Self-energy-functional approach to systems of correlated electrons

M. Potthoff[a]

Lehrstuhl Festkörpertheorie, Institut für Physik, Humboldt-Universität zu Berlin, 10115 Berlin, Germany

Abstract. The grand potential of a system of interacting electrons is considered as a stationary point of a self-energy functional. It is shown that a rigorous evaluation of the functional is possible for self-energies that are representable within a certain reference system. The variational scheme allows to construct new non-perturbative and thermodynamically consistent approximations. Numerical results illustrate the practicability of the method.

PACS. 71.10.-w Theories and models of many-electron systems – 71.15.-m Methods of electronic structure calculations – 74.20.-z Theories and models of superconducting state – 75.10.-b General theory and models of magnetic ordering – 71.30.+h Metal-insulator transitions and other electronic transitions

1 Introduction

Systems of strongly correlated electrons continue to represent a central subject of current research. Different interesting correlation phenomena, such as high-temperature superconductivity [1], Mott metal-insulator transitions [2] or itinerant ferromagnetism [3], are far from being finally clarified. Progress in this field crucially depends on the development of new theoretical methods as even highly idealized model systems pose notoriously difficult problems. There are only a few general approaches which are able to access the equilibrium thermodynamics as well as excitation properties of an extended system of correlated electrons.

General methods can be based on the Green's-function formalism of Luttinger and Ward [4] and Baym and Kadanoff: [5] Here the grand potential Ω is expressed in terms of the time- or frequency-dependent one-electron Green's function $\mathbf{G}$. The functional $\Omega[\mathbf{G}]$ can be shown to be stationary at the physical $\mathbf{G}$. In principle, this is an exact variational approach which provides information not only on static equilibrium but also on dynamic excitation properties. The functional dependence $\Omega[\mathbf{G}]$, however, is generally not known explicitly as it must be constructed by summation of an infinite series of renormalized skeleton diagrams. In the standard approximation the exact but unknown functional is replaced by an explicitly known but approximate one which is based on an incomplete summation of the diagram series. This leads to the well-known

---
[a] *Present address*: Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany
e-mail: `potthoff@physik.uni-wuerzburg.de`

perturbational ("conserving") theories [5]. Higher-order approaches as the fluctuation-exchange approximation [6] are mainly applied to discrete lattice models while for continuum systems, *e.g.* for the inhomogeneous electron gas, one has to be content with lowest-order theories as the GW method [7–9].

A second type of general methods is based on density-functional (DF) approaches [10,11]. Normally these aim at the inhomogeneous electron gas but can also be applied to Hubbard-type lattice models [12]. Compared with the Green's-function formalism, there is a conceptually similar situation for DF approaches: In the latter the ground-state energy E (or the grand potential Ω) [13] is given as a functional of the (static) density $\mathbf{n}$. The variational principle associated with the functional $E[\mathbf{n}]$ is rigorous but cannot be evaluated as $E[\mathbf{n}]$ is generally unknown. In the standard local-density approximation (LDA) the (unknown part of the) functional is replaced by an explicitly known but approximate functional which is taken from the homogeneous system. For systems with weakly varying density the LDA should be justified. Information on excitation properties is contained in dynamic response functions which are in principle accessible via time-dependent DF theory [14] where the action A is considered as a functional of the time-dependent density $\mathbf{n}$. Again, the exact but unknown functional $A[\mathbf{n}]$ is approximated to make it explicit and the variational principle is exploited afterwards.

The method proposed here rests on a variational principle which uses the electron self-energy $\mathbf{\Sigma}$ as the basic dynamic variable. A new functional $\Omega[\mathbf{\Sigma}]$ is constructed which can be shown to be stationary at the physical self-energy. The main result is that the variational principle can be exploited without any approximation of the

functional dependence. Namely, a *rigorous* evaluation of the functional $\Omega[\mathbf{\Sigma}]$ is possible on a certain subspace of trial self-energies. Trial self-energies must be representable within an exactly solvable reference system sharing the same interaction with the original system.

This result has important consequences as it opens a route for constructing a novel class of approximations. Although the self-energy essentially contains the same information as the Green's function or the (time-dependent) density, the new approach is conceptually contrary to the Green's-function approach and to the DF approach as there is no approximation to be tolerated for the central functional. Instead of approximating the functional itself, it is considered on a restricted domain. The self-energy-functional approach is completely general and yields approximations which are non-perturbative, thermodynamically consistent and systematic. Opposed to numerical techniques directly applied to systems of finite size, the self-energy-functional approach provides a variational or self-consistent embedding of finite systems and thus yields results in the thermodynamical limit. Such techniques are needed to construct phase diagrams from standard correlated lattice models. A potentially fruitful field of application are systems with competing types of order resulting from spin, charge or orbital correlations as it is typical *e.g.* for numerous transition-metal oxides [1–3].

In the present paper the approach is introduced and a number of general aspects are discussed in detail (Sect. 2). To demonstrate its usefulness, two applications will be considered for the single-band Hubbard model: In Section 3 it is shown that the dynamical mean-field theory (DMFT) [15] can be recovered within the self-energy-functional approach, namely by choosing a decoupled set of impurity Anderson models as a reference system. The DMFT generally requires the treatment of a quantum-impurity problem with an infinite number of degrees of freedom ($n_s = \infty$). In Section 4 a new approximation is discussed which is based on an impurity model with a *finite* number of degrees of freedom only and which approaches the DMFT for $n_s = \infty$. The method is closely related to the exact-diagonalization approach (ED) [16,17]. Opposed to the ED, however, thermodynamical consistency is guaranteed at any stage of the approximation. New approaches beyond the mean-field level will be discussed elsewhere. The conclusions and an outlook are given in Section 5.

2 Self-energy-functional approach

Consider a general Hamiltonian $H = H_0(\mathbf{t}) + H_1(\mathbf{U})$ with one-particle ("hopping") parameters $\mathbf{t}$ and two-particle interaction parameters $\mathbf{U}$:

$$H = \sum_{\alpha\beta} t_{\alpha\beta} c_\alpha^\dagger c_\beta + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} U_{\alpha\beta\delta\gamma} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta \,. \qquad (1)$$

Here $\alpha, \beta, ...$ refer to an orthonormal and complete set of one-particle basis states. We are interested in the equilibrium thermodynamics and in elementary one-particle

excitations of the system for temperature T and chemical potential μ. This is described by the one-particle Green's function $G_{\alpha\beta}(\mathrm{i}\omega) = \langle\langle c_\alpha ; c_\beta^\dagger \rangle\rangle$ of the imaginary fermionic Matsubara frequencies $\mathrm{i}\omega = \mathrm{i}(2n+1)\pi T$ with integer n [18]. The Green's function can be calculated from the self-energy $\Sigma_{\alpha\beta}(\mathrm{i}\omega)$ *via* the Dyson equation. Using a matrix notation, this reads as $\mathbf{G} = \mathbf{G}_0 + \mathbf{G}_0 \mathbf{\Sigma} \mathbf{G}$ where $\mathbf{G}_0 = 1/(\mathrm{i}\omega + \mu - \mathbf{t})$ is the "free" Green's function. The self-energy is given by $\mathbf{\Sigma} = \mathbf{\Sigma}[\mathbf{G}] = T^{-1}\delta\Phi[\mathbf{G}]/\delta\mathbf{G}$, where $\Phi[\mathbf{G}]$ is the so-called Luttinger-Ward functional [4,5]. This allows to derive the Green's function from a variational principle: One has $\delta\Omega[\mathbf{G}]/\delta\mathbf{G} = 0$ where $\Omega[\mathbf{G}] = \Phi[\mathbf{G}] + \mathrm{Tr}\ln(-\mathbf{G}) - \mathrm{Tr}((\mathbf{G}_0^{-1} - \mathbf{G}^{-1})\mathbf{G})$ and using the notation $\mathrm{Tr}\,\mathbf{A} = T\sum_{\omega,\alpha} A_{\alpha\alpha}(\mathrm{i}\omega)$. In general, however, the functional $\Phi[\mathbf{G}]$ is not known explicitly which prevents an evaluation of $\Omega[\mathbf{G}]$ for a given $\mathbf{G}$. So-called conserving approximations [5] provide an explicit but approximate functional $\Phi_{\mathrm{pert.}}[\mathbf{G}] \approx \Phi[\mathbf{G}]$. However, these are weak-coupling approaches where a certain subclass of Φ diagrams is summed up.

Here a different but still rigorous variational principle is proposed which is based on a functional $\mathbf{G} = \mathbf{G}[\mathbf{\Sigma}]$ defined as the inverse of $\mathbf{\Sigma} = \mathbf{\Sigma}[\mathbf{G}]$. We can assume the latter to be invertible (locally) provided that the system is not at a critical point for a phase transition (see Appendix A). Consider then:

$$\Omega_\mathbf{t}[\mathbf{\Sigma}] \equiv \mathrm{Tr}\ln(-(\mathbf{G}_0^{-1} - \mathbf{\Sigma})^{-1}) + F[\mathbf{\Sigma}] \qquad (2)$$

where $F[\mathbf{\Sigma}] \equiv \Phi[\mathbf{G}[\mathbf{\Sigma}]] - \mathrm{Tr}(\mathbf{\Sigma}\,\mathbf{G}[\mathbf{\Sigma}])$ is the Legendre transform of $\Phi[\mathbf{G}]$. The subscript $\mathbf{t}$ indicates the explicit $\mathbf{t}$ dependence of Ω due to the free Green's function $\mathbf{G}_0$. Using $T^{-1}\delta F[\mathbf{\Sigma}]/\delta\mathbf{\Sigma} = \mathbf{G}[\mathbf{\Sigma}]$, one finds that

$$\delta\Omega_\mathbf{t}[\mathbf{\Sigma}]/\delta\mathbf{\Sigma} = 0 \;\Leftrightarrow\; \mathbf{G}[\mathbf{\Sigma}] = (\mathbf{G}_0^{-1} - \mathbf{\Sigma})^{-1}\,. \qquad (3)$$

Thus $\Omega_\mathbf{t}[\mathbf{\Sigma}]$ is stationary at the exact (physical) self-energy and its value is the exact grand potential of the system. Again, the problem is that the functional $\Omega_\mathbf{t}[\mathbf{\Sigma}]$ is in general not known explicitly.

As the domain of the self-energy functional $\Omega_\mathbf{t}[\mathbf{\Sigma}]$ we define the class of all $\mathbf{t}'$ representable self-energies. $\mathbf{\Sigma}$ is termed $\mathbf{t}'$ representable, if there is a set of hopping parameters $\mathbf{t}'$ such that $\mathbf{\Sigma}$ is the exact self-energy of the model $H_0(\mathbf{t}') + H_1(\mathbf{U})$. This implies that any self-energy in the domain of $\Omega_\mathbf{t}[\mathbf{\Sigma}]$ can be parameterized as $\mathbf{\Sigma} = \mathbf{\Sigma}(\mathbf{t}')$. The interaction parameters $\mathbf{U}$ are taken to be fixed. Suppose we are interested in the model $H = H_0(\mathbf{t}) + H_1(\mathbf{U})$. Then the function $\Omega_\mathbf{t}(\mathbf{t}') \equiv \Omega_\mathbf{t}[\mathbf{\Sigma}(\mathbf{t}')]$ is stationary at $\mathbf{t}' = \mathbf{t}$. Thus $\partial\Omega_\mathbf{t}(\mathbf{t}')/\partial\mathbf{t}' = 0$ at $\mathbf{t}' = \mathbf{t}$.

It is important to note that $F[\mathbf{\Sigma}]$ is universal: The functional dependence is the same for any $\mathbf{t}$, *i.e.* it remains unchanged for an arbitrary reference system H' with the same interaction but modified hopping parameters: $H' = H_0(\mathbf{t}') + H_1(\mathbf{U})$. $F[\mathbf{\Sigma}]$ is universal as it is the Legendre transform of $\Phi[\mathbf{G}]$ which in turn is universal because it can be constructed formally as the sum of all closed, irreducible, and renormalized skeleton diagrams which, apart from $\mathbf{G}$, include the vertices $\mathbf{U}$ only.

Consequently, one has:

$$\Omega_{\mathbf{t}'}[\boldsymbol{\Sigma}] = \text{Tr} \ln(-(\mathbf{G}_0'^{-1} - \boldsymbol{\Sigma})^{-1}) + F[\boldsymbol{\Sigma}], \qquad (4)$$

for the reference system H' with $\mathbf{G}_0'^{-1} = \mathrm{i}\omega + \mu - \mathbf{t}'$.

The self-energy functionals $\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}]$, equation (2), and $\Omega_{\mathbf{t}'}[\boldsymbol{\Sigma}]$, equation (4), are different if $\mathbf{t} \neq \mathbf{t}'$, *i.e.* if the original and the reference system are different. $\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}]$ is stationary at $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathbf{t})$ while $\Omega_{\mathbf{t}'}[\boldsymbol{\Sigma}]$ is stationary at $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathbf{t}')$. The important point is that both functionals share the same (unknown but universal) part $F[\boldsymbol{\Sigma}]$ and that both functionals are defined globally. Therefore, combining equations (2) and (4), $F[\boldsymbol{\Sigma}]$ can be eliminated:

$$\begin{aligned}\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}] = \Omega_{\mathbf{t}'}[\boldsymbol{\Sigma}] &+ \text{Tr} \ln(-(\mathbf{G}_0^{-1} - \boldsymbol{\Sigma})^{-1}) \\ &- \text{Tr} \ln(-(\mathbf{G}_0'^{-1} - \boldsymbol{\Sigma})^{-1}).\end{aligned} \qquad (5)$$

This form of the self-energy functional $\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}]$ is still exact.

In the next step the functional has to be evaluated to search for a stationary point. Evaluating $\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}]$ for self-energies taken from the reference system, *i.e.* for self-energies parameterized as $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathbf{t}')$, one obtains:

$$\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}(\mathbf{t}')] = \Omega' + \text{Tr} \ln(-(\mathbf{G}_0^{-1} - \boldsymbol{\Sigma}(\mathbf{t}'))^{-1}) - \text{Tr} \ln(-\mathbf{G}').$$
$$(6)$$

Here it has been used that $\Omega_{\mathbf{t}'}[\boldsymbol{\Sigma}(\mathbf{t}')] = \Omega'$, the exact grand potential of the reference system H', and $(\mathbf{G}_0'^{-1} - \boldsymbol{\Sigma}(\mathbf{t}'))^{-1} = \mathbf{G}'$, the exact Green's function of H'. Suppose that the reference system H' is much simpler than the original system H so that it can be solved exactly for any $\mathbf{t}'$ belonging to a certain subspace of the entire space of hopping parameters. The resulting equation (6) is remarkable, as it shows that the functional $\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}]$ can be evaluated rigorously for trial self-energies $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathbf{t}')$ taken from the reference system H'.

This is the main result. Contrary to previous approaches (*e.g.* conserving theories, LDA), there is no need to approximate the functional dependence in a fundamental variational principle. Approximations are constructed by searching for a stationary point of $\Omega_{\mathbf{t}}[\boldsymbol{\Sigma}]$ on a *restricted* set of trial self-energies $\boldsymbol{\Sigma}(\mathbf{t}')$.

The stationary point is determined by the Euler equation: $\partial \Omega_{\mathbf{t}}[\boldsymbol{\Sigma}(\mathbf{t}')]/\partial \mathbf{t}' = 0$. Calculating the derivative,

$$T \sum_{\omega} \sum_{\alpha\beta} \left(\frac{1}{\mathbf{G}_0^{-1} - \boldsymbol{\Sigma}(\mathbf{t}')} - \mathbf{G}' \right)_{\beta\alpha} \frac{\partial \Sigma_{\alpha\beta}(\mathbf{t}')}{\partial \mathbf{t}'} = 0. \quad (7)$$

Note that the equation involves, apart from $\mathbf{G}_0$, quantities of the reference system H' only. The linear response of the self-energy of H' due to a change of the hopping $\mathbf{t}'$ can be calculated along the lines of reference [5]. It turns out that $\partial \boldsymbol{\Sigma}(\mathbf{t}')/\partial \mathbf{t}'$ is given by a two-particle Green's function of H'. Since $\mathbf{G}' = \mathbf{G}[\boldsymbol{\Sigma}(\mathbf{t}')]$, the exact self-energy of the system H is determined by the condition that the bracket in (7) be zero. Hence, one can consider equation (7) to be obtained from the *exact* equation that determines the "vector" $\boldsymbol{\Sigma}$ in the self-energy space through *projection* onto the hypersurface of $\mathbf{t}'$ representable trial self-energies $\boldsymbol{\Sigma}(\mathbf{t}')$ by taking the scalar product with vectors $\partial \boldsymbol{\Sigma}(\mathbf{t}')/\partial \mathbf{t}'$ tangential to the hypersurface.

An analysis of the second derivative $\partial^2 \Omega_{\mathbf{t}}[\boldsymbol{\Sigma}(\mathbf{t}')]/\partial t'_{\alpha\beta} \partial t'_{\gamma\delta}$ shows that a stationary point is not an extremum point in general. This feature is shared with the time-dependent DF approach [14], the Green's-function approach [5] and also with a recently considered variant [19]. Only in the static DF theory there is a convex (density) functional [10,11,13]. Nevertheless, the proposed self-energy-functional approach is systematic: For any sequence of reference systems H' including more and more degrees of freedom and converging to the original system H there is, from the variational principle, a corresponding sequence of grand potentials which must converge to the exact $\Omega = \Omega_{\mathbf{t}}[\boldsymbol{\Sigma}(\mathbf{t})]$ as the subspace of trial self-energies increases and eventually includes the exact self-energy $\boldsymbol{\Sigma}(\mathbf{t})$.

It is also possible to build the theory on the Green's function $\mathbf{G}$ instead of the self-energy $\boldsymbol{\Sigma}$ as the basic variable. One may start *e.g.* from the Green's-function functional $\Omega_{\mathbf{t}}[\mathbf{G}]$ of reference [19]. This can be split into a universal part and a part depending on the hopping explicitly. Introducing the concept of the reference system and using arguments similar to those that have lead to equations (4, 5), and (6), one can show that an exact evaluation of $\Omega_{\mathbf{t}}[\mathbf{G}]$ is possible on a certain subspace of $\mathbf{t}'$ representable Green's functions. Here it is preferred to use the self-energy as the basic variable, however, as it is advantageous to approximate an irreducible rather than a reducible quantity.

3 Relation to the DMFT

Given an original model H, what could a suitable reference system H' look like? Consider, for example, H to be the Hubbard model [20] which is shown in Figure 1a schematically: A filled dot represents a correlated site i with on-site Hubbard interaction U, and a line connecting two sites i and j represents the nearest-neighbor hopping $t_{i,j}$. The number of sites is $L \mapsto \infty$. Figure 1c shows a conceivable reference system H'. H' is obtained from H (Fig. 1a) by (i) adding to each correlated site i a number of $n_s - 1$ uncorrelated ("bath") sites $k = 2, ..., n_s$ (open dots) which are disconnected from the rest of the system, by (ii) switching off the hopping $t_{i,j}$ between the correlated sites and (iii) switching on a hopping $V_{i,k}$ to the bath sites. After step (i) the Hamiltonian Figure 1b (in the figure $n_s = 5$) has an enlarged Hilbert space but the same self-energy. It is important to note that steps (i) - (iii) leave the interaction part unchanged and thus preserve the functional dependence $F[\boldsymbol{\Sigma}]$. Actually, the system H' is a set of L decoupled single-impurity Anderson models (SIAM) [21] with n_s sites each. Compared to H, the problem posed by H' is strongly simplified. This is achieved at the cost of restricting the set of trial self-energies. In particular, as the correlated sites are decoupled in H', the trial self-energies are local: $\Sigma_{ij}(\mathrm{i}\omega, \mathbf{t}') \propto \delta_{ij}$. One has to consider H' for arbitrary one-particle parameters, namely the on-site energies of the correlated ("c") and of the bath sites ("a"), $\epsilon_i^{(c)}$ and $\epsilon_{i,k}^{(a)}$, respectively, and the hopping ("hybridization") $V_{i,k}$
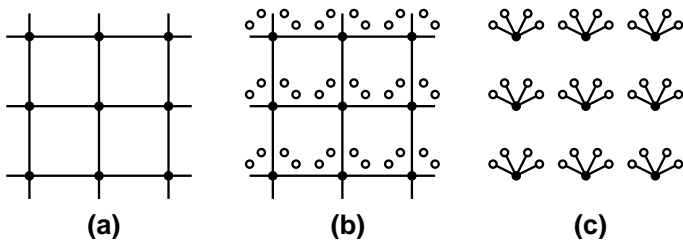
Fig. 1. Schematic representation of the Hubbard model H (a), an equivalent model (b), and a possible reference system H' (c). See text for discussion.

between them and take these as variational parameters in the principle $\delta\Omega_{\mathbf{t}}[\mathbf{\Sigma}(\mathbf{t}')] = 0$.

Let us discuss the case $n_{\mathrm{s}} \mapsto \infty$. For a homogeneous phase of the (translationally invariant) original system, $\Omega_{\mathbf{t}}[\mathbf{\Sigma}(\mathbf{t}')]$ will be stationary at a homogeneous set of variational parameters: $\mathbf{t}' = \{\epsilon_i^{(c)}, \epsilon_{i,k}^{(a)}, V_{i,k}\} = \{\epsilon^{(c)}, \epsilon_k^{(a)}, V_k\}$. Consequently, it is sufficient to consider one SIAM only. As the different equivalent SIAM's are spatially decoupled, not only the self-energy but also its linear response is local: $\partial\Sigma_{ij}(\mathbf{t}')/\partial\mathbf{t}' \propto \delta_{ij}$. To solve the Euler equation (7), it is thus sufficient to fulfill the "locally projected" equation

$$\left(\frac{1}{\mathbf{G}_0^{-1}(\mathrm{i}\omega) - \mathbf{\Sigma}(\mathrm{i}\omega)}\right)_{ii} = G'_{ii}(\mathrm{i}\omega). \qquad (8)$$

This is just the self-consistency equation of the DMFT [15]: the SIAM parameters have to be found such that the on-site ("impurity") Green's function at a correlated site i coincides with the on-site Green's function of the Hubbard model which is calculated from $\mathbf{G}_0$ and the ("impurity") self-energy of H' by means of the Dyson equation. Therefore, one can state that the DMFT (as an approximation for any finite-dimensional system or as the exact theory in infinite dimensions) is recovered as a stationary point of $\Omega_{\mathbf{t}}[\mathbf{\Sigma}]$ when restricting the search to local self-energies representable by a SIAM.

Within the DMFT the computation of the self-energy requires an iterative procedure: $\mathbf{\Sigma}_{\mathrm{old}} \mapsto \mathbf{\Sigma}_{\mathrm{new}}$. Here it turns out that this corresponds to a certain (discrete) path on the hypersurface of SIAM trial self-energies. Convergence of the iteration ($\mathbf{\Sigma}_{\mathrm{old}} = \mathbf{\Sigma}_{\mathrm{new}}$), however, is by no means guaranteed physically but depends on the contracting properties of the map $\mathbf{\Sigma}_{\mathrm{old}} \mapsto \mathbf{\Sigma}_{\mathrm{new}}$. The self-energy-functional approach offers an alternative as instead of solving equation (8) one may calculate $\Omega_{\mathbf{t}}[\mathbf{\Sigma}(\mathbf{t}')]$ by equation (6) and determine the stationary point. Hence, the DMFT can also be obtained by a more *direct* computation avoiding any iterations – similar (in this respect) to the random-dispersion approximation [22]. Note that in case of more than a single stationary point there is also an equally direct access to metastable phases.

For any inhomogeneous situation, equation (8) represents a system of self-consistency equations to fix the parameters of non-equivalent impurity models labeled by the site index i. The models can be solved independently but are coupled indirectly due to the matrix inversion in (8).

This exactly recovers the DMFT generalized to systems with reduced translational symmetry [23,24].

4 A consistent ED method

A brief discussion of two limiting cases of the Hubbard model may be instructive. Consider the band limit with $U = 0$ first. Here $H = \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma}$ describes a system of non-interacting electrons. This case is exceptional as obviously the functional $F[\mathbf{\Sigma}] \equiv 0$ and therefore $\Omega_{\mathbf{t}}[\mathbf{\Sigma}] = \mathrm{Tr}\ln(-(\mathbf{G}_0^{-1} - \mathbf{\Sigma})^{-1})$. Any valid reference system H' must have the same (*i.e.* a vanishing) interaction part as H, and thus $\mathbf{\Sigma}(\mathbf{t}') \equiv 0$ and $\Omega_{\mathbf{t}}[\mathbf{\Sigma}(\mathbf{t}')] = \mathrm{Tr}\ln(-\mathbf{G}_0^{-1})$, the exact grand potential for non-interacting electrons.

The atomic limit, $H = \sum_{i\sigma}(t_0 c_{i\sigma}^\dagger c_{i\sigma} + (U/2)n_{i\sigma}n_{i-\sigma})$ is more interesting as $\Phi[\mathbf{G}]$ and $F[\mathbf{\Sigma}]$ cannot be constructed explicitly. Within the self-energy-functional approach one has to compute $\mathbf{\Sigma}(\mathbf{t}')$, $\mathbf{G}'$, and Ω' for a suitable reference system H' and to insert into equation (6) for optimization. The only meaningful choice for the reference system is $H' = H$ in this case. Obviously, this yields the exact solution. Generally, whenever the original model H is exactly solvable, the choice $H' = H$ will do.

For a non-trivial model H, the choice $H' = H$ is useless for a practical computation. Any simplified reference system, however, yields a consistent approximation. The case of the DMFT discussed in Section 3 is an illustrative example. On the other hand, in the context of the DMFT actually both, H and H', are highly non-trivial models, and further approximations or large-scale numerics are needed to treat the reference system H'. More simple approximations for the Hubbard model which are still consistent are generated by considering reference systems with a *finite* number of degrees of freedom. The reference system of Figure 1c with $n_{\mathrm{s}} < \infty$ is an interesting example which shall be discussed in the following. For small n_{s} one can easily obtain numerical results as a complete diagonalization of H' is feasible.

Consider the Hubbard model

$$H = \sum_{ij\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + \frac{U}{2} \sum_{i\sigma} n_{i\sigma} n_{i-\sigma} \qquad (9)$$

at temperature $T = 0$ and chemical potential $\mu = U/2$. For symmetric conditions this implies half-filling. The Hamiltonian of the reference system is given by $H' = \sum_i H'(i)$ with

$$H'(i) = \sum_\sigma \epsilon_i^{(c)} c_{i\sigma}^\dagger c_{i\sigma} + \frac{U}{2} \sum_\sigma n_{i\sigma} n_{i-\sigma}$$
$$+ \sum_{\sigma,k=2}^{n_{\mathrm{s}}} \epsilon_{i,k}^{(a)} a_{ik\sigma}^\dagger a_{ik\sigma} + \sum_{\sigma,k} \left(V_{i,k} c_{i\sigma}^\dagger a_{ik\sigma} + \mathrm{h.c.}\right).$$
$$(10)$$

For the sake of simplicity we consider a homogeneous paramagnetic phase and the most simple case $n_{\mathrm{s}} = 2$ where one is left with three independent variational parameters only,
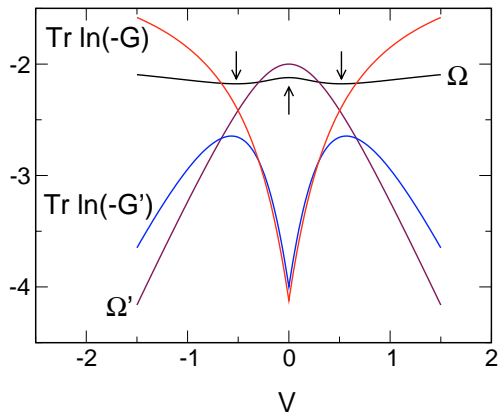
Fig. 2. Grand potential Ω (per lattice site) and the different contributions (per lattice site) according to equation (6): Ω', $\mathrm{Tr}\ln(-(\mathbf{G}_0^{-1} - \mathbf{\Sigma}(\mathbf{t}'))^{-1})$, and $\mathrm{Tr}\ln(-\mathbf{G}')$ for $U = 4$, $T = 0$, and $\mu = U/2$ (half-filling) as functions of V. Stationary points (arrows) at $V = \pm 0.519$ (metal) and at $V = 0$ (insulator). $\epsilon_c = 0$ and $\epsilon_a = 2$.

Fig. 3. Ω as a function of ϵ_c for $V = 0.519$ (metal) and $\epsilon_a = 2$. $U = 4$. Inset: impurity- and bath-orbital filling, n_c and n_a, as functions of ϵ_c.

Fig. 4. Ω and the impurity- and bath-orbital filling, n_c and n_a, as functions of ϵ_a for $V = 0.519$ and $\epsilon_c = 0$. $U = 4$.

namely $V \equiv V_{i,k=2}$, $\epsilon_a \equiv \epsilon_{i,k=2}^{(a)}$, and $\epsilon_c \equiv \epsilon_i^{(c)}$. The computation of the different contributions to the grand potential, equation (6), is straightforward: Diagonalization of H' yields the ground-state energy E_0' and $\Omega' = E_0' - \mu \langle N' \rangle$ as well as the excitation energies, the ground state and the excited states. The Green's function $\mathbf{G}'$ and the free Green's function $\mathbf{G}_0'$ can be computed from their respective Lehmann representations. The self-energy of the reference system is obtained as $\mathbf{\Sigma}(\mathbf{t}') = \mathbf{G}_0'^{-1} - \mathbf{G}'^{-1}$. Since the self-energy is local, as in the DMFT, the lattice structure enters via the free ($U = 0$) density of states only. Therefore, the $\mathbf{k}$-sum which appears in the first trace in equation (6) can be performed conveniently by a one-dimensional density-of-states integration. A semi-elliptical free density of states with the band width $W = 4$ is used for the calculations. This sets the energy scale for the results discussed below.

Figure 2 shows the grand potential Ω and the three different contributions as functions of V. The interaction is kept fixed at $U = W = 4$. The remaining variational parameters are set to $\epsilon_c = 0$ and $\epsilon_a = U/2 = 2$, as required by particle-hole symmetry. Each of the three contributions strongly depends on V and none of them has a stationary point at a finite $V \neq 0$. Two of them show a singular behavior at $V = 0$. Contrary, the resulting Ω is regular for any V and shows a much weaker V dependence. There are three stationary points which are indicated by the arrows. The maximum at $V = 0$ corresponds to an insulating phase since $\Sigma(\omega)$ for $n_s = 2$ and $V = 0$ is the Hubbard-I self-energy which implies a vanishing spectral density $-(1/\pi)\mathrm{Im}\mathbf{G}(\omega + i0^+)$ at $\omega = 0$. The minima at $V = \pm 0.519$ correspond to a metallic phase. Ω as well as the different contributions are symmetric functions of V. As $\Sigma(V, \omega) = \Sigma(-V, \omega)$, however, this symmetry is trivial and does not yield an additional physical phase (see also Appendix A). Due to the lower Ω the metallic phase is stable as compared to the insulating one.
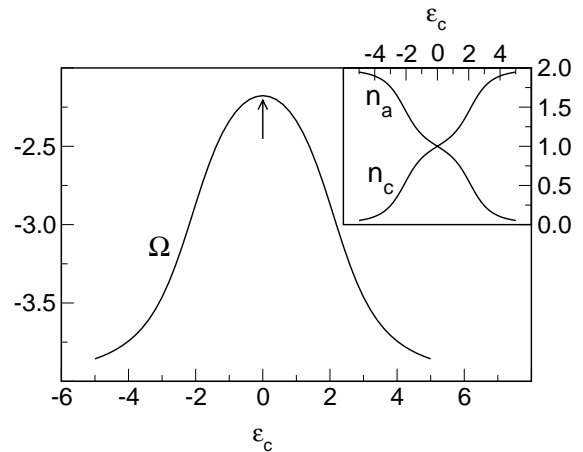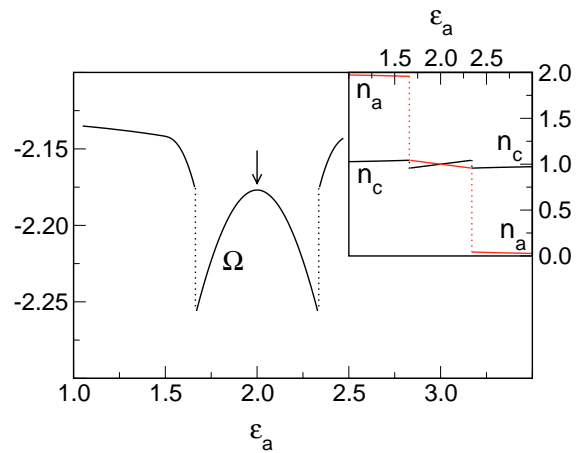
The minimum at $V = 0.519$ is actually a saddle point if the entire space of variational parameters is considered. This is demonstrated by Figure 3 which shows Ω as a function of ϵ_c for fixed $V = 0.519$ and $\epsilon_a = 2$. While $\Omega(V)$ is at a minimum for $V = 0.519$, $\Omega(\epsilon_c)$ is at a *maximum* for $\epsilon_c = 0$. In the (V, ϵ_c) space one therefore encounters a saddle point. As already noted in Section 2, there is no reason to expect an extremum in general. It is worth mentioning that stationarity at $\epsilon_c = 0$ is consistent with the requirements of particle-hole symmetry. For any $\epsilon_c \neq 0$ the impurity model is asymmetric. This can be seen from the inset where the average occupations of the impurity and of the bath site are plotted as functions of ϵ_c. The total particle number $\langle N' \rangle = \sum_\sigma (\langle c_\sigma^\dagger c_\sigma \rangle + \langle a_\sigma^\dagger a_\sigma \rangle) = n_c + n_a$ (i and $k = 2$ fixed) is constant: $\langle N' \rangle = 2$.

With respect to the third variational parameter ϵ_a, the grand potential Ω is at a maximum for $\epsilon_a = 2 = \mu$, see Figure 4. Again, this value is required by particle-hole symmetry. If ϵ_a exceeds a certain critical value (away from the stationary point), the ground-state of the reference system H' no longer lies within $N' = 2$ subspace

but is found in the $N' = 1$ or $N' = 3$ subspace, respectively (see inset of Fig. 4). While Ω' is continuous at the level crossing, the symmetry of the ground state changes. Consequently, there is a discontinuous change of the trial self-energy which implies a discontinuous change of Ω.

Consider now the *original* model at slightly modified parameters, *e.g.* U, μ, or T. Clearly, the stationary point of Ω will be expected then at slightly different values of the variational parameters V, ϵ_c, ϵ_a. This implies that all physical quantities which derive from the thermodynamical potential Ω will be continuous functions of the (original) model parameters in general – irrespective of the fact that the reference system includes a finite number of degrees of freedom only: It is a typical feature of any mean-field approach that results are directly provided for the thermodynamical limit. A discontinuous jump of Ω due to a symmetry change of the ground state of the reference system (see Fig. 4) usually occurs away from stationarity and is thus irrelevant physically. It is conceivable, however, that the stationary point moves to a point of discontinuity as a function a parameter of the original model. In this case the approach would generate an artifact which is a reminiscence of the finiteness of H'.

So far we discussed the case $U = 4 = W$ only. As a function of U the half-filled paramagnetic Hubbard model at $T = 0$ is expected to undergo a transition from a metal to a Mott-Hubbard insulator [2]. This is marked by a divergence of the effective mass or, equivalently, by a vanishing quasi-particle weight $z = 1/(1 - \mathrm{d}\Sigma(\omega)/\mathrm{d}\omega|_{\omega=0})$ as U approaches a critical value U_c from below [15,25]. The result for $z(U)$ as obtained by the use of the two-site reference system ("dynamical impurity approximation", DIA with $n_s = 2$) is shown in Figure 5. As there are less degrees of freedom contained in H', the approximation should be considered to be *inferior* as compared to the full DMFT the results of which are in Figure 5, too. It is remarkable that the simple $n_s = 2$-DIA (which requires the diagonalization of a dimer model only) yields $z(U)$ in an almost quantitative agreement with the full DMFT.

The results of the $n_s = 2$-DIA may also be compared with those of the recently developed "linearized" or "two-site" DMFT [26,28] where the Hubbard model is mapped onto the two-site SIAM (10) by means a strongly simplified self-consistency condition. As compared to the two-site DMFT, the present self-energy-functional approach not only represents a clear conceptual improvement but also improves the actual results for $z(U)$ and for U_c (see Fig. 5 and note that $U_c = 6$ within the linearized DMFT, $U_c = 5.8450$ within the DIA for $n_s = 2$, and $U_c = 5.84$ and $U_c = 5.88$ from numerical evaluations [27,29] of the full DMFT).

The self-energy-functional approach with reference system of Figure 1c and small n_s actually represents a new variant of the DMFT-exact-diagonalization method (ED) [16,17]. As compared to previous formulations of the ED, the convergence with respect to n_s appears to be faster: Compare the results for $n_s = 2$, $n_s = 4$ and $n_s = \infty$ (full DMFT) in Figure 5 with those of reference [16]. More important, however, there is no need for a fit procedure in
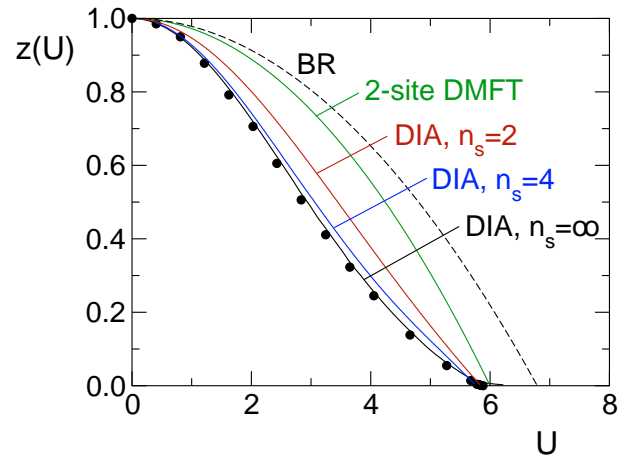
Fig. 5. U dependence of the quasi-particle weight $z = 1/(1 - \mathrm{d}\Sigma(\omega = 0)/\mathrm{d}\omega)$ within different approximations. BR: Brinkman-Rice (Gutzwiller) approach [25]. 2-site DMFT: a non-variational two-site approach [26]. DIA, $n_s = 2$: self-energy-functional approach with a reference system H' consisting of decoupled two-site impurity models. DIA, $n_s = 4$: dynamical-impurity approximation with $n_s = 4$ sites. DIA, $n_s = \infty$: DMFT limit (circles: NRG [27], line: ED using 8 sites).

the present approach; any arbitrariness in the method to find the SIAM parameters is avoided completely. Furthermore, consistent results will be obtained for any finite n_s while in the usual ED this can be expected in the DMFT limit $n_s \mapsto \infty$ only.

5 Conclusions and outlook

Concluding, the proposed self-energy-functional method is a systematic scheme for the construction of new non-perturbative and consistent approximations for extended systems of interacting fermions. For Hubbard-type lattice models with on-site interaction several relations to and generalizations of existing approaches are obtained immediately. The numerical results obtained by considering a rather simple reference system clearly demonstrate the practicability of the theory. Its generality promises that the approach may successfully be applied also in different contexts.

For a Hubbard-type system including $M > 1$ orbitals per site, a consistent DMFT can only be defined when using M baths. There is no such necessity within the self-energy-functional approach. While clearly the optimal local approximation requires M baths, any $M' < M$ will nevertheless lead to a fully consistent approximation. This represents an interesting option for numerical studies of multi-band systems. Non-local trial self-energies can be constructed by grouping the sites into identical clusters of finite size N_s, switching off the inter-cluster hopping and treating the intra-cluster hopping as variational parameters. Each of the N_s sites in a cluster can be coupled to $n_s - 1$ additional bath sites. The relation of such an approach to cluster extensions of the DMFT [30,31] and to

the cluster perturbation theory [32–34] will be the subject of a forthcoming paper. Extensions and applications of the method to continuous models (inhomogeneous electron gas) and to Bose systems deserve further investigations.

Appendix A: The functional $\mathbf{G}[\boldsymbol{\Sigma}]$

For the definition of $\mathbf{G}[\boldsymbol{\Sigma}]$, invertibility of the functional $\boldsymbol{\Sigma}[\mathbf{G}]$ is required. The *local* invertibility of $\boldsymbol{\Sigma}[\mathbf{G}]$ is controlled by the Jacobian $\Gamma_{\alpha\beta'\alpha'\beta}(i\omega, i\omega') = \delta\Sigma_{\alpha\beta}(i\omega)/\delta G_{\alpha'\beta'}(i\omega')$. The two-particle self-energy [5] $\boldsymbol{\Gamma} = \delta\boldsymbol{\Sigma}/\delta\mathbf{G}$ can be assumed to be non-singular in general.

For a further analysis we need the following *lemma:* Consider the interaction $\mathbf{U}$ to be fixed. Then two different Green's functions $\mathbf{G}_1$ and $\mathbf{G}_2$ must result from two different sets of one-particle parameters $\mathbf{t}'_1$ and $\mathbf{t}'_2$, respectively. The *proof* is straightforward: Consider the high-frequency expansion of the Green's function $G_{\alpha\beta}(\omega) = \sum_{n=1}^{\infty} M_{\alpha\beta}^{(n)}\omega^{-n}$. The coefficients are given by the moments $M_{\alpha\beta}^{(n)} = \int d\omega\ \omega^n(-1/\pi)\operatorname{Im} G_{\alpha\beta}(\omega + i0^+) = \langle[\mathcal{L}^n c_\alpha, c_\beta]_+\rangle$ with $\mathcal{L}\mathcal{O} \equiv [\mathcal{O}, H]_-$. Using the symmetry $U_{\alpha\beta\gamma\delta} = U_{\beta\alpha\delta\gamma}$ one has:

$$
\begin{aligned}
G_{\alpha\beta}(\omega) = {} & \delta_{\alpha\beta}\frac{1}{\omega} \\
& + \left(t'_{\alpha\beta} + \sum_{\gamma\delta}(U_{\alpha\gamma\beta\delta} - U_{\alpha\gamma\delta\beta})\langle c_\gamma^\dagger c_\delta\rangle\right)\frac{1}{\omega^2} \\
& + \mathcal{O}(\omega^{-3}).
\end{aligned} \tag{A.1}
$$

Now $\mathbf{t}'_1 \neq \mathbf{t}'_2$ implies the ω^{-2} coefficients to be different because

$$
\langle c_\alpha^\dagger c_\beta\rangle = -\frac{1}{\pi}\int_{-\infty}^{\infty} d\omega \frac{1}{e^{\omega/T} + 1}\operatorname{Im} G_{\beta\alpha}(\omega + i0^+). \tag{A.2}
$$

Consequently, we must have $\mathbf{G}_1 \neq \mathbf{G}_2$.

The lemma shows that the relation $\mathbf{t}' \leftrightarrow \mathbf{G}$ is one-to-one. Consequently, we can write $\boldsymbol{\Sigma}[\mathbf{G}] = \boldsymbol{\Sigma}[\mathbf{t}'[\mathbf{G}]]$ and $\boldsymbol{\Gamma} = \delta\boldsymbol{\Sigma}/\delta\mathbf{G} = \delta\boldsymbol{\Sigma}/\delta\mathbf{t}' \cdot \delta\mathbf{t}'/\delta\mathbf{G}$ with a non-singular Jacobian $\delta\mathbf{t}'/\delta\mathbf{G}$. Hence, a singular $\boldsymbol{\Gamma} = \delta\boldsymbol{\Sigma}/\delta\mathbf{G}$ implies a singular $\delta\boldsymbol{\Sigma}/\delta\mathbf{t}'$ and *vice versa*. However, $\delta\boldsymbol{\Sigma}/\delta\mathbf{t}'$ is just the "projector" in the Euler equation (7). We conclude that local non-invertibility of the functional $\boldsymbol{\Sigma}[\mathbf{G}]$ at $\mathbf{G} = \mathbf{G}(\mathbf{t}')$ is indicated by $\partial\boldsymbol{\Sigma}[\mathbf{t}']/\partial t'_{\mathbf{n}} = 0$ with $t'_{\mathbf{n}} = \mathbf{t}' \cdot \mathbf{n}$ for a certain "direction" $\mathbf{n}$ in the space of hopping parameters. For such a direction, the Euler equation (7) would be satisfied *trivially*.

Referring to the present numerical results, one can state that generally the projector $\delta\boldsymbol{\Sigma}/\delta\mathbf{t}'$ is non-singular in fact, as has been expected. There is one exception, however, namely points in the hopping-parameter space where one or more bath sites are *decoupled* from the rest of the system (vanishing hybridization V). Here the one-particle

energy of a decoupled bath site can be varied without changing the trial self-energy. Even for this case, however, there are no formal difficulties with the inverse functional $\mathbf{G}[\boldsymbol{\Sigma}]$: To ensure the local invertibility of $\boldsymbol{\Sigma}[\mathbf{G}]$, one simply has to restrict the space of variational parameters $\mathbf{t}'$ by excluding the one-particle energies of the decoupled bath sites, *i.e.* one has to focus on the physically relevant parameters. This implies a respective restriction of the space of $\mathbf{t}'$-representable Green's functions $\mathbf{G}(\mathbf{t}')$ and ensures the local invertibility of $\boldsymbol{\Sigma}[\mathbf{G}]$ on the restricted domain. Similarly, a restriction of the $\mathbf{t}'$ space becomes necessary to ensure the local invertibility of $\boldsymbol{\Sigma}[\mathbf{G}]$ in case of a system where the self-energy is trivial (as *e.g.* for a model of spinless fermions with nearest-neighbor Coulomb interaction in the limit of infinite spatial dimensions where the self-energy is given by the Hartree term, *cf.* Ref. [35]).

Finally, it should be mentioned that generally the functional $\boldsymbol{\Sigma}[\mathbf{G}]$ cannot be inverted *globally*. Consider, for example, the Hubbard model on the infinite-dimensional hypercubic lattice with nearest-neighbor hopping t at half-filling. Due to manifest particle-hole symmetry, a sign change of the hopping $t \mapsto -t$ leaves the (local) self-energy invariant but transforms (the non-local elements of) the Green's function $\mathbf{G}$. We conclude that $\mathbf{G}[\boldsymbol{\Sigma}]$ in general cannot be defined uniquely. Due to this non-uniqueness, and also due to non-linearity, there may be more than a single solution of equation (3). However, this does not cause any problem since for any $\boldsymbol{\Sigma}$ satisfying (3) we have:

$$
\begin{aligned}
& \mathbf{G}[\boldsymbol{\Sigma}] = (\mathbf{G}_0^{-1} - \boldsymbol{\Sigma})^{-1} \\
\Rightarrow\quad & \boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\mathbf{G}_0^{-1} - \boldsymbol{\Sigma})^{-1} \\
\Rightarrow\quad & \boldsymbol{\Sigma} = \boldsymbol{\Sigma}[\mathbf{G}] \quad\text{and}\quad \mathbf{G} = (\mathbf{G}_0^{-1} - \boldsymbol{\Sigma})^{-1}.
\end{aligned} \tag{A.3}
$$

This means that $\boldsymbol{\Sigma}$ is given by the (formal) sum of all skeleton diagrams built up by a propagator $\mathbf{G}$ which is constructed via the Dyson equation from the same $\boldsymbol{\Sigma}$ in turn. Hence, any stationary point should be regarded as a physically meaningful solution. Among different physical solutions (corresponding *e.g.* to different phases) the minimum grand potential selects the stable one.

References

1. J. Orenstein, A.J. Millis, Science **288**, 468 (2000)
2. F. Gebhard, *The Mott Metal-Insulator Transition* (Springer, Berlin, 1997)
3. *Band-ferromagnetism*, edited by K. Baberschke, M. Donath, W. Nolting (Springer, Berlin, 2001)
4. J.M. Luttinger, J.C. Ward, Phys. Rev. **118**, 1417 (1960)
5. G. Baym, L.P. Kadanoff, Phys. Rev. **124**, 287 (1961)
6. N.E. Bickers, D.J. Scalapino, S.R. White, Phys. Rev. Lett. **62**, 961 (1989)
7. G. Strinati, H.J. Mattausch, W. Hanke, Phys. Rev. Lett. **45**, 290 (1980)
8. G. Strinati, H.J. Mattausch, W. Hanke, Phys. Rev. B **25**, 2867 (1982)
9. F. Aryasetiawan, O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998)
10. P. Hohenberg, W. Kohn, Phys. Rev. **136**, 864 (1964)

11. W. Kohn, L.J. Sham, Phys. Rev. **140**, 1133 (1965)
12. K. Schönhammer, O. Gunnarsson, R.M. Noack, Phys. Rev. B **52**, 2504 (1995)
13. N.D. Mermin, Phys. Rev. **137**, A 1441 (1965)
14. E. Runge, E.K.U. Gross, Phys. Rev. Lett. **52**, 997 (1984)
15. A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
16. M. Caffarel, W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
17. Q. Si, M.J. Rozenberg, G. Kotliar, A.E. Ruckenstein, Phys. Rev. Lett. **72**, 2761 (1994)
18. A.A. Abrikosow, L.P. Gorkov, I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New Jersey, 1964)
19. R. Chitra, G. Kotliar, Phys. Rev. B **63**, 115110 (2001)
20. J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963); M.C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963); J. Kanamori, Prog. Theor. Phys. (Kyoto) **30**, 275 (1963)
21. P.W. Anderson, Phys. Rev. **124**, 41 (1961)
22. R.M. Noack, F. Gebhard, Phys. Rev. Lett. **82**, 1915 (1999)
23. M. Potthoff, W. Nolting, Phys. Rev. B **59**, 2549 (1999)
24. M. Potthoff, W. Nolting, Phys. Rev. B **60**, 7834 (1999)
25. W.F. Brinkman, T.M. Rice, Phys. Rev. B **2**, 4302 (1970)
26. M. Potthoff, Phys. Rev. B **64**, 165114 (2001)
27. R. Bulla, Phys. Rev. Lett. **83**, 136 (1999)
28. R. Bulla, M. Potthoff, Eur. Phys. J. B **13**, 257 (2000)
29. G. Moeller, Q. Si, G. Kotliar, M. Rozenberg, D.S. Fisher, Phys. Rev. Lett. **74**, 2082 (1995)
30. M.H. Hettler, A.N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, H.R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998)
31. G. Kotliar, S.Y. Savrasov, G. Pálsson, G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
32. C. Gros, R. Valenti, Phys. Rev. B **48**, 418 (1993)
33. D. Sénéchal, D. Pérez, M. Pioro-Ladrière, Phys. Rev. Lett. **84**, 522 (2000)
34. D. Sénéchal, D. Pérez, D. Plouffe, Phys. Rev. B **66**, 075129 (2002)
35. E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989)