tilde{\boldsymbol{k}} + \boldsymbol{K}$. This is the DCA equivalent of the CPT relation (11.20). Note that we have not used a matrix notation for G, t or Σ, as all the above quantities are all single numbers; when dealing with a multi-band model or a situation where the two spin projections are hybridized (e.g., superconductivity), then a matrix formulation is warranted. Note that the Green function (11.55) is discontinuous as a function of $\boldsymbol{k}$ because of the discrete patches $\boldsymbol{K}$ over the Brillouin zone; this can be fixed by an interpolation of the self-energy $\Sigma(\boldsymbol{K}, i\omega_n)$ between its discrete values; this accomplishes the equivalent of periodization in CDMFT or CPT (see Chap. 7.9).

It is important to realize that DCA is *not* equivalent to CDMFT with periodic clusters. Let us assume, for instance, that a periodic cluster is used in CDMFT, with a bath structure that respects translation invariance on the cluster and thus conserves cluster momentum. The CDMFT self-consistency condition is, explicitly,

$$\frac{1}{i\omega_n - \boldsymbol{t}' - \boldsymbol{\Gamma}(i\omega_n) - \boldsymbol{\Sigma}(i\omega_n)} = \frac{L}{N}\sum_{\tilde{\boldsymbol{k}}} \frac{1}{i\omega_n - \boldsymbol{t}(\tilde{\boldsymbol{k}}) - \boldsymbol{\Sigma}(i\omega_n)}. \tag{11.56}$$

This is a matrix relation that can be expressed in any basis of one-particle states.

To understand the difference between conditions (11.53) and (11.56), let us introduce the unitary matrix $\boldsymbol{U}$ that brings us from the lattice site basis $(\boldsymbol{r})$ to the momentum basis $(\boldsymbol{k})$, so that Fourier transforms on a function $f_{\boldsymbol{r}}$ take the form

$$\tilde{f}_{\boldsymbol{k}} = \sum_{\boldsymbol{r}} U_{\boldsymbol{k}\boldsymbol{r}} f_{\boldsymbol{r}} \qquad U_{\boldsymbol{k}\boldsymbol{r}} = \frac{1}{\sqrt{N}} \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{11.57}$$

Let us introduce likewise the matrix $\boldsymbol{V}$ that makes the same transition, this time between cluster positions $\tilde{\boldsymbol{r}}$ and reduced momentum $\tilde{\boldsymbol{k}}$, and the matrix $\boldsymbol{W}$ that does the same between cluster sites $\boldsymbol{r}_a$ and cluster momenta $\boldsymbol{K}$:

$$V_{\tilde{\boldsymbol{k}}\tilde{\boldsymbol{r}}} = \sqrt{\frac{L}{N}} \mathrm{e}^{-i\tilde{\boldsymbol{k}}\cdot\tilde{\boldsymbol{r}}} \qquad W_{\boldsymbol{K},a} = \frac{1}{\sqrt{L}} \mathrm{e}^{-i\boldsymbol{K}\cdot\boldsymbol{r}_a} \tag{11.58}$$

(these matrices were also introduced in Chap. 9.5). The matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ commute, since they do not act on the same spaces, and their product $\boldsymbol{S} = \boldsymbol{W} \otimes \boldsymbol{V}$, even if it performs a compound Fourier transform from direct space $(\tilde{\boldsymbol{r}}, \boldsymbol{r}_a)$ to reciprocal space $(\tilde{\boldsymbol{k}}, \boldsymbol{K})$, does not coincide with the direct Fourier transform $\boldsymbol{U}$ with $\boldsymbol{k} = \tilde{\boldsymbol{k}} + \boldsymbol{K}$. It is a straightforward matter to show that the matrices $\boldsymbol{\Lambda} \equiv \boldsymbol{U}\boldsymbol{S}^{-1}$ and $\boldsymbol{D} \equiv \boldsymbol{S}^{-1}\boldsymbol{U}$ are not trivial:

$$\Lambda_{\boldsymbol{k}\boldsymbol{k}'} = \delta_{\tilde{\boldsymbol{k}}\tilde{\boldsymbol{k}}'} \frac{1}{L}\sum_{a} \mathrm{e}^{-i\boldsymbol{r}_a\cdot(\tilde{\boldsymbol{k}}+\boldsymbol{K}-\boldsymbol{K}')} \qquad D_{\boldsymbol{r}\boldsymbol{r}'} = \delta_{ab} \frac{L}{N}\sum_{\tilde{\boldsymbol{k}}} \mathrm{e}^{i\tilde{\boldsymbol{k}}\cdot(\tilde{\boldsymbol{r}}-\tilde{\boldsymbol{r}}'-\boldsymbol{r}_a)}. \tag{11.59}$$

If $\boldsymbol{t}$ is the real-space hopping matrix on the lattice, then the dispersion relation is given by the diagonal matrix $\boldsymbol{\varepsilon} = \boldsymbol{U}\boldsymbol{t}\boldsymbol{U}^\dagger$ ($\boldsymbol{k}$ being the index along the diagonal). However, the matrix appearing in the CPT Green function on the r.h.s. of the CDMFT self-consistency condition (11.56) is $\boldsymbol{t}(\tilde{\boldsymbol{k}}) = \boldsymbol{S}\boldsymbol{t}\boldsymbol{S}^\dagger$ and is not diagonal. For instance, in the case of a two-site cluster within the one-dimensional Hubbard model with nearest-neighbor hopping only, that matrix is

$$\boldsymbol{S}\boldsymbol{t}\boldsymbol{S}^\dagger = -t \begin{pmatrix} 1+\cos(2\tilde{k}) & -i\sin(2\tilde{k}) \\ i\sin(2\tilde{k}) & -1-\cos(2\tilde{k}) \end{pmatrix}. \tag{11.60}$$

The DCA condition (11.53) would coincide with the CDMFT condition (11.56) for a periodic cluster if the lattice hopping matrix, instead of $\boldsymbol{t}$, were $\bar{\boldsymbol{t}} = \boldsymbol{D}\boldsymbol{t}\boldsymbol{D}^\dagger$. Chopping off the $\tilde{\boldsymbol{r}} \neq \tilde{\boldsymbol{r}}'$ terms in that matrix yields the cluster restriction

$$\bar{t}_{ab} = \frac{1}{N}\sum_{\boldsymbol{K}}\sum_{\tilde{\boldsymbol{k}}} \varepsilon(\tilde{\boldsymbol{k}}+\boldsymbol{K})\mathrm{e}^{i\boldsymbol{K}\cdot(\boldsymbol{r}_a-\boldsymbol{r}_b)}, \tag{11.61}$$

which is precisely the coarse-grained dispersion (11.54), but expressed in real-space indices.

This means that the DCA, contrary to the CDMFT, cannot be understood within the SFA (see Chap. 9.5) from the same lattice hopping $\boldsymbol{t}$. Otherwise the two conditions (11.53) and (11.56) would coincide for periodic clusters. This difference between the DCA and a periodic version of CDMFT disappears as the cluster size L increases: it can be shown that the mean squared off-diagonal matrix elements of $\boldsymbol{\Lambda}$ and $\boldsymbol{D}$ scale like $1/L$.

Let us now comment on the difference between the DCA and the usual implementation of CDMFT, i.e., with open boundary conditions. Here again the difference between the two approaches obviously disappear in the large cluster limit, but benchmarks conducted on systems whose exact solution is known have shown that the two approaches do not converge equally fast as $L \to \infty$, for the same quantities. In [43–45] (the last ones being comments on the previous ones) the CDMFT and DCA were applied to a simplified one-dimensional large-N model whose solution is known in the $N \to \infty$ limit. The overall conclusion of these works is that, as $L^{-1} \to 0$, DCA converges faster than CDMFT for quantities that are averages over the cluster, whereas the opposite is true for local quantities. This is due to the use of periodic boundary conditions and the absence of cluster edges in DCA. On the other hand, local quantities measured near the cluster center in CDMFT converge exponentially fast, whereas the corresponding local quantities converges like L^{-2} in DCA. More recently, the authors of [46] have done a similar analysis for the two-dimensional, saddle-point t–J model. Again, this being a static (mean-field) model, the comparative study only addresses the geometrical aspects of the two approaches. The authors conclude that (1) the four-site, plaquette cluster is a poor approximation of the thermodynamic limit for both approaches; (2) the

DCA generally converges faster for global quantities such as order parameters, but may be more erratic than CDMFT for small clusters.

In practice, the issue of the convergence rate of DCA and CDMFT is moot for small clusters like the four-site plaquette that are far enough from the thermodynamic limit: no approach is demonstrably better than the other, for any quantity. As of this writing, large-cluster simulations of the Hubbard model (necessarily with a QMC solver) are still costly and rarely performed [37, 47, 48].

Acknowledgments Discussions with A-M. Tremblay and P. Sémon are gratefully acknowledged. Computational resources for this review were provided by RQCHP and Compute Canada.

References

1. A.I. Lichtenstein, M.I. Katsnelson, Phys. Rev. B **62**(14), R9283 (2000)
2. G. Kotliar, S. Savrasov, G. Pálsson, G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
3. M. Caffarel, W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
4. E. Koch, G. Sangiovanni, O. Gunnarsson, Phys. Rev. B **78**(11), 115102 (2008)
5. A. Liebsch, N.H. Tong, Phys. Rev. B **80**(16), 165126 (2009)
6. D. Sénéchal, Phys. Rev. B **81**(23), 235125 (2010)
7. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**(25), 1445 (1968)
8. M. Capone, M. Civelli, S.S. Kancharla, C. Castellani, G. Kotliar, Phys. Rev. B **69**(19), 195105 (2004)
9. B. Kyung, G. Kotliar, A.M.S. Tremblay, Phys. Rev. B **73**(20), 205106 (2006)
10. J.E. Hirsch, R.M. Fye, Phys. Rev. Lett. **56**(23), 2521 (1986)
11. P. Werner, A. Comanac, L. de Medici, M. Troyer, A.J. Millis, Phys. Rev. Lett. **97**(7), 076405 (2006)
12. K. Haule, G. Kotliar, Phys. Rev. B **76**(10), 104509 (2007)
13. E. Gull, P. Werner, O. Parcollet, M. Troyer, Europhys. Lett. **82**(5), 57003 (2008)
14. E. Gull, P. Werner, A. Millis, M. Troyer, Phys. Rev. B **76**(23), 235123 (2007)
15. R. Blankenbecler, D.J. Scalapino, R.L. Sugar, Phys. Rev. D **24**(8), 2278 (1981)
16. R.M. Fye, Phys. Rev. B **33**(9), 6271 (1986)
17. B.B. Beard, U.J. Wiese, Phys. Rev. Lett. **77**(25), 5130 (1996)
18. N.V. Prokof'ev, B.V. Svistunov, I.S. Tupitsyn, JETP Letters **64**, 911 (1996)
19. A.N. Rubtsov, V.V. Savkin, A.I. Lichtenstein, Phys. Rev. B **72**(3), 035122 (2005)
20. P. Werner, A.J. Millis, Phys. Rev. B **74**(15), 155107 (2006)
21. K. Haule, Phys. Rev. B **75**(15), 155113 (2007)
22. A. Georges, G. Kotliar, W. Krauth, M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
23. H. Park, K. Haule, G. Kotliar, Phys. Rev. Lett. **101**(18), 186403 (2008)
24. M. Balzer, B. Kyung, D. Sénéchal, A.M.S. Tremblay, M. Potthoff, Europhys. Lett. **85**, 17002 (2009)
25. Y.Z. Zhang, M. Imada, Phys. Rev. B **76**(4), 045108 (2007)
26. B. Kyung, S. Kancharla, D. Sénéchal, A.M.S. Tremblay, M. Civelli, G. Kotliar, Phys. Rev. B **73**, 165114 (2006)
27. B. Kyung, A.M.S. Tremblay, Phys. Rev. Lett. **97**(4), 046402 (2006)
28. O. Parcollet, G. Biroli, G. Kotliar, Phys. Rev. Lett. **92**(22), 226402 (2004)
29. T. Ohashi, T. Momoi, H. Tsunetsugu, N. Kawakami, Phys. Rev. Lett. **100**(7), 076402 (2008)
30. P. Sahebsara, D. Sénéchal, Phys. Rev. Lett. **100**, 136402 (2008)
31. T. Ohashi, S.I. Suga, N. Kawakami, H. Tsunetsugu, J. Phys. Cond. Matt. **19**(14), 145251 (2007)
32. G. Sordi, K. Haule, A.M.S. Tremblay, Phys. Rev. Lett. **104**(22), 226402 (2010)

33. M. Capone, G. Kotliar, Phys. Rev. B **74**(5), 054513 (2006)
34. S.S. Kancharla, M. Civelli, M. Capone, B. Kyung, D. Sénéchal, G. Kotliar, A.M. Tremblay, Phys. Rev. B **77**, 184516 (2008)
35. D. Sénéchal, P.L. Lavertu, M.A. Marois, A.M.S. Tremblay, Phys. Rev. Lett. **94**, 156404 (2005)
36. M. Aichhorn, E. Arrigoni, M. Potthoff, W. Hanke, Phys. Rev. B **74**, 024508 (2006)
37. T.A. Maier, M. Jarrell, T.C. Schulthess, P.R.C. Kent, J.B. White, Phys. Rev. Lett. **95**(23), 237001 (2005)
38. B. Kyung, A. Georges, A.M.S. Tremblay, Phys. Rev. B **74**(2), 024501 (2006)
39. L. De Leo, M. Civelli, G. Kotliar, Phys. Rev. B **77**(7), 075107 (2008)
40. T.A. Maier, M. Jarrell, A. Macridin, C. Slezak, Phys. Rev. Lett. **92**, 027005 (2004)
41. P. Kent, M. Jarrell, T. Maier, T. Pruschke, Phys. Rev. B **72**, 60411 (2005)
42. M. Civelli, Phys. Rev. B **79**(19), 195113 (2009)
43. G. Biroli, G. Kotliar, Phys. Rev. B **65**, 155112 (2002)
44. K. Aryanpour, T. Maier, M. Jarrell, Phys. Rev. B **71**(3), 037101 (2005)
45. G. Biroli, G. Kotliar, Phys. Rev. B **71**, 037102 (2005)
46. A. Isidori, M. Capone, Phys. Rev. B **79**(11), 115138 (2009)
47. S. Biermann, A. Georges, A. Lichtenstein, T. Giamarchi, Phys. Rev. Lett. **87**(27), 276405 (2001)
48. G. Alvarez, M. Summers, D. Maxwell, M. Eisenbach, J. Meredith, J. Larkin, J. Levesque, T. Maier, P. Kent, E. D'Azevedo, T. Schulthess, in *High Performance Computing, Networking, Storage and Analysis, 2008. International Conference for SC 2008* (2008), pp. 1 –10

Chapter 12
Functional Renormalization Group for Interacting Many-Fermion Systems on Two-Dimensional Lattices

Carsten Honerkamp

Abstract Functional renormalization group (fRG) methods have become a widely used tool to investigate the low-energy properties of interacting fermions. In particular for two-dimensional lattice systems with competing ordering tendencies, they are very suitable for investigating the interplay of various possible instabilities and their mechanisms. Here we review the main elements of the fRG method, its current state and some applications to correlated electron systems where the fRG has proved useful, and outline some issues for the further development of the method.

12.1 Introduction

The low-temperature phase diagrams of modern correlated electron materials like layered copper oxides [1, 2], heavy fermions [3, 4], or iron pncitides [5–7] exhibit various competing ground states with different types of collective behavior. In most cases, these ground states can be classified by the means of order parameters and collective excitations, but there is also strong interest in the states of electron matter without long-range order, such as spin liquid [8–11] or fractionalized states [12, 13], or with unconventional excitation spectra, such as topological insulators [14]. For theory, this poses the challenge of predicting the ground state or low energy properties of a given model Hamiltonian on some lattice whose form and parameters are related to the real materials by direct comparison with experiments or with the help of ab-initio calculations.

C. Honerkamp
Institute for Theoretical Solid State Physics, RWTH Aachen, 52056 Aachen, Germany

JARA-FIT, Fundamentals of Future Information Technology, Germany
e-mail: honerkamp@physik.rwth-aachen.de

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_12,

As many of the materials mentioned show a certain degree of electron itineracy, a theoretical approach that is perturbative in the screened electron–electron interactions can be considered as a useful guide to describe competing orders. Straightforward diagrammatic perturbation theory in lower orders is often not a satisfactory description, as some diagrams in the perturbation series produce large, possible logarithmically divergent contributions at low temperatures and energies. An infinite summation of these diagrams would be preferable. However, in many situations one encounters more than one important channel in the perturbation expansion, making the summation difficult. Moreover, even if there is a clear hierarchy in the importance of the various contributions, sub-leading terms might change the effective coupling in the leading channel qualitatively, causing a single-channel-summation to be physically misleading. Here the renormalization group (RG) proves useful, as it sums up the perturbation series to infinite order in the bare interactions in the process of integrating up the RG differential equations. Furthermore in the RG the most dangerous degrees of freedom are included on equal footing and step by step. In this way possible singularities are approached in a transparent and more controlled and unbiased way. In this chapter we review how powerful functional RG (fRG) methods [15–18] have been used recently to explore possible instabilities of a number of systems. Most of the work has been performed using various simplifications, but already provides a wealth of physical information. We also mention some improvements of these approximate schemes have been tried and how these affect the physical picture.

The fRG is a broader concept rooting in the works of Wilson, Wegner, and others in the 1970s [19, 20]. In the context of what is described in this chapter, there are at least two aspects that are inherited from these seminal works. One is the idea of coarse graining to effective theories, i.e., the decimation of the degrees of freedom, and the second is the field theoretical formalism in momentum space with exact flow equations for a generating functional [20]. In other aspects, the use of the fRG for correlated electrons described here is somewhat different. In statistical physics, Wilsonian RG techniques are mainly understood as concept to explore the vicinity of critical points. These are given by fixed-points of the RG procedure. Universal behavior arises naturally and is in the center of interest. Functional RG equations, like the one for the effective average action derived by Wetterich [16] that forms the basis for the formalism described in this chapter, also prove to be very useful in the study of critical phenomena. For example, using the Wetterich equation one can compute the critical exponents of O(N) models in three dimensions with good precision without having to perform an ϵ-expansion or resummations [21, 22]. However, for the correlated electron lattice systems described in this chapter, the scope somewhat different and nonuniversal properties play a larger role. In many problems, in particular those with competing orders, the main issue is the proper identification and characterization of the low-temperature state of a system that does not have to be critical, and the mechanisms responsible for it. This set of questions includes nonuniversal properties such as energy scales, gap structures around Fermi surfaces, low- and intermediate-energy excitations, etc. Therefore, the primary goal of most of the studies described here is to determine to which effective theory the

flow goes at intermediate scales on the way down to lower energies. This question is often decided by nonuniversal features of the system like lattice type and channel coupling effects that can be irrelevant in the RG sense. Furthermore, the search for and description of fixed points is a priori somewhat less meaningful at least in the current approximations, as most interesting flows lead to strong coupling, i.e., a divergence of the interaction function occurs, and the approximations made in the scheme become invalid. The fRG approach discussed below is best understood as performing a partial trace over high-energy degrees of freedom that allows one to track the build-up of possibly long-ranged effective interactions in an unbiased way to infinite order in the bare interactions. This already provides very useful information on the potential ground state that can be compared with experiments and other theoretical approaches. More abstractly and generally, the main idea of the fRG is to derive (initially exact) flow equations for generating functionals of the theory that take one from a starting point where the correlation or vertex functions are known very well to the point of physical interest. In standard momentum-shell fRG schemes, the starting point is obtained by excluding all quantum fluctuations by a cutoff. Then the initial correlation functions can be simply obtained from the bare action. Lowering the cutoff includes quantum fluctuations step by step toward the full theory. Obviously, for many cases, one expects some freedom in both the starting point of this procedure and the flow trajectory as well. Below we will mention how this freedom in the fRG idea has been used to obtain alternative views on many-fermion lattice systems. We expect that future work will make even more use of this great flexibility. While these considerations reflect the main part of what is described in this chapter, we want to mention activities underway to extend the fRG approach also to the more precise study of (quantum) critical points in correlated electron systems, in particular to such problems where the conventional Hertz–Millis [23–25] approach becomes questionable [26].

Various works using RG methods for fermions have been inspired an earlier review article by Shankar [27]. Here, in analogy to the RG for scalar ϕ^4-models, some important examples like the flows in one dimension, the Kohn–Luttinger mechanism, and phase space arguments for the validity of Fermi liquid theory are presented in physical terms. The fRG methods described here go beyond these arguments and may represent a more general perspective and some methodical refinements that allow one to successively reduce the approximations and to address in a straightforward manner additional issues such as flows to into symmetry-broken regimes [28], boson–fermion mixtures [29–31] and in future possibly strongly coupled systems [32, 33]. We also note that fRG techniques can be used to obtain mathematically precise arguments for the validity range of Fermi liquids in two dimensions [17].

This chapter concentrates on the use of the fRG as a tool detect instabilities in weakly to moderately coupled many-fermion systems on two-dimensional (2D) lattices. The same RG equations have, however, been applied intensively in other fields in condensed matter and ultracold atom physics and currently there is an active interchange of ideas between these fields. As examples we mention a series of works on transport through nanoscale low-dimensional correlated electron systems [34,35]

and also on the BCS-BEC crossover [36] and bound-state formation [37] in ultracold atoms. Recently, the fRG formalism has been extended to the study of quantum spin systems [38]. Furthermore, nonequilibrium situations can also be tackled using fRG techniques [39–41].

12.2 Functional RG Schemes for Fermions: Exact Flow Equations and Truncations

12.2.1 Basic Elements

In the functional integral formalism for many-particle systems at nonzero temperature T, the fermionic degrees of freedom are described by Grassmann fields, $\psi(k,s)$ and $\bar{\psi}(k,s)$, where k is a multi-index containing fermionic Matsubara frequencies $\omega_n = (2n+1)\pi T$, wavevectors $\boldsymbol{k}$ in the respective Brillouin zone and possibly more, like band indices, etc. s is the spin degree of freedom, in the usual framework $s = \pm 1/2$.

Let the free part of the action of our many-fermion system be of the type

$$S_0 = \sum_{k,s} \bar{\psi}(k,s)\left[-i\,\omega_n + \epsilon(\boldsymbol{k})\right]\psi(k,s)\,. \tag{12.1}$$

Here the sum over k is understood as sum over everything contained in the multi-index. $\epsilon(\boldsymbol{k})$ is the band dispersion. We suppress possible band indices, but the extension to the multiband case is straightforward.

For spin-rotational invariance, the fermionic interaction can be written as

$$S_{\mathrm{I}} = \frac{T}{2N}\sum_{\substack{k_1,k_2,k_3\\ s,s'}} V(k_1,k_2,k_3)\bar{\psi}(k_1,s)\bar{\psi}(k_2,s')\psi(k_4,s')\psi(k_3,s) \tag{12.2}$$

N is the number of lattice sites. The fourth wavevector and the fourth Matsubara frequency in k_4 is usually fixed by the appropriate wavevector- and frequency-conservation and not summed over, but in the general multiband case, one has to sum over a fourth band index as well. For Hubbard onsite interactions in a one-band model, the coupling function is just $V(k_1,k_2,k_3) = U$, i.e., structureless in its wavevector dependence. In the multiband case this changes, as explained below for the pnictides.

The free propagator is the negative inverse of the quadratic part in S_0,

$$G_0^{(2)}(k,s) = -\langle\psi(k,s)\bar{\psi}(k,s)\rangle_0 = \frac{1}{i\,\omega_k - \epsilon(\boldsymbol{k})} \tag{12.3}$$

With the free propagator the generating functional of the theory is written as functional integral over the Grassmann fields

$$W(\eta,\bar{\eta}) = -\ln \int D\psi(k,s) D\bar{\psi}(k,s)\, \mathrm{e}^{-S(\psi,\bar{\psi},\eta,\bar{\eta})} \tag{12.4}$$

with

$$\begin{aligned} S(\psi,\bar{\psi},\eta,\bar{\eta}) = &-\sum_{k,s} \bar{\psi}(k,s) \left[G_0^{(2)}(k,s)\right]^{-1} \psi(k,s) + S_{\mathrm{I}} \\ &+ \sum_{k,s} \left[\bar{\eta}(k,s)\psi(k,s) + \bar{\psi}(k,s)\eta(k,s)\right]. \end{aligned} \tag{12.5}$$

The integral without the logarithm and the sources η, $\bar{\eta}$ set to zero is the partition function of the system. By taking derivatives with respect to the source fields $\eta(k,s)$ and $\bar{\eta}(k,s)$, one obtains the connected Green's functions of the theory, most importantly

$$G^{(2)}(k,s) = -\langle \psi(k,s)\bar{\psi}(k,s)\rangle = \frac{\delta^2}{\delta\eta(k,s)\delta\bar{\eta}(k,s)} W(\eta,\bar{\eta}) \tag{12.6}$$

and

$$\begin{aligned} G^{(4)}(k_1,s;k_2,s';k_3,s;k_4,s') &= -\langle \psi(k_1,s)\psi(k_2,s')\bar{\psi}(k_4,s')\bar{\psi}(k_3,s)\rangle \\ &= \frac{\delta^4}{\delta\eta(k_3,s)\delta\eta(k_4,s')\delta\bar{\eta}(k_2,s')\delta\bar{\eta}(k_1,s)} W(\eta,\bar{\eta}) \end{aligned} \tag{12.7}$$

It is convenient to introduce a Nambu notation where barred and nonbarred Grassmann fields are grouped in 2-vectors. With $\Psi = (\bar{\psi}, \psi)$ we write

$$\begin{aligned} S(\Psi,H) = &\frac{1}{2}\sum_{k,s} \Psi(k,s)\hat{Q}(k,s)\Psi(k,s) + S_{\mathrm{I}} \\ &+ \sum_{k,s} H(k,s)\Psi(k,s) \end{aligned} \tag{12.8}$$

with the two-by-two matrix

$$\hat{Q}(k,s) = \begin{pmatrix} 0 & Q(k,s) \\ -Q(k,s) & 0 \end{pmatrix}, \tag{12.9}$$

inverse free propagator $Q(k,s) = \left[G_0^{(2)}(k,s)\right]^{-1}$ and sources $H = (-\eta,\bar{\eta})$.

12.2.2 Functional Renormalization Group Differential Equations

To implement the renormalization group idea of deriving the an effective action for the low-energy modes, i.e., to integrate first over high-energy modes, it is now appropriate to split the propagator into high- and low-energy parts. This is done by writing

$$G_0^{(2)}(k,s) = G_0^{>,(2)}(k,s) + G_0^{<,(2)}(k,s) = \frac{C_\Lambda(k)}{i\omega_n - \epsilon(\boldsymbol{k})} + \frac{1 - C_\Lambda(k)}{i\omega_n - \epsilon(\boldsymbol{k})} \quad (12.10)$$

with $C_\Lambda(k) = 1$ if k describes a high-energy mode, and $C_\Lambda(k) = 0$ if k belongs to the low-energy modes, and $\Lambda > 0$ is an energy scale discriminating between high and low. Targeting situations with only weak to moderate interactions, it plausible to discriminate modes with respect to the bare dispersion, i.e., to write $C_\Lambda(k = \{\omega_n, \boldsymbol{k}\}) = \Theta\left(|\epsilon(\boldsymbol{k})| - \Lambda\right)$ with $\Theta(x) = 0$ for $x < 0$ and $\Theta(x) = 1$ for $x > 0$. We call this a sharp cutoff in wavevector space, but as noted above the concept of effective action be played through with other cutoffs as well, like smoother cutoffs, or involving the Matsubara frequencies.

Now let us include only the high-energy modes into the theory, i.e., let the bare propagator in the free part of the action be given by $G_0^{>,(2)}(k,s) = G_{\Lambda,0}^{(2)}(k,s) = Q_\Lambda^{-1}(k,s)$. Introducing the RG scale or flow parameter Λ into the free part of the action promotes the generating functional $W(H)$ to $W_\Lambda(H)$. For nonzero Λ, we could now compute observables of the fermionic theory in a diagrammatic approach. As the low-energy modes in the propagator $G_0^{(2)}$ are set to zero, only high-energy modes would appear in the diagrams. Studying these quantities as a function of Λ often gives more physical insights, as the full functional integral over all modes (i.e., for $\Lambda = 0$) typically develops singular contributions at low T. More generally, the strategy for the cutoff is to deform the free propagator in such a way that (a) the initial correlation functions are known, and e.g., correspond to the bare ones if all internal lines are set to zero by a high enough cutoff, and (b) the relevant observables become free of singularities. Then one continues to undo the deformation continuously such that the build-up and competition of possible singularities can be studied. This gives information on the leading instabilities of the system at low scales.

Furthermore, the connected correlation functions computed only with the modes above Λ appear in the effective action of the modes below Λ (amputated with the free propagator $G_0^{>,(2)}(k,s)$) [42, 43]. The self-energy for a wave-vector $\boldsymbol{k}$ in the low-energy sector computed at scale Λ can be interpreted as that arising by absorbing the modes above Λ, and the interaction computed with the modes above Λ on the internal lines renormalizes the interaction of the modes below Λ, leading to a Λ-dependent effective interaction.

One possibility would now be to compute the change of the connected correlation functions as a function of Λ directly from the change of the generating functional $W_\Lambda(H)$ with Λ. Formally, one can write down an exact (RG) differential equation

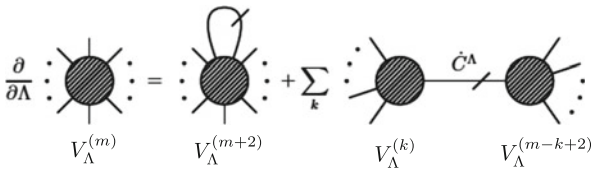

Fig. 12.1 Polchinski flow equation for the amputated connected correlation functions $V_\Lambda^{(m)}$. Nambu notation is used, therefore there are no arrows on the fermion lines, and the *round symbols* stand for vertices in their fully antisymmetric version. The flow equations for the connected correlation functions obtained from W_Λ have the same structure. On the *right-hand side*, only single propagator lines and no one-loop diagrams with two propagator lines appear. Adapted from [44]

for the Λ-dependence of $W_\Lambda(H)$. However, being a functional of Grassmann variables, $W_\Lambda(H)$ needs to be expanded in a power series in the source fields, with the connected correlation functions $G_\Lambda^{(m)}$ ("m-point functions") appearing as expansion coefficients. This leads to an infinite coupled hierarchy of flow equations (see, e.g., [44] or [18] for these equations, or Fig. 12.1 for diagrammatic expressions) for the $G_\Lambda^{(m)}$ where higher-order $G_\Lambda^{(m)}$s show up on the right-hand side of the flow equation of the lower $G_\Lambda^{(m)}$. For practical purposes, this hierarchy requires a truncation to be closed. Now, to make contact to ordinary perturbation theory and series summations like random phase approximation (RPA) or Cooper ladder summations, one would like to retrieve the corresponding diagrams of the perturbation theory in the truncation of the fRG hierarchy. This means that in particular in the flow equation for the four-point function (four external legs, the connected two-particle Green's function), one-loop diagrams made from two propagators should show up. One directly sees in the diagrammatic expression for this hierarchy in Fig. 12.1 that no such diagrams occur. These contributions are hidden further down in the hierarchy, in the feedback of the six-point correlation function on the four-point function. This means that including these important contributions would require to include the flow of the four-point function, after inserting the solution of the flow equation for the four-point function into that for the four-point function, dealing with Λ-nonlocal flow equations.

The flow equation for W_Λ can also be transformed into a flow equation for the so-called effective action, the generating functional for amputated connected correlation functions $V_\Lambda^{(m)}$. This equation was originally derived by Polchinski [15] and used by Zanchi and Schulz [45, 46] for fermions on a lattice. Here as well, the six-point contributions have to be kept in some way, e.g., by resorting to Λ-nonlocal flow equations [45–47], implying a larger numerical effort. In other works, the flow of W_Λ is derived using a cumulant expansion (see, e.g., [27, 45, 46, 48]). Then one gets the desired one-loop diagrams for the flow of the interactions, but now both internal lines have to be in the thin wavevector-shell being integrated out. While this catches the most important contributions in the particle–particle channel and in the cases of perfect nesting, it makes it still hard to include the contributions where the

two internal lines are at somewhat different absolute values of their band energies. In this sense, also these equations suffer from some inflexibility when it comes to treating 2D models with general dispersions.

Here, and also regarding some other aspects, the flow equations for the one-particle irreducible (1PI) vertex functions provide a very useful alternative. These equations [16] are derived from an exact flow equation for the functional $\Gamma_\Lambda(\Phi)$, the Legendre transform of $W_\Lambda(H)$, which in our notation reads [49]

$$\dot{\Gamma}_\Lambda(\Phi) = \frac{1}{2}\,\mathrm{tr}\left(\hat{G}^{(2)}_{0,\Lambda}\dot{\hat{Q}}_\Lambda\right) + \frac{1}{2}\left(\Phi, \dot{\hat{Q}}_\Lambda\Phi\right) - \frac{1}{2}\,\mathrm{tr}\left[\dot{\hat{Q}}_\Lambda\left(\frac{\delta^2\Gamma_\Lambda}{\delta\Phi^2}\right)^{-1}\right]. \quad (12.11)$$

Here, $\hat{G}^{(2)}_{0,\Lambda}$ is the inverse of the quadratic part in Nambu space, $\hat{Q}_\Lambda$. Γ_Λ possesses an expansion in terms of 1PI vertex function $\gamma^{(m)}_\Lambda$ (see e.g., [50] for more information on the relation of the 1PI vertices to connected correlation functions). For theories without external fermionic sources, only even orders of m are nonzero. The zeroth-order term corresponds to the grand canonical potential, $\gamma^{(2)}_\Lambda$ is the full inverse propagator, and $\gamma^{(4)}_\Lambda$ is the interaction vertex from which the two-particle Green's function and two-particle susceptibilities can be constructed, now all for the theory with quantum fluctuations included down to scale Λ. From (12.11) the flow equations for the vertex functions can be obtained by inserting the field expansion in the γ^ms. They are displayed diagrammatically in Fig. 12.2. Now the right-hand sides only contain 1PI one-loop terms. When there are two internal lines, one is a so-called

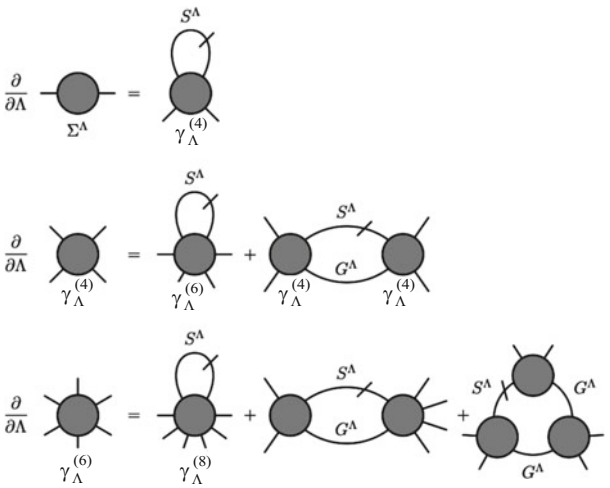

Fig. 12.2 Flow equations for the 1PI vertex functions. Σ_Λ is (up to the free propagator) the two-point 1PI vertex, $\gamma^{(4)}$ the four-point vertex etc., again using Nambu notation. The *line* with the *slanted bar* denotes a single scale propagator S_Λ, the *solid lines* are full scale-dependent Green's functions. Adapted from [44]

single-scale propagator defined below that is usually only nonzero for modes at the cutoff (at least when a sharper cutoff is used as flow parameter), while the other line can be anywhere at $|\epsilon(\boldsymbol{k})| \geq \Lambda$. Therefore also contributions with slightly different band energies are included directly at this level. The hierarchy is still coupled, but now the great advantage is that we can set – as an approximating truncation – the six-point vertex to zero without loosing the one-loop terms for the four-point vertex. The remaining two flow equations for self-energy and four-point interaction vertex are shown in the upper plot of Fig. 12.3. The truncation can be argued to be good for weakly coupled systems with two-particle interactions. Here the six-point vertex is zero in the initial conditions. Its initial growth is only in third order in the interaction, hence one can expect that for small enough initial interactions the six-point vertex remains small for some scale range. Then at low scales, near the Fermi surface, its impact becomes inhibited due to phase space constraints [49]. The remaining flow equation for the interaction corresponds to a summation of parquet diagrams [51, 52], with the advantage of being solved more easily by integrating the RG flow equations.

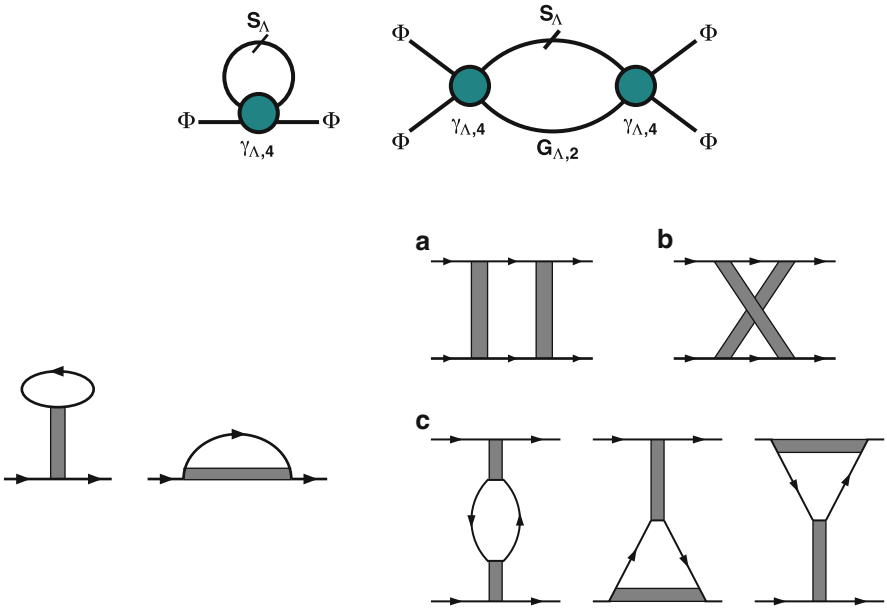

Fig. 12.3 *Upper figure*: The right-hand sides of 1PI RG equations for the self-energy and the interaction, in Nambu notation and with antisymmetric vertices, in the truncation where the six-point vertex is set to zero. *Lower plots*: The one-loop diagrams appearing on the right-hand side of the flow equation for the self-energy (two diagrams on the *left*) and for the coupling function V_Λ (five diagrams on the *right*), now with the Nambu index resolved, in the case of spin-rotational symmetry. (**a**) is the particle–particle diagram, (**b**) the crossed particle–hole diagram, (**c**) the direct particle–hole diagrams. The grey rectangles denote the coupling function V_Λ with the spin index conserved along the short edges. For details, see [49]

In the case of translational, spin-rotational and U(1)-invariance, the antisymmetric four-point vertex $\gamma^{(4)}_{\Lambda,s_1s_2s_3s_4}(k_1,k_2,k_3,k_4)$ for incoming particles k_1,s_1,k_2,s_2 and outgoing particles k_3,s_3,k_4,s_4 (the quantum numbers of particle 4 are dictated by the conservation laws, up to possible band indices) can be expressed in terms of a coupling function $V_\Lambda(k_1,k_2,k_3)$ by [49]

$$\gamma^{(4)}_{\Lambda,s_1s_2s_3s_4}(k_1,k_2,k_3,k_4) = V_\Lambda(k_1,k_2,k_3)\delta_{s_1s_3}\delta_{s_2s_4} - V_\Lambda(k_2,k_1,k_3)\delta_{s_1s_4}\delta_{s_2s_3} \tag{12.12}$$

This coupling function does not carry spin indices anymore, it is, however, understood that the spin projections of particles 1 and 3 are both s, and those particles 2 and 4 are s'. $V_\Lambda(k_1,k_2,k_3)$ carries the symmetries of the lattice (or of the Bloch functions in the multiband case) and should obey the relations $V_\Lambda(k_1,k_2,k_3) = V_\Lambda(k_2,k_1,k_4) = V_\Lambda(k_3,k_4,k_1)$. These latter relations, however, get more intricate in the N-patch discretizations described below.

With this, the flow equations for the self-energy becomes

$$\frac{\mathrm{d}}{\mathrm{d}\Lambda}\Sigma_\Lambda(k) = -\int \mathrm{d}k\, \left[2V_\Lambda(k,k',k) - V(k,k',k')\right] S_\Lambda(k')\,. \tag{12.13}$$

The equation for $V_\Lambda(k_1,k_2,k_3)$ with $k_4 = k_1+k_2-k_3$ reads [49]

$$\frac{\mathrm{d}}{\mathrm{d}\Lambda}V_\Lambda(k_1,k_2,k_3) = \mathcal{T}_{\mathrm{PP},\Lambda} + \mathcal{T}^{\mathrm{d}}_{\mathrm{PH},\Lambda} + \mathcal{T}^{\mathrm{cr}}_{\mathrm{PH},\Lambda} \tag{12.14}$$

with the one-loop particle–particle contributions $\mathcal{T}_{\mathrm{PP},\Lambda}$ and the two different particle–hole channels $\mathcal{T}^{\mathrm{d}}_{\mathrm{PH},\Lambda}$ and $\mathcal{T}^{\mathrm{cr}}_{\mathrm{PH},\Lambda}$ where

$$\begin{aligned}
&\mathcal{T}_{\mathrm{PP},\Lambda}(k_1,k_2;k_3,k_4) \\
&= \int \mathrm{d}k\, V_\Lambda(k_1,k_2,k)\, L(k,-k+k_1+k_2)\, V_\Lambda(k,-k+k_1+k_2,k_3)
\end{aligned} \tag{12.15}$$

$$\begin{aligned}
&\mathcal{T}^{\mathrm{d}}_{\mathrm{PH},\Lambda}(k_1,k_2;k_3,k_4) \\
&= \int \mathrm{d}k \Big[-2V_\Lambda(k_1,k,k_3)\, L(k,k+k_1-k_3)\, V_\Lambda(k+k_1-k_3,k_2,k) \\
&\quad + V_\Lambda(k_1,k,k+k_1-k_3)\, L(k,k+k_1-k_3)\, V_\Lambda(k+k_1-k_3,k_2,k) \\
&\quad + V_\Lambda(k_1,k,k_3)\, L(k,k+k_1-k_3)\, V_\Lambda(k_2,k+k_1-k_3,k)\Big]
\end{aligned} \tag{12.16}$$

$$\begin{aligned}
&\mathcal{T}^{\mathrm{cr}}_{\mathrm{PH},\Lambda}(k_1,k_2;k_3,k_4) \\
&= \int \mathrm{d}k\, V_\Lambda(k_1,k+k_2-k_3,k)\, L(k,k+k_2-k_3)\, V_\Lambda(k,k_2,k_3)
\end{aligned} \tag{12.17}$$

In these equations, the product of the two internal lines in the one-loop diagrams is $L(k,k') = S_\Lambda(k)G^{(2)}_\Lambda(k') + G^{(2)}_\Lambda(k)S_\Lambda(k')$ with the so-called single-scale propagator

$$S_\Lambda(k) = -G_\Lambda^{(2)}(k) \left[\frac{\mathrm{d}}{\mathrm{d}\Lambda} Q_\Lambda(k) \right] G_\Lambda^{(2)}(k) \,. \tag{12.18}$$

Here, $G_\Lambda^{(2)}(k) = C_\Lambda(k)/\left[i\omega - \epsilon(\boldsymbol{k}) - C_\Lambda(k)\Sigma_\Lambda(k)\right]$ denotes the full Green's function at RG scale Λ. If self-energy effects are ignored, the single-scale propagator is simply $S_\Lambda(k) = \dot{C}_\Lambda(k)/\left[i\omega - \epsilon(\boldsymbol{k})\right]$, with $\dot{C}_\Lambda(k)$ being confined to a momentum shell with $|\epsilon(\boldsymbol{k})| \approx \Lambda$ in the case of a momentum-shell cutoff.

12.2.3 Choice of Flow Parameter

Although we have made reference to a momentum-shell cutoff at several stages in the derivation of the 1PI fRG equation in the last subsection, it is not difficult to see that the same equations can be written down for any other continuously variable parameter in the quadratic part of the action, provided that this parameter does not appear in the interaction term. This indicates how flexible and powerful these equations are. Here we want to discuss briefly which different flow parameters have been used previously to explore the instabilities of Hubbard-type lattice system. We mention, however, that in other contexts the freedom to choose appropriate flow parameters has been exploited as well [29, 53]. The guiding principle for a good flow parameter in our context is that it should enable us (a) to approach a specific infrared singularity in a controlled way (regulator property) and (b) to include all other (possibly also singular) tendencies during the flow (unbiasedness). Then the fRG will give a more realistic picture than perturbation expansions that single out one dominant channel. In diagrammatic bubble or ladder summations, the singularities normally arise due to a pile-up of logarithms. These are roughly given by $g \log[W/\max(T, \Lambda)]$ (after Matsubara frequency summation) with a coupling constant g, bandwidth W and lower energy cutoff Λ. Hence, if we want to build in such dangerous terms step by step, we can either vary Λ, the temperature T, or the coupling g. Cutting out low Matsubara frequencies should have a similar effect as increasing T. This gives the following options.

12.2.3.1 Momentum-Shell RG

Here we introduce a cutoff $D_\Lambda[\epsilon(\boldsymbol{k})]$ into the quadratic part of the action,

$$Q_\Lambda(k) = T \sum_{ik_0,\boldsymbol{k}} D_\Lambda[\epsilon(\boldsymbol{k})]\, \bar{\psi}_k \, \left[i\omega_n - \epsilon(\boldsymbol{k})\right] \, \psi_k \,. \tag{12.19}$$

$D_\Lambda[\epsilon(\boldsymbol{k})]$ is very large for $|\epsilon(\boldsymbol{k})| \leq \Lambda$ and $D_\Lambda[\epsilon(\boldsymbol{k})] = 1$ for $|\epsilon(\boldsymbol{k})| > \Lambda$. In practice one mainly needs the inverse $D_\Lambda^{-1}[\epsilon(\boldsymbol{k})] = C_\Lambda[\epsilon(\boldsymbol{k})]$, which can be chosen conveniently as a sharp step function for analytical manipulations or as a smoothened step function for numerical treatments. Then the full Green's function, given by $G_\Lambda(k) = C_\Lambda[\epsilon(\boldsymbol{k})]/\{i\omega - \epsilon(\boldsymbol{k}) - C_\Lambda[\epsilon(\boldsymbol{k})]\Sigma_\Lambda(k)\}$, is suppressed for modes with $|\epsilon(\boldsymbol{k})| \leq \Lambda$. The same holds for the single-scale propagator, hence only high-energy modes show up in the loop diagrams. The momentum-shell RG is the widely-used standard [27]. Alternative forms can also be derived from other exact RG equations [15, 17, 45, 46].

A serious drawback of the momentum-shell cutoff RG is the nonuniformity in the RG scale at which one-loop particle–hole (PH) processes with different wavevector transfers $\boldsymbol{q}$ are included at low T. PH processes with small $\boldsymbol{q}$, i.e., a particle at wavevector $\boldsymbol{k}$ and a hole at $\boldsymbol{k} + \boldsymbol{q}$, can only occur in the vicinity of the Fermi surface. For $T \to 0$ the support of the PH bubble for $\omega = 0, \boldsymbol{q} \to 0$ shrinks to a temperature-smeared δ-function on the Fermi surface with width $\sim T$. Thus, the $\boldsymbol{q} \to 0$ PH-modes are integrated out only for cutoff $\Lambda \leq T$, even if the density of states is divergent, and these processes give a singular contribution for $T \to 0$. However, in a coupled flow with various other tendencies, the flow normally diverges *before* we get down to $\Lambda \sim T$. The flow has to be stopped and the $\boldsymbol{q} \to 0$ PH pairs did not have the chance to contribute to the flow *by construction*. Therefore, the cutoff-flow is still biased to some extent and approaches are needed to study the influence of the $\boldsymbol{q} \to 0$ particle–hole excitations.

The following three schemes work without a sharp cutoff and allow for a uniform inclusion of the important one-loop processes.

12.2.3.2 Temperature Flow

The temperature-flow scheme [56] uses the temperature T appearing in the quadratic part of the action as flow parameter. Before the T-derivative can be taken, the fermion fields need to be rescaled to remove the T-dependence appearing in front of the interaction term in the original action. Then the theory is defined at some high $T \sim T_0$ in terms of its two-point and four-point vertex. Here, we simply choose the free propagator and the local Hubbard repulsion as the initial values at T_0. This should be reasonable for sufficiently high $T_0 \sim$ bandwidth, as all perturbative corrections decay with a negative power of T. In the approximate version used e.g. in [56] without $\gamma_T^{(6)}$ and self-energy corrections, the right-hand side of the flow equation for the interaction is just given by the T-derivatives of the one-loop diagrams. This flow scheme basically reproduced the antiferromagnetic and d-wave pairing instabilities found in the t–t'-Hubbard model using the momentum-shell schemes, but also added a ferromagnetic regime near van Hove filling and triplet pairing instabilities for larger values of $|t'|$ [56].

12.2.3.3 Interaction Flow

In the interaction flow scheme [57] we first multiply Q with a scale factor $1/g$ and split it in two, yielding

$$Q_g = T \sum_{i\omega_n, \boldsymbol{k}} \bar{\psi}_k g^{-1/2} \left[i\omega_n - \epsilon(\boldsymbol{k}) \right] \psi_k g^{-1/2}, \tag{12.20}$$

g will be the flow parameter. We can absorb the factor $1/g$ in rescaled fields $\tilde{\psi}, \bar{\tilde{\psi}}$ defined as $\tilde{\psi} = g^{-1/2}\psi$. With this the interaction term gets an extra factor g^2 when written in terms of the new fields:

$$V_g^{(4)} = \frac{1}{2N} \sum_{\substack{k,k',k+q \\ s,s'}} g^2 V(k, k', k+q) \bar{\tilde{\psi}}_{k+q,s} \bar{\tilde{\psi}}_{k'-q,s'} \tilde{\psi}_{k',s'} \tilde{\psi}_{k,s} \,. \tag{12.21}$$

We observe that changing the scale factor $1/g$ in Q_g corresponds to changing the strength of the bare interactions. The rescaled fermions $\tilde{\psi}, \bar{\tilde{\psi}}$ describe a system with a bare interaction strength $g^2 V$. Now we can start at $g = 0^+$, i.e., at infinitesimally small bare interaction, and use the flow equations for the ψ-theory to integrate up to the desired bare interaction, reached at $g = 1$. We can also stop the flow at any other value of g, with the functions $g\Sigma$ and $g^2 V_g(k, k', k+q)$ being the self-energy and interacting vertex function for the bare interaction $g^2 V(k, k', k+q)$. One can call this scheme *interaction flow*, or simply flat-cutoff flow, as all modes are suppressed by the same factor g. Now singularities on the right-hand side of the flow equation are not regularized by the flow parameter. Thus, the interaction flow scheme has to be used at $T > 0$, when the one-loop diagrams are bounded. Using this scheme, the ground state phase diagram of the 2D t–t' Hubbard model comes out basically the same as in the T-flow [57].

12.2.3.4 Smooth Frequency Cutoff

Another regulator that does not ignore small-wavevector transfer processes in the particle–hole channel and is hence capable of describing ferromagnetic instabilities as well is the so-called Ω-scheme introduced by Husemann and Salmhofer [58]. Here one uses a smooth cutoff on the Matsubara frequency axis, given by $C_\Omega(\omega_n) = \omega_n^2/(\omega_n^2 + \Omega^2)$, and the flow goes from large Ω damping away all modes down to $\Omega = 0$ where the full free propagator is obtained. The benefit of the simple functional dependence on ω_n and Ω is that this cutoff function still permits one to determine the Matsubara sums in the loop diagrams analytically. The results of this cutoff-scheme together with an alternative vertex decomposition are shown in the next sections.

12.3 Implementation of the Fermionic fRG for Two Dimensional Lattices

Besides the fact that the equations above were derived from the exact flow equations for generating functionals, the name *fRG* reflects that at least for many-fermion systems with a Fermi surface, one can in general no longer reduce the parametrization of the interaction vertex to a handful of running coupling constants. Rather, the functional dependence of the interactions on the location on the Fermi surface or more generally in the Brillouin zone can turn out to be relevant, i.e., grow strongly under the RG flow. Hence keeping the functional dependence of the interaction vertex around the Fermi surface(s) is important. This leads to RG equations for functions of wavevectors, i.e., functional differential equations. The two-fermion interaction vertex, e.g., depends on three wavevectors, while the fourth is fixed by wavevector conservation on the lattice. Note that in principle also the frequency dependence of the interactions can be studied, but in the works described here the frequency dependence is mainly neglected. The wavevector dependence is typically treated by a so-called N-patch discretization, where the Brillouin zone is divided up into patches containing Fermi surface segments. This was first used in this context by Zanchi and Schulz [45, 46]. Within these patches, the wavevector dependence is ignored again, and the value for the patch region near the Fermi surface is computed and used at all wavevectors inside the patch. However, as described below, a sufficient number of patches allows one to draw a number physically relevant conclusions. The setup of the patches around the Fermi surfaces is shown in Fig. 12.4. It is guided by the idea that one wants to obtain an effective interaction for quasiparticles near the Fermi surface, and the wavevector-dependence tangential to the Fermi surface is relevant in standard examples. The dependence normal to the Fermi surface does usually not alter the leading flow. Various other patching schemes, like square patchings of the Brillouin zone have been investigated as well, without unexpected or problematic differences in the results [57, 59].

Let us now focus on the mainly used approximation which consists (besides the truncation explained earlier) in neglecting the self-energy feedback. Then only the interaction vertex is considered. Further, we ignore the frequency dependence. By the N-patch construction, the fRG equations for $\gamma_\Lambda^{(4)}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ are converted into N^3 coupled nonlinear differential equations for the components of the discretized interaction vertex. As at least the particle–particle channel contains singular diagrams for $T \to 0$, the solution of these fRG equations without self-energy corrections typically results in a *flow to strong coupling* where one or more families of interaction processes run to large absolute values. This is accompanied by the build-up of sharp structures in wavevector-space, corresponding to the formation of longer-range interactions. Then the flow has to be stopped when the coupling strength has exceeded the scale of the band width, as then self-energy corrections and possibly also the truncation error would become important. From this scale where the run-away flow occurs one can derive an estimate for critical scales for

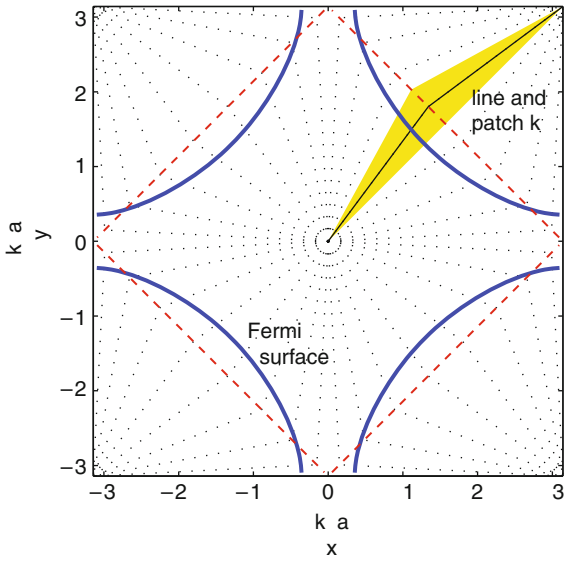

Fig. 12.4 N-patch discretization of the Brillouin zone for the one-band Hubbard model on the 2D square lattice. The *colored region* is a patch in which the coupling function is approximated as a constant

long-range ordering or gap openings. Furthermore from the analysis which class of coupling function diverges most strongly, one can infer the dominant tendencies toward long-range ordering.

An useful tool for assessing physical information, e.g., on the gap function in the case of a Cooper instability is the diagonalization of the effective interaction in the Cooper (zero total incoming wavevector) channel, $V_\Lambda(\boldsymbol{k}, -\boldsymbol{k} \to \boldsymbol{k}', -\boldsymbol{k}')$ at a scale Λ near the instability [60]. By studying the linearized gap function, one can see that the pairing gap function would adopt the $\boldsymbol{k}$-dependence of the eigenvector belonging to the most negative eigenvalue of $V_\Lambda(\boldsymbol{k}, -\boldsymbol{k} \to \boldsymbol{k}', -\boldsymbol{k}')$. This allows one to determine the structure of the pairing around the Fermi surface beyond simple symmetry statements. This reasoning can of course be extended to other order parameters.

As shown in Sect. 12.4, the flow to strong coupling typically occurs for components of the coupling function $V_\Lambda(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ with particular vales of the total wavevector $\boldsymbol{k}_1 + \boldsymbol{k}_2$ or a particular wavevector transfer $\boldsymbol{k}_2 - \boldsymbol{k}_3$. While the dependence of the coupling function of these combinations of wavevectors is rather strong, there is usually a weaker dependence on the precise location of the external legs on the Fermi surface which can be well described with lower lattice harmonics such as a constant or a $d_{x^2-y^2}$-formfactor. This led Husemann and Salmhofer [58] to propose an alternative description of the coupling function,

$$V_\Lambda(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = U - \Phi_{\mathrm{SC}}^\Lambda(\boldsymbol{k}_1,\boldsymbol{k}_3,\boldsymbol{k}_1+\boldsymbol{k}_2) + \Phi_{\mathrm{M}}^\Lambda(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3-\boldsymbol{k}_1) + \frac{1}{2}\Phi_{\mathrm{M}}^\Lambda(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_2-\boldsymbol{k}_3) - \frac{1}{2}\Phi_{\mathrm{K}}^\Lambda(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_2-\boldsymbol{k}_3), \quad (12.22)$$

where now the new functions Φ_{SC}, Φ_{M}, and Φ_{K} are treated with a fine resolution for the last entry, and where formfactors are used to describe their dependence on the first two entries. Moreover, the particle–particle channel is selected to drive the flow of the pairing part Φ_{SC}, the magnetic (crossed particle–hole) channel drives the magnetic part Φ_{M}, and the another combination of particle–hole terms drives the charge part Φ_{K}. All this does not involve any additional approximations, but for larger parameter region it turned out [58] to be an efficient additional approximation to use only s-wave (constant) and $d_{x^2-y^2}$-formfactors to describe the dependence of the Φ-functions on their two first arguments. Therefore only one wavevector dependence, namely on the last argument, for the Φ-functions needs to be treated numerically. This reduces the computational effort and allows for a better description of the strong dependence on these wavevector combinations. The results for the repulsive Hubbard model at van Hove-filling on the 2D square lattice are briefly described in Sect. 12.4. We expect that this alternative formulation will prove useful also in the study of flows into the symmetry-broken state, and to capture frequency dependences better. Very similar decompositions for the frequency structure of the vertex function have already been used in the context of impurity models [61].

12.4 Instabilities in Two-Dimensional Lattice Systems

Here we describe the results obtained by applying the N-patch fRG scheme to two interesting systems. The first system is the one-band Hubbard model on the 2D square lattice, whose phase diagram is of interest in connection with the high-T_c cuprates and other unconventional superconductors. As a second example we discuss fRG studies of multiband Hubbard-type systems on the same lattice with direct relevance to the newly discovered iron pnictide (and chalcogenide) superconductors.

12.4.1 Two-Dimensional Hubbard Model Near Half Filling

The 2D Hubbard model is one of the most-studied theoretical models in context with high-temperature superconductivity in the layered copper-oxide materials. While the original derivation of an effective one-band model for copper-oxide planes [62] implies a strong local Coulomb interaction U on the cooper $d_{x^2-y^2}$-orbitals, the model has also been investigated intensively at weaker couplings. Here we discuss the picture that is obtained by using the fRG which is perturbative in U.

Nevertheless, studying the Hubbard model at weak to moderate U is still relevant at least for cuprates in the overdoped regime. Here, experiments [63, 106, 107] exhibit a rather well-defined 2D Fermi liquid with Fermi surfaces that can be parameterized by simple one-band tight-binding prescriptions.

The Hamiltonian of the one-band Hubbard model focussed on here reads

$$H = -t \sum_{\substack{\langle i,j\rangle,\\ \text{n.n.}}} \left(c^\dagger_{i,s} c_{j,s} + c^\dagger_{j,s} c_{i,s} \right) - t' \sum_{\substack{\langle\langle i,j\rangle\rangle,\\ \text{n.n.n.}}} \left(c^\dagger_{i,s} c_{j,s} + c^\dagger_{j,s} c_{i,s} \right) + U \sum_i n_{i,\uparrow} n_{i,\downarrow} \,, \tag{12.23}$$

with t and t' being the hopping amplitudes to nearest neighbor and diagonal next-nearest neighbors on the square lattice. U is the Hubbard onsite repulsion. It serves as initial condition for the coupling function parametrizing the interaction vertex at high scales $\Lambda_0 \sim 4t$ before any modes have been integrated out, $V_{\Lambda_0}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3 = U$. Together with the chemical potential μ, the hopping parameters t and t' determine the shape of the Fermi surface. The band structure including μ is given by $\epsilon(\boldsymbol{k}) = -2t \cos k_x \cos k_y - 4t' \cos k_x \cos k_y - \mu$.

Any band structure with the inversion symmetry $\epsilon(\boldsymbol{k}) = \epsilon(-\boldsymbol{k})$ and nonzero density of states at the Fermi surface exhibits a logarithmically divergent particle–particle diagram for zero incoming wavevector and frequency for $T \to 0$. These diagrams can be summed up in a ladder summation. For repulsive onsite interactions, this does not cause a divergence, but a strong suppression of scattering processes with zero total wavevector and frequency. On the other hand, particle–hole terms are known to create attractive nonlocal interactions. Inserting these back into the particle–particle ladder into this summation will cause Cooper instabilities in unconventional (possibly very anisotropic) pairing channels if the temperature is low enough. This is the so-called Kohn–Luttinger effect [27, 64] that predicts that all inversion-symmetric metals (in at least two dimensions) have a superconducting ground state, unless another instability excludes this. In fact, mathematically it is impossible to guarantee the convergence of perturbation theory for generic systems below a certain energy scale determined by the mentioned particle–particle ladder [17]. Note, however, that the Kohn–Luttinger effect may only occur at very low T and hence may be only of academic interest for many material classes. In particular, impurities could kill many of these low-T_c anisotropic superconducting phases easily. In the 2D Hubbard model on the square lattice, however, particle–hole corrections to the pair scattering have a marked influence and lead to rather high critical scales for Cooper pairing, as described below.

Regarding the possible ground states of the one-band Hubbard model on the square lattice at weak to moderate U, the Fermi surface location and shape plays a dominant role. Let us now consider how the N-patch fRG-flows differ for different Fermi surfaces.

For $\mu = 0$ and $t' = 0$, one has a half-filled band with a perfectly nested Fermi surface. All electronic states that are occupied at $T = 0$, i.e., with $\epsilon(\boldsymbol{k}) < 0$,

can be scattered by wavevector $\boldsymbol{Q} = (\pi, \pi)$ on unoccupied states, leading to a dominance of particle–hole fluctuations with this nesting wavevector. A random-phase approximation summation of these bubbles would result in a divergent static spin susceptibility at $\boldsymbol{Q}$ for any > 0 at sufficiently low temperatures indicating the formation of an antiferromagnetic (AF) spin-density wave. The basic fRG results for low T are shown in the left half of Fig. 12.5. In the upper plot the Fermi surface is shown together with $N = 32$ discretization points. In the middle plot we show the flow of two families of coupling constants. We can see that some of

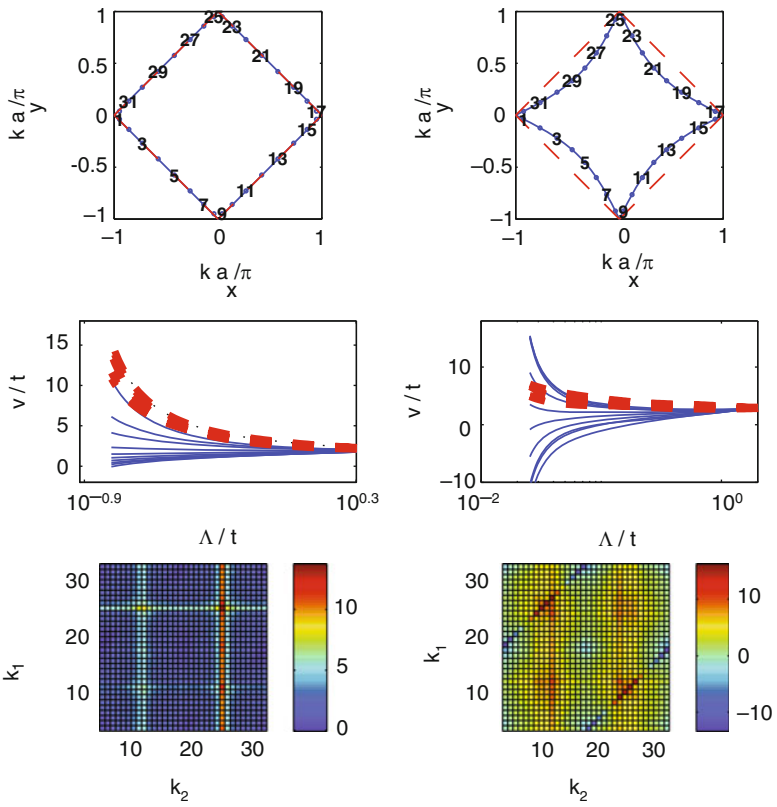

Fig. 12.5 N-patch fRG data obtained with the momentum-shell 1PI fRG scheme for the repulsive Hubbard model on the 2D square lattice. *Left plots*: $\mu = 0$, $t' = 0$ and initial $U = 2t$, *right plots*: $\mu = 1.2t$, $t' = -0.3t$, $U = 3t$. *Upper row*: Fermi surfaces for the two cases and the $N = 32$ discretization points for the two incoming $\boldsymbol{k}_1, \boldsymbol{k}_2$ and the first outgoing wavevector $\boldsymbol{k}_3$. *Middle row*: *Solid lines* show the flows of components in the coupling function $V_\Lambda(k_1, k_2, k_3)$ corresponding to Cooper pair scattering with zero total incoming wavevector, $\boldsymbol{k}_1 + \boldsymbol{k}_2 = 0$, or $|k_1 - k_2| = N/2$ in terms of patch indices. The *dashed lines* correspond to processes in the AF-SDW channel with wavevector transfer $\boldsymbol{k}_2 - \boldsymbol{k}_3$ near $\pm(\pi, \pi)$. The flow is started at $\Lambda = 4t$ and goes to the left toward smaller Λ. *Lower plots*: Snapshots of the coupling function $V_\Lambda(k_1, k_2, k_3)$ near the instability with k_1 fixed at point 1 with k_1 and k_2 moving around the Fermi surface. The *colorbars* on the *right* indicate the values of the interactions

these lines flow to large values when the scale parameter Λ is lowered. This is an example for a *flow to strong coupling*. When the interaction reaches values larger than the band width the flow has to be stopped, as explained in Sect. 12.3. Next to the divergence scale we can also read out from the data which classes of coupling constants grow most strongly. In this case, these are the dashed lines corresponding to interaction processes with wavevector transfer $\boldsymbol{Q} = (\pi, \pi)$ between $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$. The maximal coupling constant that grows most strongly also belongs to this family. All members of this family flow to strong coupling in a rather similar way when Λ is reduced toward the critical scale $\Lambda_c \sim 0.16\,t$. Other families of couplings constants grow less strongly. Shown as well are as solid lines Cooper processes with zero total incoming wavevector. These processes show some growth as well, but lag behind the leading components. The lowest plot on the left shows a snapshot of the coupling constant when the first outgoing wavevector $\boldsymbol{k}_3$ is fixed at discretization point 1 near $(-\pi, 0)$ as function of the incoming wavevectors. We can clearly see two structures: one vertical line with strongly enhanced repulsive interactions at $k_2 = 24$ (corresponding to wavevector transfer $\boldsymbol{Q} = (\pi, \pi)$ between $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$) with very little dependence on k_1, and another line at $k_1 = 24$ (corresponding to wavevector transfer $\boldsymbol{Q} = (\pi, \pi)$ between $\boldsymbol{k}_1$ and $\boldsymbol{k}_3$). These values show again only a weak dependence on k_2 and are roughly half as large as the vertical feature. Concentrating on these two features we arrive at the following effective interaction near the instability

$$H_{\Lambda}^{(\mathrm{SDW})} = \frac{V_{\mathrm{AF}}}{2} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3 \\ s,s'}} \left[2\delta(\boldsymbol{k}_2 - \boldsymbol{k}_3 \pm \boldsymbol{Q}) + \delta(\boldsymbol{k}_1 - \boldsymbol{k}_3 \pm \boldsymbol{Q})\right] c^{\dagger}_{\boldsymbol{k}_3,s} c^{\dagger}_{\boldsymbol{k}_4,s'} c_{\boldsymbol{k}_2,s'} c_{\boldsymbol{k}_1,s},$$

where $\boldsymbol{k}_4$ is understood to be given by $\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3$ modulo reciprocal lattice vectors. A simple calculation shows that this interaction is equivalent to a long-ranged AF spin–spin interaction $H = -J \sum_{\langle i,j \rangle} \mathrm{e}^{\mathrm{i}\boldsymbol{Q}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j)} \boldsymbol{S}_i \cdot \boldsymbol{S}_j$ with spin operators given by the usual fermion bilinears and $J \propto V_{\mathrm{AF}}$. This effective Hamiltonian exhibits long-ranged AF order at sufficiently low T. In this sense, the fRG flow for these parameters clearly indicate the proximity to an AF-ordered state or an AF spin-density wave (AF-SDW). Strictly speaking, this is the most we can infer at this stage. Whether long-range order actually occurs depends on the subleading terms and on the approximation errors collected on the way to this result. Note that the dynamics and interactions of emergent collective degrees of freedom are not appropriately captured in this approximation. Of course in two dimensions, a spontaneous breaking of a continuous symmetry should not occur at $T > 0$. On the other hand, most experimental systems described by our model would have additional small couplings in the third direction that change the situation in favor of our result. Moreover, as the leading instability is clearly exposed by this scheme, one could now resort to a bosonized description that allows one to treat the collective infrared physics much better (see e.g., [65] how the continuous symmetry breaking in two dimensions gets healed in the asymptotic flow).

Now let us consider a different Fermi surface with more curvature. We choose $t' = -0.3\,t$ and $\mu = -1.2$, such that the Fermi surface is now curved but still contains the van Hove point $(\pi, 0)$ and $(0, \pi)$. The flow is illustrated in the right plots in Fig. 12.5. In the middle plot on the right we now see a dominance of the Cooper scattering processes with zero incoming total wavevector. Here some lines seem to diverge to large positive values, while some other lines take very negative values. The sign structure is also visible quite nicely in the bottom figure on the right. Here the first outgoing wavevector $(k)_3$ is again fixed at point 1 (near $(-\pi, 0)$) and the dependence on the incoming wavevectors around the Fermi surface is plotted. One observes diagonal lines of enhanced interactions, corresponding to zero total incoming wavevector $\boldsymbol{k}_1 + \boldsymbol{k}_2 = 0$. This pair scattering is attractive when the incoming pair $\boldsymbol{k}_1, \boldsymbol{k}_2$ near the same saddle point $(\pm\pi, 0)$ as the outgoing pair $\boldsymbol{k}_3, -\boldsymbol{k}_3$, and repulsive, when incoming and outgoing pairs are at different saddle points. This is exactly the formfactor $d(\boldsymbol{k}) = d_0(\cos k_x - \cos k_y)$ of a $d_{x^2-y^2}$-Cooper pairing instability, where the dominant interaction is given by what is obtained by only keeping these diagonal features in $V_\Lambda(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$,

$$H_\Lambda^{(d\mathrm{SC})} = V_{d\mathrm{SC}} \sum_{\boldsymbol{k},\boldsymbol{k}'} d(\boldsymbol{k}) d(\boldsymbol{k}')\, c^\dagger_{\boldsymbol{k}',\uparrow} c^\dagger_{-\boldsymbol{k}',\downarrow} c_{-\boldsymbol{k},\downarrow} c_{\boldsymbol{k},\uparrow} \,.$$

The spin-structure can be understood by forming the antisymmetric vertex again via (12.12). As the dominant features are symmetric with respect to interchange of k_1 and k_2, the equal-spin vertex vanishes. Again, this effective Hamiltonian can be solved by mean-field theory exactly. It has a d-wave paired ground state. This d-wave pairing instability was obtained by a number of research groups with different variations of fRG approaches (e.g., [45–47, 54–56]). The parameter region for its occurrence is rather wide. It constitutes convincing evidence that the weakly coupled Hubbard model possesses a d-wave superconducting ground state. The pairing mechanism at these weaker couplings is best described as AF-spin-fluctuation induced. This is already visible in the bottom plot on the right-hand side of Fig. 12.5. Here one can see that the d-wave pairing interaction on the diagonal lines with zero total incoming wavevector crosses a region with enhanced repulsive interactions, e.g., near $k_1 \sim 8, 9$ and $k_2 \sim 24, 25$. This enhancement is the broadened, due to Fermi surface curvature, version of the vertical SDW feature in the left lower plot for the fully nested case. We see that the enhancement to positive values fits perfectly into the sign structure of the d-wave. Studying the flow as a function of the cutoff Λ, one finds that the SDW features appear first and create the attractive component in the $d_{x^2-y^2}$-pairing channel. More drastically, one can put a filter into the fRG scheme that removes these SDW features by hand, e.g., by setting these interactions to their initial values. This way one also looses the d-wave pairing instability in the same way as the AF-SDW instability, showing the strong coupling between the channels. Another way to check this view is to completely drop the particle–hole channels. In this case, no flow to strong coupling is found, as no attractive pairing components are generated, and as the SDW instability is excluded as well. If we only drop the particle–particle channel, we of course destroy the pairing instabilities, and

the SDW instability takes a wider parameter space, at even higher critical scales comparable to those found in simple RPA summations.

The fRG data discussed so far showed rather clearcut single-channel instabilities, either in the AF-SDW or in the d-wave pairing channel, provided T is low enough. For higher $T > T_c$ the flow remains finite, and the system should be in a metallic state. The critical scale at $T = 0$ is usually a good estimate for the critical temperatures T_c above which the flow remains finite, up to factors ~ 1. One can ask how the flow changes when we move from one to the other by changing the parameters. In principle there are two possibilities: One is that the flow to strong coupling changes gradually and the critical scale remains nonzero, with the SDW component getting weaker and the pairing component getting stronger continuously. The second possibility would be a quantum critical point where the critical scale for the run-away flow scale is suppressed to zero. As shown in the upper plots of Fig. 12.6 the change from AF-SDW is of the first type, while for larger t' one finds indications for quantum critical point between d-wave pairing and ferromagnetism, at least in the fermionic fRG flows, as illustrated in the lower plots of Fig. 12.6. The main factor that causes this difference between the phase changes is the overlap of the interaction processes between the two respective channels.

The AF-SDW instability and the $d_{x^2-y^2}$ both require repulsive interaction processes between the two saddle points region near $(\pi, 0)$ and $(0, \pi)$. Hence the growth of one channel also supports the other channel to a large extent. If the Fermi surface remains in the vicinity of the van Hove points for a larger range of μ, there is a rather wide crossover region where both channels, d-wave pairing and AF-SDW, grow strongly in a rather similar fashion. This is the case in the so-called saddle-point regime in the upper right plot of Fig. 12.6. Here, the mutual reinforcement of different interaction processes with initial and final states near the saddle-point regions is reminiscent of the flows to strong coupling in the half-filled two-leg Hubbard model, where d-wave pairing and AF-SDW channel are again driven in parts by the same processes. In this one-dimensional system the ground state does not select one these channels and does not develop any quasi-long-range order, but becomes a short-range correlated spin liquid with a single-particle gap [66,67]. This resemblance motivated interpretations of the flows in two dimensions as indications for a partial truncation of the Fermi surface in the saddle point regions [55, 68, 69], similar to what is argued to occur in the underdoped high-T_c cuprates. A controlled determination of the resulting strong-coupling state near the saddle-point region for these parameter regions is, however, still missing.

If the Fermi surface is further away from the saddle point regions, the strongest interaction processes with wavevector transfer (π, π) occur near the Brillouin zone diagonal where the $d_{x^2-y^2}$-formfactor is small. Then d-wave pairing and AF-SDW channels are rather weakly coupled. Correspondingly the energy scales for the two instabilities can be quite different. This is seen in the left upper plot of Fig. 12.6 for band fillings larger than one. Here the high-energy-scale AF-SDW instability gets cut off at a certain critical chemical potential and is replaced by a low-energy-scale $d_{x^2-y^2}$ pairing instability. Without further analysis it is difficult to determine the precise nature of this transition. However, as the coupling between the channels

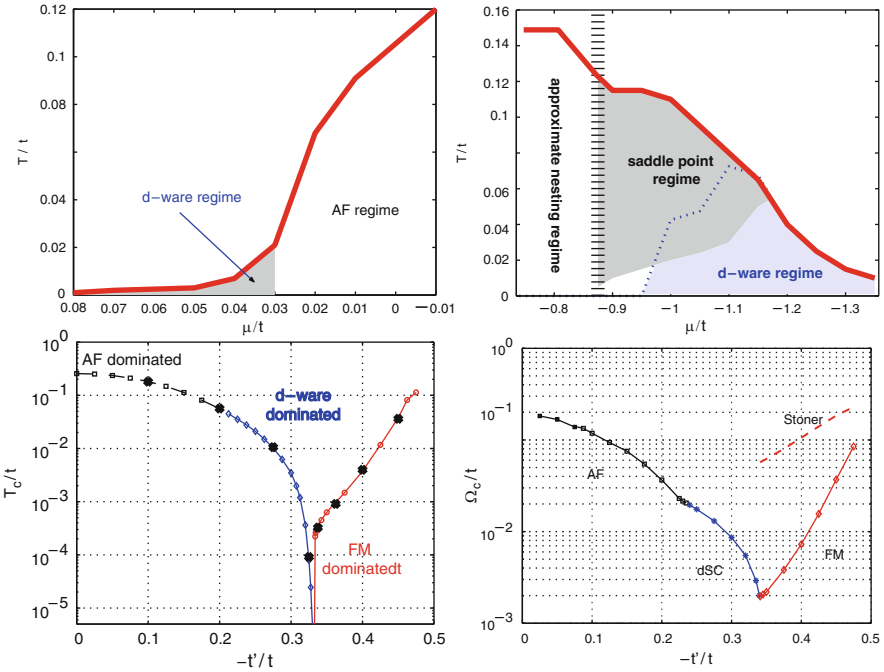

Fig. 12.6 Leading instabilities found with the N-patch fRG in the t-t'-Hubbard model. *Left upper plot*: Critical scale T_c below which the fRG flow goes to strong coupling vs. chemical potential μ for band filling larger than unity, for $t' = -0.3t$ and $U = 3t$. There is a high-energy-scale AF SDW instability with a weaker $d_{x^2-y^2}$-wave pairing instability when the AF-SDW is cut off. Data from [84]. *Right upper plot*: Data for the same t' and U on the "hole-doped" side with band fillings smaller than one, from [55]. Now there is a broad crossover "saddle point regime" between the nesting-driven AF-SDW instability and the $d_{x^2-y^2}$-wave pairing regime. *Lower left plot*: T_c vs. t' at the van Hove filling where the Fermi surface contains the points $(\pi, 0)$ and $(0, \pi)$. For large t' one finds a ferromagnetic instability. Data from [56] obtained with the T-flow scheme. *Right lower plot*: Critical scale Ω_c for the flow to strong coupling, again vs. t' at van Hove filling, now obtained with the simplified vertex parametrization of [58] and with a smooth frequency cutoff. While the results agree well with the full N-patch flow for $t' \approx 0$ and $t' \approx -0.5t$, the transition from d-wave pairing to the ferromagnetic regime does not go through a potential quantum critical point as in the *left lower plot*

is rather weak, we expect that a more accurate calculation would yield first-order transition.

The situation is again different for the transition from the d-wave pairing to the ferromagnetic instability, as illustrated in the lower plots in Fig. 12.6. Here the critical scale for the flow to strong coupling is suppressed strongly in the transition region, and the fermionic fRG flows (left lower plot) even suggest a quantum critical point. While it is difficult to fully understand these coupled nonlinear flows of N^3 components of the coupling function using a simple argument, there is now one significant difference compared to the continuous crossovers described above. Let

us again consider the interaction process that is common to both, d-wave pairing and ferromagnetic channel. The latter channel is driven by scattering processes with wavevector transfer $\boldsymbol{q}$ (between $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$) going to zero. A closer analysis shows that these processes should be repulsive to cause a ferromagnetic instability, like for the usual Stoner criterion. This, however, means that also the processes with $\boldsymbol{q} \to 0$ that have zero total momentum and enter the Cooper channel should be repulsive. A positive pair scattering with small wavevector transfer is, however, incompatible with singlet Cooper pairing that needs attractive scattering for these wavevector combinations. In this sense, the overlap between the two ordering channels can either support ferromagnetism or d-wave pairing, but not both tendencies together, and the two channels compete already in the symmetric phase. We note as well that the ferromagnetic instability is rather fragile and limited to the neighborhood of van Hove filling. Upon doping away one finds a smooth crossover to an instability with low critical scale and dominant p-wave pairing interactions [56].

The two different plot in the lower half of Fig. 12.6 are obtained for the same system parameters. The left plot is for the "full" fermionic N-patch fRG in the temperature-flow scheme [56] and shows the quantum critical point between d-wave pairing and ferromagnetic instabilities. The plot on the right-hand side is adapted from [58]. Here the fermionic vertex was decomposed into different functions in charge, spin and pairing channel and with s- and d-wave formfactors that depend only on one specific wavevector transfer or the total incoming wavevector (as discussed in Sect. 12.3. While this seems a good approximation in the parameter region further away from the transition between d-wave pairing and ferromagnetism, is seems to overestimate the critical scale in the region near $t' \approx -1/3$ and rather points to a first-order transition. Nevertheless this decomposition of the vertex can be significantly improved by also treating a remainder term capturing the previously ignored higher formfactors [58, 59] and represents a promising direction for an improved wavevector resolution of the interaction vertex.

Summarizing these results on the one-band Hubbard model we see that the fRG is capable of deriving tentative phase diagrams with detailed descriptions of the wavevector-structure (and in principle also frequency-structure) of the effective interactions. Next to establishing broader regime with clearcut instabilities toward phases with unconventional Cooper pairing, the method also shows that the effective interactions at the borders of these regimes are rather complex. To understand the physical meaning of these flows with several strong channels in the effective interaction and to relate them to observable phenomena is an interesting challenge for the future research.

12.4.2 Iron Pnictides

A new field where fRG has contributed already considerably are the newly discovered iron pnictide superconductors [5–7]. Here the applicability of a perturbative technique like the one described here may be even better, as the pncitides are

certainly not as strongly correlated as the high-T_c cuprates. This can already be inferred from the experimental phase diagram, where one only finds metallic antiferromagnetic phases (if at all), but never Mott insulating antiferromagnetism. Further, theoretical works that try to assess the value of the iron-d orbital onsite-interaction strengths find values that put the materials into the range of weak to moderate correlations [70, 71]. Regarding the electronic structure, the pnictides are, however, more complex than the cuprates. The main reason is that at least three of the five iron d-orbital have non-negligible weight near the Fermi level [72, 73], and that these d-orbitals hybridize strongly with the neighboring arsenic p-orbitals. Therefore, even if one is only interested in the vicinity of the Fermi surface, the multiband character has to be kept. The Fermi surface is divided into two hole pockets, centered around the origin of the Brillouin zone at $\boldsymbol{k} = 0$, and two electron pockets around $\boldsymbol{k} = (\pi, 0)$ and $\boldsymbol{k} = (0, \pi)$ in the unfolded zone corresponding to the small unit cell with one iron atom (or $\boldsymbol{k} = (\pi, \pi)$ in the folded zone corresponding to the large unit cell with two iron atoms). As pointed out early [72, 74], there is a potential nesting of electron and hole pockets which enhances particle–hole susceptibilities with the wavevector connecting these pockets. In addition, depending on the parameters and approximations [75], there can be a third hole pocket at (π, π) in the unfolded zone.

The first fRG work on the pncitides was performed by Wang et al. [60] for a five-band model. These authors obtained a sign-changing s-wave pairing instability driven by AF fluctuations as the dominant pairing instability. Further they found strongly anisotropic gaps around the electron pockets, with possibility of node formation. The basic structure of the phase diagram with the sign-changing pairing gap between electron- and hole-pockets can be understood already from simplified few-patch RG approaches [76]. This would, however, predict isotropic gaps around these pockets [77]. To understand the gap anisotropy one has to take into account the multiorbital nature of the electronic spectrum in the iron pnictides, as was done by the Berkeley group in their initial study [60]. Let us start with a single-particle Hamiltonian in wavevector-orbital space

$$H = \sum_{\boldsymbol{k},s,o} h(\boldsymbol{k})_{oo'} c^{\dagger}_{\boldsymbol{k},o,s} c_{\boldsymbol{k},o',s}, \tag{12.24}$$

where the matrices $h(\boldsymbol{k})_{oo'}$ take into account intra- and inter-orbital (density–density interactions and Hund's rule) terms for orbital index $o = o'$ or $o \neq o'$ respectively. s is the spin quantum number. The energy bands are obtained by a unitary transformation from orbital to band operators (index b), $c_{\boldsymbol{k},b,s} = \sum_o u_{\mathrm{bo}}(\boldsymbol{k}) c_{\boldsymbol{k},o',s}$. For the various density-functional-theory-based tight-binding parameterizations of the band structure (e.g., [71, 78–80]) of the d-dominated bands, three orbitals d_{xy}, d_{xz} and d_{yz} (in the coordinates of the small cell, with x- and y-axis pointing toward the nearest iron neighbor) contribute significantly near the Fermi level. The simplest choice for the interaction between the electrons is to introduce orbital-dependent

intra- and inter-orbital onsite repulsions, plus Hund's rule and pair hopping terms. While these local terms lead to wavevector-independent interactions in the orbital basis, parametrized by a tensor $V_{o1,o2,o3,o4}$, after the transformation to bands one arrives at a wavevector-dependent interaction function

$$V_{b1,b2,b3,b4}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) = \sum_{o1,o2,o3,o4} V_{o1,o2,o3,o4} \cdot u_{b1,o1}(\boldsymbol{k}_1)u_{b2,o2}(\boldsymbol{k}_2)u^*_{b3,o3}(\boldsymbol{k}_3)u^*_{b4,o4}(\boldsymbol{k}_4). \tag{12.25}$$

The combination of u_{bo}s behind the interaction tensor is sometimes called the "orbital make-up" [78, 79]. These prefactors cause in practice that already the initial interaction of the fRG flow exhibits a marked wavevector-structure which is then renormalized during the flow. It turns out that this orbital make-up has an essential influence on the competition between different channels in the flow and is responsible for the gap anisotropies found in the multiband fRG studies by Fa Wang and collaborators [60, 81, 82] and in subsequent studies [83]. A typical result for the predicted pairing gaps are shown in the right plot of Fig. 12.7.

Summarizing this brief section, the iron superconductors pose an interesting problem to the fRG where the main ordering tendencies have been calculated in good agreement with experiments, at least according to the currently accepted picture. For the future research, one goal should be to make the fRG a useful bridge between ab initio descriptions and experimental observables, in particular regarding materials trends in, e.g., the gap structure or the energy scales of the different systems. Furthermore, the studies should be extended to include the dispersion

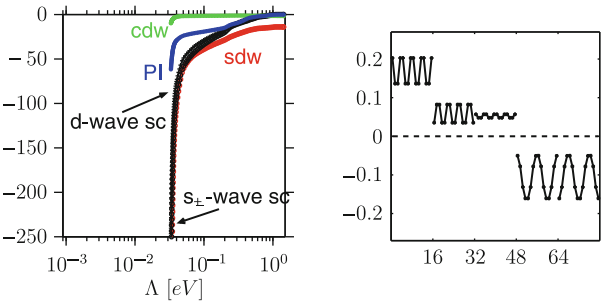

Fig. 12.7 N-patch fRG results for the five-pocket scenario for FeAs-compounds. *Left plot*: fRG flow of the leading ordering tendencies. The sign-changing s-wave channel is competing with SDW and d-wave pairing. *Right plot*: Leading eigenvector of the Cooper pair scattering around the Fermi pockets, showing the sign change between hole- (patch indices 1–48) and electron-pockets (indices 49–80) and the anisotropy of the suggested gap function. Data taken from [83], interaction parameters $U = 3.5\,\mathrm{eV}$, $U' = 2.0\,\mathrm{eV}$, $J_{\mathrm{H}} = J_{\mathrm{pair}} = 0.7\,\mathrm{eV}$

orthogonal to the iron-pnictide planes, as this would yield additional possibilities for nodes in the gap function [5–7].

12.5 Remarks on the 1PI fRG Scheme

After describing two of the main applications of the fRG for 2D correlated electron systems we discuss some formal issues to clarify the 1PI fRG scheme and to compare with other RG approaches. This will also introduce some important issues for the further development of the method.

12.5.1 Differences to Standard Wilsonian RG

In the formalism described in this chapter, only the first stage of a standard Wilsonian RG procedure, the integrating-out of modes has been performed. This first step may be considered the essential physical ingredient of all Wilson-inspired RG schemes, thus the fRG schemes described here do not fall into any different class. Regarding the other two standard stages of Wilsonian RG (see, e.g., [18, 27]), some comments are in place. The second stage usually refers to a momentum- or wavevector-rescaling. This is typically not done in the context described here. The most natural procedure for a many-fermion system would be a rescaling of the wavevector-component perpendicular to the Fermi surface. However, a simple calculation shows that this in general spoils wavevector conservation which we would like to keep, for instance to resolve the anisotropy of the effective interactions. For frequency rescaling there are no such counter-arguments, at least at zero temperature, but all the works described here keep the frequency axis unchanged during the flow. The main purpose for using the renormalization group on the works described here is to sum up the perturbation series in an unbiased way. For this, momentum or frequency-rescaling seems an unnecessary complication. Further, also the field rescaling to keep the quadratic term unchanged, commonly referred to as stage three of Wilsonian RG step, is usually not done in our context. The whole setup of the scheme in general does not depend on whether the system is close to a fixed point or certain parts of the action maintain the same form at all scales such that a field rescaling is not needed. In particular for systems with the Fermi level near a van Hove singularity, the dispersion is not scale-invariant, but the fRG equations still give a valid description of the low-energy effective action – the main target in the present context. It might be that some of these issues have to be changed if one aims at describing fixed points of the flow in more detail, and other flow schemes using all three stages are discussed in the literature [18]. Most of the applications described here concentrate on determining the leading instabilities down to a scale where the truncation error is expected to be unimportant, and for this purpose the present formalism appears to be sufficient.

12.5.2 Higher Loops

Another common question is where higher-loop contributions that one knows from other RG approaches are hidden. Note that in the exact hierarchy of flow equations for the 1PI vertices, no higher-loop diagrams occur on the right-hand side. It is, however, not difficult to see that the solution of the one-loop fRG equations contains contributions of all orders in the bare interactions, and with an arbitrarily high number of loops. To see this for the coupling function V_Λ, one has to insert the solution of the flow equation for V_Λ at scale Λ, schematically (supressing wavevector and frequency summations) given by

$$V_\Lambda = \int_{\Lambda_0}^{\Lambda} \mathrm{d}\Lambda' \, \dot{V}_{\Lambda'} = \int_{\Lambda_0}^{\Lambda} \mathrm{d}\Lambda' \sum_a V_{\Lambda'} L^a_{\Lambda'} V_{\Lambda'}$$

with the different particle–hole and particle–particle one-loop diagrams L_a for least one vertex into the "β-function" for the coupling function on the right-hand side, leading to

$$\dot{V}_\Lambda = \sum_a V_\Lambda L^a_\Lambda V_\Lambda = \int_{\Lambda_0}^{\Lambda} \mathrm{d}\Lambda' \sum_a V_\Lambda L^a_\Lambda \sum_{a'} V_{\Lambda'} L^{a'}_{\Lambda'} V_{\Lambda'} \, .$$

The resulting expression corresponds to diagrams with three vertices at different scales and four internal propagators (two coming from each L^a_Λ). These diagrams are either "chains" of two one-loop diagrams, e.g., when L^a_Λ and $L^{a'}_{\Lambda'}$ are both particle–particle diagrams, or true two-loop terms, e.g., when L^a_Λ is a particle–particle loop and $L^{a'}_{\Lambda'}$ is of particle–hole type. Of course, this procedure of reinserting integrated one-loop terms for vertices can be iterated, leading to diagrams of arbitrarily high order in bare interactions and loops. Similar procedures can be used to assess two-loop self-energy effects that otherwise would be lost due to the neglect of the frequency-dependence of the vertices, see e.g., [84–88].

12.5.3 Connection to Infinite-Order Single-Channel Summations

The integration of the 1PI equations in the present truncation contains all standard ladder and bubble summations known from perturbation theory. Hence the RG is considered to give a rather unbiased picture of the leading ordering tendencies, even in the case of competing ordering tendencies. The connection to, e.g., ladder summations in the particle–particle channel can be seen if we drop all particle–hole terms on the right-hand side of the flow equation for the interaction. Then the remaining equation has the structure

$$\dot{V}_\Lambda(k_1,k_2,k_3) = V_\Lambda(k_1,k_2,k) L^{\mathrm{PP}}_\Lambda(k,-k+k_1+k_2) V_\Lambda(k,-k+k_1+k_2,k_3)\,.$$

Again, the sum over k is not written out. For simplicity, let us assume that the initial condition $V_{\Lambda_0}(k_1,k_2,k_3)$ does not depend on the wavevectors or frequencies, and let us focus on the channel $k_1 + k_2 = 0$, i.e., zero total wavevector and zero total momentum. Near the Fermi level and at $T = 0$, $L^{\mathrm{PP}}_\Lambda(k,-k)$ typically goes like $1/\Lambda$, with $T\sum_k L^{\mathrm{PP}}_\Lambda(k,-k) \approx \rho_0/\Lambda$, as the particle–particle bubble at $k_1 + k_2 = 0$ diverges logarithmically in the infrared cutoff. These simplifications lead to the equation

$$\dot{V}_\Lambda = \frac{\rho_0}{\Lambda} V_\Lambda^2 \qquad \longrightarrow \qquad V_\Lambda = \frac{V_{\Lambda_0}}{1 + \rho_0 V_{\Lambda_0} \log \frac{\Lambda_0}{\Lambda}}$$

If the initial vertex V_{Λ_0} is attractive, this equation has a pole at $\Lambda_\mathrm{c} = \Lambda_0 \exp[-1/(\rho_0 V_{\Lambda_0}]$. Hence the RG equation exactly reproduces the well-known Cooper instability. Likewise, all other particle–particle and particle–hole instabilities obtained in ladder or bubble summations can be recovered in suitable simplifications of the full one-loop flow. Some more attention has to be paid for the case of particle–hole instabilities at small wavevector transfer, like ferromagnetic Stoner instabilities. As mentioned in Sect. 12.2.3, the "deficit by construction" in these cases can, however, be remedied by other flow parameters such as the temperature [56], interaction strength [57] or a smooth frequency cutoff [58].

12.5.4 Symmetry-Breaking: Connection to Mean-Field and Eliashberg Theory

Typically, the fRG flows in 2D lattice models lead to strong coupling, at least if the temperature is low enough and if the density of states is nonzero at the Fermi level. This means that at least one class of interaction processes becomes very strong at low scales. Very often, as discussed in Sect. 12.4, this flow to strong coupling suggests a symmetry breaking in a particular channel, accompanied with a gap opening in the single-particle spectrum. Hence, instead of just exploring which type of symmetry breaking is indicated by the flow, an improved flow would also describe the growth of the order parameter and the change in the excitation spectrum. In the fermionic models discussed here, there are various ways to pursue this idea.

One idea is *partial bosonization*, which means to perform a Hubbard–Stratonovitch transformation of the leading interaction terms (written as a square of appropriate fermionic bilinears) to an appropriate bosonic field. This field couples to the fermionic bilinear and acts like an additional self-energy accounting for the change in the dispersion relation. In a next step, the gapped fermions can be integrated out, and the resulting bosonic theory can be studied. This approach is very useful if the fermionic flow reveals a clearly defined collective boson, and if

one is primarily interested in the collective dynamics at energy scales below the fermionic gap. A more detailed account of this method is beyond the scope of this chapter and the reader is referred to original papers [30, 65, 89, 90]. Interestingly, for the bosonic sector, the full effective potential can be analyzed [91]. The bosonic 1PI flows in vertex expansion were e.g., discussed by Kopietz and collaborators [18, 29].

Another less sophisticated way to describe the symmetry-broken phase would be to stop the fRG flow at a certain scale Λ_s where the interactions have built up a definite structure. Then one can use a mean-field decoupling of the leading terms and study the self-consistent mean-field solutions for the electronic modes below Λ_s. This method has been shown to work reasonably for the repulsive Hubbard model on the square lattice [92]. One might ask whether the outcome depends strongly on the scale Λ_s where the transition to the mean-field description is made. Here one is, however, helped by the fact that, e.g., in the simple BCS model with bare attraction V_{Λ_0}, initial bandwidth Λ_0 and constant density of states ρ_0, the scale-dependent attraction grows like $V_\Lambda = V_{\Lambda_0}/\log(\Lambda/\Lambda_c)$ toward the critical scale $\Lambda_c = \Lambda_0 \exp(-1/V_{\Lambda_0}\rho_0)$. The BCS gap is then given by $\Delta = 2\Lambda_0 \exp(-1/V_{\Lambda_0}\rho_0)$ provided that the dimensionless coupling constant $V_{\Lambda_0}\rho_0$ is smaller than unity. This dependence means that the growth of V_{Λ_s}, when the bandwidth $\Lambda_0 = \Lambda_s$ is changed to a smaller $\Lambda_{s'}$, is canceled exactly by the shrinkage of the bandwidth, leaving the gap value unchanged. In more complex models like the repulsive Hubbard model, these dependences are altered by the particle–hole terms and the Λ_s-independence is no longer true. Nevertheless, in the study of [92], the Λ_s-dependence turned out to be tolerable, and the scheme gave a reasonable phase diagram with gap magnitudes that compare well with other approaches.

A more ambitious, fully fermionic approach is to allow for a tiny symmetry-breaking in the initial condition for the self-energy. Then the flow of this self-energy needs to be treated together with the flow of the interactions. The symmetry breaking might create new types of vertices. Then, at the critical scale Λ_c, the anomalous self-energy will grow rapidly and alter the flow of the interactions, preventing a true divergence. Therefore all modes can be integrated out down to $\Lambda = 0$, and the scheme provides a renormalized single-particle excitation spectrum in the symmetry-broken phase. These schemes were first tried out using momentum-shell cutoffs for simplified BCS-pairing [28] and charge-density-wave models [93] with strongly restricted interactions (e.g., only particle pairs with total momentum zero interact) for which mean-field theory and also the truncation of the fRG hierarchy is exact. It turned out that a reordering of the hierarchy for the flow of the 1PI vertex functions is essential to obtain correct results. More precisely, in the so-called Katanin-modification [94] one replaces the product of a scale-dependent Green's function and a single-scale propagator, $G_\Lambda S_\Lambda$ by $G_\Lambda \dot{G}_\Lambda$. While the single-scale-propagator only contains a scale-derivative of the cutoff-function, the latter term also contains in addition the derivative of the self-energy $\propto \dot{\Sigma}_\Lambda$. This term represents a part of the feedback of the formerly neglected six-point vertex on the flow of the four-point vertex [28]. Katanin showed that keeping these additional contributions is important to fulfill certain Ward identities better, or even exactly in the case of

simplified mean-field models. Physically, the role of these additional contributions is as follows. The flow is started with a very small initial symmetry breaking or gap parameter. This means that high-energy modes get integrated out with a gap that is different from the final one at the end of the flow. The additional $\dot{\Sigma}_\Lambda$-term contributions from all scales above Λ and corrects the contribution these of higher modes when the anomalous self-energy changes during the flow. With this Katanin-modification the "gap-flow" produces the exact BCS results when all scales are integrated out and in the limit of initial symmetry breaking going to zero. One can also nicely see diagrammatically that the fRG solution for this model is fully equivalent to solving the BCS gap equation [28]. For models with less restricted interactions, mean-field theory is no longer exact, and the fRG includes a number of corrections to mean-field theory, such as the suppression of the pairing strength through one-loop particle–hole diagrams [99]. One can also show how standard Eliashberg theory [100–102] is contained as a specific approximation in the fRG scheme (see [103] for the connection to Eliashberg theory above and at T_c, and [104] for a general argument including the symmetry-broken phase). Further, the approach in the symmetry-broken phase can be adapted to search for first-order transitions [105] that are not detected by run-away flows in the symmetric phase.

The fate of the running interactions depends on the type of symmetry-breaking. For breaking of a discrete symmetry, the interactions become smaller again at scales below the rapid growth of the self-energy (the gap opening) [93]. For continuous symmetry breaking as in the case of the BCS model, one can distinguish between an "amplitude mode" in the vertex that also flows back to smaller values below the gap opening, and a "Goldstone mode" that for $\Lambda \to 0$ basically saturates at a value inversely proportional to the initial symmetry breaking [28]. Hence, for realistic situations, these flows always produce large interactions with rather sharp momentum and frequency structures. It is by no means clear that the simple patching schemes described in this chapter are able to capture this behavior correctly, plus the truncation error might become more important. These issues have slowed down the application of these fermionic flows into the symmetry-broken phase for models where mean-field theory is not exact. Nevertheless the scheme has been tried in the attractive Hubbard model [95] and for competing orders in one dimension [96]. In both cases the results were rather promising, such that future work using this approach might provide more results.

12.5.5 Normal-State Self-Energy

In most applications of the N-patch fRG, the self-energy flow has been ignored completely. By now several aspects of the normal state self-energy $\Sigma_\Lambda(\boldsymbol{k}, \omega_n)$ have been studied. $\Sigma_\Lambda(\boldsymbol{k}, \omega_n)$ can in principle be computed from (12.13). The information contained in the result depends, however, on the approximation employed for the vertex or coupling functions appearing on the right-hand side. In most N-patch studies, the frequency dependence of the vertex is neglected and hence $\Sigma_\Lambda(\boldsymbol{k}, \omega_n)$

comes out frequency-independent and usually real. This level of approximation is still enough to obtain information on the flow of the Fermi surface. A first study was presented in [55] for the one-band Hubbard model, where the flow of the Fermi surface turned out to be rather mild with a tendency toward a flatter Fermi surface down to a scale where the flow to strong coupling occurred. While this gave some first insights, it also emphasized some formal problems that arise. First of all, the particle number usually changes when the self-energy flow is included, and for flows with fixed particle number one has to cope with a scale-dependent chemical potential. Then, in momentum-shell schemes, it might happen that the renormalized Fermi surface flows into regions that have been integrated out already. This can lead to divergences. These problems have been addressed so far on a conceptual level [97,98], but have not yet led to a clear picture on the Fermi surface renormalization in the t–t' Hubbard model. Based on the experience in [55] we do not believe that this aspect changes the obtained phase diagrams in the one-band model considerably. It might become more important in multiband models, where the orbital mixture of the bands enters the interaction functions, as explained in Sect. 12.4, and the interplay between orbital weights and Fermi surface effects is more subtle.

To obtain the imaginary part and also the frequency dependence of the self-energy, one can reinsert the solution of the one-loop flow equation for the interaction into the flow equation of the self-energy, similar to what is described in Sect. 12.5.2. This leads to a frequency-dependent two-loop contribution with nonzero imaginary part on the real frequency axis. From this, quasiparticle scattering rates, quasiparticle weights as frequency derivatives and the whole renormalization of the spectral function can be computed.

The scattering rates in the one-band Hubbard model have been addressed to some extent in [84], where the anisotropy was described as function of the Fermi surface shape and of the temperature. If the Fermi surface is located near the $(\pi, 0)$, $(0, \pi)$ points of the Brillouin zone, the scattering rate is usually strongly anisotropic and higher near $(\pi, 0)$, $(0, \pi)$ than near the Brillouin zone diagonal. Interestingly, some experiments [106, 107] point to a positive correlation between the anisotropic component of the transport scattering rate and the superconducting T_c. This trend comes out of the fRG calculation of the quasiparticle scattering rates as well [86]. The reason for this trend are again the scattering processes between the saddle points that are responsible for both the d-wave pairing and the enhanced self-energy in these regions.

The flow of the quasiparticle weight, the Z-factor, determined by the frequency derivative of the self-energy at zero frequency, has also been studied with similar techniques in at least two works [85, 109]. Here the suppression of the quasiparticle weight is strongest near the saddle points, but is not strong enough to change the leading flows to strong coupling. More drastic effects are found when the full frequency dependence of the self-energy on the real axis is considered. This has been studied in the one-band Hubbard model in [88, 108], again using two-loop self-energies. These works showed that the real part of the retarded self-energy develops an additional kink at low frequencies near the instability scale and near the saddle

points. This leads, if strong enough, to an anisotropic split-up of the quasiparticle peak. These works also showed that a parametrization of the self-energy to a Z-factor is in principle insufficient. The trends toward an anisotropic destruction of the quasiparticles are quite promising in connection with the phenomenology of high-T_c cuprates, but to date no firm connection between these weak coupling trends and observable phenomena could be established. In particular more work is needed that allows one to assess the precision of the fRG results very close to the instability. Note also that in all these works on the frequency dependence and on the imaginary part of the self-energy, the self-energy feedback on the flow of the vertex function was ignored beyond the inclusion of Z-factors in a few works [85, 87, 109]. In this regard and on a technical level, some fRG studies performed in impurity systems [61] and for quantum spin [38] are more advanced, as they include the self-energy feedback at least on some approximate level into the flow of the vertices. It is mainly due to the rich wavevector structure of the interactions that for 2D many-electron lattice systems the full self-energy feedback has not been tackled yet.

12.5.6 Refined Studies

Here we mention two refinements of the usual N-patch approach that give additional support to the picture established in the simplest approximation. The first issue is the frequency dependence of the vertex. Here, a simple patching of the Matsubara frequency axis into a number of segments has been introduced in [110, 111]. This was employed in [110] to study the impact of various phonon-mediated electron–electron interactions on the flow to strong coupling in the t–t'-Hubbard model. Regarding the basic trends toward AF-SDW and d-wave pairing instabilities, the refined treatment did not result in qualitative changes. However, it provided some information on the frequency structure of the pairing interaction. Its width as function of the frequency transfer between incoming and outgoing legs can be interpreted as characteristic frequency of the "pairing boson." This turned out to very similar to the frequency width of the spin susceptibility at wavevector (π, π), again corroborating the picture of d-wave pairing mediated by AF spin fluctuations.

Another important and very useful step to improve the previous N-patch results for the t–t'-Hubbard model was undertaken by Katanin [87]. He included into the flow of the four-point vertex the integrated solution of a term in third-order of the four-point vertex that in the usual 1PI flow hierarchy contributes to the flow of the six-point vertex. This goes beyond the usual truncation. Katanin also computed real and imaginary part of the self-energy at the Fermi surface and included the self-energy into the flow via a Z-factor. Summarizing his results, Katanin found that these corrections do not change the leading instability but may lead to a slight shift of the phase boundaries in comparison with the previous one-loop analysis. In agreement with earlier studies [85], for curved Fermi surface and not too low temperatures the quasiparticle weight was only mildly reduced so that the quasiparticles remain well defined during the fRG flow. The quasiparticle scattering

and Fermi surface shifts came out numerically small. In this sense, for most of the parameter space, the commonly used one-loop fRG approach without self-energy corrections should be expected to give a correct physical picture about the types and characteristic scales of the leading instabilities.

12.6 Conclusions and Outlook

The fRG for interacting fermions is a widely applicable and flexible method based on the safe and powerful foundation of exact flow equations for generating functionals. In this chapter we have focussed on the use of this method for investigating the leading low-temperature ordering tendencies of correlated electrons on 2D lattices. For the one-band Hubbard model, one main result of the method has been to show that there is a wide parameter regime for a d-wave Cooper-pairing instability, and to map out the other main instabilities depending on the model parameters. While most of the studies showing these instabilities still involve many approximations like the neglect of self-energy effects, all available information indicates that the obtained picture is qualitatively correct. The next steps will bring more detailed investigations of the effective interactions and of self-energy effects. Another important extension would be a detailed treatment of the symmetry-broken state for relevant models like the 2D repulsive Hubbard model, where to the flow is continued down to $\Lambda = 0$. Then the fRG can not only give information on the leading instabilities, but also provide renormalized excitation spectra and correlation functions. For pnictide superconductors, the fRG has proven to be a method that allows one study how details of the microscopic model, like the orbital compositions of the bands, lead to changes in the effective interactions and the predicted pairing gaps. Here the task will be to describe experimentally observed material-specific properties (e.g., the presence or absence of magnetically ordered phases or superconducting gap nodes and their location on the Fermi surfaces) in a reliable but still transparent way. This will again require some further developments of the fRG method and its implementation, as in these complex multiband systems the observables of interest obtained so far might still depend on the degree of approximation.

Besides these two fields of applications described here, there are many other activities in condensed matter physics using fRG methods, some of them were mentioned in the Introduction. In particular the large freedom to choose the RG flow trajectory will lead to many other occasions where the same fundamental flow equations will lead to interesting and relevant results.

References

1. A.J. Leggett, Nat. Phys. **2**, 134 (2006)
2. D.A Bonn, Nat. Phys. **2**, 159 (2006)

3. P. Coleman, in *Handbook of Magnetism and Advanced Magnetic Materials*, vol. 1, ed. by H. Kronmuller, S. Parkin (Wiley, New York, 2007), p. 95
4. C. Pfleiderer, Rev. Mod. Phys. **81**, 115 (2009)
5. M.R. Norman, Physics **1**, 21 (2008)
6. K. Ishida, Y. Nakai, H. Hosono, J. Phys. Soc. Jpn. **78**, 062001 (2009)
7. P.J. Hirschfeld, D.J. Scalapino, Physics **3**, 64 (2010)
8. P.W. Anderson, Mater. Res. Bull. **8**, 153 (1973)
9. P. Fazekas, P.W. Anderson, Philos. Mag. **30**, 432 (1974)
10. P.W. Anderson, Science **235**, 1196 (1987)
11. P.A. Lee, N. Nagaosa, X.G. Wen, Rev. Mod. Phys. **78**, 17 (2006)
12. T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, M.P.A. Fisher, Science **303**, 1490 (2004)
13. T. Senthil, S. Sachdev, M. Vojta, Phys. Rev. Lett. **90**, 216403 (2003)
14. X.L. Qi, S.C. Zhang, arXiv:1008.2026; M.Z. Hasan, C.L. Kane, Rev. Mod. Phys. **82**, 3045 (2010)
15. J. Polchinski, Nucl. Phys. B **231**, 269 (1984)
16. C. Wetterich, Phys. Lett. B **301**, 90 (1993)
17. M. Salmhofer, *Renormalization. An Introduction* (Springer, Berlin, 1999)
18. P. Kopietz, L. Bartosch, F. Schütz, *Introduction to the Functional Renormalization Group*. Springer Lecture Notes in Physics 798 (Springer, Heidelberg, 2010)
19. K.G. Wilson, J. Kogut, Phys. Rep. C **12**, 75 (1974)
20. F.J. Wegner, A. Houghton, Phys. Rev. A **8**, 401 (1973)
21. H. Ballhausen, J. Berges, C. Wetterich, Phys. Lett. B **582**, 144 (2004)
22. D.F. Litim, Dario Zappala, Phys. Rev. D **83**, 085009 (2011)
23. J.A. Hertz, Phys. Rev. B **14**, 1165 (1976)
24. A.J. Millis, Phys. Rev. B **48**, 7183 (1993)
25. H.v. Löhneysen, A. Rosch, M. Vojta, P. Wölfle, Rev. Mod. Phys. **79**, 1015 (2007)
26. P. Strack, S. Takei, W. Metzner, Phys. Rev. B **81**, 125103 (2010)
27. R. Shankar, Rev. Mod. Phys. **66**, 129 (1994)
28. M. Salmhofer, C. Honerkamp, W. Metzner, O. Lauscher, Prog. Theor. Phys. **112**, 943 (2004)
29. F. Schütz, L. Bartosch, P. Kopietz, Phys. Rev. B **72**, 035107 (2005)
30. P. Strack, R. Gersch, W. Metzner, Phys. Rev. B **78**, 014522 (2008)
31. L. Mathey, S.W. Tsai, A.H. Castro-Neto, Phys. Rev. Lett. **97**, 030601 (2006)
32. T. Machado, N. Dupuis, Phys. Rev. E **82**, 041128 (2010)
33. A. Isidori, D. Roosen, L. Bartosch, W. Hofstetter, P. Kopietz, Phys. Rev. B **81**, 235120 (2010)
34. V. Meden, S. Andergassen, T. Enss, H. Schoeller, K. Schoenhammer, New J. Phys. **10**, 045012 (2008)
35. S. Andergassen, V. Meden, H. Schoeller, J. Splettstoesser, M. R. Wegewijs, Nanotechnology **21**, 272001 (2010)
36. S. Diehl, S. Floerchinger, H. Gies, J. M. Pawlowski, C. Wetterich, Ann. Phys. (Berlin) **522**, 615 (2010)
37. S. Moroz, S. Floerchinger, R. Schmidt, C. Wetterich, Phys. Rev. A **79**, 042705 (2009)
38. J. Reuther, P. Wölfle, Phys. Rev. B **81**, 144410 (2010)
39. R. Gezzi, Th. Pruschke, V. Meden, Phys. Rev. B **75**, 045324 (2007)
40. S.G. Jakobs, V. Meden, H. Schoeller, Phys. Rev. Lett. **99**, 150603 (2007)
41. J. Berges, G. Hoffmeister, Nucl. Phys. B **813**, 383 (2009)
42. T. Morris, Int. J. Mod. Phys. **9**, 2411 (1994)
43. M. Kinza, J. Ortloff, C. Honerkamp, Phys. Rev. B **82**, 155430 (2010)
44. Tilman Enss, Ph.D. Thesis, University of Stuttgart, arXiv:cond-mat/0504703
45. D. Zanchi, H.J. Schulz, Europhys. Lett. **44**, 235 (1998)
46. D. Zanchi, H.J. Schulz, Phys. Rev. B. **61**, 13609 (2000)
47. S.-W. Tsai, J.B. Marston, Can. J. Phys. **79**, 1463 (2001); arXiv:cond-mat/0010300 (unpublished)
48. C. Bourbonnais, B. Guay, R. Wortis, in *Theoretical Methods for Strongly Correlated Electrons*. CRM Series in Mathematical Physics (Springer, New York, 2003)

49. M. Salmhofer, C. Honerkamp, Prog. Theor. Phys. **105**, 1 (2001)
50. J.W. Negele, H. Orland, *Quantum Many-particle systems*. Advanced Book Classics, ISBN: 0738200522, Westview Press (1998)
51. I.E. Dzyaloshinskii, V.M. Yakovenko, Sov. Phys. JETP **67**, 844 (1988); Int. J. Mod. Phys. B **2**, 667 (1988)
52. A.T. Zheleznyak, V.M. Yakovenko, I.E. Dzyaloshinskii, Phys. Rev. B **55**, 3200 (1997)
53. S.G. Jakobs, V. Meden, H. Schoeller, Phys. Rev. Lett. **99**, 150603 (2007)
54. C.J. Halboth, W. Metzner, Phys. Rev. B **61**, 7364 (2000); Phys. Rev. Lett. **85**, 5162 (2000)
55. C. Honerkamp, M. Salmhofer, N. Furukawa, T.M. Rice, Phys. Rev. B **63**, 035109 (2001)
56. C. Honerkamp, M. Salmhofer, Phys. Rev. B **64**, 184516 (2001); Phys. Rev. Lett. **87**, 187004 (2001)
57. C. Honerkamp, D. Rohe, S. Andergassen, T. Enss, Phys. Rev. B **70**, 235115 (2004)
58. C. Husemann, M. Salmhofer, Phys. Rev. B **79**, 195125 (2009)
59. J. Ortloff, C. Husemann, C. Honerkamp, M. Salmhofer (in preparation)
60. F. Wang, H. Zhai, Y. Ran, A. Vishwanath, D.-H. Lee, Phys. Rev. Lett. **102**, 047005 (2009)
61. C. Karrasch, R. Hedden, R. Peters, Th. Pruschke, K. Schönhammer, V. Meden, J. Phys. Condens. Matter **20**, 345205 (2008)
62. F.C. Zhang, T.M. Rice, Phys. Rev. B **37**, 3759 (1988)
63. C. Proust, E. Boaknin, R.W. Hill, L. Taillefer, A.P. Mackenzie, Phys. Rev. Lett. **89**, 147003 (2002)
64. W. Kohn, J.H. Luttinger, Phys. Rev. Lett. **15**, 524 (1965)
65. T. Baier, E. Bick, C. Wetterich, Phys. Rev. B **70**, 125111 (2004)
66. H.H. Lin, L. Balents, M.P.A. Fisher, Phys. Rev. B **56**, 6569 (1997); H.H. Lin, L. Balents, M.P.A. Fisher, Phys. Rev. B **58**, 1794 (1998)
67. M.P.A. Fisher, in *Topological Aspects of Low Dimensional Field Theories*, ed. by A. Comtet, T. Jolicoeur, S. Ouvry, F. David. Les Houches Lecture Session LXIX (Springer, New York, 1999)
68. N. Furukawa, T.M. Rice, M. Salmhofer, Phys. Rev. Lett. **81**, 3195 (1998)
69. A. Läuchli, C. Honerkamp, T.M. Rice, Phys. Rev. Lett. **92**, 037006 (2004)
70. V. I. Anisimov, Dm.M. Korotin, M. A. Korotin, A.V. Kozhevnikov, J. Kunes, A.O. Shorikov, S.L. Skornyakov, S.V. Streltsov, J. Phys. Condens. Mat. **21**, (2009) 075602
71. T. Miyake, K. Nakamura, R. Arita, M. Imada, J. Phys. Soc. Jpn. **79**, 044705 (2010)
72. I.I. Mazin, D.J. Singh, M.D. Johannes, M.H. Du, Phys. Rev. Lett. **101**, 057003 (2008)
73. M. Daghofer, A. Nicholson, A. Moreo, E. Dagotto, Phys. Rev. B **81**, 014511 (2010)
74. K. Kuroki, S. Onari, R. Arita, H. Usui, Y. Tanaka, H. Kontani, H. Aoki, Phys. Rev. Lett. **101**, 087004 (2008)
75. H. Ikeda, R. Arita, J. Kunes, Phys. Rev. B **81**, 054502 (2010)
76. A.V. Chubukov, D.V. Efremov, I. Eremin, Phys. Rev. B **78**, 134512 (2008)
77. Ch. Platt, C. Honerkamp, W. Hanke, New J. Phys. **11**, 055058 (2009)
78. T.A. Maier, S. Graser, D.J. Scalapino, P.J. Hirschfeld, Phys. Rev. B **79**, 224510 (2009)
79. S. Graser, T.A. Maier, P.J. Hirschfeld, D.J. Scalapino, New J. Phys. **11**, 025016 (2009)
80. K. Kuroki, H. Usui, S. Onari, R. Arita, H. Aoki, Phys. Rev. B **79**, 224511 (2009)
81. F. Wang, H. Zhai, D.-H. Lee, Europhys. Lett. **85**, 37005 (2009)
82. F. Wang, H. Zhai, D-H. Lee, Phys. Rev. B **81**, 184512 (2010)
83. R. Thomale, Ch. Platt, W. Hanke, B.A. Bernevig, arXiv:1002.3599
84. C. Honerkamp, Eur. Phys. J. B **21**, 81 (2001)
85. C. Honerkamp, M. Salmhofer, Phys. Rev. B **67**, 174504 (2003)
86. M. Ossadnik, C. Honerkamp, T.M. Rice, M. Sigrist, Phys. Rev. Lett. **101**, 256405 (2008)
87. A.A. Katanin, Phys. Rev. B **79**, 235119 (2009)
88. A.A. Katanin, A.P. Kampf, Phys. Rev. Lett. **93**, 106406 (2004)
89. H.C. Krahl, J.A. Müller, C. Wetterich, Phys. Rev. B **79**, 094526 (2009)
90. S. Friederich, H.C. Krahl, C. Wetterich, Phys. Rev. B **81**, 235108 (2010)
91. P. Jakubczyk, W. Metzner, H. Yamase, Phys. Rev. Lett. **103**, 220602 (2009)
92. J. Reiss, D. Rohe, W. Metzner, Phys. Rev. B **75**, 075110 (2007)

93. R. Gersch, C. Honerkamp, D. Rohe, and W. Metzner, Eur. Phys. J. B **48**, 349 (2005)
94. A.A. Katanin, Phys. Rev. B **70**, 115109 (2004)
95. R. Gersch, C. Honerkamp, W. Metzner, New J. Phys. **10**, 045003 (2008)
96. M. Ossadnik, C. Honerkamp, arXiv:0911.5047
97. M. Salmhofer, Annalen Phys. **16**, 171 (2007)
98. A.A. Katanin, arXiv:1009.6044
99. L.P. Gorkov, T.K. Melik-Barkhudarov, Sov. Phys. JETP **18**, 1031 (1964)
100. G.M. Eliashberg, Sov. Phys. JETP **11**, 696 (1960)
101. D.J. Scalapino, J.R. Schrieffer, J.W. Wilkins, Phys. Rev. **148**, 263 (1966)
102. J.P. Carbotte, Rev. Mod. Phys. **62**, 1028 (1990)
103. S.-W. Tsai, A.H. Castro Neto, R. Shankar, D.K. Campbell, Phil. Mag. **86**, 2631 (2006)
104. C. Honerkamp, M. Salmhofer, Prog. Theor. Phys. **113**, 1145 (2005)
105. R. Gersch, J. Reiss, C. Honerkamp, New J. Phys. **8**, 320 (2006)
106. M. Abdel-Jawad, M.P. Kennett, L. Balicas, A. Carrington, A.P. Mackenzie, R.H. McKenzie, N.E. Hussey, Nat. Phys. **2**, 821 (2006)
107. M. Abdel-Jawad, J.G. Analytis, L. Balicas, A. Carrington, J.P.H. Charmant, M.M.J. French, A.P. Mackenzie, N.E. Hussey, Phys. Rev. Lett. **99**, 107002 (2007)
108. D. Rohe, W. Metzner, Phys. Rev. B **71**, 115116 (2005)
109. D. Zanchi, Europhys. Lett. **55**, 376 (2001)
110. C. Honerkamp, H.C. Fu, D.-H. Lee, Phys. Rev. B **75**, 014503 (2007)
111. K.-M. Tam, S.-W. Tsai, D.K.Campbell, A.H. Castro Neto, Phys. Rev. B **75**, 161103(R) (2007)

Chapter 13
Two-Particle-Self-Consistent Approach for the Hubbard Model

André-Marie S. Tremblay

Abstract Even at weak to intermediate coupling, the Hubbard model poses a formidable challenge. In two dimensions in particular, standard methods such as the random phase approximation are no longer valid since they predict a finite temperature antiferromagnetic phase transition prohibited by the Mermin–Wagner theorem. The two-particle-self-consistent (TPSC) approach satisfies that theorem as well as particle conservation, the Pauli principle, the local moment and local-charge sum-rules. The self-energy formula does not assume a Migdal theorem. There is consistency between one- and two-particle quantities. The internal accuracy checks allow one to test the limits of validity of TPSC. Here I present a pedagogical review of TPSC along with a short summary of existing results and two case studies: (a) the opening of a pseudogap in two dimensions when the correlation length is larger than the thermal de Broglie wavelength and (b) the conditions for the appearance of d-wave superconductivity in the two-dimensional Hubbard model.

13.1 Introduction

Very few models can describe complex behavior observed in nature with an economy of parameters. The Hubbard model is in this category. It has become the cornerstone of correlated electron physics. On the down side, it is extremely difficult to solve. While it was proposed in 1963 [1–3], the only exact results that we know are in one dimension [4] and in infinite dimension [5]. A variety of approximate approaches to solve this model exist, as can be checked from the table of contents of this volume. The two-particle-self-consistent (TPSC) [6–8]

A.-M.S. Tremblay (✉)
Département de physique, and RQMP, Université de Sherbrooke, Sherbrooke, QC, Canada J1K 2R1

Canadian Institute for Advanced Research, Toronto, ON, Canada
e-mail: tremblay@physique.usherbrooke.ca

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6_13,

approach that is described in the present chapter is in the category of nonperturbative semianalytical approaches. By semianalytical, I mean that while it is possible to find many analytical results, numerical integrations are necessary in the end to obtain quantitative results.

Why should you bother to learn yet another approach? Because in its known regime of applicability it is extremely reliable, as can be judged by benchmark quantum Monte Carlo (QMC) calculations; because it satisfies a number of exact results that control the quality of the approximation and make it physically appealing; and because it gives physical insight into many questions related to the two-dimensional Hubbard model relevant for the high-temperature superconductors and many other materials. As a case study, I discuss in this chapter the physics of pseudogap induced by precursors to long-range order. We will see that this describes the physics of the pseudogap in electron-doped high-temperature superconductors where predictions of TPSC have been verified experimentally. Pseudogap phenomena include the appearance of a minimum in the single-particle spectral weight and density of state at the Fermi level. I will be more precise later.

This chapter offers to the reader a simple pedagogical introduction to this approach along with the case studies mentioned above and a guide to various other problems that have been, or have not yet been, solved with TPSC.

I assume familiarity with the basics of many-body theory, i.e., the canonical formalism, second quantization, many-body Green's functions, response functions and with the Matsubara formalism for finite temperature calculations. Knowledge of functional derivative approaches would be useful for some of the more advanced topics, but it is not essential to learn the important results.

Before you read on, you might be interested to know a little more about the method to decide whether it is worth the effort. TPSC is designed to study the one-band Hubbard model

$$H = -\sum_{ij\sigma} t_{i,j} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}, \tag{13.1}$$

where the operator $c_{i\sigma}$ destroys an electron of spin σ at a site i. Its adjoint $c^{\dagger}_{i\sigma}$ creates an electron and the number operator is defined by $n_{i\sigma} = c^{\dagger}_{i\sigma} c_{i\sigma}$. The symmetric hopping matrix $t_{i,j}$ determines the band structure, which here can be arbitrary. The screened Coulomb interaction is represented by the energy cost U of double occupation. It is also possible to generalize to cases where near-neighbor interactions are included. We work in units where $k_{\mathrm{B}} = 1$, $\hbar = 1$ and the lattice spacing is also unity, $a = 1$. In all numerical calculations, we take as unit of energy the nearest-neighbor hopping $t = 1$.

One of the first concepts that is discussed with the Hubbard model is that of the Mott transition [9]. When dimension is larger than unity, at "strong coupling," large U/t, the states are localized, but at "weak coupling,"small U/t, the states are delocalized. The Mott transition is quite subtle and has been the subject of many

papers. It is discussed in the chapters of Jarrell, Potthoff, Sénéchal and others in this volume.

TPSC is valid from weak to intermediate coupling. Hence, on the negative side, it does not describe the Mott transition. Nevertheless, there is a large number of physical phenomena that it allows us to study. An important one is antiferromagnetic fluctuations in two- or higher-dimensional lattices. A standard random phase approximation (RPA) calculation of the spin susceptibility signals a finite temperature phase transition to antiferromagnetic long-range order. This is prohibited by the Mermin–Wagner theorem [10, 11] that states that in two dimensions you cannot break a continuous symmetry at finite temperature. It is extremely important physically that in two dimensions there is a wide range of temperatures where there are huge antiferromagnetic fluctuations in the paramagnetic state. The standard way to treat fluctuations in many-body theory, RPA, misses this. As we will see, the RPA also violates the Pauli principle in an important way. The composite operator method (COM), described in this volume by Avella and Mancini (see Chap. 3.7), is another approach that satisfies the Mermin–Wagner theorem and the Pauli principle [12–14]. What other approaches satisfy the Mermin–Wagner theorem at weak coupling? The fluctuation exchange approximation (FLEX) [15, 16], and the self-consistent renormalized theory of Moriya–Lonzarich [17–19]. Each has its strengths and weaknesses, as discussed in [7, 20]. Weak coupling renormalization group approaches[1] become uncontrolled when the antiferromagnetic fluctuations begin to diverge [21–24]. Other approaches include the effective spin-Hamiltonian approach [25].

In summary, the advantages and disadvantages of TPSC are as follows. Advantages:

- There are no adjustable parameters.
- Several exact results are satisfied: conservation laws for spin and charge, the Mermin–Wagner theorem, the Pauli principle in the form $\left\langle n_{\uparrow}^{2}\right\rangle = \left\langle n_{\uparrow}\right\rangle$, the local moment and local-charge sum-rules and the f-sum rule.
- Consistency between one- and two-particle properties serves as a guide to the domain of validity of the approach. (Double-occupancy obtained from sum-rules on spin and charge equals that obtained from the self-energy and the Green function.)
- Up to intermediate coupling, TPSC agrees within a few percent with QMC calculations. Note that QMC calculations can serve as benchmarks since they are exact within statistical accuracy, but they are limited in the range of physical parameter accessible because of the sign problem.
- We do not need to assume that Migdal's theorem applies to be able to obtain the self-energy.

[1] See the contribution of Honerkamp in this volume.

The main successes of TPSC that I will discuss include:

- Understanding the physics of the pseudogap induced by precursors of a long-range ordered phase in two dimensions. For this understanding, one needs a method that satisfies the Mermin–Wagner theorem to create a broad temperature range where the antiferromagnetic correlation length is larger than the thermal de Broglie wavelength. That method must also allow one to compute the self-energy reliably. Only TPSC does both.
- Explaining the pseudogap in electron-doped cuprate superconductors over a wide range of dopings.
- Finding estimates of the transition temperature for d-wave superconductivity that were found later in agreement with quantum cluster approaches such as the dynamical cluster approximation.
- Giving quantitative estimates of the range of temperature where quantum critical behavior can affect the physics.

The drawbacks of this approach, that I explain as we go along, are that

- It works well in two or more dimensions, not in one dimension[2] [27].
- It is not valid at strong coupling, except at very high temperature where it recovers the atomic limit [28].
- It is not valid deep in the renormalized classical regime [6].
- For models other than the one-band Hubbard model, one usually runs out of sum-rules and it is in general not possible to find all parameters self-consistently. With nearest-neighbor repulsion, it has been possible to find a way out as I will discuss below.

For detailed comparisons with QMC calculations, discussions of the physics and detailed comparisons with other approaches, you can refer to [7, 20]. You can read [29] for a review of the work related to the pseudogap and superconductivity up to 2005 including detailed comparisons with quantum cluster approaches in the regime of validity that overlaps with TPSC (intermediate coupling).

Section 13.2 introduces TPSC in the simplest physically motivated way and demonstrates the various results that are exactly satisfied. Section 13.3 presents two case studies: the pseudogap and d-wave superconductivity. Many more known results and extensions are summarized in Sect. 13.4. The attractive Hubbard model is in the next to the last Sect. 13.5. I conclude with some open problems in Sect. 13.6.

13.2 The Method

In the first part of this section, I present TPSC as if we were discussing in front of a chalkboard. More formal ways of presenting the results come later.

[2]Modifications have been proposed in zero dimension to use as impurity solver for DMFT [26].

13.2.1 Physically Motivated Approach, Spin and Charge Fluctuations

As basic physical requirements, we would like our approach to satisfy (a) conservation laws, (b) the Pauli principle and (c) the Mermin–Wagner–Hohenberg–Coleman theorem. The standard RPA approach satisfies the first requirement but not the other two. Let us see this. With the charge and spin given by

$$n_i \equiv n_{i\uparrow} + n_{i\downarrow} \tag{13.2}$$

$$S_i^z \equiv n_{i\uparrow}(\tau) - n_{i\downarrow}(\tau), \tag{13.3}$$

the RPA spin and charge susceptibilities in the one-band Hubbard model are given, respectively, by

$$\chi_{\rm sp}(q) = \frac{\chi_0(q)}{1 - \frac{1}{2}U\chi_0(q)} \; ; \; \chi_{\rm ch}(q) = \frac{\chi_0(q)}{1 + \frac{1}{2}U\chi_0(q)} \tag{13.4}$$

with q a short-hand for both wave vector $\mathbf{q}$ and Matsubara frequency and where $\chi_0(q)$ is the Lindhard function that in analytically continued retarded form is, for a discrete lattice of N sites,

$$\chi^{0\rm R}(\mathbf{q},\omega) = -\frac{2}{N}\sum_{\mathbf{k}} \frac{f(\varepsilon_{\mathbf{k}}) - f\left(\varepsilon_{\mathbf{k}+\mathbf{q}}\right)}{\omega + i\eta + \varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}+\mathbf{q}}}. \tag{13.5}$$

In this expression, assuming periodic boundary conditions,

$$\varepsilon_{\mathbf{k}} = \left(-\sum_j e^{i\mathbf{k}\cdot(\mathbf{r}_i - \mathbf{r}_j)} t_{i,j}\right) - \mu \tag{13.6}$$

with the sum over j running over all neighbors of any of the sites i. The chemical potential μ is chosen so that we have the required density.

It is known on general grounds [30] that RPA satisfies conservation laws, but it is easy to check that for a special case. Since spin and charge are conserved, then the equalities $\chi^{\rm R}_{\rm sp}(\mathbf{q}=\mathbf{0},\omega) = 0$ and $\chi^{\rm R}_{\rm ch}(\mathbf{q}=\mathbf{0},\omega) = 0$ for $\omega \neq 0$ follow from the corresponding equality for the non-interacting Lindhard function $\chi^{0\rm R}(\mathbf{q}=\mathbf{0}, \omega) = 0$.

To check that RPA violates the Mermin–Wagner theorem, it suffices to note that if U is larger than $U_c = 2/\chi^{\rm R}_0(\mathbf{q}_{\rm max}, \omega = 0)$, then the denominator $1 - \frac{1}{2}U\chi_0(q)$ of $\chi_{\rm sp}(q)$ can diverge at some wave vector $\mathbf{q}_{\rm max}$ and temperature.

The violation of the Pauli principle requires a bit more thinking. We derive a sum-rule that rests on the use of the Pauli principle and checks that it is violated by RPA to second order in U. First note that if we sum the spin and charge susceptibilities

over all wave vectors $\mathbf{q}$ and all Matsubara frequencies $i\omega_n$,[3] we obtain local, equal-time correlation functions, namely,

$$\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\chi_{\rm sp}(\mathbf{q},i\omega_n)=\left\langle\left(n_\uparrow-n_\downarrow\right)^2\right\rangle=\langle n_\uparrow\rangle+\langle n_\downarrow\rangle-2\langle n_\uparrow n_\downarrow\rangle \tag{13.7}$$

and

$$\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\chi_{\rm ch}(\mathbf{q},i\omega_n)=\left\langle\left(n_\uparrow+n_\downarrow\right)^2\right\rangle-\langle n_\uparrow+n_\downarrow\rangle^2=\langle n_\uparrow\rangle+\langle n_\downarrow\rangle+2\langle n_\uparrow n_\downarrow\rangle-n^2, \tag{13.8}$$

where on the right-hand side, we used the Pauli principle $n_\sigma^2=\left(c_\sigma^\dagger c_\sigma\right)\left(c_\sigma^\dagger c_\sigma\right)=c_\sigma^\dagger c_\sigma-c_\sigma^\dagger c_\sigma^\dagger c_\sigma c_\sigma=c_\sigma^\dagger c_\sigma=n_\sigma$ that follows from $c_\sigma^\dagger c_\sigma^\dagger=c_\sigma c_\sigma=0$. This is the simplest version of the Pauli principle. Full antisymmetry is another matter [31,32]. We call the first of the above displayed equations the local-spin sum-rule and the second one the local-charge sum-rule. For RPA, adding the two sum-rules yields

$$\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\left(\chi_{\rm sp}(\mathbf{q},i\omega_n)+\chi_{\rm ch}(\mathbf{q},i\omega_n)\right) \tag{13.9}$$

$$=\frac{T}{N}\sum_{q}\left(\frac{\chi_0(q)}{1-\frac{1}{2}U\chi_0(q)}+\frac{\chi_0(q)}{1+\frac{1}{2}U\chi_0(q)}\right)=2n-n^2. \tag{13.10}$$

Since the noninteracting susceptibility $\chi_0(q)$ satisfies the sum-rule, we see by expanding the denominators that in the interacting case it is violated already to second order in U because $\chi_0(q)$ being real and positive (see (13.22)), the quantity $\sum_q\chi_0(q)^3$ cannot vanish.

How can we go about curing this violation of the Pauli principle while not damaging the conserving aspects? The simplest way is to proceed in the spirit of Fermi liquid theory and assume that the effective interaction (irreducible vertex in the jargon) is renormalized. This renormalization has to be different for spin and charge so that

$$\chi_{\rm sp}(q)=\frac{\chi^{(1)}(q)}{1-\frac{1}{2}U_{\rm sp}\chi^{(1)}(q)} \tag{13.11}$$

$$\chi_{\rm ch}(q)=\frac{\chi^{(1)}(q)}{1+\frac{1}{2}U_{\rm ch}\chi^{(1)}(q)}. \tag{13.12}$$

[3] In other references we often use iq_n instead of $i\omega_n$ to denote the Matsubara frequency corresponding to the wave vector $\mathbf{q}$.

In practice $\chi^{(1)}(q)$ is the same[4] as the Lindhard function $\chi_0(q)$ (13.5) for $U = 0$ but, strictly speaking, there is a constant self-energy term that is absorbed in the definition of μ [20]. We are almost done with the collective modes. Substituting the above expressions for $\chi_{\rm sp}(q)$ and $\chi_{\rm ch}(q)$ in the two sum-rules, local-spin and local-charge appearing in (13.7) and (13.8), we could determine both $U_{\rm sp}$ and $U_{\rm ch}$ if we knew $\langle n_\uparrow n_\downarrow \rangle$. The following *ansatz*

$$U_{\rm sp} \langle n_\uparrow \rangle \langle n_\downarrow \rangle = U \langle n_\uparrow n_\downarrow \rangle \tag{13.13}$$

gives us the missing equation. Now notice that $U_{\rm sp}$, or equivalently $\langle n_\uparrow n_\downarrow \rangle$ depending on which of these variables you want to treat as independent, is determined self-consistently. That explains the name of the approach, "TPSC." Since the sum-rules are satisfied exactly, when we add them up the resulting equation, and hence the Pauli principle, will also be satisfied exactly. In other words, in (13.10) that follows from the Pauli principle, we now have $U_{\rm sp}$ and $U_{\rm ch}$ on the left-hand side that arrange each other in such a way that there is no violation of the principle. In standard many-body theory, the Pauli principle (crossing symmetry) is achieved in a much more complicated way by solving parquet equations [31, 32].

The ansatz (13.13) is inspired from the work of Singwi [33, 34] and was also found independently by Hedeyati and Vignale [35]. The whole procedure is not as arbitrary as it may seem and we justify this in more detail in Sect. 13.2.5 with the formal derivation. For now, let us just add a few physical considerations.

Remark 13.1. Since $U_{\rm sp}$ and $U_{\rm ch}$ are renormalized with respect to the bare value, one might have expected that one should use the dressed Green's functions in the calculation of $\chi_0(q)$. It is explained in Appendix A of [7] that this would lead to a violation of the results $\chi^{\rm R}_{\rm sp}(\mathbf{q} = \mathbf{0}, \omega) = 0$ and $\chi^{\rm R}_{\rm ch}(\mathbf{q} = \mathbf{0}, \omega) = 0$. In the present approach, the f-sum rule

$$\int \frac{{\rm d}\omega}{\pi} \omega \chi''_{\rm ch,sp}(\mathbf{q}, \omega) = \lim_{\eta \to 0} T \sum_{i\omega_n} \left({\rm e}^{-{\rm i}\omega_n \eta} - {\rm e}^{{\rm i}\omega_n \eta} \right) i\omega_n \chi_{\rm ch,sp}(\mathbf{q}, i\omega_n) \tag{13.14}$$

$$= \frac{1}{N} \sum_{\mathbf{k}\sigma} \left(\epsilon_{\mathbf{k}+\mathbf{q}} + \epsilon_{\mathbf{k}-\mathbf{q}} - 2\epsilon_{\mathbf{k}} \right) n_{\mathbf{k}\sigma} \tag{13.15}$$

is satisfied with $n_{\mathbf{k}\sigma} = n^{(1)}_{\mathbf{k}\sigma}$, which is the same as the Fermi function for the noninteracting case since it is computed from $G^{(1)}$.[5]

Remark 13.2. $U_{\rm sp} \langle n_\uparrow \rangle \langle n_\downarrow \rangle = U \langle n_\uparrow n_\downarrow \rangle$ can be understood as correcting the Hartree–Fock factorization to obtain the correct double-occupancy. Expressing the

[4] The meaning of the superscripts differs from that in [7]. Superscripts (2) (1) here correspond, respectively, to (1) (0) in [7].

[5] For the conductivity with vertex corrections [36], the f-sum rule with $n_{\mathbf{k}\sigma}$ obtained from $G^{(2)}$ is satisfied.

irreducible vertex in terms of an equal-time correlation function is inspired by the approach of Singwi and Tosi [33] to the electron gas. But TPSC is different since it also enforces the Pauli principle and connects to a local correlation function, namely, $\langle n_\uparrow n_\downarrow \rangle$.

13.2.2 Mermin–Wagner, Kanamori–Brueckner and Benchmarking Spin and Charge Fluctuations

The results that we found for spin and charge fluctuations have the RPA form but the renormalized interactions $U_{\rm sp}$ and $U_{\rm ch}$ must be computed from

$$\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\frac{\chi^{(1)}(q)}{1-\frac{1}{2}U_{\rm sp}\chi^{(1)}(q)} = n - 2\langle n_\uparrow n_\downarrow\rangle \tag{13.16}$$

and

$$\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\frac{\chi^{(1)}(q)}{1+\frac{1}{2}U_{\rm ch}\chi^{(1)}(q)} = n + 2\langle n_\uparrow n_\downarrow\rangle - n^2. \tag{13.17}$$

With the *ansatz* (13.13), the above system of equations is closed and the Pauli principle is enforced. The first of the above equations is solved self-consistently with the $U_{\rm sp}$ ansatz. This gives the double-occupancy $\langle n_\uparrow n_\downarrow\rangle$ that is then used to obtain $U_{\rm ch}$ from the next equation. The fastest way to numerically compute $\chi^{(1)}(q)$ is to use fast-Fourier transforms [36].

These TPSC expressions for spin and charge fluctuations were obtained by enforcing the conservations laws and the Pauli principle. In particular, TPSC satisfies the f-sum rule (13.15). But we obtain for free a lot more, namely, Kanamori–Brueckner renormalization and the Mermin–Wagner theorem.

Let us begin with Kanamori–Brueckner renormalization of U. Many years ago, Kanamori in the context of the Hubbard model [2], and Brueckner in the context of nuclear physics, introduced the notion that the bare U corresponds to computing the scattering of particles in the first Born approximation. In reality, we should use the full scattering cross-section and the effective U should be smaller. From Kanamori's point of view, the two-body wave function can minimize the effect of U by becoming smaller to reduce the value of the probability that two electrons are on the same site. The maximum energy that this can cost is the bandwidth since that is the energy difference between a one-body wave function with no nodes and one with the maximum allowed number. Let us see how this physics comes out of our results. Far from phase transitions, we can expand the denominator of the local moment sum-rule equation to obtain

$$\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\chi^{(1)}(q)\left(1+\frac{1}{2}U_{\rm sp}\chi^{(1)}(q)\right) = n - 2\frac{U_{\rm sp}}{U}\langle n_\uparrow\rangle\langle n_\downarrow\rangle. \tag{13.18}$$

Since $\frac{T}{N}\sum_{\mathbf{q}}\sum_{i\omega_n}\chi^{(1)}(q) = n - 2\langle n_\uparrow\rangle\langle n_\downarrow\rangle$, we can solve for $U_{\rm sp}$ and obtain[6]

$$U_{\rm sp} = \frac{U}{1+\Lambda U} \tag{13.19}$$

$$\Lambda \equiv \frac{1}{n^2}\frac{T}{N}\sum_{i\omega_n}\sum_{\mathbf{q}}\left(\chi^{(1)}\right)^2(\mathbf{q},i\omega_n)\,. \tag{13.20}$$

We see that at large U, $U_{\rm sp}$ saturates to $1/\Lambda$, which in practice we find to be of the order of the bandwidth. For those that are familiar with diagrams, note that the Kanamori–Brueckner physics amounts to replacing each of the interactions U in the ladder or bubble sum for diagrams in the particle–hole channel by infinite ladder sums in the particle–particle channel [37]. This is not quite what we obtain here since $\left(\chi^{(1)}\right)^2$ is in the particle–hole channel, but in the end, numerically, the results are close and the physics seems to be the same. One cannot make strict comparisons between TPSC and diagrams since TPSC is nonperturbative.

While Kanamori–Brueckner renormalization, or screening, is a quantum effect that occurs even far from phase transitions, when we are close we need to worry about the Mermin–Wagner theorem. To satisfy this theorem, approximate theories must prevent $\langle S_z^2\rangle$ from becoming infinite, which is equivalent to stopping $\langle n_\uparrow n_\downarrow\rangle$ from taking unphysical values. This quantity is positive and bounded by its value for $U=\infty$ and its value for noninteracting systems, namely, $0 \le \langle n_\uparrow n_\downarrow\rangle \le n^2/4$. Hence, the right-hand side of the local-moment sum-rule (13.16) is contained in the interval $\left[n, n-\frac{1}{2}n^2\right]$. To see how the Mermin–Wagner theorem is satisfied, write the self-consistency condition (13.16) in the form

$$\frac{T}{N}\sum_q \frac{\chi^{(1)}(q)}{1-\frac{1}{2}U\frac{\langle n_\uparrow n_\downarrow\rangle}{\langle n_\uparrow\rangle\langle n_\downarrow\rangle}\chi^{(1)}(q)} = n - 2\langle n_\uparrow n_\downarrow\rangle. \tag{13.21}$$

Consider increasing $\langle n_\uparrow n_\downarrow\rangle$ on the left-hand side of this equation. The denominator becomes smaller, hence the integral larger. To become larger, $\langle n_\uparrow n_\downarrow\rangle$ has to decrease on the right-hand side. There is thus a negative feedback in this equation that will make the self-consistent solution finite. This, however, does not prevent the expected phase transition in three dimensions [38]. To see this, we need to look in more details at the phase space for the integral in the sum-rule.

As we know from the spectral representation for χ,

$$\chi_{\rm ch,sp}(\mathbf{q},i\omega_n) = \int\frac{{\rm d}\omega'}{\pi}\frac{\chi''_{\rm ch,sp}(\mathbf{q},\omega')}{\omega'-i\omega_n} = \int\frac{{\rm d}\omega'}{\pi}\frac{\omega'\chi''_{\rm ch,sp}(\mathbf{q},\omega')}{(\omega')^2+(\omega_n)^2}. \tag{13.22}$$

[6] There is a misprint of a factor of 2 in [7]. It is corrected in [28].

The zero Matsubara frequency contribution is always the largest. There, we find the so-called Ornstein–Zernicke form for the susceptibility.

13.2.2.1 Ornstein–Zernicke Form

Lets us focus on the zero Matsubara frequency contribution and expand the denominator near the point, where $1 - \frac{1}{2}U_{\rm sp}\chi^{(1)}(\mathbf{Q},0) = 0$. The wave vector $\mathbf{Q}$ is that where $\chi^{(1)}$ is maximum. We find [38],

$$\chi_{\rm sp}(\mathbf{q}+\mathbf{Q}, i\omega_n) \simeq \frac{\chi^{(1)}(\mathbf{Q},0)}{1 - \frac{1}{2}U_{\rm sp}\chi^{(1)} - \frac{1}{4}U_{\rm sp}\frac{\partial^2\chi^{(1)}}{\partial\mathbf{Q}^2}q^2 - \frac{1}{2}U_{\rm sp}\frac{\partial\chi^{(1)}}{\partial(i\omega_n)}i\omega_n}$$
$$\sim \frac{\xi^2}{1+\xi^2q^2 + i\omega_n/\omega_{\rm sp}}, \tag{13.23}$$

where all functions and derivatives in the denominator are evaluated at $(\mathbf{Q}, 0)$ and where, on dimensional grounds,

$$\frac{-\frac{1}{4}U_{\rm sp}\frac{\partial^2\chi^{(1)}(\mathbf{Q},0)}{\partial\mathbf{Q}^2}}{\left(1 - \frac{1}{2}U_{\rm sp}\chi^{(1)}(\mathbf{Q},0)\right)} \tag{13.24}$$

scales (noted $\sim$) as the square of a length, ξ, the correlation length. That length is determined self-consistently. Since $\omega_{\rm sp} \sim \xi^{-2}$, all finite Matsubara frequency contributions are negligible if $2\pi T/\omega_{\rm sp} \sim 2\pi T\xi^2 \gg 1$. That condition in the form $\omega_{\rm sp} \ll T$ justifies the name of the regime we are interested in, namely, the renormalized classical regime. The classical regime of a harmonic oscillator occurs when $\omega \ll T$. The regime here is "renormalized" classical because at temperatures above the degeneracy temperature, the system is a free classical gas. As temperature decreases below the Fermi energy, it becomes quantum mechanical, then close to the phase transition, it becomes classical again.

Substituting the Ornstein–Zernicke form for the susceptibility in the self-consistency relation (13.16), we obtain

$$T\int \frac{\mathrm{d}^d\mathbf{q}}{(2\pi)^d}\frac{1}{q^2+\xi^{-2}} = \widetilde{C}, \tag{13.25}$$

where $\widetilde{C}$ contains non-zero Matsubara frequency contributions as well as $n-2\langle n_\uparrow n_\downarrow\rangle$. Since $\widetilde{C}$ is finite, this means that in two dimensions ($d = 2$), it is impossible to have $\xi^{-2} = 0$ on the left-hand side otherwise the integral would diverge logarithmically. This is clearly a dimension-dependent statement that proves the Mermin–Wagner theorem. In two dimensions, we see that the integral gives a logarithm that leads to

$$\xi \sim \exp\left(\frac{C'}{T}\right),$$

where in general, C' can be temperature dependent [38]. When C' is not temperature dependent, the above result is similar to what is found at strong coupling in the nonlinear sigma model. The above dimensional analysis is a bit expeditive. A more careful analysis [39, 40] yields prefactors in the temperature dependence of the correlation length.

The set of TPSC equations for spin and charge fluctuations (13.16), (13.17), (13.13) is rather intuitive and simple. The agreement of calculations with benchmark QMC calculations is rather spectacular, as shown in Fig. 13.1. There, one can see the results of QMC calculations of the structure factors, i.e., the Fourier transform of the equal-time charge and spin correlation functions, compared with the corresponding TPSC results. This figure allows one to watch the Pauli principle in action. At $U = 4t$, Fig. 13.1a shows that the charge structure factor does not have a monotonic dependence on density. This is because, as we approach half-filling, the spin fluctuations are becoming so large that the charge fluctuations have to decrease so that the sum still satisfies the Pauli principle, as expressed by (13.10).

More comparisons may be found in [29] and [6,7,41,42]. This kind of agreement is found even at couplings of the order of the bandwidth and when second-neighbor hopping t' is present [43, 44].

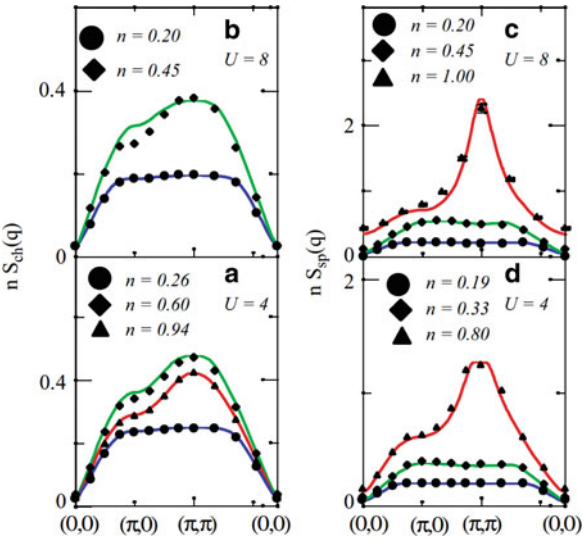

Fig. 13.1 Wave vector (**q**) dependence of the spin and charge structure factors for different sets of parameters. *Solid lines* are from TPSC and *symbols* are QMC data. Monte Carlo data for $n = 1$ and $U = 8t$ are for 6×6 clusters and $T = 0.5t$; all other data are for 8×8 clusters and $T = 0.2t$. Error bars are shown only when significant. (a) and (b) are for the charge structure factor and (c) (d) for the spin structure factor as defined in [6] where the original figure appears

Remark 13.3. Even though the entry in the renormalized classical regime is well described by TPSC [45], (13.13) for $U_{\rm sp}$ fails deep in that regime because $\Sigma^{(1)}$ becomes too different from the true self-energy. At $n = 1$, $t' = 0$, deep in the renormalized classical regime, $U_{\rm sp}$ becomes arbitrarily small, which is clearly unphysical. However, by assuming that $\langle n_\uparrow n_\downarrow \rangle$ is temperature independent below T_X, a property that can be verified from QMC calculations, one obtains a qualitatively correct description of the renormalized-classical regime. One can even drop the ansatz and take $\langle n_\uparrow n_\downarrow \rangle$ from QMC on the right-hand side of the local moment sum-rule (13.16) to obtain $U_{\rm sp}$.

13.2.3 Self-Energy

Collective charge and spin excitations can be obtained accurately from Green's functions that contain a simple self-energy, as we have just seen. Such modes are determined more by conservation laws than by details of the self-energy, especially at finite temperature where the lowest fermionic Matsubara frequency is not zero. The self-energy on the other hand is much more sensitive to collective modes since these are important at low frequency. The second step of TPSC is thus to find a better approximation for the self-energy. This is similar in spirit to what is done in the electron gas [8] where plasmons are found with noninteracting particles and then used to compute an improved approximation for the self-energy. This two-step process is also analogous to renormalization group calculations where renormalized interactions are evaluated to one-loop order and quasiparticle renormalization appears only to two-loop order [46–48].

The method to derive the result is justified using the formal derivation [20] that appears in Sect. 13.2.5. If you are familiar with diagrams, you can understand physically the result by looking at Fig. 13.2 that shows the exact diagrammatic expressions for the three-point vertex (green triangle) and self-energy (blue circle) in terms of Green's functions (solid black lines) and irreducible vertices (red boxes). The bare interaction U is the dashed line. One should keep in mind that we are not using perturbation theory despite the fact that we draw diagrams. Even within an exact approach, the quantities defined in the figure have well-defined meanings. The numbers on the figure refer to spin, space and imaginary time coordinates. When there is an overbar, there is a sum over spin and spatial indices and an integral over imaginary time.

In TPSC, the irreducible vertices in the first line of Fig. 13.2 are local, i.e., completely momentum and frequency independent. They are given by $U_{\rm sp}$ and $U_{\rm ch}$. If we set point 3 to be the same as point 1, then we can obtain directly the TPSC spin and charge susceptibilities from that first line. In the second line of the figure, the exact expression for the self-energy is displayed.[7] The first term on the right-hand

[7] In the Hubbard model the Fock term cancels with the same-spin Hartree term.

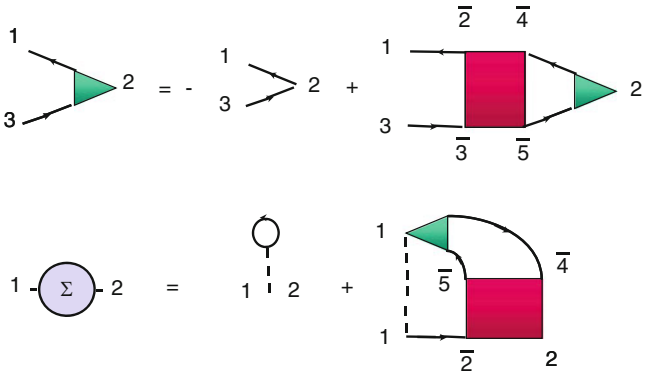

Fig. 13.2 Exact expression for the three-point vertex (*green triangle*) in the first line and for the self-energy in the second line. Irreducible vertices are the *red boxes* and Green's functions *solid black lines*. The numbers refer to spin, space and imaginary time coordinates. *Symbols* with an over-bard are summed/integrated over. The self-energy is the *blue circle* and the bare interaction U the *dashed line*

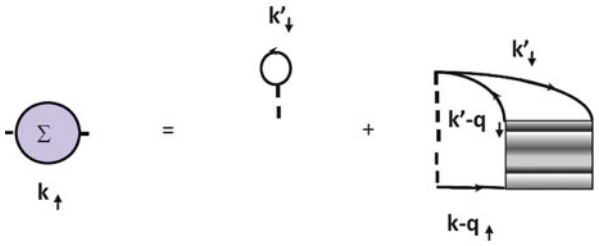

Fig. 13.3 Exact self-energy in terms of the Hartree–Fock contribution and of the fully reducible vertex Γ represented by a textured box

side is the Hartree–Fock contribution. In the second term, one recognizes the bare interaction U at one vertex that excites a collective mode represented by the green triangle and the two Green's functions. The other vertex is dressed, as expected. In the electron gas, the collective mode would be the plasmon. If we replace the irreducible vertex using $U_{\rm sp}$ and $U_{\rm ch}$ found for the collective modes, we find that here, both types of modes, spin and charge, contribute to the self-energy [41].

There is, however, an ambiguity in obtaining the self-energy formula [49]. Within the assumption that only $U_{\rm sp}$ and $U_{\rm ch}$ enter as irreducible particle–hole vertices, the self-energy expression in the transverse spin-fluctuation channel is different. What do we mean by that? Consider the exact formula for the self-energy represented symbolically by the diagram of Fig. 13.3. In this figure, the textured box is the fully reducible vertex $\Gamma\,(q, k-k', k+k'-q)$ that depends in general on three momentum-frequency indices. The longitudinal version of the self-energy corresponds to expanding the fully reducible vertex in terms of diagrams that are irreducible in the longitudinal (parallel spins) channel illustrated in Fig. 13.2. This

takes good care of the singularity of Γ when its first argument q is near (π, π) . The transverse version [20,49] does the same for the dependence on the second argument $k-k'$, which corresponds to the other (antiparallel spins) particle–hole channel. But the fully reducible vertex obeys crossing symmetry. In other words, interchanging two fermions just leads to a minus sign. One then expects that averaging the two possibilities gives a better approximation for Γ since it preserves crossing symmetry in the two particle–hole channels [49]. By considering both particle–hole channels only, we neglect the dependence of Γ on $k + k' - q$ because the particle–particle channel is not singular. The final formula that we obtain is [49]

$$\Sigma_\sigma^{(2)}(k) = U n_{-\sigma} + \frac{U}{8}\frac{T}{N}\sum_q \left[3U_{\rm sp}\chi_{\rm sp}(q) + U_{\rm ch}\chi_{\rm ch}(q)\right] G_\sigma^{(1)}(k+q), \quad (13.26)$$

where $n_{-\sigma}$ is the average single-spin occupation. The superscript (2) reminds us that we are at the second level of approximation. $G_\sigma^{(1)}$ is the same Green's function as that used to compute the susceptibilities $\chi^{(1)}(q)$. Since the self-energy is constant at that first level of approximation, this means that $G_\sigma^{(1)}$ is the noninteracting Green's function with the chemical potential that gives the correct filling. That chemical potential $\mu^{(1)}$ is slightly different from the one that we must use in $\left(G^{(2)}\right)^{-1} = i\omega_n + \mu^{(2)} - \varepsilon_{\mathbf{k}} - \Sigma^{(2)}$ to obtain the same density [50]. Estimates of $\mu^{(1)}$ may be found in [20, 50]. Further justifications for the above formula are given below in Sect. 13.2.4.

But before we come up with more formalism, we check that the above formula is accurate by comparing in Fig. 13.4 the spectral weight (imaginary part of the Green's function) obtained from (13.26) with that obtained from QMC calculations. The latter are exact within statistical accuracy and can be considered as benchmarks.

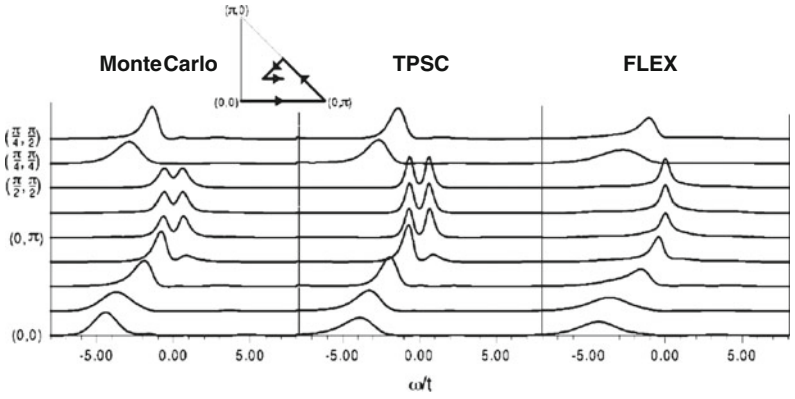

Fig. 13.4 Single-particle spectral weight $A(\mathbf{k}, \omega)$ for $U = 4$, $\beta = 5$, $n = 1$, and all independent wave vectors $\mathbf{k}$ of an 8×8 lattice. Results obtained from maximum entropy inversion of quantum Monte Carlo data on the *left panel*, from TPSC in the *middle panel* and from the FLEX approximation on the *right panel* (Relative error in all cases is about 0.3%.) From [49]

The meaning of the curves are detailed in the caption. The comparison is for half-filling in a regime where the simulations can be done at very low temperature and where a non-trivial phenomenon, a pseudogap, appears. This all important phenomenon is discussed further below in Sect. 13.2.6 and in the first case study, Sect. 13.3.1. In the third panel, we show the results of another popular many-body approach, the FLEX [15]. It misses [51] the physics of the pseudogap in the single-particle spectral weight because it uses fully dressed Green's functions and assumes that Migdal's theorem applies, i.e., the vertex does not need to be renormalized consequently [7, 52]. The same problem exists in the corresponding version of the GW approximation. [53]

Remark 13.4. The dressing of one vertex in the second line of Fig. 13.2 means that we do not assume a Migdal theorem. Migdal's theorem arises in the case of electron–phonon interactions [54]. There, the small ratio m/M, where m is the electronic mass and M the ionic mass, allows one to show that the vertex corrections are negligible. This is extremely useful in formulating the Eliashberg theory of superconductivity.

Remark 13.5. In [7, 49] we used the notation $\Sigma^{(1)}$ instead of $\Sigma^{(2)}$. The notation of the present chapter is the same as that of [20]

13.2.4 Internal Accuracy Checks

How can we make sure that TPSC is accurate? We have shown sample comparisons with benchmark QMC calculations, but we can check the accuracy in other ways. For example, we have already mentioned that the f-sum rule (13.15) is exactly satisfied at the first level of approximation (i.e., with $n_{\mathbf{k}}^{(1)}$ on the right-hand side). Suppose that on the right-hand side of that equation, one uses $n_{\mathbf{k}}$ obtained from $G^{(2)}$ instead of the Fermi function. One should find that the result does not change by more than a few percent. This is what happens when agreement with QMC is good.

When we are in the Fermi liquid regime, another way to verify the accuracy of the approach is to verify if the Fermi surface obtained from $G^{(2)}$ satisfies Luttinger's theorem very closely.

Finally, there is a consistency relation between one- and two-particle quantities (Σ and $\langle n_\uparrow n_\downarrow \rangle$). The relation

$$\frac{T}{N}\sum_{\mathbf{k}}\sum_{n}\Sigma(\mathbf{k},i\omega_n)G(\mathbf{k},i\omega_n)\mathrm{e}^{-\mathrm{i}\omega_n 0^-} = \frac{1}{2}\mathrm{Tr}\,(\Sigma G) = U\langle n_\uparrow n_\downarrow\rangle \qquad (13.27)$$

should be satisfied exactly for the Hubbard model. This result follows from the definition of self-energy and is derived in (13.40) below. In standard many-body books [54], it is encountered in the calculation of the free energy through a

coupling-constant integration. In TPSC, it is not difficult[8] to show that the following equation

$$\frac{1}{2}\mathrm{Tr}\left(\Sigma^{(2)}G^{(1)}\right) = U\left\langle n_\uparrow n_\downarrow\right\rangle \tag{13.28}$$

is satisfied exactly with the self-consistent $U\left\langle n_\uparrow n_\downarrow\right\rangle$ obtained with the susceptibilities.[9] An internal accuracy check consists in verifying by how much $\frac{1}{2}\mathrm{Tr}\left(\Sigma^{(2)}G^{(2)}\right)$ differs from $\frac{1}{2}\mathrm{Tr}\left(\Sigma^{(2)}G^{(1)}\right)$. Again, in regimes where we have agreement with QMC calculations, the difference is only a few percent.

The above relation between Σ and $\left\langle n_\uparrow n_\downarrow\right\rangle$ gives us another way to justify our expression for $\Sigma^{(2)}$. Suppose one starts from Fig. 13.2 to obtain a self-energy expression that contains only the longitudinal spin fluctuations and the charge fluctuations, as was done in the first papers on TPSC [6]. One finds that each of these separately contributes an amount $U\left\langle n_\uparrow n_\downarrow\right\rangle/2$ to the consistency relation (13.28). Similarly, if we work only in the transverse-spin channel [20, 49] we find that each of the two transverse-spin components also contributes $U\left\langle n_\uparrow n_\downarrow\right\rangle/2$ to $\frac{1}{2}\mathrm{Tr}\left(\Sigma^{(2)}G^{(1)}\right)$. Hence, averaging the two expressions also preserves rotational invariance. In addition, one verifies numerically that the exact sum-rule ([7], Appendix A)

$$-\int\frac{\mathrm{d}\omega'}{\pi}\Sigma''^{R}_{\sigma}\left(\mathbf{k},\omega'\right) = U^2 n_{-\sigma}\left(1-n_{-\sigma}\right) \tag{13.29}$$

determining the high-frequency behavior is satisfied to a higher degree of accuracy with the symmetrized self-energy expression (13.26).

Equation (13.26) for $\Sigma^{(2)}$ is different from so-called Berk–Schrieffer type expressions [56] that do not satisfy[10] the consistency condition between one- and two-particle properties, $\frac{1}{2}\mathrm{Tr}\left(\Sigma G\right) = U\left\langle n_\uparrow n_\downarrow\right\rangle$.

Remark 13.6. Schemes, such as FLEX, that use on the right-hand side $G^{(2)}$ are thermodynamically consistent (Sect. 13.4.4.1) and might look better. However, as we just saw, in Fig. 13.4, FLEX misses some important physics. The reason [7] is that the vertex entering the self-energy in FLEX is not at the same level of approximation as the Green's functions. Indeed, since the latter contain self-energies that are strongly momentum and frequency dependent, the irreducible vertices that can be derived from these self-energies should also be frequency and momentum dependent, but they are not. In fact they are the bare vertices. It is as if the quasiparticles had a lifetime while at the same time interacting with each other with the bare interaction. Using dressed Green's functions in the susceptibilities with momentum and frequency independent vertices leads to problems as well. For example, the conservation law $\chi_{\mathrm{sp,ch}}\left(\mathbf{q}=\mathbf{0},i\omega_n\right) = 0$ is violated in that case, as

[8] Appendix B of [7].

[9] FLEX does not satisfy this consistency requirement. See Appendix E of [7]. In fact double-occupancy obtained from ΣG can even become negative [55].

[10] (See [7], Appendix E).

shown in Appendix A of [7]. Further criticism of conserving approaches appears in Appendix E of [7] and in [20].

13.2.5 A More Formal Derivation

Details of a more formal derivation may be found in [57]. For completeness we repeat some of the derivation. The reader more interested in the physics may skip that section. The first two subsections present some general formalism that is then used in the following two subsections to derive TPSC.

13.2.5.1 Single-Particle Properties

Following functional methods of the Schwinger school[30, 58, 59], we begin with the generating function with source fields ϕ_σ and field destruction operators ψ in the grand canonical ensemble

$$\ln Z\left[\phi\right] = \ln \mathrm{Tr}\left[\mathrm{e}^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)}\mathrm{T}_\tau\left(\mathrm{e}^{-\psi_{\overline{\sigma}}^{\dagger}(\overline{1})\phi_{\overline{\sigma}}(\overline{1},\overline{2})\psi_{\overline{\sigma}}(\overline{2})}\right)\right]. \tag{13.30}$$

We adopt the convention that 1 stands for the position and imaginary time indices $(\mathbf{r}_1, \tau_1)$. The overbar means summation over every lattice site and integration over imaginary-time from 0 to β, and $\overline{\sigma}$ summation over spins. T_τ is the time-ordering operator.

The propagator in the presence of the source field is obtained from functional differentiation

$$G_\sigma\left(1,2;\{\phi\}\right) = -\left\langle\psi_\sigma\left(1\right)\psi_\sigma^{\dagger}\left(2\right)\right\rangle_\phi = -\frac{\delta \ln Z\left[\phi\right]}{\delta\phi_\sigma\left(2,1\right)}. \tag{13.31}$$

From now on, *the time-ordering operator in averages,* $\langle\rangle$, *is implicit*. Physically relevant correlation functions are obtained for $\{\phi\} = 0$ but it is extremely convenient to keep finite $\{\phi\}$ in intermediate steps of the calculation.

Using the equation of motion for the field ψ and the definition of the self-energy, one obtains the Dyson equation in the presence of the source field [60]

$$\left(G_0^{-1} - \phi\right)G = 1 + \Sigma G; \quad G^{-1} = G_0^{-1} - \phi - \Sigma, \tag{13.32}$$

where from the commutator of the interacting part of the Hubbard Hamiltonian H, one obtains

$$\Sigma_\sigma\left(1,\overline{1};\{\phi\}\right)G_\sigma\left(\overline{1},2;\{\phi\}\right) = -U\left\langle\psi_{-\sigma}^{\dagger}\left(1^+\right)\psi_{-\sigma}\left(1\right)\psi_\sigma\left(1\right)\psi_\sigma^{\dagger}\left(2\right)\right\rangle_\phi. \tag{13.33}$$

The imaginary time in 1^+ is infinitesimally larger than in 1.

13.2.5.2 Response Functions

Response (four-point) functions for spin and charge excitations can be obtained from functional derivatives ($\delta G/\delta\phi$) of the source-dependent propagator. Following the standard approach and using matrix notation to abbreviate the summations and integrations we have,

$$GG^{-1} = 1 \tag{13.34}$$

$$\frac{\delta G}{\delta\phi}G^{-1} + G\frac{\delta G^{-1}}{\delta\phi} = 0. \tag{13.35}$$

Using the Dyson equation (13.32) $G^{-1} = G_0^{-1} - \phi - \Sigma$ this may be rewritten as

$$\frac{\delta G}{\delta\phi} = -G\frac{\delta G^{-1}}{\delta\phi}G = G\hat{\cdot}G + G\frac{\delta\Sigma}{\delta\phi}G, \tag{13.36}$$

where the symbol $\hat{\cdot}$ in $G\hat{\cdot}G$ reminds us that the neighboring labels of the propagators have to be the same as those of the ϕ in the functional derivative. If perturbation theory converges, we may write the self-energy as a functional of the propagator. From the chain-rule, one then obtains an integral equation for the response function in the particle–hole channel that is the analog of the Bethe–Salpeter equation in the particle–particle channel

$$\frac{\delta G}{\delta\phi} = G\hat{\cdot}G + G\left[\frac{\delta\Sigma}{\delta G}\frac{\delta G}{\delta\phi}\right]G. \tag{13.37}$$

The labels of the propagators in the last term are attached to the self-energy, as in (13.36).[11] Vertices appropriate for spin and charge responses are given, respectively, by

$$U_{\rm sp} = \frac{\delta\Sigma_\uparrow}{\delta G_\downarrow} - \frac{\delta\Sigma_\uparrow}{\delta G_\uparrow} \quad ; \quad U_{\rm ch} = \frac{\delta\Sigma_\uparrow}{\delta G_\downarrow} + \frac{\delta\Sigma_\uparrow}{\delta G_\uparrow}. \tag{13.38}$$

13.2.5.3 TPSC First Step: Two-Particle Self-Consistency for $G^{(1)}$, $\Sigma^{(1)}$, $\Gamma_{\rm sp}^{(1)} = U_{\rm sp}$ and $\Gamma_{\rm ch}^{(1)} = U_{\rm ch}$

In conserving approximations, the self-energy is obtained from a functional derivative $\Sigma[G] = \delta\Phi[G]/\delta G$ of Φ the Luttinger–Ward functional, which is itself computed from a set of diagrams. To liberate ourselves from diagrams, we start

[11] To remind ourselves of this, we may also adopt an additional vertical matrix notation convention and write (13.37) as $\frac{\delta G}{\delta\phi} = G\hat{\cdot}G + G\left[\frac{\frac{\delta\Sigma}{\delta G}}{\frac{\delta G}{\delta\phi}}\right]G$.

instead from the exact expression for the self-energy, (13.33) and notice that when label 2 equals 1^+, the right-hand side of this equation is equal to double-occupancy $\langle n_\uparrow n_\downarrow \rangle$. Factoring as in Hartree–Fock amounts to assuming no correlations. Instead, we should insist that $\langle n_\uparrow n_\downarrow \rangle$ be obtained self-consistently. After all, in the Hubbard model, there are only two local four-point functions: $\langle n_\uparrow n_\downarrow \rangle$ and $\langle n_\uparrow n_\uparrow \rangle = \langle n_\downarrow n_\downarrow \rangle$. The latter is given exactly, through the Pauli principle, by $\langle n_\uparrow n_\uparrow \rangle = \langle n_\downarrow n_\downarrow \rangle = \langle n_\uparrow \rangle = \langle n_\downarrow \rangle = n/2$, when the filling n is known. In a way, $\langle n_\uparrow n_\downarrow \rangle$ in the self-energy equation (13.33) can be considered as an initial condition for the four-point function when one of the points, 2, separates from all the others which are at 1. When that label 2 does not coincide with 1, it becomes more reasonable to factor *à la* Hartree–Fock. These physical ideas are implemented by postulating

$$\Sigma_\sigma^{(1)}\left(1,\overline{1};\{\phi\}\right) G_\sigma^{(1)}\left(\overline{1},2;\{\phi\}\right) = A_{\{\phi\}} G_{-\sigma}^{(1)}\left(1,1^+;\{\phi\}\right) G_\sigma^{(1)}\left(1,2;\{\phi\}\right), \quad (13.39)$$

where $A_{\{\phi\}}$ depends on the external field and is chosen such that the exact result[12]

$$\Sigma_\sigma\left(1,\overline{1};\{\phi\}\right) G_\sigma\left(\overline{1},1^+;\{\phi\}\right) = U\left\langle n_\uparrow(1)\, n_\downarrow(1)\right\rangle_\phi \quad (13.40)$$

is satisfied. It is easy to see that the solution is

$$A_{\{\phi\}} = U\frac{\left\langle n_\uparrow(1)\, n_\downarrow(1)\right\rangle_\phi}{\left\langle n_\uparrow(1)\right\rangle_\phi \left\langle n_\downarrow(1)\right\rangle_\phi}. \quad (13.41)$$

Substituting $A_{\{\phi\}}$ back into our *ansatz* (13.13) we obtain our first approximation for the self-energy by right-multiplying by $\left(G_\sigma^{(1)}\right)^{-1}$:

$$\Sigma_\sigma^{(1)}\left(1,2;\{\phi\}\right) = A_{\{\phi\}} G_{-\sigma}^{(1)}\left(1,1^+;\{\phi\}\right)\delta(1-2). \quad (13.42)$$

We are now ready to obtain irreducible vertices using the prescription of the previous section, (13.38), namely, through functional derivatives of Σ with respect to G. In the calculation of U_{sp}, the functional derivative of $\langle n_\uparrow n_\downarrow \rangle / \left(\langle n_\uparrow \rangle \langle n_\downarrow \rangle\right)$ drops out, so we are left with[13]

$$U_{\mathrm{sp}} = \left.\frac{\delta\Sigma_\uparrow^{(1)}}{\delta G_\downarrow^{(1)}}\right|_{\{\phi\}=0} - \left.\frac{\delta\Sigma_\uparrow^{(1)}}{\delta G_\uparrow^{(1)}}\right|_{\{\phi\}=0} = A_{\{\phi\}=0} = U\frac{\langle n_\uparrow n_\downarrow \rangle}{\langle n_\uparrow \rangle \langle n_\downarrow \rangle}. \quad (13.43)$$

[12] See footnote (14) of [20] for a discussion of the choice of limit 1^+ versus 1^-.

[13] For $n > 1$, all particle occupation numbers must be replaced by hole occupation numbers.

The renormalization of this irreducible vertex may be physically understood as coming from Kanamori–Brueckner screening [7]. This completes the derivation of the *ansatz* that was missing in our first derivation in Sect. 13.2.1.

The functional-derivative procedure generates an expression for the charge vertex $U_{\rm ch}$ which involves the functional derivative of $\langle n_\uparrow n_\downarrow \rangle / \left(\langle n_\uparrow \rangle \langle n_\downarrow \rangle \right)$ which contains six-point functions that one does not really know how to evaluate. But, if we again assume that the vertex $U_{\rm ch}$ is a constant, it is simply determined by the requirement that charge fluctuations also satisfy the fluctuation–dissipation theorem and the Pauli principle, as in (13.17).

Note that, in principle, $\Sigma^{(1)}$ also depends on double-occupancy, but since $\Sigma^{(1)}$ is a constant, it is absorbed in the definition of the chemical potential and we do not need to worry about it in this case. That is why the noninteracting irreducible susceptibility $\chi^{(1)}(q) = \chi_0(q)$ appears in the expressions for the susceptibility, even though it should be evaluated with $G^{(1)}$ that contains $\Sigma^{(1)}$. A rough estimate of the renormalized chemical potential (or equivalently of $\Sigma^{(1)}$) is given in the Appendix of [20]. One can check that spin and charge conservations are satisfied by our susceptibilities.

13.2.5.4 TPSC Second Step: An Improved Self-Energy $\Sigma^{(2)}$

Collective modes are emergent objects that are less influenced by details of the single-particle properties than the other way around. We thus wish now to obtain an improved approximation for the self-energy that takes advantage of the fact that we have found accurate approximations for the low-frequency spin and charge fluctuations. We begin from the general definition of the self-energy (13.33) obtained from Dyson's equation. The right-hand side of that equation can be obtained either from a functional derivative with respect to an external field that is diagonal in spin, as in our generating function (13.30), or by a functional derivative of $\left\langle \psi_{-\sigma}(1)\, \psi^\dagger_\sigma(2) \right\rangle_{\phi_t}$ with respect to a transverse external field ϕ_t.

Working first in the longitudinal channel, the right-hand side of the general definition of the self-energy (13.33) may be written as

$$\Sigma_\sigma\left(1,\overline{1}\right) G_\sigma\left(\overline{1},2\right) = -U\left[\left. \frac{\delta G_\sigma\left(1,2;\{\phi\}\right)}{\delta\phi_{-\sigma}\left(1^+,1\right)} \right|_{\{\phi\}=0} - G_{-\sigma}\left(1,1^+\right) G_\sigma\left(1,2\right) \right]. \tag{13.44}$$

The last term is the Hartree–Fock contribution. It gives the exact result for the self-energy in the limit $\omega \to \infty$ [7]. The $\delta G_\sigma / \delta\phi_{-\sigma}$ term is thus a contribution to lower frequencies and it comes from the spin and charge fluctuations. Right-multiplying the last equation by G^{-1} and replacing the lower energy part $\delta G_\sigma / \delta\phi_{-\sigma}$ by its general expression in terms of irreducible vertices (13.37), we find

$$\Sigma_\sigma^{(2)}(1,2) = UG_{-\sigma}^{(1)}\left(1,1^+\right)\delta(1-2) - UG_\sigma^{(1)}\left(1,\overline{3}\right)\left[\left.\frac{\delta\Sigma_\sigma^{(1)}\left(\overline{3},2;\{\phi\}\right)}{\delta G_{\overline{\sigma}}^{(1)}\left(\overline{4},\overline{5};\{\phi\}\right)}\right|_{\{\phi\}=0} \times \left.\frac{\delta G_{\overline{\sigma}}^{(1)}\left(\overline{4},\overline{5};\{\phi\}\right)}{\delta\phi_{-\sigma}\left(1^+,1\right)}\right|_{\{\phi\}=0}\right]. \tag{13.45}$$

Every quantity appearing on the right-hand side of that equation has been taken from the TPSC results. This means in particular that the irreducible vertices $\delta\Sigma_\sigma^{(1)}/\delta G_{\sigma'}^{(1)}$ are at the same level of approximation as the Green functions $G_\sigma^{(1)}$ and self-energies $\Sigma_\sigma^{(1)}$. In approaches that assume that Migdal's theorem applies to spin and charge fluctuations, one often sees renormalized Green functions $G^{(2)}$ appearing on the right-hand side along with unrenormalized vertices, $\delta\Sigma_\sigma/\delta G_{\sigma'} \to U$. In terms of $U_{\rm sp}$ and $U_{\rm ch}$ in Fourier space, the above formula [41] reads,

$$\Sigma_\sigma^{(2)}(k)_{\rm long} = Un_{-\sigma} + \frac{U}{4}\frac{T}{N}\sum_q\left[U_{\rm sp}\chi_{\rm sp}^{(1)}(q) + U_{\rm ch}\chi_{\rm ch}^{(1)}(q)\right]G_\sigma^{(1)}(k+q). \tag{13.46}$$

The approach to obtain a self-energy formula that takes into account both longitudinal and transverse fluctuations is detailed in [20]. Crossing symmetry, rotational symmetry and sum-rules and comparisons with QMC dictate the final formula for the improved self-energy $\Sigma^{(2)}$ as we have explained in Sect. 13.2.3.

13.2.6 Pseudogap in the Renormalized Classical Regime

When we compared TPSC with QMC simulations and with FLEX in Fig. 13.4 above, perhaps you noticed that at the Fermi surface, the frequency dependent spectral weight has two peaks instead of one. In addition, at zero frequency, it has a minimum instead of a maximum. That is called a pseudogap. A cartoon explanation [29] of this pseudogap is given in Fig. 13.5. At high temperature we start from a Fermi liquid, as illustrated in panel I. Now, suppose the ground state has long-range antiferromagnetic order as in panel III, in other words at a filling between half-filling and n_c. In the mean-field approximation we have a gap and the Bogoliubov transformation from fermion creation–annihilation operators to quasiparticles has weight at both positive and negative energies. In two dimensions, because of the Mermin–Wagner theorem, as soon as we raise the temperature above zero, long-range order disappears, but the antiferromagnetic correlation length ξ remains large so we obtain the pseudogap illustrated in panel II. As we will explain analytically below, the pseudogap survives as long as ξ is much larger than the thermal de Broglie wavelength $\xi_{\rm th} \equiv v_{\rm F}/(\pi T)$ in our usual units. At the crossover temperature T_X, the relative size of ξ and $\xi_{\rm th}$ changes and we recover the Fermi liquid.

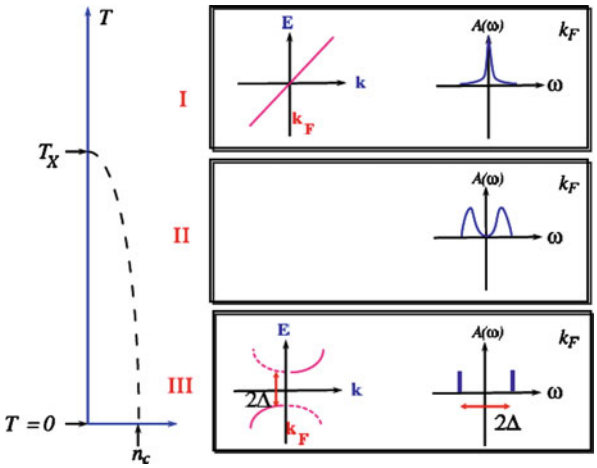

Fig. 13.5 Cartoon explanation of the pseudogap due to precursors of long-range order. When the antiferromagnetic correlation length ξ becomes larger than the thermal de Broglie wavelength, there appears precursors of the $T = 0$ Bogoliubov quasiparticles for the long-range ordered antiferromagnet. This can occur only in the renormalized classical regime, below the *dashed line* on the *left* of the figure

We now proceed to sketch analytically where these results come from starting from finite T. Details and more complete formulae may be found in [6, 7, 41, 61].[14] We begin from the TPSC expression (13.26) for the self-energy. Normally one has to do the sum over bosonic Matsubara frequencies first, but the zero Matsubara frequency contribution has the correct asymptotic behavior in fermionic frequencies $i\omega_n$ so that, as in Sect. 13.2.2, one can once more isolate on the right-hand side the contribution from the zero Matsubara frequency. In the renormalized classical regime then, we have

$$\Sigma(\mathbf{k}_F, i\omega_n) \propto T \int q^{d-1} \mathrm{d}q \frac{1}{q^2 + \xi^{-2}} \frac{1}{i\omega_n - \varepsilon_{\mathbf{k}_F+\mathbf{Q}+\mathbf{q}}}, \tag{13.47}$$

where $\mathbf{Q}$ is the wave vector of the instability.[15] This integral can be done analytically in two dimensions [7, 64]. But it is more useful to analyze limiting cases [41]. Expanding around the points known as hot spots where $\varepsilon_{\mathbf{k}_F+\mathbf{Q}} = 0$, we find after analytical continuation that the imaginary part of the retarded self-energy at zero frequency takes the form

[14] Note also the following study from zero temperature [62].

[15] This formula is similar to one that appeared in [63].

$$\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},0) \propto -\pi T \int \mathrm{d}^{d-1}q_{\perp}\mathrm{d}q_{||}\frac{1}{q_{\perp}^2+q_{||}^2+\xi^{-2}}\delta(v'_{\mathrm{F}}q_{||}) \tag{13.48}$$

$$\propto \frac{\pi T}{v'_{\mathrm{F}}}\xi^{3-d}. \tag{13.49}$$

In the last line, we just used dimensional analysis to do the integral.

The importance of dimension comes out clearly [41]. In $d = 4$, $\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},0)$ vanishes as temperature decreases, $d = 3$ is the marginal dimension and in $d = 2$ we have that $\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},0) \propto \xi/\xi_{\mathrm{th}}$ that diverges at zero temperature. In a Fermi liquid the quantity $\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},0)$ vanishes at zero temperature, hence in three or four dimensions one recovers the Fermi liquid (or close to one in $d = 3$). But in two dimensions, a diverging $\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},0)$ corresponds to a vanishingly small $A(\mathbf{k}_{\mathrm{F}}, \omega = 0)$ as we can see from

$$A(\mathbf{k},\omega) = \frac{-2\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},\omega)}{(\omega-\varepsilon_{\mathbf{k}}-\Sigma'^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},\omega))^2+\Sigma''^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},\omega)^2}. \tag{13.50}$$

Figure 31 of [29] illustrates graphically the relationship between the location of the pseudogap and large scattering rates at the Fermi surface. At stronger U the scattering rate is large over a broader region, leading to a depletion of $A(\mathbf{k},\omega)$ over a broader range of $\mathbf{k}$ values.

Remark 13.7. Note that the condition $\xi/\xi_{\mathrm{th}} \gg 1$, necessary to obtain a large scattering rate, is in general harder to satisfy than the condition that corresponds to being in the renormalized classical regime. Indeed, $\xi/\xi_{\mathrm{th}} \gg 1$ corresponds $T/v_{\mathrm{F}} \gg \xi^{-1}$ while the condition $\omega_{\mathrm{sp}} \ll T$ for the renormalized classical regime corresponds to $T \gg \xi^{-2}$, with appropriate scale factors, because ω_{sp} scales as ξ^{-2} as we saw in (13.23) and below.

To understand the splitting into two peaks seen in Figs. 13.4 and 13.5 consider the singular renormalized contribution coming from the spin fluctuations in (13.47) at frequencies $\omega \gg v_{\mathrm{F}}\xi^{-1}$. Taking into account that contributions to the integral come mostly from a region $q \leq \xi^{-1}$, one finds

$$\Sigma'^{\mathrm{R}}(\mathbf{k}_{\mathrm{F}},\omega) = \left(T\int q^{d-1}\mathrm{d}q\frac{1}{q^2+\xi^{-2}}\right)\frac{1}{ik_n-\varepsilon_{\mathbf{k}_{\mathrm{F}}+\mathbf{Q}}}$$

$$\equiv \frac{\Delta^2}{\omega-\varepsilon_{\mathbf{k}_{\mathrm{F}}+\mathbf{Q}}}, \tag{13.51}$$

which, when substituted in the expression for the spectral weight (13.50), leads to large contributions when

$$\omega-\varepsilon_{\mathbf{k}}-\frac{\Delta^2}{\omega-\varepsilon_{\mathbf{k}_{\mathrm{F}}+\mathbf{Q}}} = 0 \tag{13.52}$$

or, equivalently,

$$\omega = \frac{(\varepsilon_{\mathbf{k}} + \varepsilon_{\mathbf{k}_F+\mathbf{Q}}) \pm \sqrt{(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}_F+\mathbf{Q}})^2 + 4\Delta^2}}{2}, \tag{13.53}$$

which, at $\omega = 0$, corresponds to the position of the hot spots.[16] At finite frequencies, this turns into the dispersion relation for the antiferromagnet [66].

It is important to understand that analogous arguments hold for any fluctuation that becomes soft because of the Mermin–Wagner theorem [7, 67], including superconducting ones [7, 50, 57]. The wave vector $\mathbf{Q}$ would be different in each case.

To understand better when Fermi liquid theory is valid and when it is replaced by the pseudogap instead, it is useful to perform the calculations that lead to $\Sigma''^{\mathrm{R}}(\mathbf{k}_F, 0) \propto \xi/\xi_{\mathrm{th}}$ in the real frequency formalism. The details may be found in Appendix D of [7].

13.3 Case Studies

In this short pedagogical review it is impossible to cover all topics in depth. This section will nevertheless expand a bit on two important contributions of TPSC to problems of current interest, namely, the pseudogap of cuprate superconductors and superconductivity induced by antiferromagnetic fluctuations.

13.3.1 Pseudogap in Electron-Doped Cuprates

High-temperature superconductors are made of layers of CuO_2 planes. The rest of the structure is commonly considered as providing either electron or hole doping of these planes depending on chemistry. At half-filling, or zero doping, the ground state is an antiferromagnet. As one dopes the planes, one reaches a doping, so-called optimal doping, where the superconducting transition temperature T_c is maximum. Let us start from optimal hole or electron doping and decrease doping towards half-filling. That is the underdoped regime. In that regime, one observes a curious phenomenon, the pseudogap. What this means is that as temperature decreases, physical quantities behave as if the density of states near the Fermi level were decreasing. Finding an explanation for this phenomenon has been one of the major challenges of the field [68, 69].

To make progress, we need a microscopic model for high-temperature superconductors. Band structure calculations [70, 71] reveal that a single band crosses

[16] For comparisons with paramagnon theory see [65].

the Fermi level. Hence, it is a common assumption that these materials can be modeled by the one-band Hubbard model. Whether this is an oversimplification is still a subject of controversy [72–77]. Indeed, spectroscopic studies [72, 78] show that hole doping occurs on the oxygen atoms. The resulting hole behaves as a copper excitation because of Zhang–Rice [79] singlet formation. In addition, the phase diagram [80–85] and many properties of the hole-doped cuprates can be described by the one-band Hubbard model. Typically, the band parameters that are used are: nearest-neighbor hopping $t = 350$–400 meV and next-nearest-neighbor hopping $t' = -0.15$ to $-0.3\,t$ depending on the compound [70,71]. A third-nearest-neighbor hopping $t'' = -0.5\,t'$ is sometimes added to fit finer details of the band structure [71]. The second-neighbor hopping breaks particle–hole symmetry at the band structure level.

In electron-doped cuprates, the doping occurs on the copper, hence there is little doubt that the single-band Hubbard model is even a better starting point in this case. Band parameters [86] are similar to those of hole-doped cuprates. It is sometimes claimed that there is a pseudogap only in the hole-doped cuprates. The origin of the pseudogap is indeed probably different in the hole-doped cuprates. But even though the standard signature of a pseudogap is absent in nuclear magnetic resonance (NMR) [87] there is definitely a pseudogap in the electron-doped case as well [88], as can be seen in optical conductivity [89] and in angle resolved photoemission spectroscopy (ARPES) [90]. As we show in the rest of this section, in electron-doped cuprates strong evidence for the origin of the pseudogap is provided by detailed comparisons of TPSC with ARPES as well as by verification with neutron scattering [91] that the TPSC condition for a pseudogap, namely, $\xi > \xi_{\rm th}$, is satisfied. The latter length makes sense from weak to intermediate coupling when quasiparticles exist above the pseudogap temperature. In strong coupling, i.e., for values of U larger than that necessary for the Mott transition, there is evidence that there is another mechanism for the formation of a pseudogap. This is discussed at length in [92, 93].[17] The recent discovery [94] that at sufficiently large U there is a first order transition in the paramagnetic state between two kinds of metals, one of which is highly anomalous, gives a sharper meaning to what is meant by strong-coupling pseudogap.

Let us come back to modeling of electron-doped cuprates. Evidence that these are less strongly coupled than their hole-doped counterparts comes from the fact that (a) The value of the optical gap at half-filling, $\sim$1.5 eV, is smaller than for hole doping, $\sim$2.0 eV [95]. (b) In a simple Thomas–Fermi picture, the screened interaction scales like $\partial\mu/\partial n$. Quantum cluster calculations [92] show that $\partial\mu/\partial n$ is smaller on the electron-doped side, hence U should be smaller. (c) Mechanisms based on the exchange of antiferromagnetic fluctuations with U/t at weak to intermediate coupling [16, 42] predict that the superconducting $T_{\rm c}$ increases with U/t. Hence $T_{\rm c}$ should decrease with increasing pressure in the simplest model where pressure increases hopping t while leaving U essentially unchanged. The opposite behavior,

[17] See also conclusion of [29].

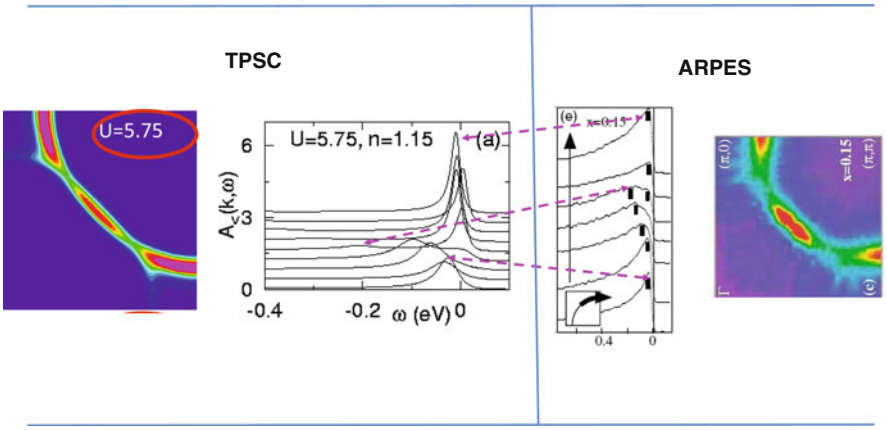

Fig. 13.6 On the *left*, results of TPSC calculations [29, 97] at optimal doping, $x = 0.15$, corresponding to filling 1.15, for $t = 350\,\text{meV}$, $t' = -0.175\,t$, $t'' = 0.05\,t$, $U = 5.75\,t$, and $T = 1/20$. The *left-most panel* is the magnitude of the spectral weight times a Fermi function, $A(\mathbf{k}, \omega)\, f(\omega)$ at $\omega = 0$, so-called momentum-distribution curve (MDC). *Red* (*dark black*) indicates larger value and purple (*light grey*) smaller value. The next panel is $A(\mathbf{k}, \omega)\, f(\omega)$ for a set of fixed $\mathbf{k}$ values along the Fermi surface. These are so-called energy-dispersion curves (EDC). The two panels to the *right* are the corresponding experimental results [90] for $Nd_{2-x}Ce_xCuO_4$. *Dotted arrows* show the correspondence between TPSC and experiment

expected at strong coupling where $J = 4t^2/U$ is relevant [85, 96], is observed in the hole-doped cuprates. (d) Finally and most importantly, we have shown detailed agreement between TPSC calculations [29, 93, 97] and measurements such as ARPES [90, 98], optical conductivity [89] and neutron scattering [91].

To illustrate the last point, consider Fig. 13.6 that compares TPSC calculations with experimental results for ARPES. Apart from a tail in the experimental results. the agreement is striking.[18] In particular, if there were no interaction, the Fermi surface would be a line (red) on the momentum distribution curve (MDC). Instead, it seems to disappear at symmetrical points displaced from $(\pi/2, \pi/2)$. These points, so-called hot spots, are linked by the wave vector (π, π) to other points on the Fermi surface. This is where the antiferromagnetic gap would open first if there were long-range order. The pull back of the weight from $\omega = 0$ at the hot spots is close to the experimental value: 100 meV for the 15% doping (show), and 300 meV for 10% doping (not shown). More detailed ARPES spectra and comparisons with experiment are shown in [29]. The value of the temperature T^* at which the pseudogap appears [97] is also close to that observed in optical spectroscopy [89]. In addition, the size of the pseudogap is about ten times T^* in the calculation as well as in the experiments. For optical spectroscopy, vertex corrections (see Sect. 13.4.5)

[18] Such tails tend to disappear in more recent laser ARPES measurements on hole-doped compounds [99].

have to be added to be more quantitative. Experimentally, the value of T^* is about twice the antiferromagnetic transition temperature up to $x = 0.13$. That can be obtained [97] by taking $t_z = 0.03\,t$ for hopping in the third direction. Recall that in strictly two dimensions, there is no long-range order. Antiferromagnetism appears on a much larger range of dopings for electron-doped than for hole-doped cuprates.

These TPSC calculations have predicted the value of the pseudogap temperature at $x = 0.13$ before it was observed experimentally [98] by a group unaware of the theoretical prediction. In addition, the prediction that ξ should scale like ξ_{th} at the pseudogap temperature has been verified in neutron scattering experiments [91] in the range $x = 0.04$–0.15. At that doping, which corresponds to optimal doping, T^* becomes of the order of 100 K, more than four times lower than at $x = 0.04$. The antiferromagnetic correlation length ξ beyond optimal doping begins to decrease and violate the scaling of ξ with ξ_{th}. In that doping range, T^* and the superconducting transition temperature are close. Hence it is likely that there is interference between the two phenomena [100], an effect that has not yet been taken into account in TPSC.

An important prediction that one should verify is that inelastic neutron scattering will find over-damped spin fluctuations in the pseudogap regime and that the characteristic spin fluctuation energy will be smaller than $k_{\text{B}}T$ whenever a pseudogap is present. Equality should occur above T^*.

Finally, note that the agreement found in Fig. 13.6 between ARPES and TPSC is for $U \sim 6\,t$. At smaller values of U the antiferromagnetic correlations are not strong enough to produce a pseudogap in that temperature range. For larger U, the weight near $(\pi/2, \pi/2)$ disappears, in disagreement with experiments. The same value of U is found for the same reasons in strong coupling calculations with cluster perturbation theory (CPT) [92] and with slave boson methods [101]. Recent first principle calculations [102] find essentially the same value of U. In that approach, the value of U is fixed, whereas in TPSC it was necessary to increase U by about 10% moving towards half-filling to get the best agreement with experiment. In any case, it is quite satisfying that weak and strong coupling methods agree on the value of U for electron-doped cuprates. This value of U is very near the critical value for the Mott transition at half-filling [103]. Hence, antiferromagnetic fluctuations at finite doping can be very well described by Slater-like physics (nesting) in electron-doped cuprates.

For recent calculations including the effect of the third dimension on the pseudogap, see [104]. Finally, note that the analog of the above mechanism for the pseudogap has also been seen in two-dimensional charge-density wave dichalcogenides [105].

13.3.2 d-Wave Superconductivity

In the BCS theory of superconductivity, pairs of electrons form because of an effective attraction mediated by phonons. The pairs then condense in a coherent

state. The suggestion that superconductivity could arise from purely repulsive forces goes back to Kohn and Luttinger who showed that for pairs of sufficiently high angular momentum, the *screened* Coulomb interaction in an electron gas could be attractive [106]. Just before the discovery of high-temperature superconductors, an extension of that idea was proposed [107–109]. The suggestion was that antiferromagnetic fluctuations present in the Hubbard model could replace the phonons in BCS theory and lead to d-wave superconductivity. This is difficult to prove beyond any doubt since superconductivity in this case does not arise at the mean-field level. Mean-field on the Hubbard model gives antiferromagnetism near half-filling but not superconductivity. In the high-temperature superconductors, the situation is made even more difficult because of Mott Physics. Nevertheless, the question is well posed and, as we just saw, Mott physics might be less important in the electron-doped superconductors.

To investigate how the pairing susceptibility is influenced by antiferromagnetic fluctuations in TPSC we proceed as follows [42, 110]. The reader who did not go through the formal Sect. 13.2.5 may skip the next paragraph without loss of continuity to read the physical results below. The few equations that appear below give details that are missing in the literature.

13.3.2.1 Some Details of the Derivation

We work in Nambu space and add an off-diagonal source field θ and θ^* in the generating function (13.30). The transverse spin fluctuations are included by working with four by four matrices. The pair susceptibility in the normal state can be obtained from the second functional derivative of the generating function with respect to the off-diagonal source field, evaluated at zero source field [42]. In more detail, one proceeds as in the formal derivation in Sect. 13.2.5. The expressions for the spin and charge susceptibilities are not modified. Once the two-particle quantities have been found as above, the second step of the approach [7,49] consists in improving the approximation for the single-particle self-energy by starting from the *exact* expression where the high-frequency Hartree–Fock behavior is explicitly subtracted

$$\begin{aligned}\boldsymbol{\Sigma}\left(1,2\right) = -U\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\frac{\delta\mathbf{G}\left(1,\overline{3}\right)}{\delta\phi_{\downarrow}\left(1^{+},1\right)}\mathbf{G}^{-1}\left(\overline{3},2\right)\\ +U\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\frac{\delta\mathbf{G}\left(1,\overline{3}\right)}{\delta\phi_{\uparrow}\left(1^{+},1\right)}\mathbf{G}^{-1}\left(\overline{3},2\right).\end{aligned} \tag{13.54}$$

The bold face objects are matrices in Nambu space. To be able to express the right-hand side of the above equation in terms of irreducible vertices, susceptibilities and powers of G, one differentiates $\mathbf{G}\mathbf{G}^{-1} = \mathbf{I}$ to obtain $\left(\delta\mathbf{G}/\delta\phi\right)\mathbf{G}^{-1} = -\mathbf{G}\left(\delta\mathbf{G}^{-1}/\delta\phi\right)$. With the help of Dyson's equation on the right-hand side of the last equation as well as the chain-rule, one finds an expression where one can replace

every term by their value at the first step, namely, $U_{\rm sp}$ and $U_{\rm ch}$ for the irreducible low-frequency vertices as well as $G_\sigma^{(1)}(k+q)$ and $\chi_{\rm sp}(q)$, $\chi_{\rm ch}(q)$. For the diagonal piece of the self-energy at the second step, one then obtains (13.26) above or equivalently (3) of [20,49] by considering both longitudinal and transverse channels and requiring crossing symmetry of the fully reducible vertex in the two particle–hole channels as well as consistency with the sum-rule $\mathrm{Tr}\left(\Sigma^{(2)}G^{(1)}\right) = 2U\left\langle n_\uparrow n_\downarrow\right\rangle$ [7]. The off-diagonal piece of the exact self-energy (13.54), on the other hand, reads

$$\Sigma_{12}^{(2)} = -UG_{11}^{(1)}\left(4,\overline{3}\right)\frac{\delta\Sigma_{12}^{(1)}\left(\overline{3},5\right)}{\delta\phi_\downarrow\left(4^+,4\right)} + UG_{12}^{(1)}\left(4,4^+\right)\delta\left(4-5\right) - UG_{12}^{(1)}\left(4,\overline{3}\right)\frac{\delta\Sigma_{22}^{(1)}\left(\overline{3},5\right)}{\delta\phi_\downarrow\left(4^+,4\right)}. \quad (13.55)$$

The pairing susceptibility mediated by spin fluctuations may now be computed from the derivative with respect to the source field

$$\left.\frac{\delta G_{12}^{(2)}\left(1,3\right)}{\delta\theta\left(2,4\right)}\right|_{\theta=0} = G_{11}^{(2)}\left(1,2\right)G_{22}^{(2)}\left(4,3\right) + G_{11}^{(2)}\left(1,\overline{4}\right)\left.\frac{\delta\Sigma_{12}^{(2)}\left(\overline{4},\overline{5}\right)}{\delta G_{12}^{(1)}\left(\overline{6},\overline{7}\right)}\right|_{\theta=0}\left.\frac{\delta G_{12}^{(1)}\left(\overline{6},\overline{7}\right)}{\delta\theta\left(2,4\right)}\right|_{\theta=0}G_{22}^{(2)}\left(\overline{5},3\right) \quad (13.56)$$

with the irreducible vertex $\delta\Sigma_{12}^{(2)}/\delta G_{12}^{(1)}$ obtained by functional differentiation of (13.55). Neglecting $\delta/\delta\phi_\downarrow\left(\delta\Sigma_{12}^{(1)}/\delta G_{12}^{(1)}\right)$, which represents the influence of spin fluctuations on the local piece of the irreducible vertex, and including the transverse component, we find for the d-wave pair susceptibility, $\chi_{\rm d}$, the expression that appears in (1) of [42]. The TPSC expression for the pair susceptibility $\chi_{\rm d}$ contains the bubble part and the first term of what would be an infinite series if $\Sigma_{12}^{(2)}$ in the irreducible vertex could be differentiated with respect to $G_{12}^{(2)}$ instead of $G_{12}^{(1)}$.

Why should we trust the results for the d-wave susceptibility obtained for TPSC? Let us look again at benchmarks. Figure 13.7a displays the d-wave susceptibility obtained from QMC calculations shown as symbols and from TPSC as lines. Because of the sign problem, it is not practical to do the QMC calculations at lower temperatures. Nevertheless, the temperatures are low enough that we see a nontrivial effect, the appearance of a maximum in susceptibility at finite doping and a substantial increase with decreasing temperatures. The agreement between QMC and TPSC is found to be within a few percent and improves for lower values of U. When the interaction strength reaches the intermediate coupling regime, $U = 6$, deviations of the order of 20–30% may occur but the qualitative dependence on temperature and doping remains accurate. In TPSC the pseudogap is the key ingredient that leads to a decrease in $\chi_{\rm d}$ in the underdoped regime. This is easy

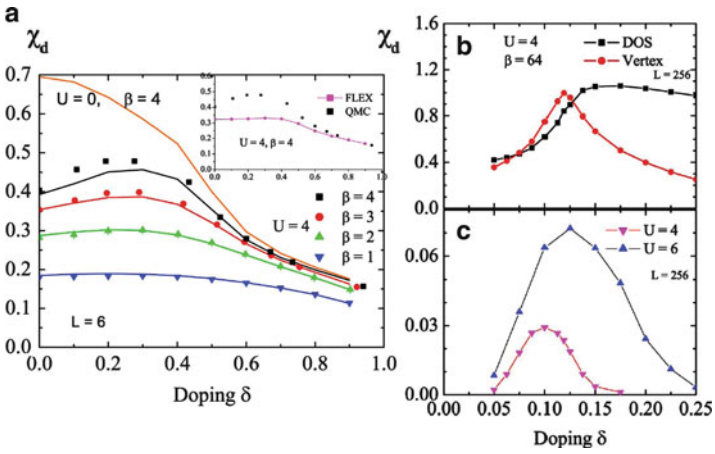

Fig. 13.7 (Color online) From [42]. (**a**) Comparisons between the $d_{x^2-y^2}$ susceptibility obtained from QMC simulations and from the approach described in the present work. QMC *error bars* are smaller than the *symbols*. Analytical results are joined by *solid lines*. Both calculations are for $U = 4$, a 6×6 lattice, and four different temperatures. The case $U = 0$, $\beta = 4$ is shown for reference. The size dependence of the results is small at these temperatures. The inset compares QMC and FLEX at $U = 4$, $\beta = 4$. (**b**) Contributions from the bubble (DOS) represented by *squares* and vertex represented by *circles*. (**c**) Estimate of T_c using the Thouless criterion for $U = 4$ and $U = 6$, $t' = t'' = 0$

to understand since the pseudogap leaves fewer states for pairing at the Fermi level. Another way to say this is that the strong inelastic scattering that leads to the pseudogap is pair breaking. The inset shows that previous spin-fluctuation calculations (FLEX) in two dimensions [16, 111] deviate both qualitatively and quantitatively from the QMC results. More specifically, in the FLEX approach χ_d does not show a pronounced maximum at finite doping. This is because, as we have shown in Fig. 13.4, in FLEX there is no pseudogap in the single-particle spectral weight at the Fermi surface [49, 51].

In two dimensions, superconductivity is in the Kosterlitz–Thouless universality class. Vortex physics that is absent in TPSC is important to understand the precise value of the transition temperature. Nevertheless, a necessary condition for this transition to occur is that there is a higher temperature at which pairs form, a sort of mean-field T_c. That T_c can be obtained from the so-called Thouless criterion, i.e., from the temperature at which the d-wave susceptibility diverges. This divergence occurs because of growing vertex corrections. Since TPSC contains only the first term in what would be an infinite ladder sum, we take T_c as the temperature at which the bubble (that we call DOS) and the first term of the series (that we call vertex) become equal. This is illustrated in Fig. 13.7b. At $\beta = 1/64$, the doping range between the intersection of the two curves is below T_c. The resulting T_c versus doping is shown in Fig. 13.7c. In that calculation, $\langle n_\uparrow n_\downarrow \rangle$ is fixed at its value at T^*.

There is clearly an additional approximation in finding T_c with TPSC that goes beyond what can be checked with QMC calculations. How can we be sure that this is

correct? First there is consistency with other weak-coupling approaches. For example, Ref, [112] has shown consistency with cutoff renormalization group technique [24] for competing ferromagnetic, antiferromagnetic and d-wave superconductivity. Second, there is consistency as well with quantum cluster approaches that are best at strong coupling. Indeed, an extensive study as a function of system size by the group of Jarrell [81] has shown that for 10% doping, T_c is in the range 0.02, not far from 0.03 that can be read from Fig. 13.7c.

One of the major theoretical questions in the field of high-temperature superconductivity has been, "Is there d-wave superconductivity in the two-dimensional Hubbard model?" TPSC has contributed to answer this question. One of the most discouraging early results was that QMC simulations showed that the pairing susceptibility was smaller at finite U than at $U = 0$. This is clearly seen in Fig. 13.7a. TPSC allows us to understand why. At $\beta = 4$, the bubble largely dominates and the effect can only be pair breaking because of the inelastic scattering. The vertex, representing exchange of antiferromagnetic fluctuations analogous to the phonons in ordinary BCS theory, contributes 22% at most at zero doping and much less at larger doping. Nevertheless, it clearly increases the pair susceptibility to bring it in closer agreement with QMC. TPSC allows us to do the calculation at temperatures much lower than QMC and to verify that indeed the vertex eventually grows large enough to lead to a transition. We understand also that in this parameter range the dome shape comes from the fact that antiferromagnetic fluctuations can both increase pairing through the vertex and be detrimental through the pseudogap produced by the large self-energy. Antiferromagnetic fluctuations can both help and hinder d-wave superconductivity.

Another question is whether the presence of a quantum critical point below the maximum of the superconducting dome plays a role in superconductivity. In the case we discussed above, long-range antiferromagnetic order appears at $T = 0$ (not at finite T because of Mermin–Wagner) up to doping $\delta = 0.17$ for $U = 4$ [42] and $\delta = 0.205$ for $U = 6$ [36]. In this case, then, according to Fig. 13.7c the quantum critical point is far to the right of the maximum T_c but superconductivity can exist to the right of that point, the more so when U is larger.

How general are the above results? This and many more questions on the conditions for magnetically mediated superconductivity were studied with TPSC on the square lattice *at half-filling* for second-neighbor hopping t' different from zero [110]. At $t' = 0$ at half-filling, there is a pseudogap on the whole Fermi surface because of perfect nesting, so T_c vanishes. When t' is increased from zero, the pseudogap is not complete at half-filling and T_c is different from zero. In addition, for t' larger than 0.71, the Fermi surface topology changes and the dominant magnetic fluctuations are near $(0, \pm\pi)$, $(\pm\pi, 0)$.

Additional conclusions of the TPSC study of [110] as a function of t' and U at half-filling are as follows. First some qualitative conclusions that could be found from just the BCS gap equation with an interaction potential given by the static component of the spin susceptibility [113]: The symmetry of the d-wave order parameter is determined by the wave vector of the magnetic fluctuations.

Those that are near (π, π) lead to $d_{x^2-y^2}$-wave (B_{1g}) superconductivity while those that are near $(0, \pi)$ induce d_{xy}-wave (B_{2g}) superconductivity. The dominant wave vector for magnetic fluctuations is determined by the shape of the Fermi surface, so $d_{x^2-y^2}$-wave superconductivity occurs for values of t' that are relatively small while d_{xy}-wave superconductivity occurs for $t' > 1$. Second, the maximum value that T_c can take as a function of t' increases with interaction strength. TPSC cannot reach the strong-coupling regime where T_c should decrease with U.

One also finds [110] that, contrary to what is expected from BCS, the noninteracting single-particle density of states does not play a dominant role. At small t', T_c is reduced by self-energy effects as discussed above and for intermediate values of t' the magnetic fluctuations are smaller and incommensurate so no singlet superconductivity appears. Hence at fixed U, *there is an optimal value of* t' *(frustration)* for superconductivity. For $d_{x^2-y^2}$ superconductivity in under-frustrated systems (small t') T_c occurs below the temperature T_X where the crossover to the renormalized classical regime occurs. In other words, in under-frustrated systems at T_c the antiferromagnetic correlation length is much larger than the thermal de Broglie wavelength and the renormalized classical spin fluctuations dominate. The opposite relationship between these lengths occurs for over-frustrated systems (t' larger than optimal) where T_c is larger than T_X and hence occurs in a regime where renormalized classical fluctuations do not dominate. The two temperatures, T_c and T_X, are comparable for optimally frustrated systems. In all cases, at T_c the antiferromagnetic correlation length is larger than the lattice spacing.

The superconductivity induced by antiferromagnetic fluctuations in weak to intermediate coupling has also been studied by Moriya and Ueda [17] with the self-consistent renormalized approach that also satisfies the Mermin–Wagner theorem. However, in that approach there are adjustable parameters and no guarantee that the Pauli principle is satisfied so one cannot be certain this is an accurate solution to the Hubbard model.

That there is d-wave superconductivity in the two-dimensional Hubbard model has by now been seen by a number of different approaches: variational[19] [114, 115], various quantum cluster approaches[20] [80–85] functional renormalization group [116], and even at asymptotically small U by renormalization group [117]. The retardation that can be observed even tells us that spin fluctuations remain important for d-wave superconductivity even at strong coupling [77, 118, 119]. The most serious objection to the existence of d-wave superconductivity comes from a variational and a gaussian QMC approach in [120]. It could be that d-wave superconductivity in the two-dimensional Hubbard model is not the absolute minimum but only a local one. If this were the case, one could conclude that a small interaction term is missing in the Hubbard model to make the d-wave state the ground state. All other studies show that the physical properties of that state are very close than those of actual materials.

[19] See contribution of Randeria in this volume.

[20] See contribution of Sénéchal in this volume.

13.4 More Insights on the Repulsive Model

The following two sections of this chapter give a short summary of other results obtained with TPSC. The purpose is to show what has already been done, leading to the last section that contains a few open problems that could possibly be treated with TPSC. These two sections have some of the flavor of a review article, not that of a pedagogical introduction. In addition, an important aspect of a real review article is missing: there are very few references to the rest of the literature on any given topic. For anyone interested in pursuing some of these problems, the citation index is highly recommended.

13.4.1 Critical Behavior and Phase Transitions

The self-consistent renormalized approach of Moriya–Lonzarich–Taillefer [18, 19] was one of the first ones to treat the Hubbard model in two dimensions in a way that satisfies the Mermin–Wagner theorem. Other approaches exist for the half-filling case: Schwinger bosons [121] or constrained spin-waves [122]. The drawback of Moriya's approach is that it contains several fitting parameters. Kanamori screening, discussed in Sect. 13.2.2, is put by hand, as is the value of the mode coupling constant. In addition, nothing guarantees the Pauli principle. In other words, Moriya's approach has much of the same physics as TPSC but it cannot be considered an accurate solution to the Hubbard model. There is also no prescription to compute the self-energy in a way that is consistent with double-occupancy.

More generally, the question that arises with TPSC is whether it predicts the correct universality class. It was shown in [38] that its results are in the universality class of the spherical model, namely, $O(N=\infty)$ instead of $O(N=3)$ as it should be for the Hubbard model with spin–rotation invariance. This result is not surprising since the self-consistency condition on double-occupancy found from the local moment sum-rule (13.16) is very similar to the self-consistency condition for the spherical model. With the standard convention for critical exponents, one finds $\gamma/\nu = 2, z = 2$, and for dimension d such that the condition $2 < d < 4$ is satisfied, we find $\nu = 1/(d-2)$. This gives in $d = 3, \nu = 1, \gamma = 2, \beta = 1/2, \eta = 0$ and $\delta = 5$. This should be compared with numerical results [123] for the 3D Heisenberg ($n = 3$) model, $\nu = 0.7$ and $\gamma = 1.4$. Clearly, too close to the critical point, or too deep in the renormalized classical regime in $d = 2$, TPSC looses its accuracy. Results in three dimensions can be found in [55, 124].

13.4.1.1 Crossover to 3D [38]

The crossover from two- to three-dimensional critical behavior of nearly antiferromagnetic itinerant electrons was also studied in a regime where the interplane single-particle motion of electrons is quantum mechanically incoherent because of

thermal fluctuations. The universal renormalized classical crossover function from $d = 2$ to $d = 3$ for the susceptibility has been explicitly computed, as well as a number of other properties such as the dependence of the Néel temperature on the ratio between hopping in the plane $t_{\|}$ and the hopping perpendicular to it, $t_{\perp}$,

$$\frac{1}{T_N} \sim \frac{T_N^2}{U_{mf,c}^2} \left| \ln \left(\frac{t_{\|}}{t_{\perp}} \right) \right| \tag{13.57}$$

with $U_{mf,c} \equiv 2/\chi\left(\mathbf{Q}_{d=2}, 0\right)$ at the $d = 2$ pseudogap temperature [38].

13.4.1.2 Quantum Critical Behavior

At $T = 0$, half-filling, the ground state has long-range antiferromagnetic order. As one dopes, the order becomes incommensurate and eventually disappears at a critical point that is called "quantum critical" because it occurs at $T = 0$. Such quantum critical points are common in heavy-fermion systems for example. One of the surprising things about this critical point is that it affects the physics at surprisingly large T.

The quantum critical behavior of TPSC in $d = 2$ is in the $z = 2$ universality class, like the self-consistent renormalized theory of Moriya et al. [39]. Like that theory, it includes some of the logarithmic corrections found in the renormalization group approach [125]. In addition, TPSC can be quantitative and answer the question, "How far in T does the influence of that point extend?" It was found [40, 126] by explicit numerical calculations away from the renormalized classical regime of the $d = 2$ Hubbard model that logarithmic corrections are not really apparent in the range $0.01t < T < t$ and that the maximum static spin susceptibility in the (T, n)-plane obeys quantum critical scaling. However, near the commensurate–incommensurate crossover, one finds obvious nonuniversal T and filling n dependence. Everywhere else, the (T, n)-dependence of the nonuniversal scale factors is relatively weak. Strong deviations from scaling occur at T of order t, the degeneracy temperature. That high temperature limit should be contrasted with $J/2$ found in the strong coupling case [127]. In generic cases the upper limit $T \sim t$ is well-above room temperature. In experiment, however, the non-universality due to the commensurate–incommensurate crossover may make the identification of quantum critical scaling difficult. In addition, note that properties other than the maximum spin susceptibility may deviate from quantum critical scaling at a lower temperature [36].

13.4.2 Longer Range Interactions

Suppose one adds nearest-neighbor repulsion V to the Hubbard model. The TPSC ansatz (13.39) can be generalized [128, 129]. Then one needs to compute the

effect of functional derivatives of the pair-correlation functions that appear, as in Sect. 13.2.5.3, in the calculation of spin and charge irreducible vertices. Since the Pauli principle and local-spin and local charge sum-rules do not suffice, the functional derivatives are evaluated assuming particle–hole symmetry, which remains approximately true when the physics is dominated by states close to the Fermi surface. The resulting theory, called ETPSC, for extended TPSC, satisfies conservation laws and the Mermin–Wagner theorem and is in agreement with benchmark QMC results. This approach allows us to reliably determine the crossover temperatures toward renormalized-classical regimes, and hence, the dominant instability as a function of U and V. Contrary to RPA, even the spin fluctuations are modified by the presence of V. Phase diagrams have been calculated. In the presence of V, charge order will generally compete with spin order [128, 129].

13.4.3 Frustration

The ETPSC formalism outlined in the previous section is particularly important to treat interesting problems such as that of the sodium cobaltates. These compounds are often modeled in an over-simplified way by the two-dimensional Hubbard model on the triangular lattice. To account for charge fluctuations, one must also include nearest-neighbor repulsion V. Even with this complication this is an oversimplified model.

The density- and interaction-dependent crossover diagram for spin- and charge-density wave instabilities of the normal state at arbitrary wave vector has been computed [67]. When U dominates over V and electron filling is large, instabilities are mostly in the spin sector and are controlled by Fermi surface properties. Increasing V eventually leads to charge instabilities where it is mostly the wave vector dependence of the vertex that determines the wave vector of the instability rather than Fermi surface properties. At small filling, nontrivial instabilities appear only beyond the weak coupling limit. Charge-density wave instabilities are favored over a wide range of dopings by large V at wave vectors corresponding to $\sqrt{3}\times\sqrt{3}$ superlattice in real space. Commensurate fillings do not play a special role for this instability. Increasing U leads to competition with ferromagnetism. At negative values of U or V, neglecting superconducting fluctuations, one finds that charge instabilities are favored. In general, the crossover diagram presents a rich variety of instabilities. Thermal charge-density wave fluctuations in the renormalized-classical regime can open a pseudogap in the single-particle spectral weight, just as spin or superconducting fluctuations [67].

13.4.4 Thermodynamics, Conserving Aspects

Conserving approaches are very popular. FLEX [15, 16] is an example. These approaches are attractive because they guarantee that if one evaluates the same

physical quantity directly from the Green function or from a derivative of the free energy, the answer will be identical. All that is needed for a "conserving" approximation is that the self-energy be generated from a Luttinger–Ward functional [130] that enters the expression for the free energy. In addition, conservation laws will be satisfied in transport if irreducible vertices are obtained from functional derivatives of the self-energy. This gives a so-called Φ derivable theory. Since in perturbation theory there is an infinite number of possible such Luttinger–Ward functionals, depending on which closed two-particle irreducible diagrams constructed from G and the bare interaction one wishes to keep, the constraint of being conserving is not a very restrictive one. Conserving approximations do not satisfy the Pauli principle in general and they sometimes give negative values of double-occupancy [55]. Various other limitations of conserving approaches are discussed in Appendix E of [7].

TPSC is not obtained from a functional derivative of a Luttinger–Ward functional. Is this a drawback? We have seen that it satisfies conservation laws for spin and charge at the first step. The question is whether one can find a unique free energy that is consistent with the one-particle Green function and collective modes that TPSC focuses on. This question was addressed in the MSc [131] and PhD thesis [126] of Roy. The results are summarized below. We conclude with an example of thermodynamic calculation in the context of cold atoms.

13.4.4.1 Thermodynamic Consistency

We should really distinguish conservation laws and thermodynamic consistency. These two notions are sometimes confused, as outlined in the previous paragraph. We call an approach thermodynamically consistent when all possible ways of computing the same thermodynamic quantity give the same result.

Obtaining the self-energy from a functional derivative of the Luttinger–Ward functional leads to thermodynamic consistency. In TPSC, there is a change in perspective. Instead of looking for an approximation for the free energy and then deducing everything else consistently, we find a single-particle Green's function $G^{(2)}$, or equivalently $\Sigma^{(2)}$, as well as double-occupancy

$$\frac{1}{2}\mathrm{Tr}\left(\Sigma^{(2)}G^{(2)}\right) = U\left\langle n_\downarrow n_\uparrow\right\rangle^{(2)}. \tag{13.58}$$

and deduce everything else: the free energy from integration and the irreducible vertices for transport quantities from functional derivatives of $\Sigma^{(2)}$(Sect. 13.4.5).

There are three ways to extract the free energy from integration: (a) coupling constant integration of double-occupancy, (b) integration of $\mu(n)$ and (c) integration of the specific heat calculated from the total energy. With the free energy, all thermodynamic quantities can be obtained. If we make sure that the three ways to compute the free energy give the same result, then there is thermodynamic consistency.

With the above expression for double-occupancy, (13.58), the three different ways of obtaining the thermodynamic quantities are all based on the same object $G^{(2)}$. If $G^{(2)}$ were the exact solution, they would have to be consistent. However $G^{(2)}$ is approximate. Since $G^{(2)}$ satisfies all the requirements for a physical Green function, it is likely to be the exact solution of some Hamiltonian H that is close to, but slightly different from, the Hubbard model. For example, H could have longer-range interactions. In deriving the formulas for the free energy, we assume that we are working with the Hubbard model. Hence, there is no guarantee that all three methods of obtaining F will give the same result. One can check this numerically in principle. A simpler test, admittedly less stringent, is to compare $n(T, \mu, U)$ obtained from derivatives of the three different F's. For the nearest-neighbor hopping model with $\beta = 10$, $U = 4$, for example, the results are identical in the percent range, except deep in the renormalized classical regime close to half-filling where TPSC anyway fails [126, 131].

The specific heat was calculated for the nearest-neighbor hopping Hubbard model at half-filling as a function of temperature. TPSC reproduces the peak observed at small temperature in QMC [132]. It is associated with the entrance in the renormalized classical regime. Physically, the low-temperature peak is a remnant of the specific heat jump that would occur at finite temperature in the mean-field theory.

13.4.4.2 Cold Atoms, Entropy

In the context of cold atoms on optical lattices, adiabatic cooling can be used to reach interesting low T regimes such as the pseudogap or ordered phases by manipulating the scattering length or the strength of the laser-induced lattice potential. TPSC has been used [28], and compared with QMC calculations, to provide isentropic curves for the two- and three-dimensional Hubbard models at half-filling. Since double-occupancy D is extremely accurate in TPSC, the entropy S was computed by integrating the Maxwell relation

$$\left(\frac{\partial S}{\partial U}\right)_{T,n} = -\left(\frac{\partial D}{\partial T}\right)_{U,n} \tag{13.59}$$

with $S(T, U = 0)$, the known constant of integration.

The main findings are that adiabatically turning on the interaction in $d = 2$ to cool the system is not very effective. In three dimensions, adiabatic cooling to the antiferromagnetic phase can be achieved in such a manner, although the cooling efficiency is not as high as initially suggested by dynamical mean-field theory [133]. Adiabatic cooling by turning off the repulsion beginning at strong coupling is possible in certain cases.

13.4.5 Vertex Corrections and Conservation Laws

Using the functional derivative formalism of Baym and Kadanoff [58], it is possible to find an expression within TPSC for the optical conductivity that satisfies conservation laws and hence the f-sum rule. In other words, one can include vertex corrections. Note that the f-sum rule this time involves the momentum distribution $n_{\mathbf{k}}^{(2)}$ obtained from the best self-energy.

The two types of vertex corrections that are found [36] are the antiferromagnetic analogs of the Aslamasov–Larkin and Maki–Thompson contributions of superconducting fluctuations to the conductivity but, contrary to the latter, they include nonperturbative effects. The calculations are impossible unless a number of advanced numerical algorithms are used. Take the case with nearest-neighbor hopping only [36]. In the pseudogap regime induced by two-dimensional antiferromagnetic fluctuations, the effect of vertex corrections is dramatic. Without vertex corrections the resistivity increases as we enter the pseudogap regime. Instead, vertex corrections lead to a drop in resistivity, as observed in a number of high-temperature superconductors. At high temperature, the resistivity naturally saturates at the Ioffe–Regel limit. At the quantum critical point and beyond, the resistivity displays both linear and quadratic temperature dependence. The disappearance of superconductivity in the over-doped regime is correlated with the disappearance of the linear term in the T dependence of the resistivity [134, 135]. The relation to the physics of hot spots and results for other band structures ($t' \neq 0$) should appear soon.

13.5 Attractive Hubbard Model

Working in Nambu space and following a formal procedure analogous to that explained in Sect. 13.2.5, one can derive TPSC for the attractive Hubbard model [20, 57, 136]. The irreducible vertex U_{pp} in the particle–particle singlet channel is given by

$$U_{\mathrm{pp}} \langle n_\downarrow \rangle \langle (1 - n_\uparrow) \rangle = U \langle n_\downarrow (1 - n_\uparrow) \rangle \tag{13.60}$$

and is determined self-consistently at the two-particle level by the local-pair sum-rule

$$\langle n_\downarrow n_\uparrow \rangle = \langle \Delta^\dagger \Delta \rangle = \frac{T}{N} \sum_q \chi_{\mathrm{p}}^{(1)}(q) \exp(-\mathrm{i}\omega_n 0^-) \tag{13.61}$$

with

$$\chi_{\mathrm{p}}^{(1)}(q) = \frac{\chi_0^{(1)}(q)}{1 + U_{\mathrm{pp}} \chi_0^{(1)}(q)} \tag{13.62}$$

and the irreducible particle–particle susceptibility

$$\chi_0^{(1)}(q) = \frac{T}{N}\sum_k G_\sigma^{(1)}(q-k)G_{-\sigma}^{(1)}(k)\ . \tag{13.63}$$

Again the Pauli principle and a number of crucial sum-rules are satisfied. So is the Mermin–Wagner theorem.

In the second step of the approximation, an improved expression for the self-energy is obtained by using the results of the first step in an exact expression for the self-energy, to obtain,

$$\Sigma_\sigma^{(2)}(k) = Un_{-\sigma} - U\frac{T}{N}\sum_q U_{\rm pp}\chi_{\rm p}^{(1)}(q)G_{-\sigma}^{(1)}(q-k), \tag{13.64}$$

where $q = (i\omega_n, \mathbf{q})$. This is a cooperon-like formula. The required vertex corrections are included as required by the absence of a Migdal theorem. Comparisons with other approaches can be found in [137].

13.5.1 Pseudogap from Superconductivity in Attractive Hubbard Model

Using TPSC for the attractive Hubbard model, quantitative agreement with Monte Carlo calculations is obtained for both single-particle and two-particle quantities [50]. As discussed for the repulsive case in Sect. 13.2.6 one obtains a pseudogap in both the density of states and the single-particle spectral weight[21] below some characteristic temperature T^*. It was even checked in QMC calculations that the ratio of the thermal de Broglie wavelength to the pairing correlation length must be smaller than unity to observe the pseudogap [140]. The pseudogap, also found in [141] for example, reflects precursors of Bogoliubov quasiparticles that are not local pairs, contrary to what is often discussed in the context of the crossover from BCS to Bose-Einstein condensation [142].

With increasing temperature the spectral weight fills in the pseudogap instead of closing it [50]. This type of behavior is obtained in high-temperature superconductors. The pseudogap appears earlier in the density of states than in the spectral function. A characteristic behavior observed at strong coupling appears already in TPSC at weak to intermediate coupling, namely, small temperature changes around T^* can modify the spectral weight over frequency scales much larger than temperature [50].

[21] For a pseudogap in the single-particle spectral weight, it is important not to assume a Migdal theorem [138, 139], and include vertex corrections [7].

Our earlier discussion about Kosterlitz–Thouless physics in Sect. 13.3.2 is valid in this case as well. In the attractive Hubbard model, the superconducting transition temperature has a dome shape because at half-filling the symmetry is SO(3) so the T_c there vanishes while it is given by the finite Kosterlitz–Thouless T_c [143] elsewhere. The pseudogap temperature on the other hand decreases monotonically from half-filling where it is largest. This exemplifies the fact that symmetry and dimension are important to understand pseudogap physics at weak to intermediate coupling [140].

13.6 Open Problems

At weak to intermediate coupling, TPSC gives the best agreement with benchmark QMC methods. Its strength, compared with all other methods, resides in a non-perturbative treatment of the Hubbard model that satisfies the Pauli principle and the Mermin–Wagner theorem, in addition to a number of other exact constraints. Also, one works in the infinite size limit so the effect of collective fluctuations is not limited to a small lattice like in dynamical mean-field theory [26] and its generalizations[22] [144, 145].

The main weakness of TPSC is the difficulty to extend the method beyond the one-band Hubbard model. One needs to find enough sum-rules to determine the irreducible vertices. This can be seen as a challenge and an opportunity for creativity.

For example, to include nearest-neighbor Coulomb repulsion, one needs a way to evaluate functional derivatives of pair correlation functions to obtain irreducible vertices. It has been possible achieve this [129], as discussed in Sect. 13.4.2, but every new problem is different. As another example, take the case of more than one band. Then the irreducible vertices become a matrix in band index and one does not have enough obvious sum-rules to evaluate all the matrix elements [146].

In the presence of, say, antiferromagnetism the number of irreducible vertices also multiplies and one faces the same type of challenge. To treat long-range ordered states with TPSC, it might be easier to start with simpler broken symmetries such as ferromagnetism or the Pomeranchuk instability [147]. The interest of treating long range order is clear. For example, the renormalized classical regime of antiferromagnetic fluctuations is presently inaccessible if T is much smaller than T^*. Starting from the ordered state may offer an alternative [62].

The question of the interplay of disorder and interactions is a difficult but topical one. Far from the Anderson disorder-induced metal–insulator transition, the impurity averaging technique [148] may prove a useful way to introduce disorder in TPSC. One may then answer the question of what happens to the $\xi > \xi_{th}$ criterion

[22] See contributions of Vollhardt, Sénéchal, Potthoff and Jarrell in this volume.

for pseudogap when the mean-free path becomes shorter than the thermal de Broglie wavelength.

Climbing the ladder of difficult problems, the case of strong coupling [26, 144, 145] is a real challenge. At strong coupling the self-energy is singular at small frequencies. In fact it diverges as $1/\omega$ at half-filling. This is inconsistent with the starting point of TPSC where the self-energy is constant. Perhaps there is a way to start from the self-energy in the atomic limit inspired by methods that allow for multiple poles to zeroth order [12], or some other way [149], but it is an unsolved problem for now.

Some problems, by contrast, appear straightforward but they can be very tedious. For example, in the presence of incommensurate magnetic fluctuations where singlet d-wave pairing does not occur, can triplet pairing take over? One can proceed along the lines of the derivation for d-wave superconductivity [110] but with matrix source fields to generate triplet pairing in the Nambu formalism. The irreducible pairing vertex would again be obtained from functional derivatives of a matrix $\Sigma^{(2)}$.

Paring in the attractive Hubbard model is much more straightforward, as we have seen. It appears at the first step, without the need to generate irreducible vertices from $\Sigma^{(2)}$. Nevertheless, to study the triplet channel in the attractive Hubbard model, one needs to introduce near-neighbor attraction V. That leads to the problems mentioned above in the repulsive case. As a curiosity, one could also investigate whether functional derivatives of $\Sigma^{(2)}$ in the attractive Hubbard model can mediate the formation of order in some particle–hole channel. This might be a first step towards developing a method to take into account the different channels on the same footing in TPSC [150], as is done in renormalization group approaches [151, 152]. It has been found recently within the renormalization group that in quasi one-dimensional systems there is a strong interference between antiferromagnetism and unconventional superconductivity [100].

Proceeding along the lines of [36] for the conductivity, it is also clearly possible to compute other transport quantities, such as the thermopower, but this is a serious computational challenge.

There are roadblocks, but there are also opportunities for original solutions and breakthroughs.

Acknowledgments I am indebted to Yury Vilk who was at the origin of most of the ideas on TPSC and to Bumsoo Kyung who has collaborated on numerous TPSC project. In addition, many students, postdocs and colleagues over the years have worked extremely hard and made use of their creative powers and originality to extend and apply this approach to a large number of problems. Students and postdocs include in chronological order Liang Chen, Alain Veilleux, Anne-Marie Daré, Steve Allen, Hugo Touchette, Samuel Moukouri, Bumsoo Kyung, François Lemay, David Poulin, Jean-Sébastien Landry, Vasyl Hankevych, Bahman Davoudi, Syed Raghib Hassan, Sébastien Roy, Charles Brillon and Dominic Bergeron. I am also indebted to my colleagues David Sénéchal, Claude Bourbonnais, and collaborators Gilbert Albinet and Anne-Marie Daré. I am grateful to D. Sénéchal and D. Bergeron for critical comments on this work. This work was partially supported by NSERC (Canada) and by the Tier I Canada Research Chair Program.

References

1. J. Hubbard, Proc. Roy. Soc. London A **276**, 238 (1963)
2. J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963)
3. M.C. Gutzwiller, Phys. Rev. Lett. **10**(5), 159 (1963)
4. E.H. Lieb, F.Y. Wu, Phys. Rev. Lett. **20**(25), 1445 (1968)
5. W. Metzner, D. Vollhardt, Phys. Rev. Lett. **62**(3), 324 (1989)
6. Y. Vilk, L. Chen, A.-M.S. Tremblay, Phys. Rev. B **49**, 13267 (1994)
7. Y. Vilk, A.-M.S. Tremblay, J. Phys I (France) **7**, 1309 (1997)
8. G.D. Mahan, *Many-Particle Physics* (Kluwer, New-York, 2000)
9. N. Mott, Proc. Phys. Soc. A **62**, 416 (1949)
10. N.D. Mermin, H. Wagner, Phys. Rev. Lett. **17**(22), 1133 (1966)
11. P.C. Hohenberg, Phys. Rev. **158**(2), 383 (1967)
12. A. Avella, F. Mancini, Adv. Phys. **53**(5–6) (2004)
13. A. Avella, F. Mancini, Condens. Matter Phys. **9**(3), 569 (2006)
14. A. Avella, F. Mancini, Euro. Phys. J. B **36**(1), 37 (2003)
15. N. Bickers, D. Scalapino, Ann. Phys. (USA) **193**(1), 206 (1989)
16. N. Bickers, D. Scalapino, Phys. Rev. Lett. **62**, 961 (1989)
17. T. Moriya, K. Ueda, Rep. Prog. Phys. **66**(8), 1299 (2003)
18. G.G. Lonzarich, L. Taillefer, J. Phys. C **18**, 4339 (1985)
19. T. Moriya, *Spin Fluctuations in Itinerant Electron Magnetism* (Springer, Heidelberg, 1985)
20. S. Allen, A.-M.S. Tremblay, Y.M. Vilk, in *Theoretical Methods for Strongly Correlated Electrons*, CRM Series in Mathematical Physics, (Springer, New York, 2003), p. 341–355, ed. by D. Sénéchal, C. Bourbonnais, A.M. Tremblay (2003)
21. I. Dzyaloshinskii, Sov. Phys. JETP **66**, 848 (1987)
22. H.J. Schulz, Europhys. Lett. **4**, 609 (1987)
23. P. Lederer, G. Montambaux, D. Poilblanc, J. Phys. **48**, 1613 (1987)
24. C. Honerkamp, M. Salmhofer, Phys. Rev. Lett. **87**, 187004 (2001)
25. M.A. Tusch, Y.H. Szczech, D.E. Logan, Phys. Rev. B **53**(9), 5505 (1996)
26. A. Georges, G. Kotliar, W. Krauth, M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
27. H. Nélisse, C. Bourbonnais, H. Touchette, Y. Vilk, A.-M.S. Tremblay, Euro. Phys. J. B (France) **12**, 351 (1999)
28. A.M. Daré, L. Raymond, G. Albinet, A.-M.S. Tremblay, Phys. Rev. B **76**(6), 064402 (2007)
29. A.-M.S. Tremblay, B. Kyung, D. Sénéchal, Low Temp. Phys. **32**(4–5), 424 (2006)
30. G. Baym, Phys. Rev. **127**, 1391 (1962)
31. N.E. Bickers, S.R. White, Phys. Rev. B **43**(10), 8044 (1991)
32. V. Janis, Condens. Matter Phys. **9**, 499 (2006)
33. K.S. Singwi, M.P Tosi, *Solid State Physics*, ed. by H. Ehrenreich, F. Seitz, D. Turnbull (Academic, New York, 1981)
34. S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982)
35. M.R. Hedeyati, G. Vignale, Phys. Rev. B **40**, 9044 (1989)
36. D. Bergeron, B. Kyung, V. Hankevych, A.-M.S. Tremblay, arXiv:1101.4037 and D. Bergeron, PhD thesis, Université de Sherbrooke (2011) Phys. Rev. B, in press
37. L. Chen, C. Bourbonnais, T. Li, A.-M.S. Tremblay, Phys. Rev. Lett. **66**, 369 (1991)
38. A.M. Daré, Y. Vilk, A.-M.S. Tremblay, Phys. Rev. B **53**, 14236 (1996)
39. T. Moriya, Y. Takahashi, K. Ueda, J. Phys. Soc. Jpn. **59**, 2905 (1990)
40. S. Roy, A.-M.S. Tremblay, EPL (Europhys. Lett.) **84**(3), 37013 (6pp) (2008)
41. Y. Vilk, A.-M.S. Tremblay, Europhys. Lett. **33**, 159 (1996)
42. B. Kyung, J.S. Landry, A.-M.S. Tremblay, Phys. Rev. B **68**, 174502 (2003)
43. A. Veilleux, Master's thesis, Université de Sherbrooke (1994)
44. A. Veilleux, A.M. Daré, L. Chen, Y. Vilk, A.-M.S. Tremblay, Phys. Rev. B **52**, 16255 (1995)
45. B. Kyung, J. Landry, D. Poulin, A.-M.S. Tremblay, Phys. Rev. Lett. **90**, 099702 (2003)
46. N. Menyhard, J. Solyom, J. Low Temp. Phys. **12**, 529 (1973)

47. C. Bourbonnais, Mol. Cryst. Liq. Cryst. **119**, 11 (1985)
48. D. Zanchi, Euophys. Lett. **55**, 376 (2001)
49. S. Moukouri, S. Allen, F. Lemay, B. Kyung, D. Poulin, Y. Vilk, A.-M.S. Tremblay, Phys. Rev. B **61**, 7887 (2000)
50. B. Kyung, S. Allen, A.-M.S. Tremblay, Phys. Rev. B **64**, 075116 (2001)
51. J.J. Deisz, D.W. Hess, J.W. Serene, Phys. Rev. Lett. **76**(8), 1312 (1996)
52. P. Monthoux, Phys. Rev. B **68**(6), 064408 (2003)
53. L. Hedin, J. Phys. Condens. Matter **11**(42), R489 (1999)
54. G. Mahan, *Many-Particle Physics*, 3rd edn., Sect. 6.4.4 (Kluwer, New York, 2000)
55. R. Arita, S. Onoda, K. Kuroki, H. Aoki, J. Phys. Soc. Jpn. **69**(3), 785 (2000)
56. N.F. Berk, J.R. Schrieffer, Phys. Rev. Lett. **17**(8), 433 (1966)
57. S. Allen, A.-M.S. Tremblay, Phys. Rev. B **64**, 075115 (2001)
58. G. Baym, L.P. Kadanoff, Phys. Rev. **124**(2), 287 (1961)
59. P.C. Martin, J. Schwinger, Phys. Rev. **115**(6), 1342 (1959)
60. L. Kadanoff, G. Baym, *Quantum Statistical Mechanics* (Benjamin, Menlo Park, 1962)
61. Y. Vilk, A.-M.S. Tremblay, J. Phys. Chem. Solid (UK) **56**(12), 1769 (1995)
62. K. Borejsza, N. Dupuis, Phys. Rev. B **69**(8), 085119 (2004)
63. P.A. Lee, T.M. Rice, P.W. Anderson, Phys. Rev. Lett. **31**(7), 462 (1973)
64. Y.M. Vilk, Phys. Rev. B **55**(6), 3870 (1997)
65. T. Saikawa, A. Ferraz, Euro. Phys. J. B Condens. Matter Complex Syst. **20**, 65 (2001)
66. J.R. Schrieffer, X.G. Wen, S.C. Zhang, Phys. Rev. B **39**(16), 11663 (1989)
67. B. Davoudi, S.R. Hassan, A.-M.S. Tremblay, Phys. Rev. B **77**(21), 214408 (2008)
68. T. Timusk, B. Statt, Rep. Progr. Phys. **62**(1), 61 (1999)
69. M. Norman, D. Pines, C. Kallin, cond-mat/050731, Adv. Phys. **54**(8), 715–733 (2005)
70. O. Andersen, A. Liechtenstein, O. Jepsen, F. Paulsen, J. Phys. Chem. Solid **56**, 1573 (1995)
71. E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, O.K. Andersen, Phys. Rev. Lett. **87**(4), 047003 (2001)
72. D.C. Peets, D.G. Hawthorn, K.M. Shen, Y.J. Kim, D.S. Ellis, H. Zhang, S. Komiya, Y. Ando, G.A. Sawatzky, R. Liang, D.A. Bonn, W.N. Hardy, Phys. Rev. Lett. **103**(8), 087402 (2009)
73. A. Liebsch, Phys. Rev. B **81**(23), 235133 (2010)
74. P. Phillips, M. Jarrell, Phys. Rev. Lett., vol. 105, 199701 (2010)
75. A. Shekhter, C.M. Varma, Phys. Rev. B **80**(21), 214501 (2009)
76. A. Macridin, T.A. Maier, M.S. Jarrell, G. Sawatzky, Phys. Rev. B **71**, 134527 (2005)
77. W. Hanke, M. Kiesel, M. Aichhorn, S. Brehm, E. Arrigoni, The European Physical Journal - Special Topics, **188**, p.15 (2010)
78. C.T. Chen, F. Sette, Y. Ma, M.S. Hybertsen, E.B. Stechel, W.M.C. Foulkes, M. Schulter, S.W. Cheong, A.S. Cooper, L.W. Rupp, B. Batlogg, Y.L. Soo, Z.H. Ming, A. Krol, Y.H. Kao, Phys. Rev. Lett. **66**(1), 104 (1991)
79. F. Zhang, T.M. Rice, Phys. Rev. B **37**, 3759 (1988)
80. D. Sénéchal, P.L. Lavertu, M.A. Marois, A.-M.S. Tremblay, Phys. Rev. Lett. **94**, 156404 (2005)
81. T.A. Maier, M. Jarrell, T.C. Schulthess, P.R.C. Kent, J.B. White, Phys. Rev. Lett. **95**(23), 237001 (2005)
82. M. Aichhorn, E. Arrigoni, M. Potthoff, W. Hanke, Phys. Rev. B **74**(23), 235117 (2006)
83. M. Aichhorn, E. Arrigoni, Z.B. Huang, W. Hanke, Phys. Rev. Lett. **99**(25), 257002 (2007)
84. K. Haule, G. Kotliar, Phys. Rev. B (Condens. Matter Mater. Phys.) **76**(10), 104509 (2007)
85. S.S. Kancharla, B. Kyung, D. Sénéchal, M. Civelli, M. Capone, G. Kotliar, A.-M.S. Tremblay, Phys. Rev. B (Condens. Matter Mater. Phys.) **77**(18), 184516 (2008)
86. S. Massidda, N. Hamada, J. Yu, A. Freeman, Phys. C Supercond. **157**(3), 571 (1989)
87. G.q. Zheng, T. Sato, Y. Kitaoka, M. Fujita, K. Yamada, Phys. Rev. Lett. **90**(19), 197005 (2003)
88. N.P. Armitage, P. Fournier, R.L. Greene, Rev. Mod. Phys. **82**(3), 2421 (2010)
89. Y. Onose, Y. Taguchi, K. Ishizaka, Y. Tokura, Phys. Rev. Lett. **87**, 217001 (2001)
90. N. Armitage et al., Phys. Rev. Lett. **88**, 257001 (2002)
91. E. Motoyama, G. Yu, I. Vishik, O. Vajk, P. Mang, M. Greven, Nature **445**, 186 (2007)

92. D. Sénéchal, A.-M.S. Tremblay, Phys. Rev. Lett. **92**, 126401 (2004)
93. V. Hankevych, B. Kyung, A.M. Daré, D. Sénéchal, A.-M.S. Tremblay, J. Phys. Chem. Solid **67**, 189 (2006)
94. G. Sordi, K. Haule, A.-M.S. Tremblay, Phys. Rev. Lett. **104**(22), 226402 (2010)
95. Y. Tokura, S. Koshihara, T. Arima, H. Takagi, S. Ishibashi, T. Ido, S. Uchida, Phys. Rev. B **41**(16), 11657 (1990)
96. G. Kotliar, J. Liu, Phys. Rev. Lett. **61**, 1784 (1988)
97. B. Kyung, V. Hankevych, A.M. Daré, A.-M.S. Tremblay, Phys. Rev. Lett. **93**, 147004 (2004)
98. H. Matsui, K. Terashima, T. Sato, T. Takahashi, S.C. Wang, H.B. Yang, H. Ding, T. Uefuji, K. Yamada, Phys. Rev. Lett. **94**, 047005 (2005)
99. J.D. Koralek, J.F. Douglas, N.C. Plumb, Z. Sun, A.V. Fedorov, M.M. Murnane, H.C. Kapteyn, S.T. Cundiff, Y. Aiura, K. Oka, H. Eisaki, D.S. Dessau, Phys. Rev. Lett. **96**(1), 017005 (2006)
100. C. Bourbonnais, A. Sedeki, Phys. Rev. B **80**(8), 085105 (2009) and arXiv:1108.1842
101. Q. Yuan, F. Yuan, C.S. Ting, Phys. Rev. B **72**(5), 054504 (2005)
102. H.K.K.G. Weber, C., Nat. Phys. **6**(8), 574 (2010)
103. H. Park, K. Haule, G. Kotliar, Phys. Rev. Lett. **101**(18), 186403 (2008)
104. T.A. Sedrakyan, A.V. Chubukov, Phys. Rev. B **81**(17), 174536 (2010)
105. S.V. Borisenko, A.A. Kordyuk, A.N. Yaresko, V.B. Zabolotnyy, D.S. Inosov, R. Schuster, B. Büchner, R. Weber, R. Follath, L. Patthey, H. Berger, Phys. Rev. Lett. **100**(19), 196402 (2008)
106. W. Kohn, J.M. Luttinger, Phys. Rev. Lett. **15**(12), 524 (1965)
107. M. Beal-Monod, C. Bourbonnais, V. Emery, Phys. Rev. B **34**, 7716 (1986)
108. D. Scalapino, J. Loh, E., J. Hirsch, Phys. Rev. B **34**, 8190 (1986)
109. K. Miyake, S. Schmitt-Rink, C. Varma, Phys. Rev. B **34**, 6554 (1986)
110. S.R. Hassan, B. Davoudi, B. Kyung, A.-M.S. Tremblay, Phys. Rev. B **77**(9), 094501 (2008)
111. C.H. Pao, N.E. Bickers, Phys. Rev. B **51**(22), 16310 (1995)
112. V. Hankevych, B. Kyung, A.-M.S. Tremblay, Phys. Rev. B **68**, 214405 (2003)
113. D. Scalapino, Phys. Rep. **250**(6), 329 (1995)
114. T. Giamarchi, C. Lhuillier, Phys. Rev. B **43**(16), 12943 (1991)
115. A. Paramekanti, M. Randeria, N. Trivedi, Phys. Rev. B **70**(5), 054504 (2004)
116. R.T. Honerkamp, Physica C: Superconductivity and its Applications **388–389**, 11 (2003)
117. S. Raghu, S.A. Kivelson, D.J. Scalapino, Phys. Rev. B **81**(22), 224505 (2010)
118. T.A. Maier, D. Poilblanc, D.J. Scalapino, Phys. Rev. Lett. **100**(23), 237001 (2008)
119. B. Kyung, D. Sénéchal, A.-M.S. Tremblay, Phys. Rev. B (Condens. Matter Mater. Phys.) **80**(20), 205109 (2009)
120. T. Aimi, M. Imada, J. Phys. Soc. Jpn. **76**(11), 113708 (2007)
121. A. Auerbach, D.P. Arovas, Phys. Rev. Lett. **61**(5), 617 (1988)
122. M. Takahashi, Phys. Rev. Lett. **58**(2), 168 (1987)
123. P. Pfeuty, G. Toulouse, *Introduction to the Renormalization Group and to Critical Phenomena* (Wiley, London, 1977), p. 103
124. A.M. Daré, G. Albinet, Phys. Rev. B **61**(7), 4567 (2000)
125. H.v. Löhneysen, A. Rosch, M. Vojta, P. Wölfle, Rev. Mod. Phys. **79**(3), 1015 (2007)
126. S. Roy, Le modèle de Hubbard bidimensionnel à faible couplage: thermodynamique et phénomènes critiques. Ph.D. thesis, Université de Sherbrooke (2008)
127. A. Kopp, S. Chakravarty, Nat. Phys. **1**, 53 (2005)
128. B. Davoudi, A.-M.S. Tremblay, Phys. Rev. B **74**(3), 035113 (2006)
129. B. Davoudi, A.-M.S. Tremblay, Phys. Rev. B **76**(8), 085115 (2007)
130. J.M. Luttinger, J.C. Ward, Phys. Rev. **118**(5), 1417 (1960)
131. S. Roy, Calculs thermodynamiques et leur uniformité pour le modèle de Hubbard bidimensionnel. Master's thesis, Université de Sherbrooke (2002)
132. T. Paiva, R.T. Scalettar, C. Huscroft, A.K. McMahan, Phys. Rev. B **63**, 125116 (2001)
133. F. Werner, O. Parcollet, A. Georges, S.R. Hassan, Phys. Rev. Lett. **95**(5), 056401 (2005)
134. R.A. Cooper, Y. Wang, B. Vignolle, O.J. Lipscombe, S.M. Hayden, Y. Tanabe, T. Adachi, Y. Koike, M. Nohara, H. Takagi, C. Proust, N.E. Hussey, Science **323**(5914), 603 (2009)

135. N. Doiron-Leyraud, P. Auban-Senzier, S. René de Cotret, C. Bourbonnais, D. Jérome, K. Bechgaard, L. Taillefer, Phys. Rev. B **80**(21), 214531 (2009)
136. S. Allen, Approximation auto-cohérente à deux particules, pseudogap et supraconductivité dans le modèle de Hubbard attractif. Ph.D. thesis, Université de Sherbrooke (2000)
137. S. Verga, R.J. Gooding, F. Marsiglio, Phys. Rev. B **71**(15), 155111 (2005)
138. S. Fujimoto, J. Phys. Soc. Jpn. **71**(5), 1230 (2002)
139. Y. Yanase, J. Phys. Soc. Jpn. **73**(4), 1000 (2004)
140. S. Allen, H. Touchette, S. Moukouri, Y.M. Vilk, A.-M.S. Tremblay, Phys. Rev. Lett. **83**(20), 4128 (1999)
141. D. Rohe, W. Metzner, Phys. Rev. B **63**(22), 224509 (2001)
142. M. Randeria, N. Trivedi, A. Moreo, R.T. Scalettar, Phys. Rev. Lett. **69**(13), 2001 (1992)
143. A. Moreo, D.J. Scalapino, S.R. White, Phys. Rev. B **45**(13), 7544 (1992)
144. T. Maier, M. Jarrell, T. Pruschke, M.H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
145. G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, C.A. Marianetti, Rev. Mod. Phys. **78**(3), 865 (2006)
146. A.M. Daré, Unpublished
147. C.J. Halboth, W. Metzner, Phys. Rev. Lett. **85**(24), 5162 (2000)
148. A.A. Abrikosov, L.P. Gorkov, I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice Hall, NJ, 1963)
149. T. Saso, J. Phys. Soc. Jpn. **69**(12), 3912 (2000)
150. R. Frésard, A. Ruckenstein, Phys. B Condens. Matter **281–282**, 890 (2000)
151. C. Honerkamp, M. Salmhofer, N. Furukawa, T.M. Rice, Phys. Rev. B **63**(3), 035109 (2001)
152. C. Honerkamp, M. Salmhofer, Prog. Theor. Phys. **105**, 1 (2001)

Index

A. Avella and F. Mancini (eds.), *Strongly Correlated Systems*, Springer Series in Solid-State Sciences 171, DOI 10.1007/978-3-642-21831-6,